Essays on the Economic and Political Effects of Immigration

by

Marco Tabellini

M.Sc. Economics and Social Sciences, Università Bocconi (2013)
B.Sc. Economics and Social Sciences, Università Bocconi (2011)

Submitted to the Department of Economics
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2018

Signature redacted

Author..
Department of Economics
May 15, 2018

Signature redacted

Certified by..
Daron Acemoglu
Elizabeth and James Killian Professor of Economics
Thesis Supervisor

Signature redacted

Certified by.........
....
Heidi Williams
Associate Professor of Economics
Thesis Supervisor

Signature redacted

Accepted by........
....
Ricardo J. Caballero
Ford International Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Essays on the Economic and Political Effects of Immigration

by

Marco Tabellini

Submitted to the Department of Economics
on May 15, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis consists of three chapters on the economic and political effects of in-migration.

In the first chapter, I show that political opposition to immigration can arise even when immigrants bring significant economic prosperity to receiving areas. I exploit exogenous variation in European immigration to US cities between 1910 and 1930 induced by World War I and the Immigration Acts of the 1920s, and instrument immigrants' location decision relying on pre-existing settlement patterns. Immigration increased natives' employment and occupational standing, and fostered industrial production and capital utilization. However, despite these economic benefits, it triggered hostile political reactions, such as the election of more conservative legislators, higher support for anti-immigration legislation, and lower public goods provision. Stitching the economic and the political results together, I provide evidence that natives' backlash was, at least in part, due to cultural differences between immigrants and natives, suggesting that diversity might be economically beneficial but politically hard to manage.

The second chapter asks the following question: is racial heterogeneity responsible for the distressed financial conditions of US central cities and for their limited ability to provide even basic public goods? If so, why? I study these questions exploiting the movement of more than 1.5 million African Americans from the South to the North of the United States during the first wave of the Great Migration (1915-1930). Black immigration and the induced white outmigration ("white flight") are both instrumented for using, respectively, pre-migration settlements and their interaction with MSA geographic characteristics that affect the cost of moving to the suburbs. The inflow of African Americans imposed a strong, negative fiscal externality on receiving places by lowering property values and, mechanically, reducing tax revenues. Unable or unwilling to raise tax rates, cities cut public spending, especially in education, to meet a tighter budget constraint. While the fall in tax revenues was partly offset by higher debt, this strategy may, in the long run, have proven unsustainable, contributing to the financially distressed conditions of several US central cities today.

The third chapter, coauthored with Michela Carlana, studies the effects of immigration on natives' marriage, fertility, and family formation across US cities between 1910 and 1930. Instrumenting immigrants' location decision by interacting pre-existing ethnic settlements with aggregate migration flows, we find that immigration raised marriage rates, fertility, and the propensity to leave the parental house for young native men and women. We show that these effects were driven by the large and positive impact of immigration on native men's employment and occupational standing, which increased the supply of "marriageable men". We also explore alternative mechanisms – changes in sex ratios, natives' cultural responses, and displacement effects of immigrants on female employment – and provide evidence that none of them can account for a quantitatively relevant fraction of our results.

Thesis supervisor: Daron Acemoglu
Title: Elizabeth and James Killian Professor of Economics

Thesis supervisor: Heidi Williams
Title: Associate Professor of Economics

Chapter 1. Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration

1.1 Introduction

The recent migration waves to Europe and the US have generated a heated political debate.[1] Support for right-wing, populist parties is increasing, and proposals to introduce or tighten immigration restrictions are becoming more and more common. The mounting anti-immigration rhetoric rests on two grounds – one economic and one cultural. First, immigrants are blamed for increasing labor market competition and reducing natives' employment. Recently, some prominent scholars have pushed this argument one step further, suggesting that the deteriorating quality of immigrants may slow down productivity in receiving countries (Collier, 2013; Borjas, 2016). Second, immigrants' cultural diversity is viewed as a major obstacle to their assimilation, and is often perceived as a threat to the values and the social cohesion of receiving countries (see, for instance, the discussion in Baker et al., 2015, and in Abramitzky and Boustan, 2016).

In American history, this is not the first time that immigration is such a relevant and controversial issue. Between 1850 and 1915, during the Age of Mass Migration, more than 30 million people moved from Europe to the United States (Abramitzky and Boustan, 2016), and the share of immigrants in the US population was even higher than it is today (Figure 1.1).[2] Also at that time, anti-immigration sentiments were widespread, and the introduction of immigration restrictions was advocated on both economic and cultural grounds. After 1915, World War I and the Immigration Acts (1921 and 1924) put an end to the Age of Mass Migration, and, crucially, affected migration flows from different sending regions to different degrees. Since immigrants tend to cluster along ethnic lines (Card, 2001), the differential effect of these shocks across European countries generated significant variation in the number as well as in the mix of immigrants received by US cities over time.

Leveraging this variation, I investigate the economic and political effects of immigration across US cities between 1910 and 1930, and study whether political discontent reflects or runs counter to the economic consequences of immigration. The key econometric challenge to my analysis is that cities receiving more immigrants were not randomly selected. On the one hand, immigrants may have moved to places with better employment opportunities and with more appealing tax-public spending bundles. On the other, they could have settled in

[1] See, for instance, Porter (2017).

[2] The total number of foreign-born residents is, however, higher today. Also, contemporary immigration is underestimated because of the presence of large numbers of undocumented immigrants (see the dashed line in Figure 1.1 and Borjas, 2016).

otherwise declining cities which had lower house prices.

To overcome these and similar concerns, I construct a "leave-out" version of the shift-share instrument commonly adopted in the literature (Card, 2001). The shift-share instrument rests on the empirical regularity that immigrants cluster geographically in receiving countries, and newcomers tend to settle where their ethnic community is larger, due to family ties and social networks, and not because of local economic conditions (Stuart and Taylor, 2016). Starting from this observation, I predict the number of immigrants received by US cities over time by interacting 1900 settlements with subsequent migration flows from each sending region, net of the individuals that eventually settled in a given city's metropolitan statistical area (MSA).[3]

The validity of this instrument hinges on one critical assumption: the city-specific characteristics that attracted early movers from any given ethnic group must not be affecting the evolution of local economic and political conditions in subsequent decades.[4] Below, I perform a number of checks - including testing for pre-trends and interacting year dummies with pre-migration city characteristics - to assess the validity of the instrument. I also deal with the concern that aggregate migration flows (by ethnic group) may be endogenous to local economic conditions in US cities using two alternative strategies. First, I replace actual migration flows (from each sending region) with variation solely induced by World War I and the Immigration Acts. Second, similarly to Sequeira et al. (2017), I construct a measure of predicted immigration determined uniquely by temperature and precipitation shocks in origin countries. In both cases, my findings remain qualitatively unchanged.

I find that immigration had a positive and significant effect on natives' employment as well as on their occupational standing. My estimates suggest that, for every 10 new immigrants, two more natives found a job. Since no comprehensive data on wages is available for this period, as commonly done in the literature (e.g. Abramitzky et al., 2012, 2014), I proxy for natives' income using (log) occupational scores.[5] Consistent with immigrants improving natives' occupational mobility, I find a large and positive effect of immigration on natives' occupational scores. Moreover, using data digitized from the Census of Manufactures, I show that, even in a heavily exposed sector like manufacturing, there was no significant reduction in wages.[6] These results were made possible by two mechanisms. First, immigration

[3]In my baseline specification, I consider only immigration from Europe (see appendix Table A1.1), but results are robust to extending the analysis to all other non-European countries (see appendix A1).

[4]This assumption would be violated if, for instance, immigrants in 1900 settled in a given city in anticipation of subsequent economic growth.

[5]As discussed below, occupational scores assign to an individual the median income of his job category in 1950, and can thus be used as a proxy for lifetime earnings (Abramitzky et al., 2014).

[6]Wage data, digitized from the Census of Manufactures, do not distinguish between immigrants and natives. Since new immigrants were closer substitutes for previously arrived immigrants than for natives,

increased firms' incentives to invest, generating an outward shift in labor demand. Second, because of complementarity, natives moved away from occupations that were more exposed to immigrants' competition and specialized in jobs where they had a comparative advantage and, because of discrimination, immigrants did not have access to. In appendix B1, I present a simple model of directed technical change that captures both capital adjustments and natives' occupational upgrading, and that can explain my empirical results.

However, despite these positive economic effects, immigrants triggered widespread and hostile political reactions. First, cities cut public spending and taxes in response to immigration. The reduction in tax revenues was entirely driven by declining tax rates, while the fall in public goods provision was concentrated in categories where either inter-ethnic interactions are likely to be more salient (e.g. education) or poorer immigrants would get larger implicit transfers (e.g. sewerage, garbage collection). Second, immigration reduced the pro-immigrant party's (i.e., Democrats) vote share, and was associated with the election of more conservative representatives. Third, members of the House representing cities more exposed to immigration were significantly more likely to support the National Origins Act of 1924, which put an end to the era of unrestricted immigration to the US, and governed American immigration policy until 1965.

In the last part of the paper, stitching together the economic and the political effects of immigration, I show that political discontent was increasing in the cultural distance between immigrants and natives, suggesting that backlash had, at least in part, non-economic foundations.[7] These findings are consistent with a long-standing idea in the literature that diversity can be economically beneficial because of gains from specialization and complementarity (Alesina and La Ferrara, 2005; Alesina et al., 2016), but may be politically hard to manage, resulting in lower preferences for redistribution (Dahlberg et al., 2012), more limited public spending (Alesina et al., 1999), and higher conflict (Bazzi and Gudgeon, 2016).[8]

My work is also related to three other strands of the literature. First, a growing set of studies has investigated the effects of immigration on electoral outcomes in receiving countries (Barone et al., 2016; Dustmann et al., 2016; Halla et al., 2017).[9] In addition to providing evidence from a different historical context, I complement this literature in

these results should be interpreted as a lower bound for the negative effect (if any) of immigration on natives' earnings,.

[7]I proxy for cultural diversity with religion and linguistic distance. The use of religion is motivated by the historical evidence that, at that time, nativism often resulted in anti-Semitism and anti-Catholicism - to the point that the revival of the KKK in the 1920s rested on an openly anti-Catholic and anti-Semitic rhetoric (e.g. Higham, 1955).

[8]Nekby and Pettersson-Lidbom (2017) revisit the work by Dahlberg et al. (2012), and argue that findings in the latter paper might be sensitive to the sample used and to measurement of preferences for redistribution.

[9]See also Mayda et al. (2016) for a recent review.

two ways. On the one hand, I document that political discontent over immigration can arise even when immigrants bring diffused economic prosperity to natives, suggesting that cultural considerations are likely to be as important as economic ones in shaping natives' reactions (see also Card et al., 2012, and Sniderman et al., 2004). On the other, I study the effects of immigration on key policy variables, such as tax rates and public spending – outcomes for which, as noted in Card (2009) and Borjas (2016) among others, despite the large debate on the consequences of immigration, little is known.[10]

Second, my paper is related to the vast literature on how immigration affects natives' labor market outcomes.[11] My results are in contrast with the negative effects estimated by Borjas (2003) and Dustmann et al. (2017) among others, and somewhat different from the zero effect found by several cross-city studies for the contemporary period (e.g. Card, 2001 and Card, 2005). Relying on the largest episode of immigration in American history, I show that, under certain conditions, immigrants can provide substantial economic benefits to native workers. First, in line with Clemens et al. (2017), Lafortune et al. (2016), and Lewis (2011), firms' investment and technology adoption can absorb the immigration-induced labor supply shock, and, in some cases, even increase labor demand for both high and low skilled natives. Second, consistent with Peri and Sparber (2009) and Foged and Peri (2016) for the contemporary period, because of complementarity, immigrants may benefit natives by inducing them to specialize in more skill-intensive tasks and by favoring occupational upgrading. In line with a story of complementarity, immigration did not increase, and in some cases lowered, employment for groups highly substitutable for new immigrants and with limited opportunities for skill upgrading, namely previous immigrants and African Americans.[12]

Finally, several papers have investigated the selection and the assimilation of European immigrants during the Age of Mass Migration (Abramitzky et al., 2012, 2014, 2016), as well as their impact on contemporaneous and long-run economic development (Ager and Hansen, 2017; Lafortune et al., 2016; Sequeira et al., 2017), and on the adoption of state-level compulsory schooling laws (Bandiera et al., 2017). However, to the best of my knowledge, this is the first paper that exploits the dramatic cross-city variation induced by World War I and the Immigration Acts to study the economic and political consequences of European immigration in a unified empirical framework.

[10]In chapter 2 of my dissertation, I study how the migration of southern born African Americans affected public goods provision and government finances in northern cities during the first wave of the Great Migration.

[11]See Lewis and Peri (2015) for a recent review.

[12]Put differently, also in this context, immigration created winners and losers. What is striking, however, is that political insiders (i.e. native whites) were not harmed by immigration, and yet, they strongly opposed it.

The remainder of the chapter is organized as follows. Section 1.2 describes the historical background. Section 1.3 presents the data. Section 1.4 lays out the empirical strategy, constructs the instrument for immigration, and presents first stage results. Section 1.5 investigates the effects of immigration on natives' employment and on economic activity. Section 1.6 studies how immigrants affected tax revenues, public spending, electoral outcomes, and congressmen ideology as well as their voting behavior on the 1924 National Origins Act. Section 1.7 shows that the political effects of immigration depended on the cultural distance between immigrants and natives and on immigrants' ethnic diversity. Section 1.8 concludes.

1.2 Historical Background

1.2.1 The Age of Mass Migration

Between 1850 and 1915, more than 30 million people moved from Europe to the US. Until 1890, most immigrants came from the British Isles, Germany, and Scandinavia, but, from the late 1880s, immigration from Southern and Eastern Europe increased steadily, as the costs of migration fell with the advent of steam technology (Keeling, 1999). In 1870, almost 90% of the foreign born came from Northern and Western Europe, whereas less than 5% of immigrants had arrived from Southern and Eastern Europe (Figure 1.2). By 1920, however, the situation had changed dramatically, with the share of migrant stock from new source countries being as high as 40%. Europeans from new regions were culturally farther from natives and significantly less skilled than those from old sending regions (Hatton and Williamson, 1998, 2006). For instance, while literacy rates of immigrants that entered the US between 1900 and 1910 were very close to one for all old sending countries, they were significantly lower for new source regions (appendix Figures A1.1 and A1.2).

The shift in the composition of immigrants and concerns over their assimilation induced Congress to establish a commission that, between 1907 and 1911, studied the economic and social conditions of immigrants (Higham, 1955). In 1911, the Immigration Commission recommended the introduction of immigration restrictions, and in 1917, after decades of heated political debate, Congress passed a literacy test requiring that all immigrants entering the United States had to be able to read and write (Goldin, 1994).

Even before the adoption of the literacy test, in 1914, the Age of Mass Migration came to an abrupt end due to the onset of World War I, which drastically reduced European immigration between 1915 and 1919 (Figure 1.3). In 1920, despite the literacy test, migration flows increased again to their 1910 levels, fueling nativist movements and generating even stronger political pressure to adopt more effective measures to curb immigration. Figure A1.3 plots trends of migration flows (right axis) and of the number of articles in local newspapers referring to immigration (left axis) over time, and shows that both fell dramatically during WWI, but then increased again once the war was over. In response to the growing demand for immigration restrictions, in 1921 and 1924 Congress finally passed the Immigration Acts to limit the number of immigrants that could enter the United States in a given year by introducing country-specific quotas based on 1890 immigrants' population.[13]

Both World War I and the Immigration Acts affected different sending countries in different ways. In particular, quotas were set so as to limit the inflow of immigrants from new sending regions, while favoring that from old sources such as the UK, Germany, and Scandinavia. Figure A1.4 shows the changing composition of immigrants entering the United States during the previous decade between 1900 and 1930. Until 1920, the majority of recent immigrants came from Eastern and Southern Europe, but this trend was abruptly reversed in the 1920-1930 decade, when the share of Anglo-Saxon and Scandinavian immigrants increased as a result of the Immigration Acts. Since immigrants tend to cluster along ethnic lines (Card, 2001), the post-1915 events generated substantial variation in the number as well as in the mix of immigrants received by US cities over time (Figures A1.5 and A1.6). This is the variation I exploit in my empirical analysis.

1.2.2 Immigrants and the US Economy

Historical accounts tend to view immigrants as one of the key determinants of American industrialization and economic development during the Age of Mass Migration. When describing the economic impact of European immigrants, historian Maldwyn Jones wrote that "The realization of America's vast economic potential has...been due in significant measure to the efforts of immigrants. They supplied much of the labor and technical skill needed to tap the underdeveloped resources of a virgin continent" (Jones, 1992, pp. 309-310). Similarly, John F. Kennedy argued that "every aspect of the American economy has profited from the contribution of immigrants" (Kennedy, 1964, p. 88).

During the Age of Mass Migration, the US economy had large potentials for growth. Economic historians argue that, in this context, immigrants provided a cheap and unskilled supply of labor which could not only be absorbed, but that may have even allowed industries to expand (Foerster, 1924), in turn creating new job opportunities for native workers. Even though some studies have found a negative effect of immigrants on wages (Goldin, 1994), labor shortage was a recurring theme in this historical period.[14] For instance, in a 1906

[13]With the 1924 National Origins Act, the total number of immigrants that could be admitted in a given year was capped at 150,000. In 1921, quotas were specified reflecting the 1910 composition of immigrants. However, they were rapidly changed to 1890 to limit immigration from new sending countries even further (Goldin, 1994).

[14]Due to data limitation, Goldin (1994) could not distinguish earnings of immigrants from those of natives.

article, the New York Times was reporting that "Need of labor is the universal cry. Demand in all parts of the country is greater than supply. Not enough immigrants. Statements from agents show that men are scarce in all the States".

Moreover, since immigrants, especially from Eastern and Southern Europe, were unskilled and had low levels of English proficiency, they may have benefitted natives because of complementarity and gains from diversity (Alesina and La Ferrara, 2005; Foged and Peri, 2016). Along these lines, in his 1971 *The Transformation of the American Economy*, economic historian Robert Higgs argues that "the rapid pace of industrial expansion has increased the number of skilled and supervisory positions so fast that practically all the English speaking employees have had the opportunity to rise on the scale of occupations" (Higgs, 1971, p. 420).

1.2.3 Immigration and Natives' Backlash

Despite the positive views on the contribution of immigrants to the American economy expressed by economic historians, Europeans, especially from new sending countries, faced strong political opposition. Natives' backlash culminated in the passage of the literacy test of 1917 and, more importantly, of the Immigration Acts of 1921 and 1924, which were explicitly introduced to shut down immigration from "undesirable sources". Goldin (1994) argues that concerns about unemployment and labor market competition were the main motivation for the immigration restrictions of the 1920s. Undoubtedly, the coincidence of large immigration flows with the severe macroeconomic recessions of 1907, 1913-1914, and 1919 increased the perception among native workers that immigrants were threatening American standards of living.

However, while economic considerations certainly played a role, anti-immigration sentiments tended to have deep cultural roots (Higham, 1955; Abramitzky and Boustan, 2016). This idea is very effectively summarized in a 1921 statement by Irving Fisher, who argued that "If we could leave out of account the question of race and eugenics...I should, as an economist, be inclined to the view that unrestricted immigration...is economically advantageous...the core of the problem of immigration is...one of race and eugenics" (Leonard, 2005). On a similar vein, in 1896, the first president of the American Economic Association, Francis A. Walker, claimed that the American standard of living and the quality of American citizenship had to be protected "from degradation through the tumultuous access of vast throngs of ignorant and brutalized peasantry from the countries of Eastern and Southern Europe" (Greenwood and Ward, 2015).[15]

[15] Consistent with this qualitative evidence, using data from local newspapers, D'Amico and Tabellini (2017) find that immigration not only increased the frequency of generic terms related to immigration, but

Anti-immigration sentiments were most often directed towards two groups. First, Jews and Catholics, whose values were perceived as being different from the Puritan tradition prevailing in the US at that time.[16] Second, immigrants from Eastern and Southern Europe, who were culturally and linguistically distant from natives and, because of their lower socio-economic status, were regarded as belonging to inferior races. Countless statements by politicians and newspapers articles provide examples of how Eastern and Southern European immigrants were perceived at the time. For instance, in 1916, congressman Thomas Abercrombie claimed that "The color of thousands of them [i.e. the new immigrants: Mediterraneans, Slavs, Jews] differs materially from that of the Anglo-Saxons" (Higham, 1955), while the editor of the *Saturday Evening Post*, Kenneth Roberts, in a 1920 article wrote that "if a few more million members of the Alpine, Mediterranean and Semitic races are poured among us, the result must inevitably be a hybrid race of people as worthless and futile as the good-for-nothing mongrels of Central America and Southeastern Europe".[17]

1.3 Data

My analysis relies on a balanced panel of 180 US cities for the three Census years 1910, 1920, and 1930. The sample includes all cities with at least 30,000 residents in each of the three censuses, and where at least some Europeans were living in 1900 (see Figure A1.7 and Table A1.2 for the complete list of cities).[18] To study the economic and political effects of immigration, I combine data from several sources.

Immigration and city population. Data on city population and on the number of immigrants by country of origin at the city and at the national level were taken from the decennial US Census of Population, made available by IPUMS (Ruggles et al., 2015).[19] For 1900, I use the 5% sample, while for 1910, 1920, and 1930, I rely on the full count census datasets.

Natives' labor market outcomes. Restricting the sample to native men in working age, I compiled data on employment, literacy, and occupation from the US Census of Popu-

also, induced newspapers to adopt more racist terms when referring to the foreign born.

[16] Around the time of World War I, Jews were deemed responsible for promoting the war in order to make profits out of it. For example, in 1915 Henry Ford claimed he knew "who caused the war: German-Jewish bankers" (Watts, 2009, p. 383). During the Red Scare, and in the inter-war period more generally, Jews were often blamed for being at the origin of Bolshevism and the worldwide diffusion of Communism.

[17] Again in 1896, Francis A. Walker defined immigrants from Eastern and Southern Europe "beaten men from beaten races; representing the worst failures in the struggle for existence" (Leonard, 2005).

[18] I restrict the attention to cities with at least 30,000 residents because below this population threshold data on public spending and government finances were not reported.

[19] See Table A1.1 for the list of European countries used in my work. To classify individuals based on their country of origin, I followed the classification made by IPUMS (Ruggles et al., 2015).

lation.[20] Since until 1940 wage data are not available, I proxy for natives' income using (log) occupational scores, as commonly done in the literature (e.g. Abramitzky et al., 2012 and Abramitzky et al., 2014). Occupational scores assign to an individual the median income of his job category in 1950 and, as discussed in Abramitzky et al. (2014), represent a proxy for lifetime earnings.

Economic activity. I digitized city-level data from the quinquennial Census of Manufactures between 1904 and 1929 for the following variables for the manufacturing sector: value added by manufacture, value of products, establishment size, capital utilization (proxied by horsepower), total employment, and average wages.[21] Wage data is a potentially valuable piece of information, since, as noted above, the US Census of Population did not collect income data prior to 1940. While manufacturing wages were not separately reported for immigrants and natives, they can nonetheless be used to complement results on employment, skill ratios, and natives' occupational scores.

Public spending and government finances. Data on public spending and city finances were digitized from the Financial Statistics of Cities for years 1906, 1910, 1919, and 1930.[22] These are annual reports, available from 1906 onwards for cities with population above 30,000 (until 1934) or 100,000 (from 1934 onwards). From the Financial Statistics of Cities, I obtained data on land area, total and property tax revenues, property values, property tax rates, and public spending (total and by category).

Presidential elections. Data on electoral returns (votes shares and turnout) for Presidential elections come from Clubb et al. (1990). Since these data are available only at the county level, I aggregated them up to the MSA, fixing boundaries to 1940, and performed the analysis using MSA-level immigration, matching cities to the corresponding MSA.[23] Because Presidential elections are held every four years, I computed the average between the closest two elections after each Census year. That is, for 1910 and 1930, I averaged electoral results from 1912 and 1916 and from 1932 and 1936 respectively, while for 1920, I considered 1920 and 1924. Results are unchanged when taking the average from the two closest election years, i.e. 1908 and 1912 for 1910, and 1928 and 1932 for 1930 (see appendix A1).

[20] In my analysis, I focus on the age range 15-65, but results are unchanged when selecting different age combinations. In 1920, the US Census did not report employment status, but rather only an indicator for holding any gainful occupation. For this year, I imputed values from the latter to proxy for employment. I also report results based solely on labor force participation rather than employment.

[21] I use 1909, 1919, and 1929 data to proxy for 1910, 1920, and 1930 respectively. I make use of 1904 data to test if pre-period changes in outcomes are correlated with subsequent changes in predicted immigration.

[22] Since data for 1920 is missing, I digitized the 1919 and 1921 volumes. Results are robust to using 1921 in place of 1919, but 1919 is preferable since 1921 data was not reported for several cities. Data for 1906 is used to test the validity of the empirical strategy.

[23] Matching cities to MSAs lowers the number of units from 180 to 127. However, data on Presidential elections are not available for Washington DC, further reducing the number of MSAs to 126.

Legislators' ideology. I collected data on congressmen ideology between 1910 and 1930 from Voteview, for Congresses 61, 66, and 71 respectively.[24] Following Autor et al. (2016) as well as a vast political science literature, I proxy for politicians' ideology using the first dimension of the Poole-Rosenthal DW Nominate scores, which rank congressmen on an ideological scale from liberal to conservative using voting behavior on previous roll-calls (Poole and Rosenthal, 1985; McCarty et al., 2006). To exploit local geographic variation, I restrict my attention to the House of Representatives, and use digital boundary definitions of US congressional districts from Lewis et al. (2013) to match cities to their corresponding district in any given year.

When constructing this dataset, two problems must be dealt with. First, boundaries of congressional districts vary over time. Second, a single congressional district may represent multiple cities, while the same city may belong to more than one district. To address these issues, I follow Autor et al. (2016) and conduct the analysis at the city by congressional district level. The city-to-congressional district mapping is almost identical for the 66^{th} and the 71^{st} Congress, but redistricting between the 61^{st} and the 66^{th} Congress, especially in Massachusetts and Pennsylvania, prevents the construction of a balanced panel which includes all the cities in my sample. Below, I present results both for the unbalanced panel and for the balanced panel of cities whose congressional districts were unchanged.[25]

Representatives' voting behavior. Data on voting patterns on the National Origins Act of 1924 come from Swift et al. (2000). This dataset includes the name, the district represented, the main demographic characteristics, and the voting behavior on any rollcall of each representative in all US Congress between 1789 and 1989. As for congressmen ideology, I focus on the House of Representatives and conduct the analysis at the city by congressional district, matching each representative to the corresponding city (or cities) in my sample in the 68^{th} Congress (when the National Origins Act was passed).[26]

Table 1.1 reports the summary statistics for the main variables used in my analysis. City population ranges from more than 6.9 million (New York City in 1930) to as little as 30,200 (Pasadena in 1910). There is also wide variation in the fraction of immigrants across cities and over time, which was higher in the northeastern states of New Jersey, New York, Connecticut, and Massachusetts, and lower in the US South. As already discussed in Section 1.2, immigration fell significantly between 1910 and 1930, because of both World War I and the Immigration Acts: in 1910, the fraction of immigrants over city population

[24] To assess the validity of the empirical strategy, I also compiled data for the 56^{th} Congress.

[25] The unbalanced and the balanced panels are composed respectively of 157 and 146 city to congressional district units.

[26] Whenever multiple congressmen represent the same city, I average their votes on the Immigration Act to create a unique value, which is then assigned to that city.

was, on average, 0.18, but this number fell to 0.12 in 1930. Even starker was the decline in the fraction of foreign born that entered the United States in the previous decade, which moved from an average of 0.08 in 1910 to 0.02 in 1930.

Immigration and most of the fiscal data are available for all the 540 city-year observations in my sample. However, employment outcomes were missing for Sacramento (CA) and New Bedford (MA) in 1920, whereas data from the Census of Manufactures were not reported for a handful of cities, leaving me with 538 and 525 observations respectively.[27] Finally, aggregating cities to MSAs (for Presidential elections) and to congressional districts (for legislators' ideology) reduces the number of observations to 378 and 470 respectively.

1.4 Empirical Strategy

In this section, I first introduce the baseline estimating equation (Section 1.4.1) and construct the instrument for immigration (Section 1.4.2). Next, I present an alternative specification which explicitly relies on the variation induced by WWI and the Immigration Acts (Section 1.4.3). Finally, I report first stage results (Section 1.4.4).

1.4.1 Baseline Estimating Equation

The goal of the paper is to investigate the economic and political effects of immigration across US cities between 1910 and 1930. To do so, stacking the data for the three Census years 1910, 1920, and 1930, I estimate

$$y_{cst} = \gamma_c + \delta_{st} + \beta Imm_{cst} + u_{cst} \tag{1}$$

where y_{cst} is the outcome for city c in state s in Census year t, and Imm_{cst} refers to the fraction of immigrants received by city c in the previous decade, over city population. γ_c and δ_{st} are city and state by year fixed effects, implying that β is estimated from changes in the fraction of immigrants within the same city over time, compared to other cities in the same state in a given year. Since city population could itself be an outcome of immigration, the number of immigrants is scaled by predicted (rather than actual) city population, constructed by multiplying 1900 population by average urban growth in the US, excluding that of the Census division where the city is located. Below, I also report results obtained when scaling immigration by 1900 population.[28] Standard errors are clustered at the MSA level, and MSA

[27] Data from the Census of Manufactures were not available for Superior (WI), Washington DC in 1909 and 1919, and for Flint (MI), Galveston (TX), Huntington (WV), Lexington (KY), McKeesport (PA), Pueblo (CO), Quincy (IL), and Roanoke (VA) in 1929.

[28] Appendix A1 also presents results using the fraction of immigrants over actual population.

boundaries are fixed to 1940 in order to keep geography constant.

In my baseline specification, I restrict the attention to European immigrants that entered the United States during the previous decade. I do so because, at that time, immigrants could apply for citizenship after 5 years (Shertzer, 2016). While historical accounts suggest that after 1910 immigrants' political engagement fell steadily (Kleppner, 1982), focusing on recently arrived immigrants allows me to more confidently interpret my findings on political outcomes as natives' reactions, rather than as the direct effect of immigrants' preferences. As a robustness check, however, I repeat the analysis considering immigrants' stock, and results remain very similar to those obtained in my baseline specification (see Table A1.19).

1.4.2 Instrument for Immigration

A priori, we may expect immigrants to be attracted to cities with better job opportunities, or with more appealing tax-public spending bundles. Alternatively, immigrants might settle in otherwise declining cities, where house prices are lower. In either case, OLS estimates of equation (1) will likely be biased. To deal with this endogeneity problem, I construct a modified version of the shift-share instrument (Card, 2001). The instrument predicts the number of immigrants received by US cities over time by interacting 1900 settlements of different ethnic groups with subsequent migration flows from each sending region, excluding individuals that eventually settled in a given city's MSA. Formally, Imm_{cst} in (1) is instrumented with

$$Z_{cst} = \frac{1}{\hat{P}_{cst}} \sum_{j} \alpha_{jc} O_{jt}^{-M} \tag{2}$$

where $\hat{P}_{cst}$ is predicted city population; α_{jc} is the share of individuals of ethnic group j living in city c in 1900; and O_{jt}^{-M} is the number of immigrants from country j that entered the US between t and $t-1$, net of those that eventually settled in city c's MSA.[29]

1.4.2.1 A Graphical Example

The instrument constructed in equation (2) exploits two sources of variation: first, cross-sectional variation in the share of individuals from each ethnic group living in different US cities in 1900 (α_{jc}); second, time-series variation induced by changes in the total number of immigrants from any sending region entering the United States in a given decade (O_{jt}^{-M}). Figure 1.4 presents an example for three cities (Chicago, Milwaukee, and San Francisco) and two ethnic groups (Italians and Germans) to illustrate the variation underlying the instrument.

[29] A similar "leave-out" strategy is used in Burchardi et al. (2016).

Between 1910 and 1930, Italian immigration fell monotonically, while German immigration declined between 1910 and 1920 due to WWI, but rebounded after 1920, as the quotas were quite generous with respect to Germany. Chicago (Panel A) had large Italian and German communities in 1900. In line with the aggregate flows, both the actual (straight lines) and the predicted (dotted lines) number of Italians (yellow lines) and Germans (blue lines) arriving in Chicago fell between 1910 and 1920. However, after 1920, while Italian immigration continued its decline, Chicago experienced a positive immigration shock from Germany.

Milwaukee, instead, had a relatively large German community, but almost no Italians in 1900. Thus, as shown in Panel B, variation in immigration for this city resulted from changes in German, and not Italian, immigration. Finally, while very few Germans were living in San Francisco in 1900, Italian settlements were fairly large in this city. As documented in Panel C, the actual and predicted immigration shock for San Francisco was due to the decline in Italian immigration, and only marginally to the inflow of Germans after 1920.

The instrument in (2) extends this example to many cities and many ethnic groups, but the logic behind it can be grasped by looking at the patterns in Panels A to C of Figure 1.4.

1.4.2.2 Geographic Variation in Immigrants' Settlements

The cross-sectional variation underlying the instrument in equation (2) is based on the idea that immigrants cluster geographically and their settlements are highly persistent due to social networks and family ties, and not because of local pull factors (Card, 2001; Stuart and Taylor, 2016). As documented in Sequeira et al. (2017), the gradual expansion of the railroad network during the second half of the nineteenth century is a strong predictor of the geographic distribution of immigrants in the US: places that gained access to the railroad just before an immigration boom received more immigrants in the following decade. Moreover, upon arrival, early settlers tended to locate in places that were relatively more attractive at that time. Since the timing of outmigration varied widely across European countries, depending on local political and economic conditions (Hatton and Williamson, 1998), different US regions were populated by different ethnic groups before 1900. Early settlers then acted as a catalyst for subsequent migrants from the same ethnic group (Lafortune and Tessada, 2014).

The geographic concentration of Europeans in the United States during the Age of Mass Migration is discussed, among others, by Abramitzky and Boustan (2016). For instance, Italians clustered in the north-eastern states of New York, Pennsylvania, and New Jersey, and in California, whereas Germans and Scandinavians settled mainly in the lower and in the upper Midwest respectively. Figure 1.5 visually confirms these patterns in my data by plotting the share of individuals from different European regions living in selected US cities in 1900. While almost 4% of Swedes living in the US in 1900 were settled in Minneapolis, less than 1% of them were located in north-eastern cities like Philadelphia or Boston. Conversely, while Italian communities were present in Boston, Philadelphia, and San Francisco, they were practically non-existent in Minneapolis. Even more emblematic is the example of Eastern Europeans: in 1900, more than 8% of them were living in Cleveland, while their share in the other cities displayed in Figure 1.5 was well below 1%. Figure A1.8 presents a similar example for Ohio, and shows that differences in immigrants' settlements existed also within the same state. This is important, for otherwise the instrument in (2) would not have power, since my empirical strategy exploits only within state variation in immigration.

1.4.2.3 Identifying Assumptions and Instrument Validity

The key identifying assumption behind the instrument is that cities receiving more immigrants (from each sending area) before 1900 must not be on different trajectories for the evolution of economic and political conditions in subsequent decades. Said differently, outmigration from European regions must be independent of cross-city pull factors systematically related to 1900 settlers' country of origin. For example, between 1910 and 1920, immigration to the US was higher from Italy than from Sweden. The exclusion restriction would be violated if this happened because cities that in 1900 had attracted more Italians were growing more than cities where more Swedes had moved to in 1900.

Another threat to the validity of the identifying assumption is that the characteristics of cities that attracted early immigrants might have persistent, confounding effects on migration patterns as well as on changes in the outcomes of interest. It is possible, for instance, that larger urban centers attracted more immigrants in the nineteenth century, and that these cities kept growing more also in subsequent decades, introducing a spurious correlation between, e.g. economic activity and immigration. Similarly, one may be worried that the industry mix of cities affected both the location decision of early settlers and subsequent changes in economic and political conditions. To deal with these and similar issues, I perform several robustness checks, which I describe below when presenting my main results. Appendix A1 further explores the robustness of my findings.

1.4.3 WWI and Quotas: First and Long Difference Specifications

As discussed in Section 1.2, WWI and the Immigration Acts induced large and exogenous variation both in the number and in the ethnic composition of immigrants received by the United States over time. In this section, I explicitly rely on such variation to deal with the potential concern that aggregate migration flows by country of origin, O_{jt}^{-M}, might be endogenous to city-specific pull factors - something that would invalidate the instrument constructed in equation (2). An additional advantage of exploiting WWI and the Immigration Acts is that these shocks lowered substantially the serial correlation in migration flows to US cities over time. This is desirable since, as noted by Ruist et al. (2017), one potential threat to shift-share instruments for the contemporaneous period is precisely the high persistence of migration flows.

I start by taking (stacked) first differences of equation (1). Next, I construct two separate instruments for the decadal change (1910 to 1920 and 1920 to 1930) in the number of immigrants received by a given city in the previous ten years. These instruments (ΔZW_{cs} and ΔZQ_{cs} in equations (3) and (4)) replace actual migration flows with a measure of predicted immigration from each sending region constructed exploiting directly WWI and the Immigration Acts.[30]

Formally, the 1910-1920 and the 1920-1930 changes in immigration are instrumented with, respectively,

$$\Delta ZW_{cs} = \frac{1}{\hat{P}_{cs,1920}} \sum_j \alpha_{jc} \left(1\left[Allies_j\right] \cdot O_{j,1910} - O_{j,1910}\right) \tag{3}$$

and

$$\Delta ZQ_{cs} = \frac{1}{\hat{P}_{cs,1930}} \sum_j \alpha_{jc} \left(Q_j - O_{j,1920}\right) \tag{4}$$

The term $O_{j,1910}$ (resp. $O_{j,1920}$) is the number of immigrants from country j that entered the US between 1900 and 1910 (resp. 1910 and 1920). $1\left[Allies_j\right]$ in (3) is a dummy equal to 1 if sending country j belongs to the Allies in WWI, and zero otherwise. Finally, Q_j in (4) is the sum of the yearly quota for country j specified by the Immigration Acts of 1921 and 1924.

The intuition behind equation (3) is that, if a country was not part of the Allies, its immigration was completely shut down between 1910 and 1920. If, instead, the country belonged to the Allies, there was no change in immigration from that specific country over this period. To visually depict this intuition, Figure A1.9 plots the number of immigrants that entered the United States in the previous decade (relative to 1910) from Germany (dashed blue line) and the UK (red line). While WWI reduced immigration for both countries, the drop in German immigration was twice as large (relative to 1910) as that in immigration from Great Britain.

[30]Similarly to Sequeira et al. (2017), in appendix A1, I also construct a measure of predicted immigration that only exploits temperature and precipitation shocks in origin countries.

Interacting (3) and (4) with year (i.e. 1920 and 1930) dummies, I re-estimate equation (1) in stacked first differences with 2SLS. In formulas, the second and the first stage equations become

$$FDy_{cs\tau} = \xi_{s\tau} + \beta_S FDImm_{cs\tau} + FDu_{cs\tau} \tag{5}$$

and

$$FDImm_{cs\tau} = \xi_{s\tau} + \beta_{FW}\left(\Delta ZW_{cs} \cdot \tau\right) + \beta_{FQ}\left(\Delta ZQ_{cs} \cdot \tau\right) + \varepsilon_{cs\tau} \tag{6}$$

where FD refers to the first difference for period τ, and $\xi_{s\tau}$ includes interactions between period dummies and state dummies.[31] Variables ΔZW_{cs} and ΔZQ_{cs} in (6) are the instruments constructed in (3) and (4) above, and are both interacted with a full set of year dummies (τ). While being econometrically more demanding, this strategy allows me to perform an important placebo check. Effectively, in (6) there are four instruments, but only two of them, i.e. the interactions between ΔZW_{cs} (resp. ΔZQ_{cs}) and the 1920 (resp. 1930) dummy, should be statistically significant. In Section 1.4.4 below, I explicitly test this implication, and show that, reassuringly, the WWI (resp. the quota) instrument predicts changes in immigration only when interacted with the 1920 (resp. 1930) dummy.

As a further robustness check, below, I also report results from a long differences specification:

$$\Delta y_{cs} = \gamma_s + \beta_L \Delta Imm_{cs} + \Delta u_{cs} \tag{7}$$

where Δ is the 1910-1930 change, γ_s refers to state fixed effects, and the first stage equation is given by

$$\Delta Imm_{cs} = \gamma_s + \beta_W \Delta ZW_{cs} + \beta_Q \Delta ZQ_{cs} + \Delta\varepsilon_{cs} \tag{8}$$

1.4.4 First Stage Results

Table 1.2 presents first stage results for the relationship between actual and predicted immigration, after partialling out city and state by year fixed effects. In column 1, the dependent variable is the fraction of immigrants over actual city population, and the regressor of interest is the baseline instrument constructed in equation (2). Columns 2 and 3 replicate column 1 by dividing the actual and the predicted number of immigrants by, respectively, 1900 and predicted population. In all cases, the F-stat is very high, and there is a strong and significant relationship between the fraction of immigrants and the instrument.[32]

Figure 1.6 reports the graphical analogue of column 3, plotting the relationship between

[31] Note that, now, there are two time periods, 1920-1910 and 1930-1920, and all variables refer to the change during that period.

[32] Results, not reported for brevity, are very similar to those presented in column 1 when including only city and year fixed effects.

the fraction of immigrants and the instrument, after partialling out city and state by year fixed effects. As it appears, the city of Passaic (NJ) experienced a large drop in immigration between 1910 and 1930, and one may be concerned that, for this reason, it influences the strength of the first stage. However, omitting this city barely affects the slope of the regression line (see red dashed line in Figure 1.6). Appendix A1 replicates Table 1.2, and shows that none of the results is significantly affected when excluding Passaic from the analysis (see Table A1.11 and Figure A1.19).

From column 3 onwards, Table 1.2 presents estimates for specifications where both the actual and the predicted number of immigrants are scaled by predicted city population. Column 4 shows that the estimates are barely affected when aggregating the unit of analysis from cities to MSAs. Next, columns 5 and 6 report results for the stacked first differences and for the long differences specifications, i.e. equations (6) and (8) respectively. At the bottom of the table, columns 5 and 6 also present the p-value for the test of overidentifying restrictions. Reassuringly, in both cases, not only the F-stat is well above conventional levels, but also, the null hypothesis of overidentifying restrictions cannot be rejected. Furthermore, in column 5, the interaction between year dummies and the WWI instrument is significant only for 1920, while that between year dummies and the quota instrument is significant only for 1930. Conversely, when interacting the WWI and the quota instruments with, respectively, the 1930 and the 1920 dummies, coefficients are never statistically significant and, especially for WWI, an order of magnitude smaller.[33]

Finally, in columns 7 and 8, I augment the baseline specification by including interactions between year dummies and the 1900 (log of) city and immigrants' population, and the 1904 (log of) value added by manufacture. Not surprisingly, the F-stat falls relative to column 1, but remains well above conventional levels. Similarly, even though the magnitude of coefficients becomes somewhat lower, neither their economic nor their statistical significance is affected. As discussed in appendix A1, results are also unchanged when interacting year dummies with several other 1900 city characteristics, such as skill ratios, the share of African Americans, or the employment share in manufacturing.

1.5 The Economic Effects of Immigration

In this section, I show that immigration increased natives' employment and their occupational standing, and that, even in a sector heavily exposed to immigrants' competition like

[33] While only the interaction between the 1930 dummy and the quota instrument is statistically significant, the coefficient is not statistically different from that on the interaction between the quota istrument and the 1920 dummy. One possible explanation is that the literacy test introduced in 1917 was more binding for immigrants from Southern and Eastern Europe - groups also more penalized during the 1920s because of the Immigration Acts (see Section 1.2).

manufacturing, there was no significant reduction in either employment or wages (Section 1.5.1). In Section 1.5.2, I provide evidence that this was made possible by two mechanisms: first, because of complementarity, natives specialized in occupations where they had a comparative advantage relative to immigrants; second, firms' investment and industrial expansion absorbed the supply shock brought about by immigration, and provided natives with opportunities for skill upgrading.

1.5.1 Natives' Employment

1.5.1.1 Main Results

In Table 1.3, I study the effects of immigration on employment outcomes of native men.[34] Throughout the paper, I always report the mean of the dependent variable at baseline as well as the F-stat associated with first stage results shown in Table 1.2. The dependent variable is the employment to population ratio for native males of working age in Panel A, and the log of natives' occupational scores in Panel B. OLS estimates of equation (1) are reported in column 1, while column 2 presents 2SLS results from my baseline specification, where the fraction of immigrants (over predicted population) is instrumented with the leave-out shift-share instrument constructed in equation (2).

Starting from employment, both OLS and 2SLS coefficients are positive and significant, with the latter being slightly larger than the former. The point estimate in column 2 implies that a one standard deviation increase in the fraction of immigrants (0.05) raises natives' employment probability by 1.5% relative to its 1910 mean. Said differently, for every 10 new immigrants, two more natives found a job. Panel B documents that immigration had a strong, positive effect on natives' log occupational scores. Since occupational scores measure cross-occupational changes in earnings, these findings suggest that the employment effects reported in Panel A likely came from occupational and skill upgrading.[35] Differently from Panel A, in Panel B, 2SLS estimates are an order of magnitude larger than OLS. One possible explanation for this pattern is that OLS is downward biased because immigrants tended to move to places with fewer opportunities for skill upgrading.

Subsequent columns of Table 1.3 explore the robustness of the main results presented in column 2.[36] First, to test for pre-trends, the 1900 to 1910 change in employment and in log occupational scores is regressed against the 1910 to 1930 instrumented change in

[34] In my baseline specification, I consider men in the age range 15 to 65, but results are robust to the use of different age thresholds (see also chapter 3).

[35] Very similar results (not reported for brevity) are obtained when state by year fixed effects are replaced by either region by year or year fixed effects.

[36] Additional checks are reported in appendix A1.

immigration (column 3). Reassuringly, in both cases, the coefficient on immigration is not statistically significant, very imprecisely estimated, and quantitatively different from the estimates reported in column 2. Figures A1.10 and A1.11 provide residual scatterplots for the reduced form estimates of specifications presented in columns 2 and 3 of Panel A, and visually confirm the pattern emerging from Table 1.3. Passaic (NJ) negatively influences the slope of the regression line in Figure A1.10, and the effects of immigration on natives' employment become somewhat larger and more precisely estimated when omitting this city (see dashed line in Figure A1.10 and Table A1.13).

In column 4, I document that scaling both the actual and the instrumented number of immigrants by 1900, rather than predicted, population does not alter my findings in a significant way. In addition, to (indirectly) address the potential concern that estimates in column 2 may be partly due to natives' geographic mobility (Borjas, 2016), I replicate the analysis aggregating the unit of analysis to the MSA level (column 5).[37] Reassuringly, results remain quantitatively very similar to those reported in column 2, even though the coefficient in Panel B is no longer significant.[38]

Next, columns 6 and 7 replace the baseline instrument from (2) with that constructed exploiting directly variation induced by WWI and by the Immigration Acts (i.e., equations (3) and (4) in Section 1.4.3). Column 6 reports results for the stacked first differences regression (equation (5)), and column 7 estimates the long differences specification (equation (7)). Coefficients from the long and the stacked first differences regressions bound respectively from above and from below those obtained using the standard shift-share instrument, and results always remain statistically significant and in line with those reported in column 2.

Finally, I replicate the analysis interacting year dummies with, respectively, the (log of) 1900 city and immigrants' population, and the (log of) 1904 value added by manufacture (columns 8 and 9). This exercise is performed to check if the characteristics of cities that may have attracted more immigrants before 1900 also had persistent effects on the evolution of the economic environment. In either case, results are barely affected: the effects of immigration on natives' employment and occupational scores remain statistically significant and quantitatively similar to the baseline estimates reported in column 2. As discussed in appendix A1, results are also robust to interacting year dummies with 1900 skill ratios, value of industrial production, the employment share in manufacturing, and the fraction of blacks

[37] Historical accounts suggest that, differently from what happened with the Great Migration of blacks from the South to the North of the United States (Boustan, 2010), natives did not systematically leave cities in response to European immigration. Moreover, in line with this idea, in appendix A1 I show that, if anything, immigration promoted internal in-migration (Table A1.23).

[38] The lower precision of these estimates should not be surprising, given that when aggregating observations up to the MSA level the number of units moves from 180 to 127.

(see Tables A1.16 to A1.18).

1.5.1.2 Placebo Checks and Manufacturing Wages

I present additional results for the effects of immigration on natives' employment in Table A1.4, reporting OLS and 2SLS estimates in Panels A and B respectively. Consistent with findings discussed above, immigration had a positive and significant effect both on the fraction of natives holding any gainful occupation (column 1) and on the ratio of high to low skill natives (column 2).[39] Columns 3 and 4 perform a falsification exercise and show that immigration did not lead to employment gains for either illiterate natives or African Americans, two groups for which leaving unskilled occupations, where most recent immigrants were employed, would have been extremely difficult. Also, and reassuringly, immigration had a negative, although not statistically significant, effect on employment of previously arrived immigrants, which vanished for those that had spent at least 20 years in the United States (Figure A1.21).

Similarly, the inflow of immigrants did not significantly increase employment for natives working as manufacturing laborers (Table A1.4, column 5), an occupation highly exposed to immigrants competition (Table A1.3).[40] It is worth noting, though, that even in this heavily exposed occupation, immigration did not lead to employment losses among natives, possibly because manufacturing was able to expand, in turn absorbing the immigration-induced supply shock. In line with this idea, total employment in manufacturing increased almost one for one with immigration (Table A1.4, column 6).

Even if immigration had a positive effect on natives' employment, and no negative effect even for natives working in highly exposed sectors, it is nonetheless possible that it lowered wages at least for some workers. Unfortunately, the US Census of Population did not collect income or wage data prior to 1940, and so, this issue cannot be directly addressed using census data. While occupational scores can be used to proxy for natives' income, they may not capture short-run, within occupation changes in earnings.

To partly overcome this limitation, in column 7 of Table A1.4, I estimate the effects of immigration on (log) average manufacturing wages, digitized from the Census of Manufactures. These data do not distinguish between natives and immigrants. Since new immigrants were

[39]The skill ratio in column 2 is measured as the log of natives holding skilled occupations to the log of natives holding unskilled jobs. To classify workers across skill categories, I follow Katz and Margo (2014). As for occupational scores, also for skill ratios, OLS estimates are an order of magnitude smaller than 2SLS ones.

[40]In 1910, recent immigrants were twice as likely as natives to be employed in unskilled occupations. Similarly, while around 21% of natives were working in manufacturing, almost 45% of immigrants were employed in this sector.

closer substitutes for previously arrived immigrants than for natives, and because manufacturing was one of the most exposed sectors to immigrants' competition, one can confidently interpret these results as a lower bound for the impact of immigration on natives' earnings.

The coefficient in column 7 is negative but not statistically significant, and standard errors are very large. In addition to being very noisy, the implied magnitude is also relatively small: according to the coefficient in column 7, a five percentage points (equivalent to a one standard deviation) increase in the fraction of immigrants lowers wages in manufacturing by less than 1%. Based on this evidence, one cannot conclude that, even in a heavily exposed sector, immigration lowered wages in receiving cities. This finding is somewhat in contrast with Goldin (1994), who finds that European immigration had a negative effect on earnings of workers in selected industries between 1890 and 1915. This discrepancy may result from the fact that Goldin focuses on a slightly earlier period and on a different sample of cities, or from differences in the empirical strategy.

1.5.2 Mechanisms

The positive employment effects estimated in Table 1.3 are in contrast with some results from the contemporary immigration literature such as Borjas (2003), Borjas and Katz (2007), and Dustmann et al. (2017) among others, who find a negative and significant effect of immigration on natives' labor market outcomes. My findings are also somewhat different from those of a number of contemporaneous cross-city studies that estimate a zero effect of immigration on natives' wages (e.g. Card, 2001, 2005). However, they are consistent with a recent body of the literature which documents a positive impact of immigrants on natives' wages and occupational mobility (e.g. Ottaviano and Peri, 2012; Foged and Peri, 2016). In particular, the increase in occupational scores and skill ratios is in line with Peri and Sparber (2009) and Foged and Peri (2016) for the contemporaneous period in the US and Denmark respectively.

Appendix B1 lays out a theoretical framework that builds on a standard model of directed technical change (Acemoglu, 2002) where the direct, negative effect of immigration on labor market outcomes of unskilled natives is counterbalanced, and potentially reversed, by two forces. First, firms' incentives to invest in capital increase with immigration, raising demand for both unskilled and skilled workers. Second, complementarity between immigrants and natives induces the latter to reallocate their labor from unskilled to skilled occupations, where they might have a comparative advantage. In what follows, I provide evidence consistent with both mechanisms.

1.5.2.1 Occupational Upgrading

I start by investigating the possibility that, because of complementarity, immigration fostered natives' occupational mobility. In particular, in Table 1.4, I study the effects of immigration on the fraction of natives employed in specific occupations, exploiting the granularity of full count census data. I proxy for the degree of exposure to immigrants' competition using the ratio of the probability that natives and immigrants held a given occupation in 1910, reported at the bottom of Table 1.4: values below (resp. above) 1 indicate that immigrants were over (resp. under) represented relative to natives (see also Table A1.3).

Columns 1 to 3 consider three occupations that were heavily exposed to immigrants' competition and required relatively low skills as well as language proficiency (manufacturing laborers, waiters, and blacksmiths). While the coefficient is statistically significant only in column 3, the point estimates are consistently negative, suggesting that natives responded to immigration by moving away from these occupations. In line with this interpretation, columns 4 to 6 document a significant increase in the fraction of natives employed in more skilled and less exposed occupations such as manufacturing foremen (column 4), electricians (column 5), and engineers (column 6). These findings can be effectively summarized using the words of Jewish-American economist and statistician Isaac Hourwich who, in 1912, noted that "the effect of immigration upon the occupational distribution of industrial wage earners has been the elevation of the English-speaking workmen to the status of an aristocracy of labor, while the immigrants have been employed to perform the rough work of all industries" (Meyer, 1981).

Among the occupations considered in Table 1.4, manufacturing foremen experienced the largest percent increase relative to the 1910 mean in response to immigration (Figure A1.13). This seems plausible for two reasons. First, becoming supervisors or floor managers did not require significant investment in education, and so even natives that were already in the labor force could be employed there relatively quickly. Second, as I show below, immigration promoted the expansion of manufacturing, not only allowing to absorb the supply shock, but also creating new job opportunities for natives.[41]

If immigration induced natives to specialize in more skilled occupations because of complementarity, this effect should be stronger when skill differences between immigrants and natives were larger. Classifying immigrants as linguistically close and far from natives using the measure constructed by Chiswick and Miller (2005), I indeed find that occupational upgrading occurred only when immigrants were linguistically far from English (Figure A1.20).

[41] Figure A1.14 replicates Figure A1.13 focusing on immigrants arrived in the United States more than 10 years before (rather than natives), and reassuringly shows that immigration did not favor occupational upgrading for this group.

1.5.2.2 Firms' Investment and Industrial Expansion

As noted above, for natives' employment to increase, immigration must have also stimulated economic activity, inducing firms to create new jobs. Otherwise, absent changes in labor demand, it would be hard to reconcile the labor supply shock induced by immigration with the positive employment effects estimated above. To test this idea, in Table 1.5, I investigate the impact of immigration on (the log of) value added per establishment and (the log of) establishment size in Panels A and B respectively. The structure of the table mirrors that of Table 1.3: columns 1 and 2 report results from the baseline specification for OLS and 2SLS, while columns 3 to 9 repeat the same checks performed for Table 1.3.[42]

2SLS estimates are positive, statistically significant, and economically large. Coefficients in column 2 imply that a one standard deviation increase in immigration raised industrial production and establishment size by approximately 10%.[43] Figure A1.12 presents the residual scatterplot corresponding to the reduced form estimates of Panel A (column 2), and confirms visually the strong relationship between (predicted) immigration and value added per establishment. Reassuringly, there is no correlation between pre-migration changes in economic activity and subsequent (predicted) changes in immigration (column 3), and results are robust to all checks discussed above (columns 4 to 9).[44]

Consistent with the strong industrial expansion documented in Table 1.5, I also find that immigration had a large effect on capital utilization (Table A1.5, column 4) and on firms' productivity (Table A1.5, column 5).[45] In appendix A1, I provide suggestive evidence that these effects might have been partly due to firms' increased propensity to adopt new technologies that made intensive use of electricity, e.g. the assembly line (Table A1.22). This, in turn, might have raised the demand for managers and supervisors, and for high skilled workers such as electricians (Goldin and Katz, 2009; Katz and Margo, 2014).

My findings are in line with the historical evidence reviewed in Section 1.2 and, importantly, can explain the positive employment effects brought about by immigration. First, the industrial expansion allowed the economy to absorb the large supply shock by creating new jobs for both high and low skilled workers. Second, it provided natives with opportunities for

[42]As for Table 1.3, additional robustness checks are presented in appendix A1.

[43]As in Sequeira et al. (2017), who use a very different estimation strategy, OLS estimates are somewhat lower than 2SLS. One possible reason for this pattern is that OLS is downward biased as immigrants endogenously selected places with lower growth potential because of congestion or natives' discrimination.

[44]Results in Table 1.5 are also robust to using different proxies for economic activity (Table A1.5, columns 1 to 3).

[45]Consistent with the literature, I proxy for capital utilitazion using the log of horsepower (results are robust to using the log of horsepower per capita or per establishment). To estimate the effects of immigration on productivity, I assumed a Cobb-Douglas production function with two factors of production, capital and (homogeneous) labor.

skill upgrading. For instance, when describing the internal organization of production in the booming auto industry, Stephen Meyer writes that "an ethnic division of labor prevailed that relied on assumed stereotypical traits of different ethnic groups. The most skilled positions were reserved for native-born Americans...The laborers and unskilled workers were mostly the newer immigrants from southern and eastern Europe...".[46]

Further supporting the idea that immigration brought economic prosperity to US cities in this period, in chapter 3, I document that the inflow of immigrants had a large, positive effect on marriage rates of both native women and native men, as well as on fertility and on the probability that young adults left the parental house earlier. In appendix A1, I also show that immigration spurred internal migration of native men, again suggesting that by increasing labor market opportunities for natives, immigrants made cities economically more attractive (Table A1.23).

1.6 The Political Effects of Immigration

In this section I show that, despite its large economic benefits, immigration triggered hostile political reactions. First, cities receiving more immigrants cut tax rates and public spending, especially in categories where either inter-ethnic interactions are more salient (education) or poorer immigrants would get larger implicit transfers (garbage collection, sewerage), suggesting that immigration lowered natives' demand for redistribution (Section 1.6.1). Second, the inflow of immigrants reduced support for the pro-immigration party (i.e., Democrats) and increased the Republican-Democrat vote margin (Section 1.6.2). Third, immigration was associated with the election of more conservative representatives who were, in turn, more likely to vote in favor of the 1924 National Origins Act (Section 1.6.3).

1.6.1 Tax Revenues and Public Spending

At least until the Great Depression, US cities were responsible for the provision of public goods such as education, police, and spending on welfare or on infrastructure (e.g. roads, sewerage, etc.), while the federal (or the state) government played only a marginal role (Monkkonen, 1990). Also, since federal and state transfers were very limited, cities had to independently raise funds to finance their expenditures. More than 75% of cities' resources came from local taxes, with property taxes accounting for around 90% of total tax revenues (Fisher, 1996). Even though cities could issue debt, property tax rates represented the key (fiscal) policy variable at disposal of local public officials.[47] It follows that, if immigration

[46] See http://www.autolife.umd.umich.edu/Labor/L_Overview/L_Overview2.htm.

[47] Different from today, at the time, spending or tax limits were very rare in US cities.

lowered the desired level of redistribution and natives' utility from public goods' consumption, one would expect to find larger reductions in tax revenues, and in particular in tax rates, in cities that received more immigrants.

Motivated by this discussion, in Table 1.6, I study the effects of immigration on tax rates (Panel A) and public spending per capita (Panel B). As for the key economic outcomes (Tables 1.3 and 1.5), columns 1 and 2 report results from the baseline specification for OLS and 2SLS respectively, while columns 3 to 8 repeat all the checks performed for Tables 1.3 and 1.5.[48] Immigration is associated with a significant decline in both tax rates and public spending per capita, suggesting that the inflow of immigrants lowered (natives') demand for public services. Coefficients in column 2 of Panels A and B imply that a one standard deviation increase in the fraction of immigrants (0.05) reduced property tax rates and public spending per capita by, respectively, 7.5% and 5% relative to their 1910 average.

Reassuringly, the 1906 to 1910 change in neither tax rates nor public spending is correlated with the 1910 to 1930 change in (instrumented) immigration (column 3). Moreover, coefficients in column 3 are close to zero and imprecisely estimated.[49] When performing the additional checks, in columns 4 to 8, the precision of the estimates for the tax rate deteriorates, but their magnitude remains in line with that reported in column 2. Likewise, the relationship between public spending per capita and immigration is quantitatively similar to that estimated in column 2 and always statistically significant. In column 8, which includes interactions between year dummies and the 1904 value added by manufacture, the point estimate is twice as large (in absolute value) as that in column 2. This pattern, however, is due to the slightly different sample for which industrial data were reported in 1904 (Table A1.14).

Table A1.6 documents that the inflow of immigrants reduced total and property tax revenues per capita (columns 1 and 2). Not surprisingly, since most local government revenues came from property taxes, coefficients in columns 1 and 2 are very similar to each other. 2SLS results (Panel B) are close to OLS ones (Panel A), and imply that a one standard deviation (0.05) increase in the fraction of immigrants lowered property tax revenues per capita by 5% relative to the 1910 mean. Consistent with a net reduction in tax revenues, lower tax rates were not compensated by a significant increase in either property values (columns 3

[48]Data on property tax rates was not reported for the city of Pittsfield (MA) in 1930: for this reason, the number of observations in Panel A is 539, rather than 540 as in Panel B. Relative to Tables 1.3 and 1.5, Table 1.6 does not replicate results aggregating the unit of analysis at the MSA level, since tax rates and public spending are the by-product of the political process taking place at the city level. See appendix A1 for additional robustness checks.

[49]1906 is used because this is the first year for which the Financial Statistics of Cities collected data in a way that is comparable to subsequent years. Figures A1.15 and A1.16 plot the residual scatterplots of the reduced form estimates of columns 2 and 3 (Panel A).

and 4) or in business taxes per capita (column 5).[50]

Finally, Table A1.7 breaks down total expenditures across categories, and shows that spending cuts were larger for education (column 1) and sanitation, sewerage and garbage collection (column 5) where inter-racial interactions are likely to be more salient and poorer immigrants would get larger implicit transfers. Similarly, even if the coefficient for spending on charities and hospitals (column 4) is not significant, the point estimate is quite large, relative to its baseline mean.[51]

Taken together, these findings suggest that immigration lowered (natives') demand for redistribution and induced cities to cut tax rates. This interpretation is consistent with several historical accounts (e.g. Higham, 1955; Leonard, 2016). For example, in 1907, Prescott Hall, one of the founders of an influential anti-immigration movement, the *Immigration Restriction League*, stated that America was "receiving a great many immigrants who are not only worth nothing to the country, but are a positive [public] expense". The inflow of immigrants may have reduced natives' desired level of public spending for two related reasons. First, most immigrants, at least until 1920, came from relatively poor countries, and may have thus been perceived as a fiscal burden by natives. Second, ethnic diversity brought about by immigration might have lowered preferences for redistribution among natives (Easterly and Levine, 1997; Alesina et al., 1999). In Section 1.7 and in appendix A1, I return to this issue and, exploiting variation in immigrants' backgrounds, show that higher cultural and ethnic diversity were associated with larger reductions in tax revenues and in public spending.

1.6.2 Presidential Elections

I now investigate how immigration affected electoral outcomes in receiving places. Since prior to 1951 systematic data on municipal elections do not exist (see de Benedictis-Kessner and Warshaw, 2016), I focus on Presidential elections between 1910 and 1930, using data from Clubb et al. (1990). Because electoral results are only available at the county level, I aggregate them at the MSA level, using 1940 MSA definitions.[52] In Panel A of Table 1.7, I focus on the Democrats' vote share, reporting OLS and 2SLS estimates from the baseline specification in columns 1 and 2, and additional robustness checks in columns 3 to 8.

[50] In chapter 2, I find that the migration of southern born African Americans lowered tax revenues in northern cities, but that this happened through a reduction in property values, which resulted from whites' decision of moving to the suburbs (see Boustan, 2010).

[51] An alternative interpretation for the reduction in public spending on education is that immigration increased the number of Catholic schools, in turn reducing demand for public schools. Data limitations prevent me from testing this possibility.

[52] As discussed in Section 1.3, since Presidential elections are held every four years, I computed the average between the closest two elections after each Census year. Results are unchanged when taking the average from the two closest election years (Table A1.21).

The inflow of immigrants had a negative and statistically significant effect on support for Democrats, which was also economically relevant. In particular, the 2SLS coefficient in column 2 implies that a one standard deviation increase in the fraction of immigrants reduced the Democrats' vote share by approximately 5% relative to its 1910 mean. Reassuringly, no such relationship is found between the 1900-1910 change in the Democrats' vote share and the 1910-1930 change in the instrument (column 3). Subsequent columns of Table 1.7 (Panel A) document that results are qualitatively unchanged for most robustness checks. However, the coefficient drops to zero when either using the stacked first difference specification (column 5) or interacting year dummies with 1900 city and immigrants population (column 7). Somewhat reassuringly, though, this pattern seems to be confined to these two specifications (see additional results in appendix A1), and does not emerge when considering other political outcomes (see Section 1.6.3 and Panel B of Table 1.7).

As I show in Table A1.8, the negative effect of immigration on the Democrats' vote share was accompanied by increasing support for third parties (column 2) and, to a lesser extent, for Republicans (column 1). Even if the coefficient in column 1 is not statistically significant, immigration had a very strong, negative effect on the Democrats-Republicans margin (column 3). Specifically, the estimates in column 3 (Panel B) imply that a one standard deviation increase in the fraction of immigrants reduced the Democrats-Republicans margin by approximately 12% relative to its 1910 mean - a sizeable effect.

While both Republicans and Democrats tried to win immigrants' support, between 1890 and 1940, most naturalized immigrants tended to vote for the Democratic party (Shertzer, 2016).[53] The Irish are probably the most emblematic example, but this was true also of other ethnic groups such as Italians (Luconi, 1996).[54] I examined the voting behavior of members of the House who represented the 180 cities in my sample between 1910 and 1930, finding that Democrats were significantly less likely to vote in favor of both the literacy test of 1917 and the Immigration Acts of 1921 and 1924. Even after controlling for state fixed effects, immigration, and a number of 1900 city characteristics, Democratic legislators were 20 percentage points more likely to vote against the immigration restrictions.

One possible interpretation for my results is that immigration triggered natives' political backlash, and reduced support for the pro-immigrant party, i.e. Democrats. These ideas are corroborated by historical accounts, which document that, during the Progressive Era, political reformers were often openly racists and directly involved in the eugenic society (Leonard, 2005, 2016). The policy platform of Progressives was centered on radical urban

[53] Shertzer (2016) notes that the Democratic party was particularly appealing to foreign born because of its support for ethnic parochial schools and its opposition to the prohibition of alcohol.

[54] Similarly, Kleppner (1979) estimates that more than 80% of Catholics in Iowa voted for the Democratic party by the end of the nineteenth century.

reforms aimed at dismantling the political machines, whose main supporters were precisely the foreign born (e.g. Erie, 1990; Menes, 1999). Since data on votes by ethnicity (or place of birth) are not available, these conjectures cannot be tested directly. However, they are consistent with results obtained for the contemporaneous period by Mayda et al. (2016), Dustmann et al. (2016), and Becker and Fetzer (2016) in the US, Denmark, and the UK respectively.[55]

1.6.3 Congressmen Ideology and Voting Behavior

1.6.3.1 Legislators' Ideology

In Panel B of Table 1.7, I estimate the effects of immigration on the ideology of members of the House that represented the 180 cities in my sample in each Congress corresponding to the three Census years considered in my analysis, i.e. Congress 61 (1909-1911), Congress 66 (1919-1921), and Congress 71 (1929-1931). As discussed in Section 1.3, following Autor et al. (2016), I proxy for Congress members' ideology using the first dimension of the DW Nominate scores (Poole and Rosenthal, 1985; McCarty et al., 2006), and conduct the analysis at the city by congressional district level.[56] While most of the city-congressional district combinations did not change between 1910 and 1930, redistricting between the 61^{st} and the 66^{th} Congress prevents the construction of a balanced panel including all cities in my sample. For this reason, I present results for both the unbalanced panel (Table 1.7) and the balanced panel that includes only cities not affected by redistricting between 1910 and 1920 (Table 1.8, column 2).[57]

In what follows, I focus on the 2SLS baseline specification, reported in column 2 of Table 1.7 (Panel B), but, as it appears from subsequent columns, results are robust to all the checks discussed extensively above for other variables. Immigration had a positive and significant effect on legislators' Nominate scores.[58] Quantitatively, this effect is large, and not very different from that in Autor et al. (2016) for the impact of import competition. Specifically, a one standard deviation increase in the fraction of immigrants increases Nominate scores by approximately 0.25 standard deviations.[59] Similarly, Autor et al. (2016) estimate that a

[55] See also Barone et al. (2016) for Italy, and Halla et al. (2017) for Austria.

[56] DW Nominate scores rank Congress members on an ideological scale from liberal to conservative using voting behavior on previous roll-calls, with higher (lower) values indicating a more conservative (liberal) ideology.

[57] To ease comparisons, column 1 of Table 1.8 replicates the baseline specification of Table 1.7 (Panel B).

[58] The difference between OLS (column 1) and 2SLS (column 2) estimates is consistent with immigrants endogenously choosing to locate in cities with a less hostile political environment. Column 2 of Table 1.8 confirms that results are similar when restricting the analysis to the balanced panel of cities to congressional districts.

[59] This number is obtained by multiplying the coefficient in column 2 (Panel B) by a one standard deviation

one standard deviation increase in trade exposure raises Nominate scores by 0.36 standard deviations.

Since the analysis is conducted at decennial frequency, most of the effect of immigration on legislators' Nominate scores comes from the election of new, more conservative representatives, rather than from changes in the ideology of incumbent politicians.[60] Note that the increase in Nominate scores can come from the election of either more moderate (i.e. less liberal) Democrats or more conservative (i.e. less moderate) Republicans. Moreover, since immigration had a strong impact on the Republican-Democrat vote margin (Table A1.8, column 3), the rise in Nominate scores may simply reflect a shift from moderate Democrats to moderate Republicans.

Columns 3 to 6 of Table 1.8 address these issues by studying if immigration affected the probability of electing, respectively, a liberal Democrat (column 3), a moderate Democrat (column 4), a moderate Republican (column 5), or a conservative Republican (column 6). Liberal (resp. moderate) Democrats are defined as legislators with a Nominate score below (resp. above) the median score for Democrats in the 61^{st} Congress. Likewise, a Republican legislator is classified as moderate (resp. conservative) if his Nominate score is below (resp. above) the median score for Republicans in the 61^{st} Congress. Similar results are obtained when classifying legislators relative to the four quartiles of the overall 1910 distribution of Nominate scores.

The replacement of more liberal Democrats with more moderate Democrats is not responsible for the rise in Nominate scores estimated in Table 1.7. In fact, even though the point estimate is not significant at conventional levels, there is a negative and quantitatively large effect of immigration on the probability of electing a moderate Democrat (column 4). Also, moderate Democrats are not replaced by moderate Republicans (column 5), but rather by conservative Republicans (column 6). Figure A1.17 visually displays this pattern, by plotting 2SLS coefficients reported in columns 3 to 6 of Table 1.8. Interpreting the magnitude of these results, a one standard deviation increase in immigration raises the probability of electing a conservative Republican by 12 percentage points relative to its 1910 mean. This effect is, once again, close to that estimated in Autor et al. (2016), who find that a one standard deviation increase in trade exposure increases the probability of electing a conservative Republican by 17.5 percentage points.

Results presented in Tables 1.7 and 1.8 are in line with those from a number of recent studies documenting that the waves of refugees have increased support for right-wing, pop-

increase in immigration (0.05), and dividing it through the 1910 standard deviation in the DW Nominate scores (0.372).

[60] Indeed, only in six cases, the same congressman in office in 1910 was also in office in 1930.

ulist parties as well as political polarization in Europe (Dustmann et al., 2016; Halla et al., 2017). They are also consistent with the idea advanced by McCarty et al. (2006) that immigration could be responsible for the rise in political polarization experienced by the US in the past three decades. However, politicians' ideology, measured on a liberal-conservative scale, may be only an indirect proxy for anti-immigration sentiments. For this reason, in the next section, I explicitly investigate the voting behavior of legislators on the 1924 National Origins Act, the bill that ultimately put an end to the era of unrestricted immigration to the US, and that governed American immigration policy for more than 40 years.

1.6.3.2 Legislators' Voting Behavior and the National Origins Act

The National Origins Act, approved in 1924 as part of the Johnson-Reed Act, was the last of a series of attempts undertaken by the US Congress to restrict immigration in the early twentieth century, and remained in place until 1965. While Congress approved the literacy test in 1917 and the Emergency Quota Act in 1921, it was not until the passage of the National Origins Act that the inflow of immigrants, especially from Eastern and Southern Europe, was effectively and permanently shut down. On the one hand, even though the literacy test was accompanied by a heated political debate (Goldin, 1994), by the time of its approval it was no longer binding. On the other, the Emergency Quota Act introduced only temporary measures, which were then made permanent (and more stringent) with the National Origins Act of 1924.[61] For these reasons, I focus on the 1924 Immigration Act, and not on its predecessors.

As for Section 1.6.3.1, the analysis is conducted at the city by congressional district level, and the attention is restricted to members of the House who represented the 180 cities in my sample during the Congress that approved the National Origins Act, i.e. Congress 68. Since I examine voting behavior at a specific point in time, redistricting is no longer an issue. However, precisely because of the cross-sectional nature of the analysis, results should be interpreted as suggestive. With this caveat in mind, columns 7 and 8 of Table 1.8 document a positive and significant relationship between a legislator's propensity to vote in favor of the National Origins Act and the 1910 to 1920 change in the fraction of immigrants received by the city (or cities) he represented. Column 7 only includes state fixed effects, while column 8 also controls for a number of 1900 characteristics, such as the fraction of Europeans and

[61] The 1921 Emergency Quota Act temporarily limited the number of immigrants from any given country that could enter the United States to 3% of the 1910 population of each ethnic group. With the 1924 National Origins Act, which made the 1921 Immigration Act permanent, the ceiling was lowered to 2% and the "base" year was moved to 1890. These two changes were undertaken to shut down the inflow of immigrants from "undesired" sources, such as Eastern and Southern Europe. As the *Saturday Evening Post* put it, "if there is one thing we need more than another it is a little discrimination in our immigration policy" (Spiro, 2009).

of African Americans, as well as congressmen party of affiliation. Even if the magnitude of the coefficient in column 8 is somewhat lower, the association between immigration and representatives' voting behavior remains positive and significant.

To indirectly gauge the size and the direction of the potential bias of results in columns 7 and 8 due to the impossibility of including city (and state by year) fixed effects, Table A1.15 replicates findings in columns 1 to 6 of Table 1.8 using cross-sectional regressions. To mirror as closely as possible the specification reported in columns 7 and 8, in Table A1.15, the 1920 DW Nominate score is regressed on the (instrumented) 1910 to 1920 change in immigration and on state fixed effects. Reassuringly, results remain statistically significant and quantitatively close to those reported in the main text.

As in all other columns of Table 1.8, in columns 7 and 8, OLS estimates are lower (in absolute value) than 2SLS, consistent with immigrants endogenously selecting cities with a more friendly political environment. To interpret the magnitude of these results, the coefficient in column 8 implies that, when comparing cities at the 25^{th} and 75^{th} percentiles of immigration, legislators representing the more exposed city were more likely to vote in favor of the National Origins Act by approximately 10 percentage points. While this is a large effect, it does not seem unreasonable, given that immigration was at that time (as it is today) at the forefront of the political debate. Moreover, these findings are quantitatively in line with those in Mian et al. (2010), who show that a one standard deviation increase in the mortgage default rate during the 2007 Great Recession increases legislators' propensity to support the American Housing Rescue and Foreclosure Prevention Act by 12.6 percentage points.

1.6.4 Interpretation of Results

Taken together, results in Sections 1.6.1 to 1.6.3 suggest that immigration triggered widespread political reactions. First, immigration reduced tax rates and public spending, possibly by lowering natives' demand for redistribution. Second, the inflow of immigrants was associated with a fall in the Democrats' vote share and an increase in the Republican-Democrat vote margin. Third, and most importantly, cities receiving more immigrants elected more conservative members of the House of Representatives who were in turn more likely to vote in favor of the 1924 National Origins Act.

While the evidence in Section 1.6.1 is consistent with the idea that immigration triggered natives' backlash reactions and lowered their demand for public goods provision, there exist a few alternative interpretations. First, at that time, after five years immigrants could apply for citizenship, becoming eligible to vote (Shertzer, 2016). If immigrants had different preferences relative to natives, changes in public spending and in tax rates may have resulted from the direct effect of immigrants' preferences rather than from natives' reactions. This idea, however, is inconsistent with electoral results presented in Sections 1.6.2 and 1.6.3, and with the historical literature documenting that, after 1910, the political involvement of foreign born fell steadily.[62] Moreover, it seems somewhat implausible that poorer immigrants, who would have benefitted from higher redistribution, voted in favor of lower tax rates and public spending.

A second interpretation is that immigration altered the income distribution in receiving cities and, for reasons unrelated to political backlash, shifted natives' preferences towards a lower tax rate-public spending bundle. In particular, since immigrants fostered economic activity and increased natives' occupational standing, it is possible that the (native) median voter became richer, in turn voting to cut taxes and limit redistribution (e.g. Meltzer and Richard, 1981). Lack of systematic income or wage data, unfortunately, prevents me from testing this interpretation in detail.

However, two pieces of evidence provided in my work suggest that this mechanism alone cannot explain the negative effects of immigration on public spending and tax rates estimated in Table 1.6. First, legislators representing cities that received more immigrants were more likely to support the passage of the immigration restrictions (Section 1.6.3). Second, as I show below, immigration had heterogeneous effects on taxes and spending, which depended on immigrants' religious affiliation, and more generally, on the cultural distance between immigrants and natives (Section 1.7).

Yet another possibility is that immigration increased demand for housing and thus rents. While this might have benefitted homeowners, it might have nonetheless increased the cost of living for natives who did not own a house.[63] Two pieces of evidence seem to weigh against this idea. First, as discussed in Section 1.6.1, immigration did not have any significant effect on property values. Second, to more directly investigate the possibility that higher rents fueled natives' discontent, in appendix A1, I also check that immigration was not correlated with rents paid by natives (Figure A1.22). One possible explanation for this pattern is that immigrants represented a production amenity, but were perceived by natives as a consumption disamenity.[64]

Finally, it is possible that, even if immigration had aggregate positive economic effects, some natives were made worse off, at least in the short run (e.g. Goldin, 1994). Again, lack of detailed earnings data does not allow me to completely rule out this interpretation, but the fact that neither wages nor employment in manufacturing were significantly affected

[62] For this reason, Kleppner (1982) refers to this historical period as the "Demobilization Era".

[63] In 1910, only 40% of natives living in the cities in my sample were homeowners.

[64] This idea is consistent with findings of several papers for both Europe and the US today (e.g. Card et al., 2012; Saiz and Wachter, 2011).

(Table A1.4, columns 5 and 7) seems to weigh against it. As noted above, manufacturing was the sector most exposed to immigrants' labor market competition, and wage data digitized from the Census of Manufactures do not distinguish between immigrants and natives. Thus, if immigration had any negative effects on natives' income or employment, this is precisely where one would expect to find them. Moreover, as argued below, the heterogeneous effects of immigration, which depended on the cultural distance between immigrants and natives, suggest that natives' responses were not driven only by economic, but also by cultural considerations.

1.7 Backlash, Cultural Distance, and Ethnic Diversity

If immigration was economically beneficial and did not reduce employment even for natives in highly exposed occupations, why did backlash emerge? In this section, exploiting variation in the "mix" of immigrants received by US cities over time, I show that cultural differences between immigrants and natives were responsible, at least in part, for natives' anti-immigration reactions.

1.7.1 Cultural Distance: Religious Affiliation

The historical evidence reviewed in Section 1.2.3 suggests that opposition to immigration during the Age of Mass Migration tended to have deep cultural roots. Anti-immigration sentiments were often directed towards Jews and Catholics, whose values were perceived as a threat to the Puritan tradition prevailing in the US at that time (Higham, 1955; Spiro, 2009). One of the best examples for the strength of these sentiments is the revival of the Ku Klux Klan in the 1920s, which openly embraced an anti-Catholic and anti-Semitic ideology. Similarly, immigrants from non Anglo-Saxon and non English-speaking countries were the main target of the anti-immigration rhetoric at that time (Abramitzky et al., 2016; Leonard, 2016).

Motivated by these observations, I proxy for cultural distance between natives and immigrants using, respectively, religion and linguistic distance from English. Starting from religion, I estimate

$$y_{cst} = \gamma_c + \delta_{st} + \beta_1 Imm_{cst}^{Non-Prot} + \beta_2 Imm_{cst}^{Prot} + u_{cst} \quad (9)$$

where $Imm_{cst}^{Non-Prot}$ (resp. Imm_{cst}^{Prot}) is the fraction of Jews or Catholics (resp. Protestant) immigrants. In practice, equation (9) is estimated using two separate instruments, one for each religious group, constructed by summing predicted immigration from each sending region (see (2) in Section 1.4.2) across non-Protestant and Protestant countries respectively.

Results are reported in Table 1.9, for both OLS (Panel A) and 2SLS (Panel B). Immigration had a negative and significant effect on taxes and spending only when immigrants came from non-Protestant countries (columns 1 to 4), whereas the coefficient on Protestant immigrants is quantitatively very small (or even positive, as in columns 1 and 2) and never statistically significant.[65] Turning to electoral outcomes, even though both non-Protestant and Protestant immigrants seem to reduce the Democrat-Republican vote margin, results are statistically significant only for the former (column 5).

To more directly investigate the rise of anti-Catholic sentiments, in column 6, I study if the 1910-1930 (instrumented) change in Catholic and Protestant immigration had an effect on the percent of votes received by Alfred Smith in 1928 Presidential elections.[66] Smith was the first Roman Catholic to run for presidency for the Democratic party, and historical accounts consider his religious affiliation one of the main reasons for his defeat (Slayton, 2001). Since results in column 6 are obtained from cross-sectional regressions, they should be interpreted with some caution. However, the strong, negative association between Catholic (but not Protestant) immigrants and the percent of votes received by Smith is consistent with the idea that immigration triggered natives' backlash in receiving areas.

Finally, column 7 indicates that the increase in legislators' ideology documented in Table 1.7 was entirely due to non-Protestant immigration, while the effect of Protestant immigrants is an order of magnitude smaller and very imprecisely estimated. Likewise, legislators' propensity to support the 1924 National Origins Act is strongly correlated with the 1910-1920 change in Catholic and Jewish immigration (column 8). Conversely, there is a negative, albeit not significant, correlation between the 1910-1920 change in the fraction of Protestant immigrants and the probability of voting in favor of the immigration restrictions.[67]

[65]Since the effects of Protestant immigrants are very imprecisely estimated, and because the AP F-stat is substantially larger for Catholic and Jewish immigration, one may be concerned that results in columns 1 to 4 are mechanically due to the fact that the latter groups are driving most of the variation in immigration between 1910 and 1930. To check that this was not the case, I re-estimated (9) replacing political outcomes with employment. Differently from Table 1.9, results for both non-Protestant and Protestant immigrants were both positive, statistically significant, and very similar in magnitude: a one standard deviation change in Protestant (resp. non-Protestant) immigration increased natives' employment probability by 0.8 (resp. 1.0) percentage points. When dropping the city of Passaic (NJ), the point estimates were exactly the same.

[66]As for other electoral outcomes, county-level returns were aggregated to the MSA level. Differently from other electoral data, however, the number of votes for specific candidates at the county level were taken from Dave Leip's Atlas of US Presidential elections.

[67]These findings are in line with results in D'Amico and Tabellini (2017), who document that only Catholic and Jewish, but not Protestant, immigration increased the frequency of racist terms in local newspapers.

1.7.2 Linguistic Distance and Ethnic Diversity

As an alternative proxy for cultural differences between immigrants and natives, I rely on the measure of linguistic distance constructed by Chiswick and Miller (2005) briefly discussed in Section 1.5.2.1.[68] First, I compute the weighted average of immigrants' linguistic distance from English, $LD_{cst} = \sum_j \left(sh^j_{cst} \cdot L^j\right)$, where sh^j_{cst} is the share of ethnic group j among the foreign born population of city c in Census year t, and L^j is the linguistic distance between country j and English. Then, I re-estimate (1) using as main regressor of interest LD_{cst}, always controlling for the (instrumented) fraction of immigrants and instrumenting the actual shares (sh^j_{cst}) with the same logic of the instrument in (2).[69] To ease the interpretation of results, presented in Table 1.10, I standardize LD_{cst} by subtracting its mean and dividing it by its standard deviation.

Consistent with the qualitative evidence discussed in Section 1.2.3, higher linguistic distance is associated with larger reductions in taxes and public spending (columns 1 to 4). Moreover, and similarly to Table A1.7, the fall in spending is concentrated in education and, even though the point estimate is not statistically significant, in categories where inter-ethnic interactions are likely to be more salient (columns 5, 7, and 8). These results seem to strongly reject the idea that natives' reactions were primarily driven by economic considerations, since it was precisely immigration from linguistically far countries that boosted natives' occupational standing (see Section 1.5.2.1 and Figure A1.20).

Findings in Table 1.10 are also robust to simultaneously including a (standardized) index of average literacy among immigrants, thus reducing concerns that they might be capturing not only cultural, but also economic attributes of the foreign born (Table A1.9).[70] Not surprisingly, since there are now three different endogenous regressors and three instruments, the precision of the estimates deteriorates relative to Table 1.10. Nonetheless, only linguistic distance has a significant effect on taxes and public spending. Moreover, except for columns 7 and 8, the coefficient on linguistic distance is an order of magnitude larger (in absolute value) than that on literacy.

Differently from what one may expect, the correlation between the fraction of non-

[68] Chiswick and Miller (2005)'s measure is an increasing function of how difficult it is for English (native) speakers to learn foreign languages. See also Bleakley and Chin (2010) for a study on the effect of English proficieny on immigrants' assimilation in more recent times.

[69] The estimated effect of immigration is not reported to save space. However, I always report the AP F-stat associated with its first stage.

[70] The literacy index was constructed as $LIT_{cst} = \sum_j \left(sh^j_{cst} \cdot Lit^j_t\right)$, where Lit^j_t is the average literacy rate of males in working age from ethnic group j who entered the US in the previous decade. To ease the interpretation of results, I multiplied LIT_{cst} by -1, so that higher values of this index can be interpreted as lower average literacy among immigrants, and can be directly compared to LD_{cst}. The correlation between LD_{cst} and LIT_{cst} is relatively low, with a value of 0.26.

Protestant immigrants and the index of linguistic distance is as low as 0.05, suggesting that findings for linguistic diversity are unlikely to merely replicate those for religious affiliation. To more directly investigate the relationship between religion and linguistic distance, in appendix A1, I replicate Table 1.9 including simultaneously both measures to run a horse-race between the two (Table A1.25).

In appendix A1, I provide additional evidence that natives' backlash was at least in part driven by non-economic concerns by showing that the (negative) effect of immigration on redistribution was larger when ethnic diversity among foreign born was higher (Table A1.26). These findings are consistent with the large literature showing that ethnic diversity is associated with lower public goods provision and with more limited redistribution (e.g. Alesina et al., 1999; Beach and Jones, 2017; Luttmer, 2001).

1.8 Conclusions

Today, immigration is at the forefront of the political debate, and immigrants are increasingly opposed on both economic and cultural grounds. In this paper, I exploit variation in the number of immigrants received by US cities between 1910 and 1930 to study the political and economic consequences of immigration. Using a leave-out version of the shift-share instrument (Card, 2001), I show that immigration had a positive and significant effect on natives' employment and occupational standing, as well as on economic activity. However, despite these economic benefits, the inflow of immigrants also generated hostile political reactions, inducing cities to cut tax rates and limit redistribution, leading to the election of more conservative legislators, and increasing support for the introduction of immigration restrictions.

Exploiting variation in immigrants' background, I document that natives' backlash was increasing in the cultural distance between immigrants and natives. These findings suggest that opposition to immigration may arise not only because of economic, but also because of cultural considerations. Moreover, they highlight the existence of a potential trade-off. Immigrants may bring larger economic gains when they are more different from natives. However, higher distance between immigrants and natives may trigger stronger political backlash. Ultimately, by retarding immigrants' assimilation, and favoring the rise of populism and the adoption of inefficient policies, natives' reactions may be economically and socially costly in the medium to long run.

Findings in this paper provide motivation for future work in several directions. First, one key question not addressed here is how the effects of immigration are mediated by the economic, political, and social environment in receiving places. To deal with the recent inflows of refugees, many European countries started to implement allocation policies, and answering this question would thus have first-order policy implications. Second, in light of the contrasting economic and political effects documented in my work, it would be interesting to investigate the intergenerational mobility consequences of immigration. On the one hand, immigration can increase natives' occupational mobility by pushing them up along the occupational ladder. On the other, by inducing receiving places to limit redistribution, immigration may widen inequality not only between natives and immigrants, but also within natives.

References

Abramitzky, R. and L. P. Boustan (2016). Immigration in american economic history. *Journal of Economic Literature Forthcoming.*

Abramitzky, R., L. P. Boustan, and K. Eriksson (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review 102*(5), 1832–56.

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

Abramitzky, R., L. P. Boustan, and K. Eriksson (2016). Cultural assimilation during the age of mass migration. Working Paper 22381, National Bureau of Economic Research.

Acemoglu, D. (2002). Directed technical change. *The Review of Economic Studies 69*(4), 781–809.

Ager, P. and C. W. Hansen (2017). Closing heaven's door: Evidence from the 1920s u.s. immigration quota acts. Working Paper.

Alesina, A., R. Baqir, and W. Easterly (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics 114*(4), 1243–1284.

Alesina, A., J. Harnoss, and H. Rapoport (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth 21*(2), 101–138.

Alesina, A. and E. La Ferrara (2005). Ethnic diversity and economic performance. *Journal of economic literature 43*(3), 762–800.

Autor, D., D. Dorn, G. Hanson, and K. Majlesi (2016). Importing political polarization? the electoral consequences of rising trade exposure. *NBER Working Paper 22637.*

Baker, S., N. Bloom, and S. Davis (2015). Immigration fears and policy uncertainty. *VoxEU, 12/22/2015.*

Bandiera, O., M. Mohnen, I. Rasul, and M. Viarengo (2017). Nation-building through compulsory schooling during the age of mass migration. Working Paper.

Barone, G., A. D'Ignazio, G. de Blasio, and P. Naticchioni (2016). Mr. rossi, mr. hu and politics. the role of immigration in shaping natives' voting behavior. *Journal of Public Economics 136*, 1–13.

Bazzi, S. and M. Gudgeon (2016). Local government proliferation, diversity, and conflict. Technical report, Boston University-Department of Economics.

Beach, B. and D. B. Jones (2017). Gridlock: Ethnic diversity in government and the provision of public goods. *American Economic Journal: Economic Policy 9*(1), 112–36.

Becker, S. O. and T. Fetzer (2016). Does migration cause extreme voting. Technical report, Competitive Advantage in the Global Economy (CAGE).

Bleakley, H. and A. Chin (2010). Age at arrival, english proficiency, and social assimilation among us immigrants. *American Economic Journal: Applied Economics 2*(1), 165–92.

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

Borjas, G. J. (2016). *We Wanted Workers: Unraveling the Immigration Narrative.* WW Norton & Company.

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

Boustan, L. P. (2010). Was postwar suburbanization" white flight"? evidence from the black migration. *The Quarterly Journal of Economics 125*, 417–443.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2016). Migrants, ancestors, and investments. Technical report, National Bureau of Economic Research.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics 19*(1), 22–64.

Card, D. (2005). Is the new immigration really so bad? *The Economic Journal 115*(507).

Card, D. (2009). Immigration and inequality. *American Economic Review 99*(2), 1–21.

Card, D., C. Dustmann, and I. Preston (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association 10*(1), 78–119.

Chiswick, B. R. and P. W. Miller (2005). Linguistic distance: A quantitative measure of the distance between english and other languages. *Journal of Multilingual and Multicultural Development 26*(1), 1–11.

Clemens, M. A., E. G. Lewis, and H. M. Postel (2017). Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion. Technical report, National Bureau of Economic Research.

Clubb, J. M., W. H. Flanigan, and N. H. Zingale (1990). *Partisan realignment: Voters, parties, and government in American history.*

Collier, P. (2013). *Exodus: How migration is changing our world.* Oxford University Press.

Dahlberg, M., K. Edmark, and H. Lundqvist (2012). Ethnic diversity and preferences for redistribution. *Journal of Political Economy 120*(1), 41–76.

D'Amico, L. and M. Tabellini (2017). Measuring attitudes towards immigration using local newspapers' data and congressional speeches. Working Paper.

de Benedictis-Kessner, J. and C. Warshaw (2016). Mayoral partisanship and municipal fiscal policy. *The Journal of Politics 78*(4), 1124–1138.

Dustmann, C., U. Schonberg, and J. Stuhler (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics 132*(1), 435.

Dustmann, C., K. Vasiljeva, and A. P. Damm (2016). Refugee migration and electoral outcomes. *CReAM DP 19*, 16.

Easterly, W. and R. Levine (1997). Africa's growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics 112*(4), 1203–1250.

Erie, S. P. (1990). *Rainbow's end: Irish-Americans and the dilemmas of urban machine politics, 1840-1985*, Volume 15. Univ of California Press.

Fisher, G. W. (1996). *The worst tax?: A history of the property tax in America.* University Press of Kansas.

Foerster, R. F. (1924). *The Italian emigration of our times*, Volume 20. Harvard University Press.

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

Goldin, C. (1994). The political economy of immigration restriction in the united states, 1890 to 1921. In *The regulated economy: A historical approach to political economy*, pp. 223–258. University of Chicago Press.

Goldin, C. D. and L. F. Katz (2009). *The race between education and technology.* Harvard University Press.

Greenwood, M. J. and Z. Ward (2015). Immigration quotas, world war i, and emigrant flows from the united states in the early 20th century. *Explorations in Economic History 55*, 76–96.

Halla, M., A. F. Wagner, and J. Zweimüller (2017). Immigration and voting for the far right. *Journal of the European Economic Association*, Forthcoming.

Hatton, T. J. and J. G. Williamson (1998). *The age of mass migration: Causes and economic impact.* Oxford University Press on Demand.

Hatton, T. J. and J. G. Williamson (2006). International migration in the long run: Positive selection, negative selection, and policy. In *Labor Mobility and the World Economy*, pp. 1–31. Springer.

Higgs, R. (1971). *The transformation of the American economy, 1865-1914: An essay in interpretation.* Wiley New York.

Higham, J. (1955). *Strangers in the land: Patterns of American nativism, 1860-1925.* Rutgers University Press.

Jones, M. A. (1992). *American immigration.* University of Chicago Press.

Katz, L. F. and R. A. Margo (2014). Technical change and the relative demand for skilled labor: The united states in historical perspective. In *Human capital in history: The American record*, pp. 15–57. University of Chicago Press.

Keeling, D. (1999). The transportation revolution and transatlantic migration, 1850-1914. *Research in Economic History 19*, 39–74.

Kennedy, P. J. F. (1964). *A nation of immigrants.* New York: Harper & Row Publishers.

Kleppner, P. (1979). *The Third Electoral System 1853-1892: Parties, Voters, and Political Cultures*, Volume 2. University of North Carolina Press.

Kleppner, P. (1982). *Who voted?: The dynamics of electoral turnout, 1870-1980.* Praeger Publishers.

Lafortune, J. and J. Tessada (2014). Smooth (er) landing? the role of networks in the location and occupational choice of immigrants. Technical report, Mimeo, Instituto de Economía, Pontificia Universidad Católica de Chile September 2013.

Lafortune, J., J. Tessada, and E. Lewis (2016). People and machines: A look at the evolving relationship between capital and skill in manufacturing 1860-1930 using immigration shocks. Technical report, National Bureau of Economic Research.

Leonard, T. C. (2005). Retrospectives: Eugenics and economics in the progressive era. *Journal of Economic Perspectives 19*(4), 207–224.

Leonard, T. C. (2016). *Illiberal reformers: race, eugenics, and American economics in the Progressive Era.* Princeton University Press.

Lewis, E. (2011). Immigration, skill mix, and capital skill complementarity. *The Quarterly Journal of Economics 126*(2), 1029–1069.

Lewis, E. and G. Peri (2015). Immigration and the economy of cities and regions. *Handbook of Regional and Urban Economics 5*, 625–685.

Lewis, J. B., B. DeVine, L. Pitcher, and K. C. Martis (2013). Digital boundary definitions of united states congressional districts, 1789-2012. *Data file and code book.*

Luconi, S. (1996). Machine politics and the consolidation of the roosevelt majority: The case of italian americans in pittsburgh and philadelphia. *Journal of American Ethnic History*, 32–59.

Luttmer, E. F. (2001). Group loyalty and the taste for redistribution. *Journal of political Economy 109*(3), 500–528.

Mayda, A. M., G. Peri, and W. Steingress (2016). Immigration to the u.s.: A problem for the republicans or the democrats? Working Paper 21941, National Bureau of Economic Research.

McCarty, N., K. T. Poole, and H. Rosenthal (2006). *Polarized America: The dance of ideology and unequal riches.* MIT Press.

Meltzer, A. H. and S. F. Richard (1981). A rational theory of the size of government. *Journal of political Economy 89*(5), 914–927.

Menes, R. (1999). The effect of patronage politics on city government in american cities, 1900-1910. Working Paper 6975, National Bureau of Economic Research.

Meyer, S. (1981). *The five dollar day: Labor management and social control in the Ford Motor Company, 1908-1921.* State University of New York Press.

Mian, A., A. Sufi, and F. Trebbi (2010). The political economy of the us mortgage default crisis. *American Economic Review 100*(5), 1967–98.

Monkkonen, E. H. (1990). *America becomes urban: the development of US cities & towns, 1780-1980.* Univ of California Press.

Nekby, L. and P. Pettersson-Lidbom (2017). Revisiting the relationship between ethnic diversity and preferences for redistribution: Comment. *The Scandinavian Journal of Economics 119*(2), 268–287.

Ottaviano, G. I. and G. Peri (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association 10*(1), 152–197.

Peri, G. and C. Sparber (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics 1*(3), 135–69.

Poole, K. T. and H. Rosenthal (1985). A spatial model for legislative roll call analysis. *American Journal of Political Science*, 357–384.

Porter, E. (2017). Can immigration hurt the economy? an old prejudice returns. The New York Times, February 14, 2017.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015). Integrated public use microdata series: Version 6.0 [dataset]. *Minneapolis: University of Minnesota.*

Ruist, J., J. Stuhler, and D. A. Jaeger (2017). Shift-share instruments and the impact of immigration.

Saiz, A. and S. Wachter (2011). Immigration and the neighborhood. *American Economic Journal: Economic Policy 3*(2), 169–188.

Sequeira, S., N. Nunn, and N. Qian (2017). Migrants and the making of america: The short- and long-run effects of immigration during the age of mass migration. Technical report, National Bureau of Economic Research.

Shertzer, A. (2016). Immigrant group size and political mobilization: Evidence from european migration to the united states. *Journal of Public Economics 139*, 1–12.

Slayton, R. A. (2001). *Empire statesman: The rise and redemption of Al Smith.* Simon and Schuster.

Sniderman, P. M., L. Hagendoorn, and M. Prior (2004). Predisposing factors and situational triggers: Exclusionary reactions to immigrant minorities. *American Political Science Review 98*(01), 35–49.

Spiro, J. (2009). *Defending the master race: conservation, eugenics, and the legacy of Madison Grant.* UPNE.

Stuart, B. A. and E. J. Taylor (2016). Social interactions and location decisions: Evidence from us mass migration. *Unpublished working paper*.

Swift, E. K., R. G. Brookshire, D. T. Canon, E. C. Fink, J. R. Hibbing, B. D. Humes, M. J. Malbin, and K. C. Martis (2000). Database of congressional historical statistics [computer file]. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research.*

Watts, S. (2009). *The people's tycoon: Henry Ford and the American century.* Vintage.

Table 1.1. Summary Statistics

VARIABLES	*Mean*	*Median*	*St. Dev.*	*Min*	*Max*	*Obs.*
			Panel A. City Demographics			
Fr. all immigrants	0.152	0.149	0.097	0.007	0.518	540
Fr. recent immigrants	0.042	0.026	0.044	0.001	0.343	540
Recent immigrants over 1900 population	0.074	0.048	0.078	0.002	0.678	540
City population (1,000s)	190.1	76.05	510.4	30.20	6,930	540
			Panel B. Economic Outcomes			
Employed	0.858	0.889	0.071	0.648	0.952	538
Log occupational scores	3.263	3.265	0.047	3.080	3.427	538
Value added per establishment	87.66	65.92	74.47	7.945	556.3	525
Establishment size	52.86	43.09	37.98	5.465	229.9	525
			Panel C. Political Outcomes			
Tax rate per 1,000$ of assessed valuation	29.42	25.78	16.48	6.450	114.3	539
Expenditures per capita	14.57	12.89	7.336	3.443	49.99	540
Democrats' vote share	0.482	0.465	0.189	0.103	0.967	378
DW Nominate Score	0.178	0.334	0.338	-0.578	0.991	470

Note: the sample includes a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. *Employed* is the employment to population ratio for native men in the age range (15-65). *Fr. all immigrants* (resp. *Fr. recent immigrants*) is the total number of European immigrants (resp. the number of European immigrants arrived in the last 10 years) divided by city population.

Table 1.2. First Stage

	Dep. Variable: *Fraction of Immigrants*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Z	0.840*** (0.056)	0.968*** (0.064)	0.999*** (0.059)	0.948*** (0.104)			0.893*** (0.091)	0.900*** (0.081)
ZW*1920					0.774*** (0.106)	0.838*** (0.067)		
ZQ*1930					0.771** (0.349)	1.236*** (0.188)		
ZW*1930					0.064 (0.082)			
ZQ*1920					0.464 (0.423)			
1900 population		X						
Predicted population			X					
MSA analysis				X				
WWI-Quotas IV					First Diff.	Long Diff.		
Year by 1900 Log							City and imm pop	Value added manuf.
F-stat	225.1	226.7	288.3	82.65	106.8	207.4	96.48	124.8
P-value Overid. Test					0.456	0.432		
Cities	180	180	180	127	180	180	180	176
Observations	540	180	540	379	360	180	540	528

Note: the sample includes a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. In Col 1 the actual number of immigrants is scaled by actual population, and the instrument is the leave-out version of the shift-share IV in equation (2) (Section 4.2). Cols 2 and 3 replicate Col 1 by scaling the actual and predicted number of immigrants by, respectively, 1900 and predicted population. From Col 3 onwards, Table 2 presents results from specifications where both the predicted and the actual number of immigrants are scaled by predicted population. Col 4 replicates the analysis aggregating the unit of analysis at the MSA level. Cols 5 and 6 estimate stacked first differences equation (6) and long differences equation (8) replacing the standard shift-share instrument with those constructed exploiting World War I and the quotas (equations (3) and (4) in Section 4.3). Cols 7 and 8 include the interaction between year dummies and, respectively, the (log of) 1900 city and immigrants population, and the (log of) 1904 value added by manufacture per establishment. F-stat refers to the K-P F-stat for weak instrument. Cols 5-6 report the p-value for the test of overidentifying restrictions. All regressions partial out city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 1.3. Immigration and Natives' Employment

	(1) OLS	(2) 2SLS	(3) Pre-Trends	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
				Panel A. Natives' Employment to Population Ratio *(1910 Mean: 0.911)*					
Fr. Immigrants	0.287***	0.299***	-0.117	0.213***	0.330***	0.213***	0.362***	0.226***	0.280***
	(0.043)	(0.064)	(0.112)	(0.048)	(0.115)	(0.043)	(0.076)	(0.061)	(0.081)
				Panel B. Natives' Log Occupational Scores *(1910 Mean: 3.245)*					
Fr. Immigrants	0.000	0.097***	0.026	0.070***	0.060	0.082**	0.124***	0.082*	0.112***
	(0.053)	(0.036)	(0.066)	(0.026)	(0.067)	(0.033)	(0.042)	(0.042)	(0.039)
F-stat		251.3	313.0	175.3	82.65	102.2	207.4	82.91	107.5
Observations	538	538	180	538	379	356	180	538	526
Covariates & sample restrictions									
Immigrants over 1900pop.				X					
MSA analysis					X				
WWI-Quotas IV						Stacked FD	Long Diff		
Year by 1900								City and immig. pop	Value added manuf.

Note: this table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930, restricting the attention to native men in the age range 15 to 65 who are not enrolled in schools. The dependent variable is natives' employment to population ratio in Panel A, and natives' log occupational scores in Panel B. Occupational scores are computed by IPUMS, and assign to an individual the median income of his job category in 1950. Col 1 and 2 present OLS and 2SLS results for the baseline specification (equation (1)). Col 3 regresses the 1900-1910 change in the outcomes against the 1910-1930 change in instrumented immigration. Cols 4 and 5 replicate Col 2 by, respectively, scaling the number of immigrants by 1900 population and aggregating the unit of analysis to the MSA level. Cols 6 and 7 estimate stacked first differences equation (5) and long differences equation (7) replacing the standard shift-share instrument with those constructed exploiting World War I and the quotas (equations (3) and (4) in Section 4.3). Cols 8 and 9 include the interaction between year dummies and, respectively, the (log of) 1900 city and immigrants population, and the (log of) 1904 value added by manufacture per establishment. F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 1.4. Immigration and Natives' Occupational Upgrading

	High Immigrants' Competition			**Low Immigrants' Competition**		
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction Natives:	Manuf. Laborers	Waiters	Blacksmiths	Manuf. Foremen	Electricians	Engineers
Panel A: OLS						
Fr. Immigrants	-0.026	-0.015	-0.008**	0.020***	0.010**	0.017*
	(0.048)	(0.011)	(0.004)	(0.005)	(0.004)	(0.010)
Panel B: 2SLS						
Fr. Immigrants	-0.057	-0.015	-0.011**	0.028***	0.011***	0.031***
	(0.037)	(0.013)	(0.005)	(0.006)	(0.004)	(0.008)
F-stat	251.3	251.3	251.3	251.3	251.3	251.3
Mean dep var	0.038	0.007	0.006	0.007	0.010	0.021
Natives/Immigrants Ratio (1910)	0.220	0.583	0.750	3.500	3.667	4.200
Observations	538	538	538	538	538	538

Note: this table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930 (see Table A2 in the appendix). The dependent variable is the fraction of native males in working age (15-65) employed in the occupation reported at the top of each column. Panels A and B report, respectively, OLS and 2SLS results. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. *Natives/Immigrants Ratio (1910)* refers to the ratio of native to immigrant workers in a given skill category or occupation in 1910. All regressions include city and state by year fixed effects. The mean of each dependent variable at baseline is shown at the bottom of the Table. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 1.5. Immigration and Economic Activity

	(1) OLS	(2) 2SLS	(3) Pre-Trends	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
				Panel A. Log Value Added per Establishment					
Fr. Immigrants	2.057***	2.889***	0.031	2.105***	4.484***	1.778***	2.277***	2.465**	2.423**
	(0.703)	(0.954)	(0.414)	(0.730)	(1.084)	(0.665)	(0.729)	(1.073)	(1.113)
				Panel B. Log Establishment Size					
Fr. Immigrants	2.195***	2.532***	0.051	1.726***	4.539***	1.983***	2.146***	1.945**	2.590***
	(0.614)	(0.815)	(0.458)	(0.596)	(0.981)	(0.596)	(0.720)	(0.931)	(0.972)
F-stat		270.5	272.6	198.2	80.23	106.0	199.4	89.38	124.7
Observations	525	525	176	525	370	347	169	525	519
Covariates & sample restrictions									
Immigrants over 1900pop.				X					
MSA analysis					X				
WWI-Quotas IV						Stacked FD	Long Diff		
Year by 1900								City and immig. pop	Value added manuf.

Note: this table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930, and for which data were reported in the Census of Manufacture between 1909 and 1929. The dependent variable is the log of value added per establishment in Panel A, and the log of establishment size in Panel B. Col 1 and 2 present OLS and 2SLS results for the baseline specification (equation (1)). Col 3 regresses the 1904-1910 change in the outcomes against the 1910-1930 change in instrumented immigration. Cols 4 and 5 replicate Col 2 by, respectively, scaling the number of immigrants by 1900 population and aggregating the unit of analysis to the MSA level. Cols 6 and 7 estimate stacked first differences equation (5) and long differences equation (7) replacing the standard shift-share instrument with those constructed exploiting World War I and the quotas (equations (3) and (4) in Section 4.3). Cols 8 and 9 include the interaction between year dummies and, respectively, the (log of) 1900 city and immigrants population, and the (log of) 1904 value added by manufacture per establishment. F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 1.6. Tax Rates and Public Spending Per Capita

	(1) OLS	(2) 2SLS	(3) Pre-Trends	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
				Panel A: Property Tax Rate *(1910 Mean: 19.75)*				
Fr. Immigrants	-28.49***	-29.44*	-4.204	-16.09	-24.29	-38.16**	-21.42	-19.38
	(10.60)	(16.95)	(8.224)	(11.56)	(19.35)	(14.88)	(21.22)	(19.73)
F-stat		292.7	320.6	230.4	106.2	204.5	97.37	124.2
Observations	539	539	179	539	359	179	539	527
				Panel B: Public Spending per Capita *(1910 Mean: 12.16)*				
Fr. Immigrants	-5.958	-8.699*	0.460	-5.794*	-5.739*	-11.34*	-12.01**	-17.18***
	(3.900)	(4.453)	(4.135)	(3.178)	(2.970)	(6.197)	(5.490)	(4.421)
F-stat		288.3	318.3	226.7	106.8	207.4	96.48	124.8
Observations	540	540	180	540	360	180	540	528
Covariates & sample restrictions								
Immigrants over 1900pop.				X				
WWI-Quotas IV					Stacked FD	Long Diff		
Year by 1900							City and immig. pop	Value added manuf.

Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. The dependent variable is the property tax rate for $1,000 of assessed valuation in Panel A, and public spending per capita in Panel B. Cols 1 and 2 present OLS and 2SLS results for the baseline specification (equation (1)). Col 3 regresses the 1906-1910 change in the outcomes against the 1910-1930 change in instrumented immigration. Col 4 replicates Col 2 by scaling the number of immigrants by 1900 population. Cols 5 and 6 estimate stacked first differences equation (5) and long differences equation (7) replacing the standard shift-share instrument with those constructed exploiting World War I and the quotas (equations (3) and (4) in Section 4.3). Cols 7 and 8 include the interaction between year dummies and, respectively, the (log of) 1900 city and immigrants population, and the (log of) 1904 value added by manufacture per establishment. F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 1.7. Presidential Elections and DW Nominate Scores

	(1) OLS	(2) 2SLS	(3) Pre-Trends	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
			Panel A: Democrats' Vote Share *(1910 Mean: 0.490)*					
Fr. Immigrants	-0.528***	-0.404***	-0.147	-0.313***	0.048	-0.606***	0.169	-0.271
	(0.119)	(0.141)	(0.157)	(0.112)	(0.162)	(0.167)	(0.271)	(0.169)
F-stat		83.14	64.54	55.42	23.43	35.76	35.64	67.73
Observations	378	378	123	378	252	126	378	378
			Panel B: DW Nominate Scores *(1910 Mean: 0.165)*					
Fr. Immigrants	0.745	1.658**	0.052	1.174**	1.908**	1.168	1.760*	2.403
	(0.514)	(0.808)	(0.909)	(0.559)	(0.779)	(0.843)	(1.025)	(1.507)
F-stat		23.11	25.92	70.30	8.571	15.39	10.75	34.13
Observations	460	460	135	460	303	146	460	451
Covariates & sample restrictions								
Immigrants over 1900pop.				X				
WWI-Quotas IV					Stacked FD	Long Diff		
Year by 1900							City and immig. pop	Value added manuf.

Note: In Panel A, the dependent variable is the Democrats' vote share in Presidential elections, and the sample includes the balanced panel of the 126 metropolitan statistical areas (MSAs) containing at least one of the 180 cities in my sample. In Panel B, the dependent variable is the first dimension of DW Nominate scores of members of the House of Representatives, for the panel of city-to-congressional district units for Congress 61, 66, and 71, for the 180 cities considered in my sample. Cols 1 and 2 present OLS and 2SLS results for the baseline specification (equation (1)). Col 3 regresses the 1900-1910 change in outcomes against the 1910-1930 change in instrumented immigration. Col 4 replicates Col 2 by scaling the number of immigrants by 1900 population. Cols 5 and 6 estimate stacked first differences equation (5) and long differences equation (7) replacing the standard shift-share instrument with those constructed exploiting World War I and the quotas (equations (3) and (4) in Section 4.3). Cols 7 and 8 include the interaction between year dummies and, respectively, the (log of) 1900 city and immigrants population, and the (log of) 1904 value added by manufacture per establishment. F-stat refers to the K-P F-stat for weak instrument. All regressions include MSA (Panel A) or congressional district to city (Panel B) and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 1.8. Congressmen Ideology and the National Origins Act of 1924

Dep. Variable:	*DW Nominate Scores*		*Pr. that Winner has Given Political Orientation*				*1[Restrict Immigration]*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: OLS								
Fr. Immigrants	0.745	0.603	-0.045	-0.804	-0.290	1.238	2.121*	2.024
	(0.514)	(0.521)	(0.317)	(0.711)	(0.991)	(1.135)	(1.189)	(1.362)
Panel B: 2SLS								
Fr. Immigrants	1.658**	1.575*	-0.601	-1.655	-0.198	2.592*	3.784**	3.365*
	(0.808)	(0.841)	(0.817)	(1.039)	(1.717)	(1.354)	(1.569)	(1.770)
F-stat	23.11	19.56	23.11	23.11	23.11	23.11	88.05	39.34
Mean dep var	0.165	0.150	0.167	0.161	0.359	0.314	0.676	0.676
Observations	470	437	470	470	470	470	155	155
Balanced Panel		X						
Political Orientation			Liberal Democrat	Moderate Democrat	Moderate Republican	Conservative Republican		

Note: Cols 1 to 6 report results for the panel of city-to-congressional district units for Congress 61, 66, and 71, for the 180 cities considered in my sample (see Table A2). Because of redistricting between the 61[st] and the 66[th] Congress, it was not possible to construct a balanced panel including all city-congressional district cells in my sample. For this reason, Col 2 restricts the attention to the balanced panel of cities (to congressional districts) that were not affected by redistricting. The unbalanced (resp. balanced) panel is composed of 157 (resp. 146) units of observations. Cols 7 and 8 present results from a cross-sectional regression for the 155 combinations of cities to congressional districts in Congress 68, for the 180 cities considered in my sample. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is the first dimension of the DW Nominate score in Cols 1 and 2, an indicator for electing a politician with a given political orientation (see bottom of the Table) in Cols 3 to 6, and an indicator for voting in favor of the 1924 National Origins Act in the House of Representatives. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. Cols 1 to 6 include city by congressional district and state by year fixed effects. Cols 7 and 8 control for state fixed effects. Col 8 also includes the 1900 log of black, immigrants, and total population, as well as the share of Democratic legislators representing the city (to congressional district) in the 68[th] Congress. Robust standard errors, clustered at the congressional district level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 1.9. Immigration and Religion

Dep. Var.	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Public spending PC	(5) Dem-Rep. margin	(6) Smith's pct. votes	(7) DW Nominate Scores	(8) 1[Restrict Immigration]
Panel A: OLS								
Fr. Non-Prot.	-13.69	-11.82	-32.53**	-8.422	-1.279***	-2.605***	1.053	2.888*
	(9.424)	(7.979)	(13.68)	(5.149)	(0.269)	(0.542)	(0.822)	(1.571)
Fr. Prot.	25.96	17.69	-4.948	9.853	1.440	2.512	-0.580	-2.655
	(23.52)	(22.64)	(50.18)	(21.75)	(1.103)	(1.819)	(1.191)	(3.487)
Panel B: 2SLS								
Fr. Non-Prot.	-13.56*	-12.73*	-32.11*	-9.645**	-0.571**	-3.027***	1.912**	4.946***
	(8.051)	(7.475)	(17.09)	(4.819)	(0.282)	(0.502)	(0.895)	(1.807)
Fr. Prot.	12.33	4.284	-6.984	-0.430	-0.593	3.711	0.394	-4.151
	(25.47)	(22.42)	(71.54)	(15.95)	(0.802)	(2.416)	(1.915)	(4.954)
KP F-stat	26.37	26.37	26.23	26.37	37.94	35.87	32.16	23.74
F-stat (Non-Prot)	115.9	115.9	118.9	115.9	53.37	40.18	85.91	69.49
F-stat (Prot)	27.53	27.53	27.39	27.53	38.95	36.58	32.27	21.68
Mean of dep var	12.76	12.10	19.75	12.16	0.180	0.398	0.165	0.676
Observations	540	540	539	540	378	126	460	155

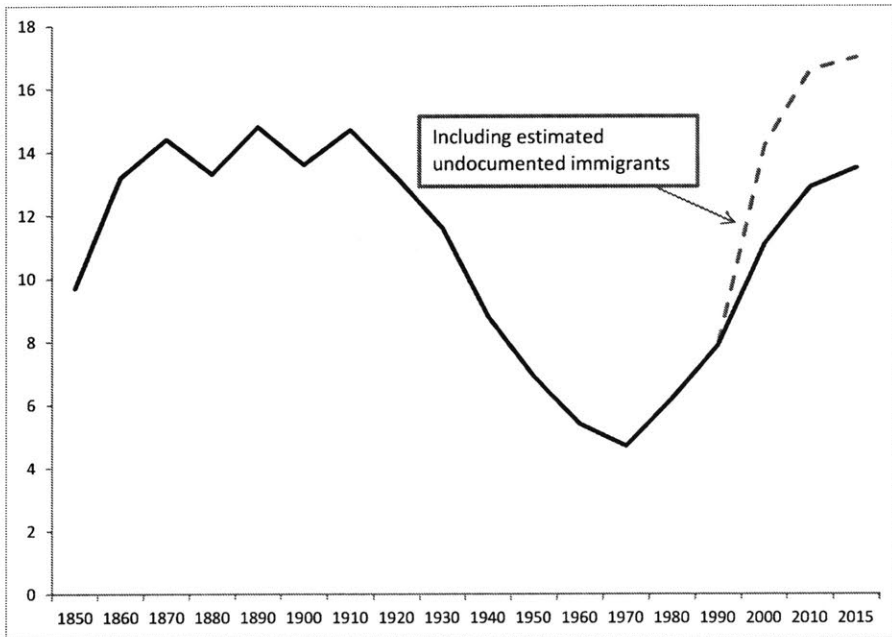
Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. The analysis is conducted at the MSA rather than at the city level, fixing boundaries using 1940 definitions in Cols 5 and 6, and at the city to congressional district level in Cols 7 and 8. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is displayed at the top of each column. *1[Restrict Immigration]* (Col 8) is an indicator for voting in favor of the 1924 National Origins Act in the House of Representatives. In Cols 1 to 5 and in Col 7, *Fr. Non-Prot.* (resp. *Prot.*) refers to the fraction of immigrants arrived in the previous decade from non-Protestant (resp. Protestant) countries, over predicted city population, for each of the three decades, 1910, 1920, and 1930. In Cols 6 and 8, *Fr. Non-Prot.* (resp. *Prot.*) is the 1910 to 1930 (1910 to 1920) change in the fraction of recent immigrants from non-Protestant (resp. Protestant) countries over predicted city population. Each endogenous regressor is instrumented with the predicted fraction immigrants (see (2) in Section 4.2), obtained by summing (predicted) immigration across non-Protestant and Protestant countries. F-stat (Non-Prot) and F-stat (Prot) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. Cols 1 to 4 (resp. 5) include city (resp. MSA) and state by year fixed effects, while Col 7 includes congressional district by city and state by year fixed effects. Cols 6 and 8 present results from a cross-sectional regression and control for state dummies. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 1.10. Linguistic Distance and Redistribution

Dep. Var.	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Public spending PC	(5) Education	(6) Police	(7) Charities and Hospitals	(8) Sanitation
Panel A: OLS								
Ling. Distance	-0.361*	-0.346	-1.485*	-0.213	-0.050	-0.032	-0.010	-0.045
	(0.205)	(0.212)	(0.840)	(0.160)	(0.060)	(0.021)	(0.039)	(0.029)
Panel B: 2SLS								
Ling. Distance	-0.875*	-0.809*	-2.308	-0.519*	-0.199*	-0.013	-0.119	-0.053
	(0.468)	(0.458)	(1.598)	(0.301)	(0.117)	(0.042)	(0.084)	(0.052)
KP F-stat	21.02	21.02	21.47	21.02	21.14	21.02	16.31	21.02
F-stat (Imm.)	123.1	123.1	124.7	123.1	106.9	123.1	101.6	123.1
F-stat (Ling.)	50.38	50.38	53.48	50.38	48.05	50.38	34.06	50.38
Mean of dep var	12.76	12.10	19.75	12.16	4.250	1.338	0.635	1.129
Observations	540	540	539	540	534	540	516	540

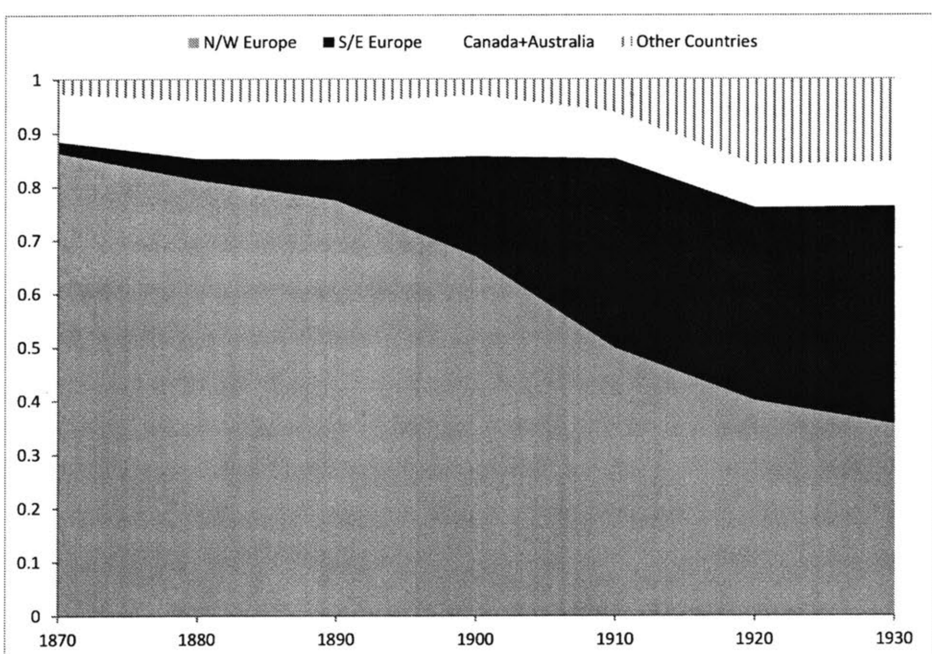
Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is displayed at the top of each column. In Cols 5 to 8, the dependent variable is spending per capita on the category listed at the top of the column. The main regressor of interest is the (standardized) weighted average linguistic distance constructed in Section 7.2, instrumented using predicted shares of immigrants from each sending region obtained from (2) in Section 4.2. F-stat is the Kleibergen-Paap F stat for joint significance of instruments. F-stat (Imm.) and F-stat (Ling.) refer to the partial F-stats for joint significance of the instruments in the two separate first-stages. All regressions include the main effect of immigration (instrumented with the baseline shift-share instrument from (2)), and control for city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Figure 1.1. Immigrants as a Percent of US Population



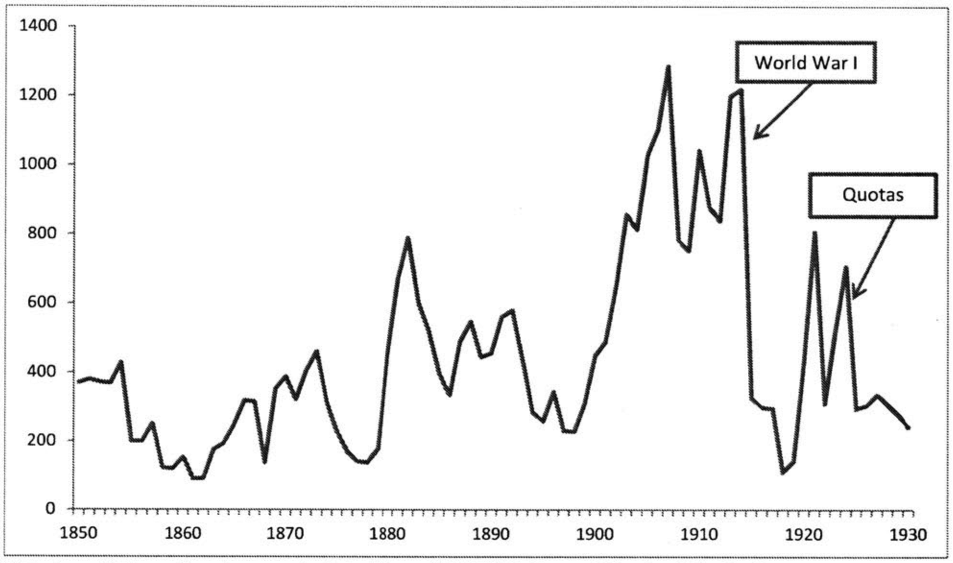
Note: The solid line shows the number of legal immigrants as a percent of US population. The dashed line includes also the estimated number of illegal immigrants, available from 2000 onwards. Source: the number of legal immigrants comes from the Migration Policy Institute, while the number of illegal immigrants was taken from the Pew Research Center tabulations

Figure 1.2. Share of Foreign Born in the United States, by Region



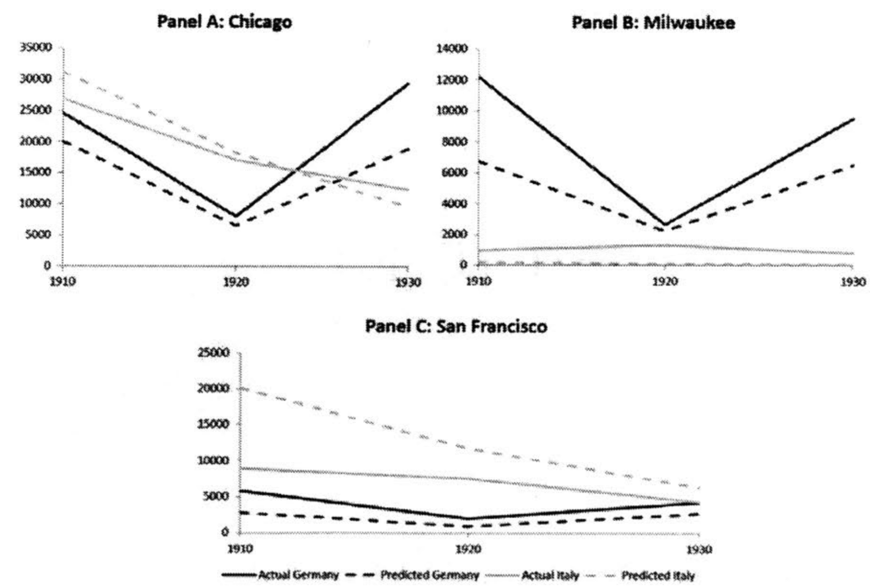
Note: Share of immigrant stock living in the United States, by sending region and by decade. Source: Author's calculations from IPUMS sample of US Census (Ruggles et al., 2015).

Figure 1.3. Total Number of Immigrants (in Thousands)



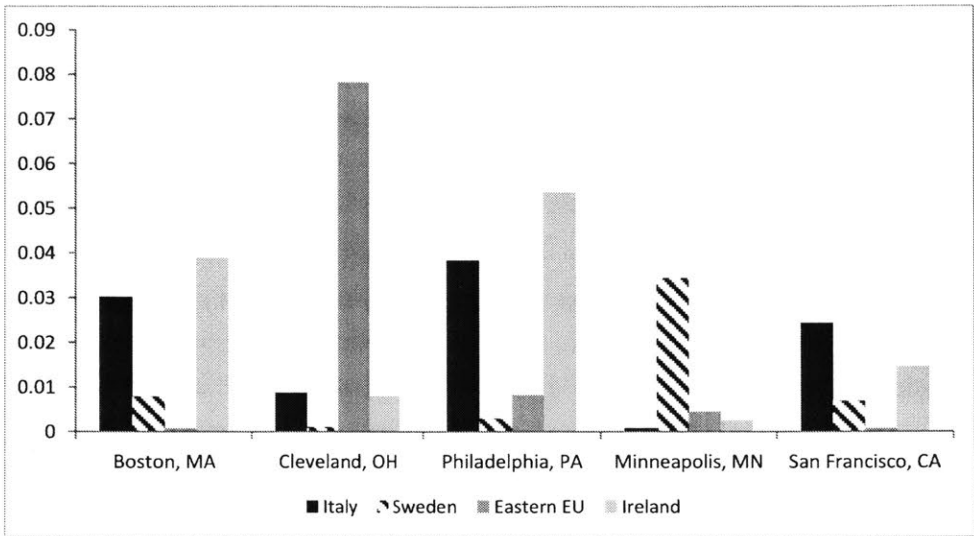
Note: Annual inflow of immigrants to the United States (1850-1930). Source: Migration Policy Institute.

Figure 1.4. A Simple Example: Actual and Predicted Immigration



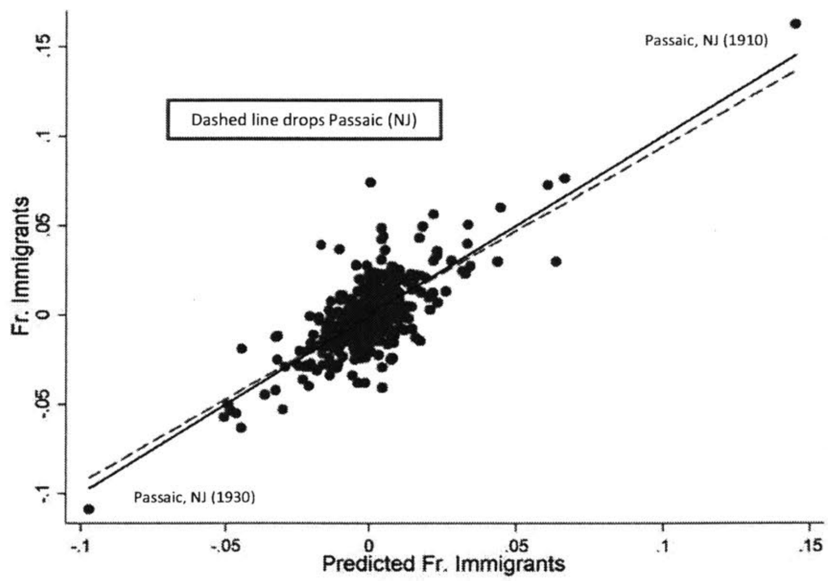
Note: This Figure reports the actual and predicted number of Italians and Germans arrived during the previous decade to Chicago (Panel A), Milwaukee (Panel B), and San Francisco (Panel C), in 1910, 1920, and 1930. Predicted immigration is obtained from the instrument constructed in equation (2) in the main text. Source: from IPUMS sample of US Census (Ruggles et al., 2015).

Figure 1.5. Share of European Immigrants in US Cities, 1900



Note: share of individuals of European ancestry living in US cities in 1900, for selected ethnic groups. Source: Author's calculations using IPUMS data.

Figure 1.6. First Stage: Actual vs Predicted Immigration



Note: the y-axis (resp. x-axis) reports the actual (resp. predicted) number of immigrants over predicted city population in each of the three Census years, 1910, 1920, and 1930. Each point in the scatter diagram represents the residual change in a city's actual and predicted fraction of immigrants after partialling out city and year by state fixed effects. The predicted number of immigrants is constructed as discussed in Section 4.2 in the text (see (2)). Predicted city population is obtained by multiplying 1900 city population with average urban growth, excluding that of the Census division where a city is located. The solid line shows the regression coefficient for the full sample (coefficient=0.999, standard error=0.059). The dotted (red) line shows the regression coefficient obtained when dropping the city of Passaic, NJ (coefficient=0.940, standard error=0.068).

Chapter 2. Racial Heterogeneity and Local Government Finances: Evidence from the Great Migration

2.1 Introduction

Central cities–suburbs inequality is a recurrent feature of US Metropolitan Statistical Areas (MSAs). Such inequality is evident along a number of dimensions: inner city residents are usually poorer, less educated, less likely to be white, and live further away from well-paying jobs relative to their suburban counterparts (see Table 2.1). Moreover, affluent suburbs often provide higher quality public goods and services, relative to those offered inside the urban ring (Boustan, 2013). In turn, disparity in the quality of and in access to key public goods such as education and health care is considered one of the main factors behind the persisting inner cities-suburbs, and the related racial, inequality (see, among others, Fryer and Katz, 2013, or Katz, 2015).[1]

One commonly proposed explanation for why US cities are chronically unfunded and often unable to provide adequate levels of public goods is racial heterogeneity. First, racial heterogeneity can reduce demand for public services either because non-coethnics have conflicting preferences over government spending or because whites' utility from public goods falls when these have to be shared with non-white individuals (Alesina et al., 1999).[2] In turn, lower demand for public spending can reduce government revenues, via cuts in the tax rate.[3]

Second, white residents can respond to racial heterogeneity by leaving the central city and relocating to richer and more racially homogeneous suburbs (Boustan, 2010). As a result, demand for housing in the inner city falls, driving down house prices and possibly encouraging a long-lasting process of urban decline (Glaeser and Gyourko, 2005). Since most local government revenues come from property taxes, reductions in property values can impose a negative fiscal externality on central cities (Boustan, 2012). Raising taxes may not be enough to compensate for the deterioration of the tax base, and cities may

[1]Very recent trends suggest that poverty and inequality might be spreading to the suburbs as well (Kneebone, 2017). In particular, while suburban areas that are further away from the central cities are still thriving, standards of living and the quality of local amenities (including schools and hospitals) are deteriorating in inner suburbs (Allard, 2017).

[2]Since non-whites are usually poorer, another explanation for lower demand for public goods is that richer whites may perceive public spending as an implicit transfer to poorer individuals.

[3]For example, Alesina et al. (1999) mention the experience of Prince George's County (Maryland) where, following a large inflow of African Americans during the 1970s, a legal ceiling on the property tax rate (TRIM) was proposed and approved by the white majority. The conventional wisdom is that TRIM is at the roots of Prince George's low quality school system.

be forced to cut spending and increase debt to deal with a tighter budget constraint.

In this paper, I investigate whether racial heterogeneity contributes to the distressed financial conditions and the poor provision of public goods in US central cities. I study the effects of the Great Migration (1915-1930) of 1.5 million southern born African Americans on government finances and public spending of US northern cities, and test the two channels (fiscal externality vs demand for public goods) discussed above.[4] When answering this question, two sources of endogeneity must be addressed. First, cities receiving more blacks were not randomly selected, as black migrants likely moved to places with better employment opportunities or with more appealing tax-public spending bundles. Second, white residents reacted to black immigration by "fleeing" to the suburbs and leaving racially diverse neighborhoods (Boustan, 2010; Shertzer and Walsh, 2016).

To deal with the first concern, I follow Boustan (2010) and, after estimating outmigration from each southern state for each decade, I assign migrants to northern cities exploiting 1900 clusters of southern born blacks.[5] To address the second source of endogeneity, predicted black inflows are interacted with a number of geographic characteristics (presence of hills, lakes, oceans, and rivers) of the area surrounding central cities that should increase the cost that whites faced when moving to the suburbs. Intuitively, black arrivals increased whites' desire to relocate to the suburbs. However, their ability to do so depended on house prices and on the costs they incurred when commuting from the suburbs to the central city.[6] For instance, in areas with more hills or water bodies, house prices are higher and commuting is more expensive, suggesting that, *ceteris paribus*, the white flight should be lower.

Black inflows had a strong, negative effect on tax revenues, which was entirely due to declining property values rather than to changes in the tax rate. My estimates suggest that, for the median city in my sample, Omaha (NE), the Great Migration reduced per capita tax revenues (resp. property values) by 6% (resp. 8%) relative to their 1910 levels. The reduction in house prices is consistent with lower demand for housing from native whites, who may have perceived the inflow of African Americans as a "disamenity" of living in central cities, and reacted by moving to the suburbs (Boustan, 2010; Shertzer

[4]Throughout the paper, when referring to northern cities, I mean cities outside the South. I follow the Census definition of the South, which includes the states of the former Confederacy, Kentucky, West Virginia, and Oklahoma, but, as in Boustan (2010), I exclude the border state of Maryland, which received substantial net immigration of blacks during the Great Migration (mostly concentrated in Baltimore).

[5]Early settlements of blacks in the North were largely determined by the railroad network, which was completed by 1890 (Collins and Wanamaker, 2015, and Black et al., 2015).

[6]Between 1910 and 1930, most jobs were still concentrated in central cities, and so individuals living in the suburbs largely needed to commute every day for work (Jackson, 1985).

and Walsh, 2016). I corroborate the interpretation that the Great Migration increased whites' demand for suburbanization by providing different pieces of suggestive evidence.

First, black immigration slowed down the construction of new housing units and the expansion of cities' boundaries. Second, in MSAs that experienced larger inflows of African Americans between 1910 and 1930, the number of local jurisdictions increased more between 1940 and 1970.[7] Third, in places that received more blacks between 1910 and 1930, more highways were planned in the Federal Highway Act (1944) and eventually built in subsequent decades. These findings are consistent with the idea that black immigration increased whites' demand for suburbanization, as proxied by highways. Once in the suburbs, whites set up their own local jurisdictions so as not to share public goods with poorer black immigrants (Alesina et al., 2004; Burns, 1994), and resisted annexation to the central city (Danielson, 1976).[8]

Having established that black immigration lowered property values and worsened city finances, I show that the Great Migration also reduced public goods provision. Consistent with changes in demand for redistribution, induced by racial or income considerations, spending cuts were larger in categories where either inter-racial interactions are more salient (education) or poorer African Americans would have obtained larger implicit transfers (sanitation and garbage collection). However, the fall in public spending was an order of magnitude smaller than that in tax revenues, as cities increased debt to make up for the lost (property) taxes.

As discussed above, a negative relationship between racial heterogeneity and public goods can be due to two mechanisms. First, a reduction in the desired level of spending; second, a negative fiscal externality, due to the deterioration of government finances induced by lower house prices. The substantial drop in property values and the related increase in public debt documented in my work are strongly consistent with a negative fiscal externality: as house prices fell, local government finances deteriorated, and cities were forced to borrow more and spend less to meet a tighter budget constraint. At the same time, the asymmetric reduction in spending (and, in education in particular) suggests that racial heterogeneity also affected the demand for public goods.

My findings speak to several strands of the literature. First, they are related to the papers that document a negative correlation between ethnic diversity and public spending across US cities (Alesina et al., 1999) and across countries (Alesina et al., 2001; Alesina

[7]Data constraints prevent me from looking at the contemporaneous change in the number of local jurisdictions: the Census of Governments started collecting this data only in 1942.

[8]Annexation was the most common for cities to expand their boundaries in this period (Jackson, 1985).

and Glaeser, 2004). My paper contributes to this literature in two ways. First, by relying on panel data, which permit the inclusion of fixed effects, and by using an instrumental variable approach, I can more confidently identify causal effects. Second, I show that racial heterogeneity can affect public spending not only by altering natives' preferences for redistribution (Luttmer, 2001) or reducing agreement over budget allocation (Beach and Jones, 2017), but also by generating fiscal externalities that, in turn, impact on governments' ability to provide public goods to their citizens. Dahlberg et al. (2012) find that the inflow of refugees to Sweden between 1985 and 1994 reduced support for redistribution in receiving municipalities.[9] I complement this work by showing that racial heterogeneity can have direct effects on actual policies and not only on natives' preferences.

Second, my paper complements the existing literature on the white flight (Boustan, 2010; Shertzer and Walsh, 2016) by documenting the negative effects that this phenomenon had on local government finances and, ultimately, on cities' ability to provide public services to their residents. Consistently with Boustan (2012), my results suggest that the white flight imposed a substantial fiscal externality, and possibly encouraged a persistent process of urban decline that may be at the roots of the current inequality between (poor) central cities and surrounding (rich) suburbs. While in the short run cities were able to partly absorb the negative shock to the available resources by increasing debt, this strategy may have not been sustainable in the medium to long run. In fact, it can even be one of the reasons for the financially distressed conditions of many US central cities today.

Because of the empirical setting, my findings are also related to the vast literature on the Great Migration. Several works have analyzed the effects of this massive episode of migration on either migrants (Collins and Wanamaker, 2014; Black et al., 2015) or residents of receiving places (Boustan, 2009, 2010, and 2017; Shertzer and Walsh, 2016). I instead focus on a set of variables (government finances and public spending) overlooked by the existing literature that likely had an effect on outcomes of both migrants and natives.

Finally, my paper speaks to the growing literature on the political consequences of immigration. Dustmann et al. (2016) and Halla et al. (2017) show that immigration increased the vote share of right wing, extremist parties in Denmark and Austria respectively. For the US, Mayda et al. (2016) find that immigration had opposing effects on the vote shares of Democrats and Republicans, depending on the fraction of naturalized

[9] Nekby and Pettersson-Lidbom (2017) revisit the work by Dahlberg et al. (2012), and argue that findings in the latter paper might be sensitive to the sample used and to measurement of preferences for redistribution.

immigrants already living in receiving areas.[10] My work complements this literature by analyzing the effects of immigration on public spending and government finances, and by suggesting avenues for future work. For instance, a policy relevant question is whether the changes in electoral outcomes documented in the aforementioned papers lead to changes in public spending or in other policies.

The remainder of the paper is organized as follows. Section 2.2 describes the historical background of the first wave of the Great Migration. Section 2.3 discusses the potential channels through which racial heterogeneity can affect government resources and public spending. Section 2.4 describes my data, and Section 2.5 lays out the empirical strategy. Section 2.6 presents results for the effects of the Great Migration on tax revenues, public spending, and debt, and explores the mechanisms discussed in Section 2.3. Section 2.7 performs several robustness checks. Section 2.8 concludes.

2.2 Historical Background: The Great Migration

Between 1915 and 1930, during the first wave of the Great Migration, more than 1.5 million African Americans left the rural South for northern cities. Such unprecedented migration wave was triggered by a number of push and pull factors (Boustan, 2017). On the one hand, World War I dramatically increased labor demand in northern industries while temporarily reducing European immigration, which was then permanently blocked by the Immigration Acts of the 1920s (Collins, 1997). Employers in the North started looking at southern born blacks as a source of cheap labor to replace European immigrants and to deal with the war-induced surge in demand. Between 1915 and 1919, more than 2 million jobs - most of them requiring minimal levels of skills - were created in northern cities, thereby increasing labor market opportunities for blacks (Boustan, 2017). On the other hand, a series of weather shocks hit the South in the early 1900s, reducing labor demand for agriculture, where most blacks were employed.[11] Racism and violence provided further incentives for African Americans to leave the South (see Tolnay and Beck, 1990, among others). Pushed by these factors and attracted by newly available jobs, blacks started moving to the North, taking advantage of the recently completed railroad network (Collins

[10]In chapter 1, I study the effects of European immigration to US cities between 1910 and 1930, exploiting variation induced by World War I and the Immigration Acts of the 1920s. I show that, even if immigrants brought economic prosperity to receiving cities, they nonetheless encountered significant political opposition.

[11]In 1892, a cotton pest - the Boll Weevil - entered in Texas and then spread throughout the South in subsequent decades, inducing substantial damages to local agriculture (Lange et al., 2009). In 1927, the Mississippi flood displaced a large number of agricultural workers in several counties of Mississippi, Louisiana, and Arkansas (Boustan et al., 2012; Hornbeck and Naidu, 2014).

and Wanamaker, 2014, and Black et al., 2015).

The combination of the factors discussed above, further reinforced by the process of chain migration, resulted in massive migration flows: between 1915 and 1930, approximately 1.5 million blacks moved from the rural South to the urban North of the United States, with the fraction of African Americans living in the North rising from 10% in 1910 to 25% in 1930 (see Figures 2.1 and 2.2). More than 60% of African Americans settled in the five most common destinations - New York, Chicago, Los Angeles, Detroit, and Philadelphia - but black immigration was a widespread phenomenon in many other large non-southern cities (see Figure 2.3). As a result, the number of blacks living in northern urban areas increased dramatically, altering the racial composition of receiving places. For instance, in Chicago, Cleveland, or Youngstown (just to mention a few), the fraction of blacks over city population moved from 2% in 1910 to more than 8% in 1930.

2.3 Racial Diversity, Government Finances, and Public Spending

Racial heterogeneity can affect government resources and public goods provision through at least two channels: first, by changing demand for public spending and preferences for redistribution; second, by reducing property values due to e.g. lower housing demand or disamenity effects and, in turn, imposing a fiscal externality on more diverse areas. In this section, I discuss each of these two channels separately, and derive some testable implications that will guide my empirical analysis.

2.3.1 Demand for Public Goods and Preferences for Redistribution

As discussed in Alesina et al. (1999) among others, racial diversity can reduce the desired level of public spending. On the one hand, people from different ethnicities and cultures may disagree on what they consider the optimal amount of government spending or its allocation across public goods (Beach and Jones, 2017). On the other, the literature has documented a strong tendency of individuals to be more altruistic with coethnics and less willing to redistribute towards non-coethnics (Luttmer, 2001; Dahlberg et al., 2012). Yet another reason for a negative effect of racial diversity on public spending is that the utility from consumption of public goods may be lower when these have to be shared with people from different races.[12] As individuals want to spend less, they also demand lower taxes, implying that public resources should be lower in more diverse communities.

[12]For instance, the utility that white parents get from sending their children to public schools may be lower when schools are more ethnically diverse (Baum-Snow and Lutz, 2011; Cascio and Lewis, 2012).

This discussion provides two testable predictions. First, if a negative relationship between ethnic diversity and government revenues were due to lower demand for public goods, one should also find a negative relationship between ethnic diversity and the tax rate. Second, the effect of racial heterogeneity on the demand for public goods (if any) should be stronger in categories where either inter-racial interactions are more salient (e.g. education) or poorer minorities would get larger implicit transfers (e.g. spending on poverty relief).

2.3.2 White Flight and Fiscal Externality

In addition to the "demand side" effects discussed in the previous paragraph, racial heterogeneity can affect government resources and public spending through the housing market. In particular, if natives have a distaste for living in racially diverse places, immigration may reduce house values and impose a negative fiscal externality on receiving areas. As documented in a number of papers for both Europe and the US, natives often respond to immigration by changing their residence in order to avoid inter-racial interactions.[13] As a result of lower demand from natives, house prices fall. Moreover, since housing is a durable good, any negative shock to demand will be amplified and will have a persistent impact on prices (Glaeser and Gyourko, 2005).

In the US, most local government revenues come from property taxes. Hence, declining house values will mechanically lower tax revenues and, potentially, impose a fiscal externality on areas experiencing in-migration (Boustan, 2012). While, in principle, the tax rate can be increased to compensate for a lower tax base, there may exist political and economic constraints that prevent municipalities from doing so. For instance, politicians may realize that, by increasing the tax rate, they would further depress housing demand and reinforce the process of urban decline. Also, precisely because of higher racial heterogeneity, voters may be reluctant to accept higher taxes. As a result, more diverse communities may be forced to cut spending and increase debt in order to meet a more binding budget constraint. Lower quality public goods and higher levels of indebtedness can also have a feedback effect on property values and trigger a self-reinforcing cycle of lower spending and worse government finances.

To sum up, if racial diversity affects public goods provision only through a fiscal externality, one should find a negative effect on house prices and no (or a positive) effect

[13]In addition to the works by Boustan (2010) and Shertzer and Walsh (2016) reviewed above, evidence on the white flight is provided by Card et al. (2008) and Saiz and Wachter (2011) for the US, and by Anderson et al. (2016) for Europe.

on the tax rate. When interpreting my results, I will use this prediction, and contrast it with those obtained above, to discriminate between the demand and the fiscal externality effects of ethnic diversity on public spending and government finances. In practice, both channels can be simultaneously at play. However, testing these predictions will allow me to shed some light on the relevance of each of the two mechanisms in my setting.

2.4 Data

My analysis is based on the three Census years that span the period of the first wave of the Great Migration, i.e. 1910, 1920, and 1930. The sample is composed of the 42 non-southern US central cities that were anchored to a MSA, had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black immigration.[14] The 42 cities in my sample (see Figure A2.1 and Table A2.1 for the complete list) absorbed more than 95% of southern black migrants that settled in northern or western urban areas during the first wave of the Great Migration, and include more than 85% of the black, white, or city population of non-southern urban areas in each decade between 1910 and 1930.

For this project, I collected, and in many cases digitized, data from two main sources. First, I used data on city population by race in each decade from the Census of Housing and Population.[15] From the Census of Housing and Population, I also collected data on the number of families, the number of dwellings, and other city-level socioeconomic and demographic characteristics used either as outcomes (in addition to public finance variables) or to perform robustness checks.[16] To construct the instrument for black migration, I used data from the 1900 5% sample and the 1910 to 1930 1% samples of the US Census of Population made available by IPUMS (Ruggles et al., 2015), and from ICPSR study 2896 (Haines et al., 2010).

Second, data on public spending and city finances were collected and digitized from the Financial Statistics of Cities. These are annual reports, available from 1906 onwards for cities with population above 30,000 or 100,000, depending on the years. Since population

[14]Consistently with Baum-Snow (2007) and Boustan (2010), central cities are defined as the largest urban center of a given MSA. My results appear robust to alternative definitions. Since I am interested in city-level fiscal outcomes, differently from Baum-Snow (2007) or Boustan (2010), I do not fix city boundaries. However, I do fix MSA boundaries as of 1940 in order to keep the geographic unit of analysis constant (results are unchanged when fixing these as of 1910 or 1970).

[15]In order to obtain more precise meaures, I used the full Census files available at http://www.census.gov/prod/www/decennial.html.

[16]Specifically, demographic and socioeconomic variables include: the fraction of the population below age 20; the fraction of foreign born whites, or with mixed parentage; the fraction of illiterate.

data are available only at decennial frequencies, data from the Financial Statistics of Cities were collected for years 1910, 1919, and 1930.[17] In my analysis, I consider public spending (total and by category), tax revenues, tax rates, property values, and debt.[18]

Table 2.2 reports the summary statistics for the main variables used in my analysis. Table A2.2 in the appendix lists all variables with their respective sources.

2.5 Empirical Strategy

In this section, I present the baseline estimating equation (Section 2.5.1), construct the instrument for black population (Section 2.5.2) and for the induced white outmigration (Section 2.5.3), and report first stage results (Section 2.5.4). The number of African Americans is predicted using the interaction of 1900 settlements of blacks born in different southern states with subsequent migration flows from each sending state. To deal with the "white flight", I instrument the number of white residents by interacting predicted black population with MSA geographic characteristics (hills and water bodies) that affected the cost faced by whites when moving to the suburbs.

2.5.1 Baseline Estimating Equation

To study the effects of the Great Migration on tax revenues, public spending, and public debt, I stack the data for the three Census years 1910, 1920, and 1930, and estimate

$$Y_{crt} = X_{crt} + \beta_1 B_{crt} + \beta_2 W_{crt} + u_{crt} \tag{10}$$

where Y_{crt} is the outcome for city c in Census region r in year t;[19] B_{crt} and W_{crt} are, respectively, the number of blacks and whites living in city c in year t; and X_{crt} includes city and year by region fixed effects, as well as additional controls described in detail below (Sections 2.5.2 and 2.5.3). Standard errors are clustered at the MSA level. Coefficients β_1 and β_2 should be interpreted as the effect of 1,000 more blacks and whites respectively in the central city.

Following the literature, and in particular Boustan (2010) and Shertzer and Walsh (2016), I adopt a specification in levels for several reasons. First, using the fraction of

[17] Since data for 1920 were missing, I used the 1919 volume - results are robust to using 1921, but 1919 is preferable because 1921 data were not reported for several cities. I digitized PDFs for the relevant years, including also 1906, which is used below to perform some robustness checks.

[18] In the Financial Statistics of Cities, property values refer to the assessed valuation of property. The tax rate on 1,000$ of such assessed value is also separately reported.

[19] When defining regions, I follow the Census classification.

blacks over city population would restrict the effect of a new black migrant moving in to be the same as the effect of a white resident moving out - a stark and, probably, unreasonable assumption in this context. Second, both the number of blacks and the number of whites in the central city are endogenous, and two separate instruments are needed.

In addition, as shown in Peri and Sparber (2011) and as discussed in Shertzer and Walsh (2016), a specification in levels performs well in Monte Carlo simulations of specification bias in displacement models. Since the white flight is one of the key channels through which black immigration may have affected city finances and public spending, it is important to minimize potential bias coming from mis-specification. Finally, using a specification in levels allows for direct comparisons of my results with those obtained in previous works on the white flight, and in particular Boustan (2010 and 2017) and Shertzer and Walsh (2016).

One possible remaining issue with a specification in levels is that it implicitly weighs larger cities more. To address this issue, similarly to Shertzer and Walsh (2016), in Section 2.7, I repeat the analysis trimming the sample at the 1^{st} and 99^{th} and at the 5^{th} and 95^{th} percentiles of black population. While, as expected, this slightly reduces the strength of the first stage, results remain in line with those from my baseline specification.

2.5.2 Instrument for Black Population

A priori, we may expect blacks to be attracted to cities with better job opportunities, or with more appealing tax-public spending bundles. Alternatively, African Americans may settle in otherwise declining cities, where house prices were lower. In either case, OLS estimates of (10) will likely be biased. To reduce such endogeneity concerns, I instrument the location decision of black migrants using a version of the "shift-share" instrument commonly adopted in the immigration literature (Card, 2001).[20]

Specifically, I first estimate outmigration from each southern state for each decade using the forward survival method (Gregory, 2005).[21] Using data for the United States as a whole, survival ratios were computed for each age-sex-race group and were then used to estimate net migration from each southern state (for each group). Next, outmigration flows are apportioned to northern cities depending on the share of southern born African Americans from each state living in each city in 1900. Formally, the number of blacks in

[20] A very similar approach is used, by Boustan (2010) and Shertzer and Walsh (2016) for the second and for the first wave of the Great Migration respectively.

[21] Before 1940, no dataset on US internal migration exists, and so migration rates must be estimated (Shertzer and Walsh, 2016). For robustness, I compared my measure of estimated outmigration with that computed in Lee et al. (1957), and the two were very similar.

the central city, B_{crt}, is instrumented with

$$Z_{crt} = \sum_{j \in South} \alpha_{jc}^{1900} O_{jt} \tag{11}$$

where α_{jc}^{1900} is the share of blacks born in southern state j residing in the non-South who were living in northern city c in 1900, and O_{jt} is the number of African Americans leaving state j between $t-1$ and t.[22]

As for other works in the literature (Card, 2001; Boustan, 2010), the instrument constructed above is based on the empirical regularity that settlement patterns of blacks were highly persistent over time. As discussed in Black et al. (2015) and Boustan (2010) among others, the railroad network was key in determining the location decision of early migrants. The stability of settlement patterns was further reinforced by chain migration: more recent migrants tended to move where other migrants from the same county (or state) had moved in the past (see, among others, Wilkerson, 2010). In addition to stability over time, settlement patterns of migrants also exhibited wide variation across both receiving cities and sending states. That is, even before 1900, several cities had received blacks from many southern states, and each southern state had sent migrants to a number of different cities.

Figure 2.4 plots the share of southern born blacks from selected states living in a number of northern cities in 1900, and confirms visually that there was indeed variation across both sending and receiving places. For instance, almost 35% of blacks born in Alabama and living in the North had moved to either St. Louis (15%) or Chicago (20%) by 1900, whereas only 3% of them were living in New York. At the same time, more than 30% of African Americans born in Florida and residing in the North in 1900 were living in New York, but less than 5% of them were living in either St. Louis or Kansas City (MO).

As discussed in Boustan (2010), the large fraction of blacks born in Alabama and living in St. Louis or Chicago was due to the presence of major railroads like the *Gulf, Mobile, and Ohio* that were connecting southern cities such as Mobile to the northern hubs of Chicago and St. Louis through another main railroad, the *Illinois Central*. The *Illinois Central*, the only inter-state railroad crossing Mississippi, was also favoring migration from the latter state to Chicago and St. Louis (see also Black et al., 2015).

The previous discussion, as well as Figure 2.4, suggests that the instrument for black

[22] I consider cities where α_{jc}^{1900} is above 0.005 for at least one southern state. Results are not sensitive to the choice of this threshold.

immigration constructed in (11) can have predictive power for actual immigrants' location. Before confirming this more formally in Section 2.5.4, I first discuss the identifying assumptions behind my instrument.

2.5.2.1 Identifying Assumptions and Instrument Validity

The instrument described above relies on one, key identifying assumption: cities receiving more blacks (from each southern state) before 1900 must not be on differential trends for the evolution of public spending and government finances. This assumption would be violated if state level outmigration was influenced by cross-city pull factors systematically related to the origin state of 1900 settlers. For example, in the 1920s, outmigration rates were higher from Alabama than from Texas. The exclusion restriction requires that this was not happening because cities that in 1900 had attracted more migrants from Alabama were growing more than cities where more blacks from Texas had moved to before 1900. To indirectly address this concern, in Section 2.7, I perform two placebo exercises, including a formal test for pre-trends, and repeat the analysis interacting year dummies with a number of pre-migration city characteristics (fraction of immigrants, skill ratios, and different measures of industrialization).

Another potential threat to the identifying assumption is that, by construction, the instrument predicts larger black in-migration to cities with relatively more black residents in 1900. While city fixed effects control for any time-invariant difference between cities, they do not address the possibility that places with a larger fraction of blacks in 1900 simultaneously attracted more migrants and experienced differential changes in e.g. government finances in subsequent decades. A related concern is that cities closer to the South were likely to receive more migrants in 1900 (Collins and Wanamaker, 2015) both in the OLS and when using 2SLS, and may have differed from other northern cities, e.g. because of their industry mix.

To deal with these issues, in my most preferred specification, I augment the vector of controls, X_{crt}, interacting year dummies with the 1900 fraction of blacks and with city geographic coordinates. This set of interactions implies that the effects of black migration are identified exploiting variation only in the (southern state) composition of African Americans' enclaves across cities, holding constant the size of their black populations and their geographic location.

2.5.3 Instrument for the White Flight

When estimating (10), it is not enough to have an instrument for black immigration, since whites often reacted to the inflow of African Americans by leaving central cities and relocating to the suburbs (Boustan, 2010; Shertzer and Walsh, 2016). To account for the white flight, I construct a second instrument by interacting predicted black immigration, Z_{crt}, with geographic characteristics of the area surrounding central cities that arguably increased the costs faced by whites when moving to the suburbs and commuting back and forth from the suburbs to the city.[23]

The intuition behind my strategy is that whites were induced to leave the central city by the arrival of blacks (Boustan, 2010), but their ability to do so depended on MSA or city characteristics, like availability of land. In places with less available land (e.g. due to water, mountains, etc.), moving out was more expensive, because of higher rents and house prices (Saiz, 2010), or higher commuting costs. For this reason, one would expect the white flight to be lower in places surrounded by less friendly geography, other things being equal.

Specifically, I consider the fraction of the area around central cities: *i*) with slope above 15% (to proxy for hills, as in Saiz, 2010); *ii*) occupied by lakes and oceans; and *iii*) occupied by rivers and streams. Data used to compute these measures were taken from USGS and from Saiz (2010).[24] Following Saiz (2010), each variable was constructed using the area of a circle drawn around the central city with a 50 km radius. As a robustness check, I experimented with radii of different size, and results always remained similar.

To reduce concerns that geography can have a direct effect on city-level outcomes, when constructing geographic variables in *i* to *iii*, I excluded the area corresponding to the central city. This should also increase the confidence that I only exploit variation in the cost of suburbanization induced by suburban (rather than urban) geography. Since maps for this period are not available, I used information on city area and computed the implied radius the city would have had, assuming a circular shape. When constructing variables *i* to *iii*, I excluded this "implied" city-area .[25]

Formally, collecting the aforementioned geographic features in a vector, R_{cr}, the instrument for white population in (10) is given by the interaction $(Z_{crt} \times R_{cr})$, and the two

[23]Before the 1960s, most jobs remained concentrated in central cities, and so suburban residents had to commute to the inner city for work every day (Jackson, 1985).

[24]For elevation, I used Digital Elevation Model (DEM) at 90-square meter cell grids (available at http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1). Data on rivers and streams were collected from http://water.usgs.gov/GIS/dsdl/ds240/, while Albert Saiz kindly shared with me data on oceans and lakes.

[25]Results are equivalent when using 1910 or 1930 city area.

first stage equations, for B_{crt} and W_{crt}, are

$$B_{crt} = X_{crt} + \delta_1^b Z_{crt} + \delta_2^b \left(Z_{crt} \times R_{cr}\right) + u_{crt}^b \tag{12}$$

$$W_{crt} = X_{crt} + \delta_1^w Z_{crt} + \delta_2^w \left(Z_{crt} \times R_{cr}\right) + u_{crt}^w \tag{13}$$

2.5.3.1 Validity of the Instrument

The main concern on the validity of the instrument just described is that geography may have direct effects on house prices, or on blacks' settlement patterns. For instance, in more hilly MSAs, house prices are likely to be higher (Saiz, 2010). This, in turn, may affect both the location decision of black migrants and government finances. As shown in (13), however, the white flight is identified only out of the interaction term between (predicted) black immigration and geography. Hence, any direct effect of geography on either second stage outcomes or settlement patterns is controlled for by the inclusion of fixed effects, as long as this is constant over time.

One possible remaining concern is that geography may have a time varying, city-specific, effect on outcomes. In particular, by limiting land available for the construction of new buildings, hills and water bodies will mechanically increase house prices as cities receive more migrants. This is because, with immigration, cities surrounded by geographic obstacles tend to become more densely populated, thereby increasing house prices. Another possible story is that, over time, mountains and oceans become valuable amenities, in turn affecting the trend in house values of different cities differently. To deal with these and similar concerns, the vector of controls X_{crt} in (10) includes a full set of interactions between year dummies and dummies indicating the presence of hills, oceans, or lakes above different thresholds.[26]

As a further robustness check, I repeated my analysis using a different instrument for white population, obtained by interacting Z_{crt} with average January temperature - a measure commonly used in the literature to proxy for city-level amenities (Glaeser and Shapiro, 2003; Glaeser and Gyourko, 2005). As shown in Section 2.7 below, results are very similar when using this alternative instrument.

[26] In my baseline specification, I define the threshold at 5%, but I experimented with higher or lower values to check the robustness of my results, which always remained very similar. The distribution of these geographic characteristics is highly skewed. For instance, 22 of the 42 cities in my sample are not surrounded by hills or oceans (implying that the median is 0), but the 75th percentile of the distribution of e.g. the lakes and oceans share of suburban area is as high as 20%. There are 15 (resp. 19) cities where the share of the suburban area with slope above 15% (resp. occupied by lakes or oceans) is higher than five percent.

2.5.4 First Stage

Table 2.3 estimates first stage equations (12) and (13). In column 1, I investigate the relationship between predicted black immigration, Z_{crt} in (11), and actual black population, B_{crt}, controlling for city and year by region fixed effects. There is a strong and significant correlation between black population and the instrument, while the F-stat is well above conventional values. The magnitude of the coefficient is smaller than that in Boustan (2010), but very similar to that in Shertzer and Walsh (2016). Boustan (2010) focuses on the second wave of the Great Migration, a period in which black migration to the North was more pervasive, and this can potentially explain the difference between her estimates and those obtained here or in Shertzer and Walsh (2016).

Next, columns 2 and 3 add interactions between predicted black migration and geography, $Z_{crt} \times R_{cr}$. The coefficient on predicted black migration in the first stage for blacks (column 2) remains strongly significant. In the first stage for whites (column 3), predicted black migration is negatively correlated with white population, consistently with findings in Boustan (2010) and Shertzer and Walsh (2016), even though the point estimate is not statistically significant. More importantly, all interactions between Z_{crt} and geography have the positive expected sign: in cities surrounded by less friendly geography, the white flight induced by black immigration was less pronounced. The interaction between Z_{crt} and lakes and oceans is always highly significant, while that between hills and rivers is always positive, but its significance varies depending on the controls included in the regression. As in column 1, the F-stat and the partial AP F-stats for the joint significance of instruments in each first stage (reported at the bottom of the Table) are above conventional levels.

A similar pattern appears also when including interactions between year dummies and geography (columns 4 and 5), even though the coefficient on predicted black immigration in the first stage for whites becomes substantially smaller (in absolute value). Figure 2.5 confirms visually the strong relationship between actual and predicted black population, by plotting the graphical analogue of column 4. In Table A2.3 in the appendix, I further test the robustness of results presented above. Columns 1 to 6 re-estimate (12) and (13) by dropping cities with each of the three geographic variables (oceans and lakes; hills; rivers) above the 95th percentile, to check that results reported in Table 2.3 are not driven by cities with extreme values of geography.[27] Reassuringly, coefficients remain very similar to those in my baseline specification. In columns 7 and 8, I show that results are not

[27] Columns 1-2 (3-4; 5-6) omit cities with values for oceans and lakes (hills; rivers) above the 95th percentile.

significantly affected when using a 30 km (rather than a 50 km) radius to construct the geographic variables included in R_{cr}.[28]

Overall, both the instrument for black population and that for the white flight - obtained by interacting predicted black immigration with geography - are strongly correlated with the corresponding endogenous regressors, and this relationship appears to be robust to different specifications.

2.6 Results

In this section, I study the effects of the Great Migration on tax revenues, public spending, and public debt in northern cities. The inflow of African Americans lowered house prices in receiving cities, in turn forcing them to lower spending and raise debt in order to meet a tighter budget constraint. Spending cuts were larger in education, where inter-racial interactions are probably more salient, suggesting that racial diversity lowered whites' utility from the consumption of public goods.

2.6.1 Tax Revenues, Tax Rates, and Property Values

At least until the Great Depression, US cities were responsible for the provision of public goods such as education, police, and spending on welfare or on infrastructure (e.g. roads, sewerage, etc.), while the federal (or the state) government played only a marginal role (Monkkonen, 1988). Since federal and state transfers were very limited, cities had to independently raise funds to finance their expenditures.[29] Local taxes represented the main source of revenues, with property taxes accounting for more than 90% of total tax revenues (Fisher, 1996). Motivated by this discussion, in Table 2.4, I analyze the impact of the Great Migration on tax revenues, tax rates, and property values.

Throughout the paper, I always report OLS and 2SLS estimates in Panels A and B respectively, and the KP and the AP F-stats at the bottom of each table. In columns 1 and 2, the dependent variables are total and property tax revenues per capita, and the vector of controls, X_{crt}, only includes city and year by region fixed effects. Both OLS and 2SLS coefficients on black population are negative and highly significant. Consistently with blacks selecting more prosperous cities, which likely had sounder government finances, OLS estimates are smaller (in absolute value) than 2SLS ones, even though the two are

[28]The only difference with Table 2.3 is that the KP F-stat becomes lower. However, the partial F-stats remain well above conventional levels, and the significance of the individual coefficients is not substantially affected.

[29]Differently from today, at the time, spending or tax limits were very rare in US cities.

not statistically different from each other. The coefficients on white population are instead positive, but an order of magnitude smaller than those on blacks.

Column 3 replicates column 2 now including in X_{crt} also the full set of interactions between year dummies and the 1900 fraction of blacks, city coordinates, and geography. Results are barely affected, and the main message is unchanged: the inflow of African Americans had a strong, negative effect on tax revenues per capita of northern cities. As discussed above, a negative relationship between black immigration and tax revenues can be due to either lower tax rates, reflecting a reduction in demand for redistribution, or falling property values, induced by the white flight (or both). Columns 4 to 7 investigate the two channels separately. The inflow of African Americans did not significantly affect tax rates (column 4), but had a strong and negative effect on property values (columns 5 to 7). Since there is no obvious way to define the tax base, property values are scaled by, respectively, 1900, 1910, and contemporaneous population.

In all cases, results point in the same direction: black immigration lowered house values in receiving cities, driving down tax revenues and imposing a negative fiscal externality. Also, the impact of black immigration on city finances was economically large. Between 1910 and 1930, the median city in my sample, Omaha (NE), received more than 6,000 southern born blacks. 2SLS coefficients on black population in columns 3 and 5 (Panel B) imply that, for this city, the Great Migration lowered per capita property tax revenues and property values by respectively 6% and 8% relative to their 1910 mean. These numbers are remarkably similar to the estimates in Boustan (2012), who finds that, during the 1970s, desegregation lowered urban house prices and rents by 6%.

Next, to test if the decline in property values and tax revenues was (at least in part) due to the white flight, in Table 2.5, I re-estimate (10) without controlling for the number of white residents in the central city. If, as it happened during the second wave of the Great Migration (Boustan, 2010), black immigration lowered whites' demand for housing in central cities, the fall in house prices should be larger (in absolute value) when the white flight is not controlled for. Consistent with this idea, the reduction in tax revenues (columns 1 and 2) and in property values (column 4) is now an order of magnitude larger than in Table 2.4.[30]

Overall, these results suggest that black immigration reduced house prices and imposed a negative fiscal externality on northern urban areas. Unable (or unwilling) to raise taxes, receiving cities may have been forced to cut spending and increase debt to meet a tighter

[30] In addition to city and year by region fixed effects and to interactions between year dummies and the 1900 fraction of blacks and city coordinates, in Table 2.5, I follow Boustan (2010) and control for total MSA population.

budget constraint, as I show in the next two sections.

2.6.2 Public Spending

Table 2.6 presents results for the effects of black immigration on public spending per capita, total (column 1) and by category (columns 2 to 8).[31] The inflow of blacks was associated with a significant reduction in per capita spending, which was larger for education (column 2) and outdoor poverty relief (column 8).[32] When comparing coefficients of Table 2.6, it immediately appears that spending cuts were not evenly distributed across categories. For instance, while spending on education fell by more than 6% relative to its 1910 mean, that on fire protection (column 4), roads (column 6), or charities and hospitals (column 7) was barely affected.[33] Also, in contrast with findings in Alesina et al. (1999), black immigration had a negative, albeit quantitatively small, effect on spending for police (column 3).

To more formally test the possibility that black immigration altered the allocation of the budget across categories, I computed spending shares by dividing each of the outcomes in columns 2 to 7 by total spending (column 1). Table 2.7 confirms the pattern that was already apparent in Table 2.6: black inflows had a negative and significant effect on the share of spending on education (column 1), police (column 2), and sewerage and garbage collection (column 4). Conversely, there is a positive relationship between the inflow of African Americans and the spending share on fire protection (column 3) and roads (column 5), even though the latter is not statistically significant.

One possible concern with the interpretation of results in Tables 2.6 and 2.7, and with the reduction in spending on education in particular, is that they were due to changes in the demographic composition of northern cities induced by the Great Migration. To address this possibility, appendix Table A2.4 repeats the analysis scaling spending on education by the number of residents in schooling age (column 1) and enrolled in public schools (column 2).[34] Again, there is a negative and significant relationship between black population and spending on education, which is quantitatively in line with that reported

[31] As before, OLS results are reported in Panel A, while 2SLS ones are shown in Panel B.

[32] Spending on outdoor poverty relief was reported only for the 14 largest cities in my sample, and was collected separately from Geddes (1937), and not from the Financial Statistics of Cities. Due to the very limited sample size, results in column 8 were obtained by only including city and year by region fixed effects, and should be interpreted with some caution.

[33] Data on spending for charities and hospitals was not reported for San Diego in 1910, for Scranton in 1910 and 1920, and for Peoria in 1930.

[34] Data on the number of students enrolled in public schools was missing for Seattle and Tacoma in 1930.

in Table 2.6.

Table A2.4 also shows that the inflow of African Americans did not lead to changes in the fraction of enrolled pupils (column 3), implying that lower spending on education resulted in fewer resources per student. Data limitations prevent me from investigating possible redistributional effects among northern residents by studying, for instance, if (whites') enrollment and resources shifted from public to private schools, or, more generally away from schools located in neighborhoods more affected by the inflow of blacks (Cascio and Lewis, 2012). Finally, in column 4, I show that, despite the fall in public spending on education, there was no significant decline in the share of literate whites. One possible reason for this is that, by 1930, almost all whites could write and read, and literacy is thus an excessively coarse proxy for education.[35]

2.6.3 Municipal Debt

While the Great Migration had a negative and significant effect on public spending per capita, the implied magnitude of coefficients in Table 2.6 is quantitatively smaller (in absolute value) than that in Table 2.4. Specifically, for the median city, 2SLS estimates imply that black immigration lowered tax revenues by more than 6% relative to their baseline value. However, public spending fell only by approximately 2%. One possible explanation for these patterns is that cities partly offset the fall in the tax base by raising debt. In Table 2.8, I directly test this idea, and show that the inflow of African Americans had a positive and significant effect on public debt per capita (column 1). Both gross and net debt went up, but black immigration had a significant impact only on the latter (columns 2 and 3).

The magnitude of the 2SLS coefficient in column 1 (Panel B) implies that, in response to the Great Migration, the level of indebtedness of the median city in my sample rose by approximately 4% relative to its 1910 level. Incidentally, this amount is close to the gap between the reduction in tax revenues and in public spending documented above, suggesting the existence of a causal link between changes in property taxes on the one hand and both public spending and debt on the other. Consistent with this conjecture, there is also a strong and positive relationship between the inflow of African Americans and the debt to tax revenues ratio (column 4), which increased by approximately 8% relative to its 1910 value because of black immigration.

[35]Before 1940, the US Census did not collect finer measures of education, like the number of years of schooling. An alternative explanation for findings in column 4 of Table A2.4 is that the reduction in spending on education is unlikely to immediately affect literacy rates.

One potential implication of findings in Table 2.8 is that, even if in the short run cities were able to partially cope with lower tax revenues by raising debt, over time this strategy may have proven unsustainable. In particular, depressed revenues and higher (and growing) levels of debt may have contributed to the distressed financial conditions and the heavy interest burdens that many US central cities systematically experienced in the past 40 years. Evaluating the long run effects of the Great Migration goes beyond the scope of this paper, but results in this section are consistent with Derenoncourt (2017), who shows that, today, intergenerational mobility is lower in MSAs that received larger inflows of African Americans between 1940 and 1970.

2.6.4 Unpacking the Effects of Black Migration

2.6.4.1 Demand for Public Goods or Fiscal Externality?

The inflow of African Americans had a negative and significant effect on tax revenues and public spending in northern cities. As shown in Tables 2.6 and 2.7, the reduction in public expenditures was not evenly distributed across categories, and was larger in education and in outdoor poverty relief. This observation is consistent with the idea that higher racial heterogeneity altered whites' demand for public goods and their preferences for redistribution (Alesina et al., 1999; Dahlberg et al., 2012). When deciding how to allocate the budget, (white) voters chose to undertake larger spending cuts in categories where inter-racial interactions are likely to be particularly salient, like education.[36]

However, even if the budget allocation decision was affected by black inflows, consistent with Boustan (2012), the evidence presented in previous sections suggests that the reduction in public spending was triggered more by a fiscal externality than by lower demand for public goods. First, Table 2.4 (columns 5 to 7) shows that the Great Migration reduced property values in receiving cities which, unable (or unwilling) to increase taxes (column 4), were forced to cut spending (Table 2.6) and increase debt (Table 2.8) to deal with a tighter budget constraint.

Second, as noted in Section 2.6.3, the reduction in tax revenues was an order of magnitude larger than that in public spending, as cities raised debt to cope with declining revenues. This observation suggests that the direction of causality runs from deteriorating finances to spending cuts rather than, on the opposite, from lower demand for public goods to lower taxes. Finally, once again consistent with a fiscal externality, the estimated

[36] The large drop in public spending on education is also in line with the existing literature on racial diversity and school choice in the US (e.g. Baum-Snow and Lutz, 2011; Cascio and Lewis, 2012).

reduction in property values, tax revenues, and public spending was larger when not accounting for the white flight (Table 2.5).

2.6.4.2 Further, Suggestive Evidence on the White Flight

The previous discussion supports the interpretation that black immigration reduced property values by either inducing whites to flee to the suburbs or simply increasing their willingness to do so. This idea is consistent with findings in Boustan (2010), who shows that the inflow of African Americans during the Great Migration induced whites to leave central cities and move to the suburbs. If more than one white left for each new black migrant, the reduction in housing demand by natives more than offset the increase in the latter due to immigration, thereby lowering house prices. Moreover, since housing is a durable good, any negative shock to housing demand will have a larger effect on house prices, relative to an equivalent positive shock (Glaeser and Gyourko, 2005).

In Table 2.9, I provide additional evidence that the Great Migration increased whites' demand for suburbanization. First, the growth in the number of dwellings (scaled by 1900 population) was lower in cities that received more blacks (column 1).[37] This pattern is consistent with Boustan (2010), who finds that the second wave of the Great Migration slowed down the construction of new housing units and increased the number of vacant dwellings in northern cities. Second, in places receiving more blacks, the increase in city-area was more limited (column 2). As explained in Jackson (1985), during this period, most cities expanded their boundaries by annexing neighboring towns and villages. If whites moved to the suburbs so as not to pay for and share public goods with blacks, they should have also tried to resist annexation by central cities. Findings in column 2 are indeed consistent with this idea.

Next, in columns 3 to 5, I study if the first wave of the Great Migration increased demand for suburbanization by analyzing its effect on the construction of highways since the 1950s (Baum-Snow, 2007). The main regressors of interest are the 1910-1930 (instrumented) change in black and white population, while the dependent variable is the number of highway rays running through the central city planned in the Federal Highway Act of 1944 (column 3) and actually built between 1950 and, respectively, 1970 (column 4) and 2000 (column 5).[38] There is a positive and significant relationship between 1910-1930 black immigration and the subsequent planning and construction of highways. While only

[37] Data on the number of dwellings was missing for Milwaukee in 1930.

[38] As in all previous Tables, these regressions also account for possible differential trends associated with the variables included in X_{crt} above, i.e., 1900 fraction of blacks; latitude and longitude; and geography. Data on highways were taken from Baum-Snow (2007), and were not available for Bridgeport (CT).

suggestive, these results are consistent with the idea that the inflow of blacks increased demand for suburbanization, ultimately leading to the construction of more highways in places where incentives to leave the central city were higher.

Finally, in columns 6 and 7, the 1940-1970 change in the number of special districts and municipalities in the MSA (collected from the Census of Governments) is regressed against the 1910-1930 (instrumented) change in black and white population.[39] As discussed in Alesina et al. (2004), racial heterogeneity may increase whites' desire for political fragmentation and, in turn, the number of local governments (see also Burns, 1994). Consistently with this idea, the 1910-1930 inflow of African Americans is positively correlated with the subsequent change in the number of special districts and municipalities. These findings are also in line with Boustan (2017), who argues that many whites were leaving central cities not necessarily to avoid inter-racial interactions in the housing or in the labor market, but in order to avoid sharing public goods with black immigrants.

2.7 Robustness Checks

In this section, I perform several robustness checks. First, I test for pre-trends and perform a placebo exercise to assess the validity the instrument for black migration (Section 2.7.1). Second, I document that results are robust to interacting year dummies with a number of 1900 city characteristics that may be correlated with pre-migration blacks' settlements and with changes in the economic and political environment (Section 2.7.2). Finally, in Section 2.7.3, I replicate the analysis trimming the sample at the 1^{st} and 99^{th} and at the 5^{th} and 95^{th} percentiles of black population, and instrumenting white population with the interaction between predicted black immigration and average low January temperature, a variable often used in the literature to proxy for city-level amenities (Glaeser and Shapiro, 2003; Glaeser and Gyourko, 2005).

2.7.1 Placebo Checks and Pre-Trends

As discussed in Section 2.5.2, one possible concern with the instrument for black population, Z_{crt} in (11), is that early migrants settled in booming northern cities that continued to grow more also in subsequent decades. To indirectly address this issue, I predicted white inflows by first estimating outmigration from each southern state for each decade using the forward survival method (exactly as done when constructing the instrument for

[39] Data limitations prevent me from examining the contemporaneous effect of black immigration on this outcome, since the number of local governments at the county level was first reported in the Census of Governments of 1942.

black immigration), and then assigning such migration flows to northern cities according to 1900 blacks' settlements, α_{jc}^{1900} in (11).

Results of this test are reported in Figure A2.2, which plots the relationship between white population in northern cities and southern whites' immigration predicted using 1900 blacks' settlements, after partialling out controls from my most preferred specification. There is no relationship between white population and white immigration from the South, predicted using African Americans' enclaves. While Figure A2.2 provides only indirect evidence on the validity of the instrument, it is reassuring to find that blacks' settlements have predictive power only for black (Figure 2.5), and not for white, immigration to northern cities.

Next, in Tables A2.5 and A2.6, I formally test if pre-migration changes in city characteristics and in the outcomes of interest predict subsequent changes in (instrumented) black immigration. In column 1 of Table A2.5, I address the possibility that the instrument for black migration may be correlated with the 1900-1910 change in European immigration. This would be problematic because the drop in European immigration after 1915 was an important determinant of the Great Migration (Collins, 1997), and one may thus be worried that results presented above are at least in part driven by changes in the number of international immigrants. However, there is no correlation between the 1900-1910 change in European immigration and the 1910-1930 change in predicted black migration.

Similarly, columns 2 to 5 show that the 1900-1910 growth in industrial production, establishment size, and wages in manufacturing is not correlated with 1910-1930 predicted black immigration. This is another important check, since most African Americans living in the North were employed in manufacturing, and one may have been worried that early movers selected cities where this sector was booming, in turn threatening the validity of the instrument constructed in Section 5.2.

Finally, in Table A2.6, I regress pre-period changes in the key outcomes considered in Section 2.6 against 1910-1930 predicted black immigration and white population. In Panel B, the dependent variable is the 1906-1910 change in per capita: total and property tax revenues (columns 1 and 2); property values (column 3); and total and education spending (columns 4 and 5).[40] To ease comparison with findings reported in Section 2.6, Panel A of Table A2.6 reports my baseline specification. The picture that emerges is reassuring: while there is a positive and significant correlation between the 1906-1910

[40]When performing this exercise, 1906 is used because this is the first year in which the Financial Statistics of Cities collected data in a way that is comparable to subsequent years.

change in tax revenues and the 1910-1930 change in white population (columns 1 and 2), there is no relationship between changes in pre-migration outcomes and subsequent changes in black immigration. Also, and importantly, magnitudes are very different from those in the baseline specification.

2.7.2 Pre-Migration City Characteristics and Differential Trends

Yet another concern is that the characteristics of cities that attracted black migrants before 1900 persisted over time, and had confounding effects on migration patterns as well as on the evolution of economic and political conditions. My most preferred specification already controls for interactions between year dummies and the 1900 fraction of blacks, latitude and longitude, and geography. However, in Table A2.7, I more systematically check that interacting year dummies with additional socio-economic variables does not alter any of the main findings presented in Section 2.6. In Panel A, I start by replacing latitude and longitude with a measure for distance from the South, and results remain very close to those from my baseline specification.[41]

Next, as noted in the previous section, cities that received more European immigrants between 1870 and 1915 may have attracted more blacks once the Immigration Acts of the 1920s were passed (Collins, 1997), and, at the same time, may have experienced differential changes in tax revenues or public goods provision (see chapter 1). For this reason, in Panel B, I include in X_{crt} interactions between year dummies and the 1900 fraction of Europeans. Similarly, one may be worried that the skill and industry composition of cities affected both blacks' settlement patterns and the evolution of city finances. To deal with these concerns, Panels C, D, and E interact year dummies with, respectively, the ratio of high to low skilled workers, value added by manufacture per establishment, and the employment share in manufacturing. Coefficients are very stable across specifications, and in all cases results are barely affected.

2.7.3 Trimmed Sample and Alternative Instrument

As discussed in Section 2.5.1, one potential remaining concern is that a specification in levels, as the one used in my paper (as well as in Boustan, 2010, and in Shertzer and Walsh, 2016), implicitly weighs larger cities more. To deal with this issue, following Shertzer and Walsh (2016), I repeated my analysis by trimming the sample at the 1^{st} and 99^{th} and the 5^{th} and 95^{th} percentiles of black population. Results are reported in Table A2.8 (Panels

[41] Distance from the South was constructed by computing the distance between each northern city in my sample and the capital of the southern state that, by 1900, had sent more blacks to that city.

A and B respectively).[42] Not surprisingly, the KP and AP F-stats are somewhat lower than in the baseline specification.

However, second stage results are in line with those reported in Section 2.6. The only difference appears to be in column 3, where the dependent variable is the property tax rate. Yet, standard errors are so large that the coefficient in either Panel A or Panel B is not statistically different from that in column 4 of Table 2.4. As a further check, I also replicated Table A2.8 dropping, in turn, each of the five largest cities in my sample (New York, Chicago, Philadelphia, Detroit, and Los Angeles), and findings were not significantly affected. Overall, while one cannot completely rule out the possibility that results may be somewhat driven by the largest cities in my sample, Table A2.8 should reduce this concern.

As a final check, I repeated my analysis using a different instrument for the white flight. Specifically, when instrumenting W_{crt} in (10), I interacted predicted black immigration with average low January temperature, rather than with the geographic variables used above.[43] As discussed in Glaeser and Gyourko (2005) among others, weather has been shown to affect residential decisions in the United States, and several papers have used it as a proxy for city-level amenities (e.g. Glaeser and Shapiro, 2003). Starting from this observation, I constructed an alternative instrument, interacting Z_{crt} with average low January temperature, to check whether findings in the paper are driven in some way by the geographic variables used in the baseline specification. Reassuringly, results, reported in Table A2.9, show that this is not the case.

2.8 Conclusions

Between 1915 and 1930, more than 1.5 million African Americans moved from the rural South to the urban North of the United States. I exploit this historical episode to study how the arrival of blacks, and the resulting increase in racial heterogeneity, affected government finances and public spending in northern cities. Instrumenting both black inflows and the induced white outflows, I find that immigration of African Americans had a strong, negative effect on tax revenues, entirely due to declining property values. Unable or unwilling to raise tax rates, cities were forced to cut spending and raise debt to cope with a tighter budget constraint. While higher debt made it possible to only partially

[42]Panels A and B omit, respectively, New York City and Duluth, and New York City, Philadelphia, Washington D.C., Duluth, San Diego, and Scranton.

[43]Data on January average temperature for each city was collected from http://www.usclimatedata.com/climate/united-states/us.

reduce public goods provision, this strategy may have proven unsustainable in the long run, contributing to the current financial distress of several US central cities.

Findings in this paper provide motivation for future work along several directions. First, I only focused on the contemporaneous effects of black immigration, but a natural extension would be that of considering the medium to long run consequences of the Great Migration on both city finances and public goods provision. Second, it may be interesting to compare the Great Migration with other episodes of (im)migration experienced by the United States, such as the Mass Migration of Europeans between 1870 and 1915 or the more recent Hispanic immigration. Finally, it would be particularly informative for the current situation in both Europe and the US to investigate if the Great Migration fueled natives' backlash and favored the election of "anti-black" mayors, and if changes in public spending documented in my work were correlated with the identity of elected politicians in northern cities.

References

[1] Alesina, A., R. Baqir, and W. Easterly (1999). "Public Goods and Ethnic Divisions". *Quarterly Journal of Economics, 114 (November)*, 1243–84.

[2] Alesina, A., R. Baqir, and C. Hoxby (2004). "Political Jurisdictions in Heterogeneous Communities". *Journal of Political Economy 112*, 348-96.

[3] Alesina, A., E. L. Glaeser, and B. Sacerdote (2001). "Why Doesn't the United States Have a European-Style Welfare State?" *Brookings Papers Econ. Activity, 2*, 187–278.

[4] Alesina, A. and E. L. Glaeser (2004). *Fighting Poverty in the US and Europe: A World of Difference.* Oxford University Press, Oxford UK.

[5] Allard, S. W. (2017). *Places in Need: The Changing Geography of Poverty.* Russell Sage Foundation.

[6] Anderson, H., H. Berg, M. Dahlberg (2016). "Migrating Natives and Foreign Immigration". Working Paper.

[7] Baum-Snow, N. (2007). "Did Highways Cause Suburbanization?". *The Quarterly Journal of Economics, 122(2)*, 775-805.

[8] Baum-Snow, N., and B. F. Lutz (2011). "School Desegregation, School Choice, and Changes in Residential Location Patterns by Race". *American Economic Review, 101(7)*, 3019-3046.

[9] Beach, B. and D. Jones (2017). "Gridlock: Ethnic Diversity in Government and the Provision of Public Goods." *American Economic Journal: Economic Policy, 9(1)*, 112-136.

[10] Black, D. A., S. G. Sanders, E. J. Taylor, and L. J. Taylor (2015). "The Impact of the Great Migration on Mortality of African Americans: Evidence from the Deep South." *American Economic Review, 105(2)*, 477-503.

[11] Boustan, L. P. (2009). "Competition in the Promised Land: Black Migration and Racial Wage Convergence in the North, 1940-1970". *The Journal of Economic History, 69(3)*, 756-783.

[12] Boustan, L. P. (2010). "Was Postwar Suburbanization 'White Flight?' Evidence from the Black Migration." *Quarterly Journal of Economics, 125(1)*, 417–443.

[13] Boustan, L. P. (2012). "School Desegregation and Urban Change: Evidence from City Boundaries." *American Economic Journal: Applied Economics, 4(1)*, 85-108.

[14] Boustan, L. P. (2013). "Local public goods and the demand for high-income municipalities." *Journal of Urban Economics, 76*, 71-82.

[15] Boustan, L. P. (2017). *Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets.* Princeton University Press.

[16] Boustan, L. P., M. E. Kahn, and P. W. Rhode (2012). "Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century." *American Economic Review, 102(3)*, 238–244.

[17] Burchfield, M., H. G. Overman, D. Puga, and M. A. Turner (2006). "Causes of Sprawl: A Portrait from Space". *The Quarterly Journal of Economics, 121 (2)*, 587-633.

[18] Burns, N. (1994). *The Formation of American Local Governments: Private Values in Public Institutions.* Oxford: Oxford University Press.

[19] Card, D. (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics, 19(1)*, 22-64.

[20] Card, D., C. Dustmann, and I. Preston (2012). "Immigration, Wages, and Compositional Amenities". *Journal of the European Economic Association, 10*, 78-119.

[21] Card, D., A. Mas, and J. Rothstein (2008). "Tipping and the Dynamics of Segregation". *Quarterly Journal of Economics, 123 (1)*, 177-218.

[22] Cascio, E. U., and E. G. Lewis (2012). "Cracks in the Melting Pot: Immigration, School Choice, and Segregation". *American Economic Journal: Economic Policy, 4(3)*, 91-117.

[23] Chay, K. and K. Munshi (2015). "Black Networks After Emancipation: Evidence from Reconstruction and the Great Migration". Working Paper.

[24] Clubb, J. M., W. H. Flanigan, and N. H. Zingale. "Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972". ICPSR08611-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-11-13. http://doi.org/10.3886/ICPSR08611.v1

[25] Collins, W. J. (1997). "When the Tide Turned: Immigration and the Delay of the Great Black Migration". *Journal of Economic History, 57(3)*, 607-632.

[26] Collins, W. J., and M. H. Wanamaker (2014). "Selection and Income Gains in the Great Migration of African Americans." *American Economic Journal: Applied, 6(1)*, 220-252.

[27] Collins, W. J., and M. H. Wanamaker (2015). "The Great Migration in Black and White: New Evidence on the Selection and Sorting of Southern Migrants". *The Journal of Economic History, 75(4)*, 947-992.

[28] Cutler, D.M., E.L. Glaeser and J.L. Vigdor (1999). "The Rise and Decline of the American Ghetto". *Journal of Political Economy, 107*, 455-506.

[29] Dahlberg, M., K. Edmark, and H. Lundqvist (2012). "Ethnic Diversity and Preferences for Redistribution". *Journal of Political Economy, 120(1)*, 41-76.

[30] Danielson, M. N. (1976). *The Politics of Exclusion.* New York: Columbia University Press.

[31] Derenoncourt, E. (2017). "Historical Origins for Intergenerational Mobility Differences across US cities Today". Working Paper.

[32] Dustmann, C., K. Vasiljeva, and A. Piil Damm (2016). "Refugee Migration and Electoral Outcomes". Discussion Paper Series, CDP 19/16.

[33] Feigenbaum, J. J., and C. Muller (2016). "Lead Exposure and Violent Crime in the Early Twentieth Century". *Explorations in Economic History*, forthcoming.

[34] Fisher, G. W. (1996). *The Worst Tax? A History of the Property Tax in America.* Lawrence: University Press of Kansas.

[35] Foged, M. and G. Peri (2016). "Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data". *American Economic Journal: Applied Economics 2015, 8(2)*, 1–34.

[36] Fryer R.G. and L. F. Katz (2013). "Achieving Escape Velocity: Neighborhood and School Interventions to Reduce Persistent Inequality". *American Economic Review Papers and Proceedings, 103(3)*, 232-37.

[37] Geddes, A. E. (1937). *Trends In Relief Expenditures, 1910-1935.* Washington: U.S. Govt. Print. Off.

[38] Glaeser, E. L. and J. Gyourko (2005). "Urban Decline and Durable Housing". *Journal of Political Economy, 113 (2)*, 345-375.

[39] Glaeser, E. L., and J. M. Shapiro (2003). "Urban Growth in the 1990s: Is City Living Back?". *Journal of Regional Science, 43(February)*, 139–65.

[40] Gregory, J. N. (2005). *The Southern Diaspora: How the Great Migrations of Black and White Southerners Transformed America.* Chapel Hill: University of North Carolina Press.

[41] Haines, M. R. (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]. doi:10.3886/ICPSR02896.v3.

[42] Halla, M., A. F. Wagner, and J. Zweimuller (2017). "Immigration and Voting for the Far Right". *Journal of the European Economic Association*, Forthcoming.

[43] Harari M. (2016). "Cities in Bad Shape: Urban Geometry in India". Working Paper.

[44] Hornbeck, R. and S. Naidu (2014). "When the Levee Breaks: Black Migration and Economic Development in the American South". *American Economic Review, 104(3)*, 963–990.

[45] Jackson, K. T (1985). *Crabgrass Frontier: The Suburbanization of the United States.* Oxford University Press.

[46] Katz, L. F. (2015). "Reducing Inequality: Neighborhood and School Interventions". *Focus 31*, 2, (Fall/Winter 2014-15), 12-17.

[47] Kneebone, E. (2017). "The changing geography of US poverty." Brookings Institutions, February 2017.

[48] Lange, F., A. L. Olmstead, and P. W. Rhode (2009). "The Impact of the Boll Weevil, 1892-1932." *Journal of Economic History, 69(3)*, 685-718.

[49] Luttmer, E. F. P (2001). "Group Loyalty and the Taste for Redistribution" *Journal of Political Economy,109 (3)*, 500-528.

[50] Mayda A. M., G. Peri, and W. Steingress (2016). "Immigration to the US: A Problem for the Republicans or the Democrats?". NBER Working Paper 21941.

[51] Monkkonen, E. (1988). *America Becomes Urban: The Development of U.S. Cities and Towns, 1780-1980.* Berkeley: University of California Press.

[52] Naidu, S. (2012). "Suffrage, Schooling, and Sorting in the Post-Bellum US South." NBER Working Paper 18129.

[53] National Research Council (1999). *Governance and Opportunity in Metropolitan America.* Washington, DC: The National Academies Press, 1999. doi:10.17226/6038.

[54] Nekby, L. and P. Pettersson-Lidbom (2017). "Revisiting the Relationship between Ethnic Diversity and Preferences for Redistribution: Comment". *The Scandinavian Journal of Economics, 119(2)*, 268-287.

[55] Ortega, F. (2005). "Immigration Quotas and Skill Upgrading". *Journal of Public Economics, 89*, 1841-1863.

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

[57] Ruggles, S., M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. Kelly Hall, M. King, and C. Ronnander (2010). Integrated Public Use Microdata Series: Version 4.0 [Machine-readable database]. Minneapolis, MN: The University of Minnesota.

[58] Saiz, A. (2010). "The Geographic Determinants of Housing Supply". *Quarterly Journal of Economics, 125 (3)*, 1253-1296.

[59] Saiz. A. and S. M. Wachter (2011). "Immigration and the Neighborhood". *American Economic Journal: Economic Policy, 3*, 169-188.

[60] Shertzer, A. and R. Walsh (2016). "Racial Sorting and the Emergence of Segregation in American Cities". NBER Working Paper 22077.

[61] Tabellini, M. (2017). "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration". Working Paper.

[62] Tolnay, S. and E. M. Beck (1990). "Black Flight: Lethal Violence and the Great Migration,1900 to 1930". Social Science History, 14(3), 347-370.

[63] Wilkerson, I. (2010). *The Warmth of Other Suns: The Epic Story of America's Great Migration.* Random House LLC, 2010.

Table 2.1. Central Cities-Suburbs Inequality (2010)

	Central Cities	Suburbs	Central City to Suburban Outcome
Non-Hispanic Whites	56.9%	77.2%	0.74
Blacks	26.7%	11.6%	2.04
Hispanics	24.8%	16.3%	1.52
Poverty rate, families	22.1%	10.6%	2.08
Unemployed men (age 25-65)	10.8%	8.6%	1.26
High school graduates (men, age 25-65)	83.4%	90%	0.92
Median wage (employed men, age 25-65)	$33,300	$40,000	0.83

Note: Author's calculation from Census Bureau (2010).

Table 2.2. Summary Statistics

VARIABLES	*Mean*	*Median*	*St. Dev.*	*Min*	*Max*	*Obs.*
City population	537,344	237,395	974,513	39,578	6,930,446	126
White population	506,340	229,755	931,592	38,465	6,587,225	126
Black population	27,537	5,568	50,851	410	327,706	126
Predicted black population	9,670	2,350	19,613	0	141,200	126
Total tax revenues PC	19.81	16.51	9.582	7.091	53.98	126
Property tax revenues PC	18.49	15.73	8.560	6.215	44.70	126
Tax rate per 1,000$ of assessed valuation	29.74	25.03	16.78	10.39	100.7	126
Property values PC	1,255	1,117	644.4	189.2	3,769	126
Expenditures PC	17.88	15.35	7.929	6.640	43.64	126
Educ. spending PC	6.199	4.830	2.947	1.900	14.76	126
Police spending PC	1.841	1.596	0.925	0.492	4.693	126
Total gross debt PC	73.43	64.22	46.56	0.125	257.4	126

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration, as discussed in the main text. *Predicted blacks* in the fourth row of the Table refers to predicted black immigration, i.e. the instrument for black population constructed in Section 5.2. All spending and revenues data are expressed in 1910 dollars. Nominal values for 1920 and 1930 are deflated using the CPI from the Minneapolis FED.

Table 2.3. First Stage

VARIABLES	(1) Blacks	(2) Blacks	(3) Whites	(4) Blacks	(5) Whites
Z	1.916***	1.552***	-1.875	1.719***	-0.118
	(0.350)	(0.556)	(1.796)	(0.514)	(1.434)
Z_water		0.047***	0.471***	0.038**	0.370***
		(0.012)	(0.034)	(0.015)	(0.083)
Z_hills		-0.037	0.483***	-0.035	0.513***
		(0.031)	(0.141)	(0.029)	(0.131)
Z_rivers		-0.340***	0.501	-0.325**	0.711
		(0.115)	(0.671)	(0.130)	(0.764)
KP F-stat	30.30	13.40	13.40	15.13	15.13
AP F-stat		189.9	606.3	73.80	147.6
Cities	42	42	42	42	42
Observations	126	126	126	126	126

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. Col 1 reports results for a regression of the number of blacks in the city on predicted black immigration (Z), city and year by region fixed effects. Cols 2 to 5 report results for regressions (4) and (5) in the main text. Cols 2 and 3 only include city and region by year fixed effects, while Cols 4 and 5 also include the full set of interactions between year dummies and geography. The dependent variable is, respectively, the number of blacks (Cols 2-4) and the number of whites (Cols 3-5) in the central city. The regressors of interest are predicted black immigration (Z) and its interaction with the share of the area around the central city: i) occupied by lakes and oceans (*Z_water*); ii) with slope above 15% (*Z_hills*); iii) occupied by rivers and streams (*Z_rivers*). AP F-stat refers to the partial F-stat for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2.4. Tax Revenues, Tax Rate, and Property Values

VARIABLES	(1) Tot. tax revenues PC	(2) Prop. tax revenues PC	(3) Prop. tax revenues PC	(4) Prop. tax rate	(5) Prop. values over 1900 population	(6) Prop. values over 1910 population	(7) Prop. values PC
Panel A: OLS							
Blacks	-0.085**	-0.067*	-0.081**	0.023	-14.51	-4.543	-1.216
	(0.041)	(0.035)	(0.041)	(0.064)	(12.23)	(4.822)	(3.643)
Whites	0.011*	0.008	0.009	-0.006	2.523	0.785	0.286
	(0.006)	(0.005)	(0.006)	(0.009)	(2.263)	(0.885)	(0.506)
Panel B: 2SLS							
Blacks	-0.101***	-0.093***	-0.110***	0.098	-18.12**	-9.059**	-5.796*
	(0.033)	(0.029)	(0.035)	(0.067)	(8.738)	(4.190)	(3.374)
Whites	0.012***	0.010***	0.011**	-0.016	1.990	0.969	0.766
	(0.004)	(0.004)	(0.005)	(0.010)	(1.246)	(0.616)	(0.503)
KP F-stat	13.40	13.40	11.44	11.44	11.44	11.44	11.44
AP (Blacks)	189.9	189.9	104.5	104.5	104.5	104.5	104.5
AP (Whites)	606.3	606.3	254.9	254.9	254.9	254.9	254.9
Mean dep var	15.63	15.16	15.16	23.12	1,308	860.1	860.1
Observations	126	126	126	126	126	126	126

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. Panels A and B report, respectively OLS and 2SLS results. The dependent variable is displayed at the top of each column, and refers to real 1910 dollars. The regressors of interest are the number of blacks (*Blacks*) and whites (*Whites*) in the central city, and are instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3 in the main text). AP (blacks) and AP (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. First stage results are reported in Table 3. All columns also report the mean of the dependent variable at baseline. All regressions control for city and year by region fixed effects. Cols 3 to 7 also include interactions between year dummies and dummies for: i) 1900 fraction of blacks; ii) latitude and longitude; and iii) the presence of geographic features around the central city. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2.5. Taxes and Spending when Omitting the White Flight

VARIABLES	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Property values over 1900 pop	(5) Total spending PC	(6) Spending on education PC
Panel A: OLS						
Blacks	-0.116**	-0.094**	0.046	-18.96	-0.061*	-0.032**
	(0.046)	(0.045)	(0.077)	(13.62)	(0.034)	(0.012)
Panel B: 2SLS						
Blacks	-0.154**	-0.169**	0.030	-60.36***	-0.131**	-0.059***
	(0.076)	(0.068)	(0.094)	(18.57)	(0.064)	(0.020)
KP F-stat	9.347	9.347	9.347	9.347	9.347	9.347
Mean of dep var	15.63	15.16	23.12	1,308	15.21	4.656
Observations	126	126	126	126	126	126

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. Panels A and B report, respectively OLS and 2SLS results. The dependent variable is displayed at the top of each column, and refers to 1910 dollars. *Blacks* is the number of blacks in the central city, and is instrumented with predicted black immigration (see Sections 5.2). KP F-stat is the Kleibergen-Paap F stat for weak instruments. First stage results are reported in Table 3. All columns also report the mean of the dependent variable at baseline. All regressions control for MSA population, city and year by region fixed effects, and include interactions between year dummies and dummies for: i) 1900 fraction of blacks; and ii) latitude and longitude. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2.6. Public Spending Per Capita

VARIABLES	(1) Total	(2) Education	(3) Police	(4) Fire	(5) Sewerage and garbage	(6) Road	(7) Charities and hospitals	(8) Outdoor poverty relief
Panel A: OLS								
Blacks	-0.053*	-0.028**	-0.001	-0.002	-0.004	-0.003	-0.011**	-0.013***
	(0.028)	(0.011)	(0.004)	(0.002)	(0.003)	(0.006)	(0.005)	(0.004)
Whites	0.007	0.003*	0.001	0.000	0.001**	-0.000	0.001	0.002**
	(0.006)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Panel B: 2SLS								
Blacks	-0.041*	-0.041***	-0.007**	-0.005	-0.008***	-0.006	0.001	-0.010*
	(0.024)	(0.012)	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.006)
Whites	0.003	0.005***	0.001***	0.000	0.002***	-0.000	-0.001	0.001
	(0.004)	(0.002)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
KP F-stat	11.44	11.44	11.44	11.44	11.44	11.44	11.99	9.327
AP (Blacks)	104.5	104.5	104.5	104.5	104.5	104.5	95.50	7.62
AP (Whites)	254.9	254.9	254.9	254.9	254.9	254.9	239.3	41.58
Mean dep var	15.21	4.656	1.812	1.801	1.802	1.043	1.892	0.144
Observations	126	126	126	126	126	126	122	42

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration, as discussed in the main text. Panels A and B report, respectively OLS and 2SLS results. The dependent variable is displayed at the top of each column, and refers to 1910 dollars. The regressors of interest are the number of blacks (*Blacks*) and whites (*Whites*) in the central city, and are instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3 in the main text). F-stat (blacks) and F-stat (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. Data on spending for outdoor poverty relief (Col 8) were digitized from Geddes (1937), and are available only for the 14 largest cities in my sample. All columns report the mean of the dependent variable at baseline. All regressions include city and year by region fixed effects, and, except for Col 8, interactions between year dummies and dummies for: i) 1900 fraction of blacks; ii) latitude and longitude; and iii) the presence of geographic features around the central city. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2.7. Spending Shares

VARIABLES	(1) Education	(2) Police	(3) Fire	(4) Sewerage and garbage	(5) Road	(6) Charities and hospitals
Panel A: OLS						
Blacks	0.001	-0.003	0.029***	-0.002	0.015	-0.035***
	(0.019)	(0.010)	(0.008)	(0.010)	(0.019)	(0.010)
Whites	-0.002	0.002	-0.002*	0.002	-0.004	0.004
	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Panel B: 2SLS						
Blacks	-0.078*	-0.029*	0.031***	-0.031*	0.012	-0.010
	(0.044)	(0.017)	(0.011)	(0.016)	(0.032)	(0.015)
Whites	0.011*	0.005*	-0.003	0.007***	-0.003	-0.002
	(0.007)	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)
KP F-stat	11.44	11.44	11.44	11.44	11.44	11.99
AP (Blacks)	104.5	104.5	104.5	104.5	104.5	95.50
AP (Whites)	254.9	254.9	254.9	254.9	254.9	239.3
Mean dep var	31.87	11.77	12.31	6.919	12.29	4.796
Observations	126	126	126	126	126	122

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. Panels A and B report, respectively OLS and 2SLS results. The dependent variable, displayed at the top of each column, is the share of spending in a given category over total current expenditures. The regressors of interest are the number of blacks (*Blacks*) and whites (*Whites*) in the central city, and are instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3). F-stat (blacks) and F-stat (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All columns also report the mean of the dependent variable at baseline. All regressions include city and year by region fixed effects and interactions between year dummies and dummies for: i) 1900 fraction of blacks; ii) latitude and longitude; and iii) the presence of geographic features around the central city. Robust standard errors, clustered at the MSA level, in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 2.8. Public Debt Per Capita

VARIABLES	(1) Total gross debt	(2) Current gross debt	(3) Net debt	(4) Debt to tax revenues ratio
Panel A: OLS				
Blacks	0.474***	0.125	0.203	0.026**
	(0.167)	(0.122)	(0.150)	(0.011)
Whites	-0.031	-0.008	-0.026	-0.002
	(0.039)	(0.017)	(0.034)	(0.002)
Panel B: 2SLS				
Blacks	0.572***	0.238	0.362**	0.050***
	(0.182)	(0.174)	(0.149)	(0.013)
Whites	-0.076***	-0.027	-0.079***	-0.007***
	(0.025)	(0.023)	(0.019)	(0.002)
KP F-stat	11.44	11.44	11.44	11.44
AP (Blacks)	104.5	104.5	104.5	104.5
AP (Whites)	254.9	254.9	254.9	254.9
Mean dep var	67.09	10.58	47.11	4.222
Observations	126	126	126	126

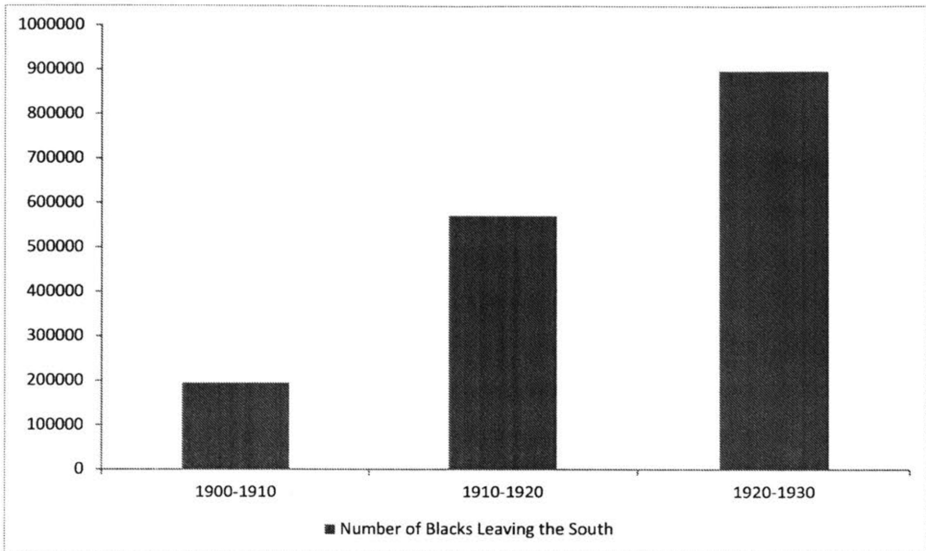
Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. Panels A and B report, respectively OLS and 2SLS results. The dependent variable, expressed in 1910 dollars, is displayed at the top of each column. All outcomes refer to per capita values. In Col 4, the dependent variable is the ratio of total debt to total tax revenues. The regressors of interest are the number of blacks (*Blacks*) and whites (*Whites*) in the central city, and are instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3). AP (blacks) and AP (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All columns also report the mean of the dependent variable at baseline. All regressions include city and year by region fixed effects and interactions between year dummies and dummies for: i) 1900 fraction of blacks; ii) latitude and longitude; and iii) the presence of geographic features around the central city. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2.9. Additional Evidence on the White Flight

VARIABLES	(1) Housing units	(2) City area	(3) Planned rays (1944)	(4) Actual rays (1970-1950)	(5) Actual rays (2000-1950)	(6) Municipalities (1970-1940)	(7) Special districts (1970-1940)
Panel A: OLS							
Blacks	-4.080	-1.464	0.016**	0.028***	0.029***	0.128	1.665***
	(2.589)	(0.945)	(0.008)	(0.008)	(0.008)	(0.148)	(0.215)
Whites	0.641	0.268	-0.001	-0.002*	-0.002*	0.001	-0.147***
	(0.448)	(0.176)	(0.001)	(0.001)	(0.001)	(0.021)	(0.032)
Panel B: 2SLS							
Blacks	-5.720**	-2.246*	0.025**	0.041***	0.044***	0.371**	1.578***
	(2.522)	(1.298)	(0.011)	(0.010)	(0.010)	(0.168)	(0.288)
Whites	0.669*	0.308	-0.002	-0.003**	-0.004***	-0.030	-0.120**
	(0.367)	(0.197)	(0.001)	(0.001)	(0.001)	(0.026)	(0.049)
KP F-stat	11.79	11.79	11.27	11.27	11.27	11.39	11.39
AP (Blacks)	102.6	102.6	63.35	63.35	63.35	66.99	66.99
AP (Whites)	207.0	207.0	197.4	197.4	197.4	216.8	216.8
Mean dep var	268.5	122.9	3.268	0.146	0.146	38	28
Observations	125	125	41	41	41	42	42

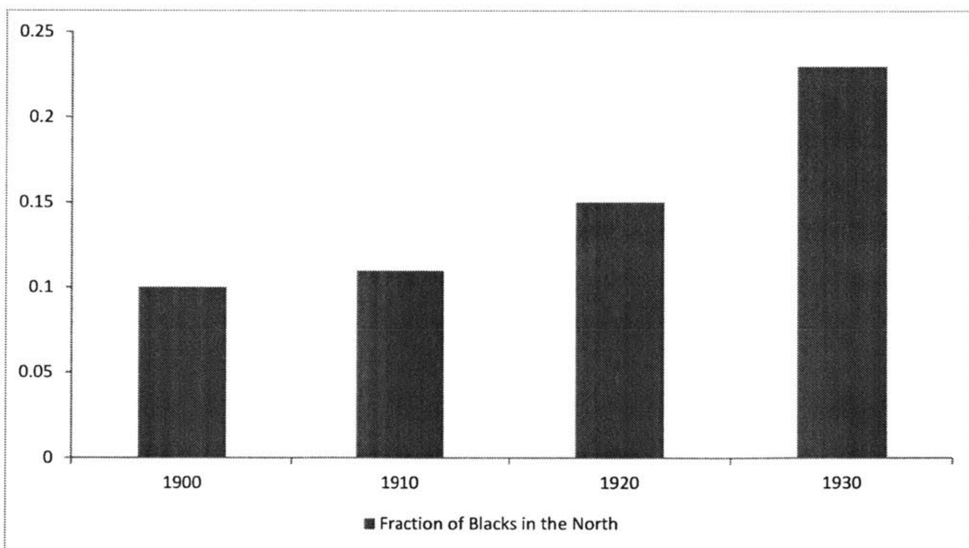
Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is: the number of housing units divided by 1900 population in Col 1; land area (in square km) in Col 2; the number of highway rays passing through the city planned in the Federal Highway Act of 1944 in Col 3; the 1950-1970 (resp. 1950-200) change in the number of highway rays passing through the city in Col 4 (resp. Col 5); the 1940-1970 change in the number of special districts (resp. municipalities) in Col 6 (resp. Col 7). The regressors of interest are the number of blacks (*Blacks*) and whites (*Whites*) in the central city in Cols 1 and 2, and their 1910-1930 change in the remaining columns. They are instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3). AP-stat (blacks) and AP-stat (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All columns also report the mean of the dependent variable at baseline. Cols 1 and 2 include city and year by region fixed effects, and interactions between year dummies and: i) 1900 fraction of blacks; ii) city coordinates; and iii) the presence of geographic features around the central city. Long difference regressions in Cols 3 to 7 partial out trends for: region; 1900 fraction of blacks; city coordinates; and geography. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Figure 2.1. Number of Black Migrants, by Decade



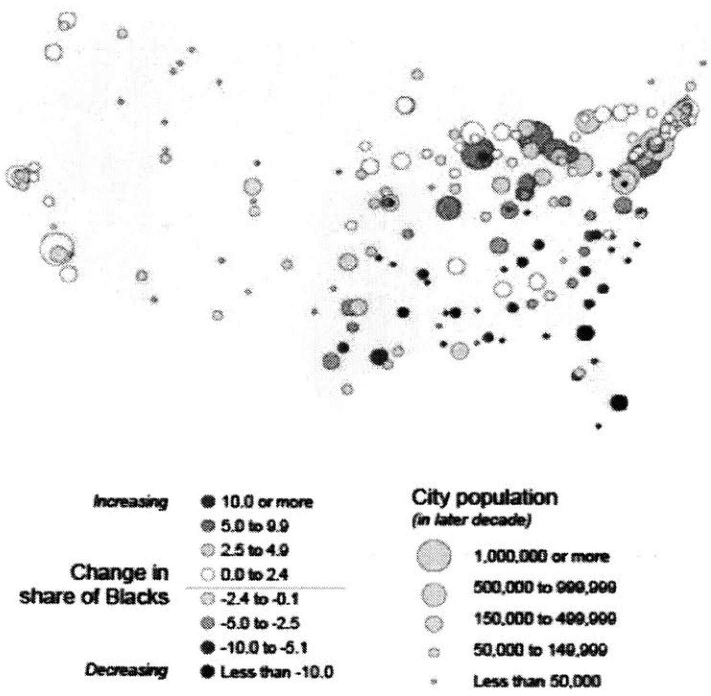
Note: Author's calculations using Census of Population data from IPUMS 1% samples for years 1900 to 1930. The number of net migrants from US Southern states is estimated using the forward survival method as in Gregory (2005). First, mortality rates are estimated by age-sex-race groups using national data from US Census of Population (1900-1930). Then, net migration for each Southern state is computed by adjusting changes in population (for each age-sex-race group) for estimated mortality rates. Finally, net migration for each southern state is aggregated for the South as a whole.

Figure 2.2. Fraction of African Americans Living in the North, by Decade



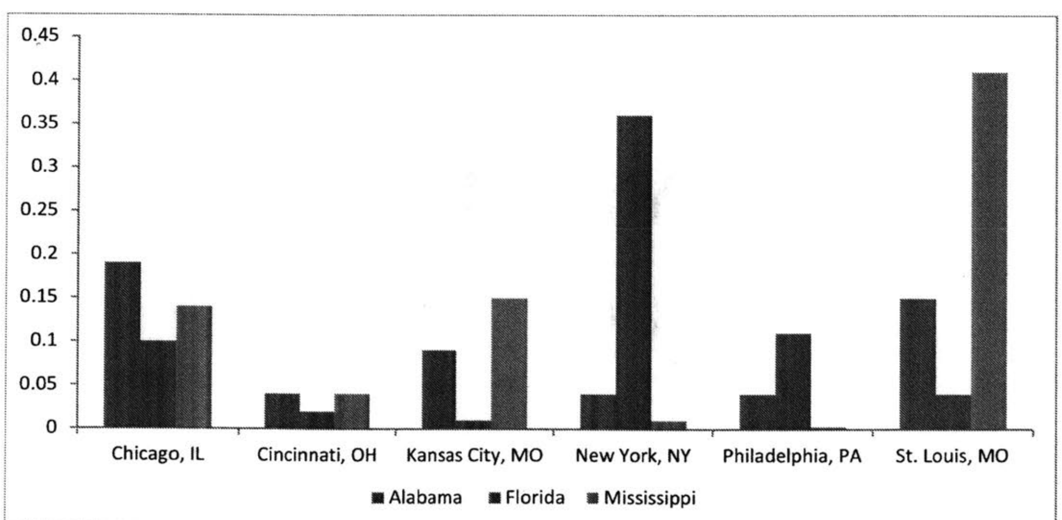
Note: The Figure plots the fraction of African Americans living in the North of the United States by decade. Author's calculations from US Census of Population (1900-1930).

Figure 2.3. The First Wave of the Great Migration



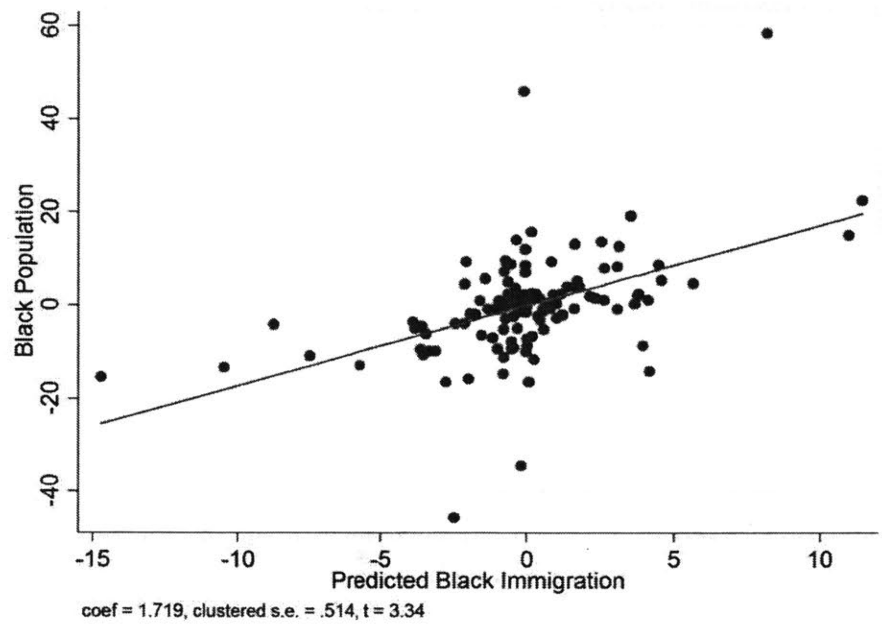
Note: The change of the share of Blacks in cities is based on the percentage point difference in the percent of the population that was black in 1940, relative to 1910. The Figure comes from the US Census Bureau.

Figure 2.4. Share of Blacks from Southern States in Northern Cities, 1900



Note: The Figure shows the fraction of southern born blacks from a given state residing in the North living in one of the selected northern cities in 1900. Author's calculation from the 1900 US Census of Population (5% sample).

Figure 2.5. First Stage: Actual vs Predicted Black Population



Note: The y-axis reports the actual number of blacks in northern cities in each decade between 1910 and 1930, and the x-axis shows the predicted number of black migrants, constructed as described in the text (Section 5.2). Each point in the scatter diagram represents the residual change in a city's actual and predicted number of blacks after partialling out: the interaction between predicted black immigration and the geographic variables listed in the main text (Section 5.3); city and year by region fixed effects; interactions between year dummies geography (see Section 5.3).

Chapter 3. Happily Ever After: Immigration, Natives' Marriage and Fertility

(Joint Work with Michela Carlana)

3.1 Introduction

Between 1970 and 2010, the number of foreign born individuals living in the United States increased from 9 to roughly 40 million, with the share of immigrants over US population skyrocketing from 4.7 to 13 percent (Figure 3.1).[1] As for previous immigration waves in American history, alongside these trends, a heated debate on the economic, social, and political consequences of immigration has emerged (Porter, 2017). A large body of the literature has investigated the economic effects of immigration, testing in particular if immigrants lower natives' wages and employment (Card, 2001; Borjas, 2003; Borjas and Katz, 2007). A more recent set of papers has studied how immigration affects political outcomes and electoral results (Mayda *et al.*, 2016; Halla *et al.*, 2017). Somewhat surprisingly, however, much less is known about the impact of immigration on key social outcomes, such as marriage rates, fertility, and family formation among natives.

In this paper, we study how the inflow of immigrants to US cities between 1910 and 1930 affected marriage and the probability of having children for young natives, as well as their decision to leave the parental house and set up an independent family unit. We provide evidence that these effects were operating mainly through changes in labor market opportunities for native men, which increased the supply of "marriageable men". A convenient feature of this setting is that, at the time, marriage markets were highly segmented along ethnic lines, and natives were unlikely to marry with foreign born individuals.[2] This, in turn, allows us to isolate the effects of immigration on natives' marriage and fertility, without confounding them with mechanical changes in the relative supply of men and women in the marriage market.

At the beginning of the twentieth century, 14 percent of the US population was foreign born, following the migration of more than 30 million Europeans between 1850 and 1915

[1] US immigration statistics are underestimated because of the presence of large numbers of undocumented immigrants. According to some recent estimates (see Pew Center 2017), if undocumented immigrants were included, the share of foreign born over US population would be at least 4 percentage points higher (i.e. around 17 percent).

[2] More than 95% of U.S. born women were married with U.S. born men in 1910. Indeed, in this historical period, according with Section 3 of the Expatriation Act of 1907, native women who married a foreigner lost their citizenship, and could get it back only when their husband was naturalized. Therefore, they were also *de iure* strongly disincentivized to marry an immigrant.

(Figure 3.2). After 1915, however, World War I and the Immigration Acts (1921 and 1924) put an end to the Age of Mass Migration and drastically reduced immigration to the US (Abramitzky and Boustan, 2017). The key feature of these shocks is that they had heterogeneous effects across European countries. Since immigrants tend to cluster geographically along ethnic lines (Card, 2001), variation across sending regions mechanically translates into variation in the number as well as in the mix of immigrants received by US cities over time.

Exploiting this variation, we construct a "leave-out" version of the classic shift-share instrument often adopted in the immigration literature (Altonji and Card, 1991; Card, 2001). In particular, we predict the number of immigrants to US cities in a given year by interacting the geographic variation in historical settlements of different ethnic groups with the time-series variation in national flows from each sending region, net of the individuals that eventually settled in a given city's metropolitan statistical area (MSA).[3] The key identifying assumption behind the instrument is that the city-specific characteristics that attracted early movers from each ethnic group must not have a time-varying effect on local economic and social conditions in subsequent decades. For instance, this assumption would be violated if immigrants in 1900 settled in a given city anticipating subsequent economic growth. Below, we perform a number of checks - including testing for pre-trends and interacting year dummies with pre-migration city characteristics - to assess the validity of the instrument, and show that our results are robust to the use of alternative specifications.

Using this empirical strategy, we find that immigration increased marriage and the probability of having children for native men and women. Our estimates are economically meaningful, and suggest that a five percentage point (equivalent to a one standard deviation) increase in immigration raised natives' marriage rates and the children-to-women ratio by 2 and 3 percent respectively. When decomposing the increase in fertility between the intensive (i.e. more children per woman) and the extensive (i.e. more women having at least one child) margin, we document that the latter was quantitatively more important than the former. Specifically, our estimates imply that for every 10 new babies born from native women, 7 were due to the extensive margin, while only 3 were due to the intensive margin.

Exploiting the granularity of full count data, we explore which age groups were responsible for the aggregate patterns just described. We show that the increase in both fertility and marriage was entirely driven by young couples, namely women (resp. men) aged 18-25 (resp. 20-27). Consistent with these findings, we also document that immigration induced young natives to leave their parental house earlier, and to set up an independent family unit.

In the second part of the paper, we investigate the mechanisms behind our main results.

[3] We focus on European immigrants (see Table A3.1) in the appendix for the complete list of sending countries), but results are robust to extending the analysis to all other non-European countries.

We provide evidence that they were driven by the large, positive impact of immigration on native men's employment, which increased the supply of "marriageable men" and made it easier for both men and women to marry, set up an independent household, and have kids. Specifically, our estimates suggest that for every ten new immigrants, one and a half more jobs were created for native men aged 20-35. Immigration also fostered natives' occupational standing, by inducing native workers to take up better jobs and move away from occupations more exposed to immigrants' competition, which tended to have lower skill requirements. Such large, positive effects on natives' employment were made possible by the fact that immigration increased firms' investment and productivity, in turn promoting industrialization and economic activity.

Next, we test a number of alternative mechanisms, and provide evidence that none of them can account for a quantitatively relevant fraction of our findings. First, we explore the possibility that immigration increased female marriage rates by altering sex ratios, i.e. the relative number of men and women, as more than 60% of immigrants entering the US at the time were young men (Figure A3.1). In contrast with this mechanism, however, immigration induced not only native women but also native men to marry more and to have more kids, suggesting that changes in sex ratios alone cannot be driving our main results. We also provide evidence that the inflow of immigrants did not raise the probability that native women married and had kids with foreign born. In fact, consistent with the idea that marriage markets were highly segmented along ethnic lines, more than 95% of U.S. born women were getting married with U.S. born men. Ethnic segmentation of marriage markets was further reinforced by the provision of Section 3 of the 1907 Expatriation Act, which mandated that native women marrying a foreign born individuals would lose their US citizenship. We also unveil some interesting heterogeneous patterns, which depended on men's parentage. Specifically, even if immigration had on average a positive effect on marriage rates of native men, this did not happen for second generation men, who were probably more exposed to immigrants' competition in the marriage market (Angrist, 2002).

Second, we rule out the possibility that higher marriage rates among US born were the result of a cultural response by native couples aimed at "preserving" their own race (Bisin and Verdier, 2000; Spolaore and Wacziarg, 2016). In particular, building on the measure of linguistic distance from Chiswick and Miller (2005), we construct an index of cultural diversity, and show that the latter did not have any effect on natives' marriage rates. Lastly, we provide evidence that direct (negative) effects of immigration on female labor force participation, which might have induced women to first leave the labor force and then get married and have kids, cannot explain our key findings. Exploiting variation across age groups, we show that the decrease in labor force participation was limited to women whose marriage rates increased in response to immigration. Given the stigma attached to the work of wives outside the home at the beginning of the twentieth century, women were likely to quit their job as a consequence of marriage (Goldin, 2006).

Our results are related to several strands of the literature. First, we complement the recent paper by Autor *et al.* (2017) by showing that a positive (rather than a negative) shock to employment opportunities of men increases (instead of reducing) marriage, fertility, and financial independence of young couples. Despite the difference in the historical context - early twentieth century vs contemporaneous period - and in the source of the income shock - immigration vs trade - comparing results in this paper with those in Autor *et al.* (2017) suggests that some key policy-relevant parameters, such as the elasticity of marriage and fertility to income, can be stable over time. At the same time, however, while our estimates on fertility are in line with those in Kearney and Wilson (2017), differently from us, the latter paper does not find a positive effect of an employment boom on marriage rates. One possible interpretation for this difference is that the cultural environment might mediate the transmission of income shocks to social outcomes.

Second, our paper is related to the vast literature on the effects of sex ratios on marriage market outcomes of men and women. Focusing on the same historical context, Angrist (2002) exploits variation in sex ratios for second generation immigrants induced by the arrival of individuals from different countries. We complement this paper by showing that immigration can impact marriage rates and fertility in receiving countries not only by altering sex ratios for second generation immigrants, but also by affecting natives' employment.[4] Moreover, the differential effect of immigration on marriage rates of native men (positive for natives with native parents, but close to zero for second generation immigrants) is consistent with findings in Abramitzky *et al.* (2011), who show that in French regions where more men died during WWI, men (resp. women) were better (resp. worse) off in the marriage market.

Lastly, our paper contributes to the literature that explores the effect of immigration on female labor force participation and fertility. Findings in Furtado and Hock (2010) and Furtado (2016) suggest that the availability of lower cost childcare opportunities offered by the inflow of immigrants in recent decades allowed college educated women to both have more children and work longer hours, attenuating the negative correlation between childbearing and labor force participation. This mechanism is unlikely to be at play in our context since, at the beginning of twentieth century, most women took care of their own children, and additional childbearing was assigned to black - and not immigrant - women. Goldin (1990,

[4] We find a positive effect of immigration on natives' employment, which in turn induced natives to marry and have kids more often. However, it is possible that, when immigration decreases natives' employment, it also lowers the probability of having children, marriage rates, and the propensity to set up independent households.

2006) shows that in this historical period native women would quit their job upon getting married and having a child. Consistently with these findings, we show that the negative effect of immigration on female labor force participation was concentrated exclusively on women in the age group that experienced an increase in fertility and in marriage.

The paper is structured as follows. Section 3.2 describes the historical background. Section 3.3 presents the data. Section 3.4 lays out our empirical strategy, constructs the instrument for immigration, and reports first stage results. Section 3.5 investigates the effects of immigration on natives' marriage, fertility, and propensity to leave the parental house. Section 3.6, explores the mechanisms. Section 3.7 concludes.

3.2 Historical Background

3.2.1 The Age of Mass Migration

Between 1850 and 1915, more than 30 million Europeans migrated to the United States. This massive migration episode took place in two waves: until 1890, most immigrants came from the British Isles, Germany, and Scandinavia; then, from the late 1880s, following the introduction of steam technology in shipping, which drastically reduced migration costs, immigration from Southern and Eastern Europe increased steadily (Keeling, 1999). In 1870, almost 90% of the foreign born came from Northern and Western European countries, whereas less than 5% of immigrants had arrived from Southern and Eastern Europe (Figure 3.2). By 1920, however, the situation had changed dramatically, and the share of immigrants born in new source countries was as high as 40%.

Europeans from new regions were culturally farther from natives and significantly less skilled than those from old sending regions (Hatton and Williamson, 1998, 2006). The shift in the composition of immigrants and concerns over their assimilation induced Congress to establish a commission that, between 1907 and 1911, studied the economic and social conditions of immigrants (Higham, 1955). In 1911, the Immigration Commission recommended the introduction of immigration restrictions, and in 1917, after decades of heated political debate, Congress passed a literacy test requiring that all immigrants entering the United States had to be able to read and write (Goldin, 1994).

Even before the adoption of the literacy test, in 1914, the Age of Mass Migration came to an abrupt end due to the onset of World War I, which drastically reduced European immigration between 1915 and 1919 (see Figure A3.2). In 1920, despite the literacy test, migration flows increased again to their 1910 levels, fueling nativist movements and generating even stronger political pressure to adopt more effective measures to curb immigration. In response to the growing demand for immigration restrictions, in 1921 and 1924 Congress finally passed the Immigration Acts to limit the number of immigrants that could enter the United States in a given year by introducing country-specific quotas based on 1890 immigrants' population.[5]

Both World War I and the Immigration Acts affected different sending countries in different ways. In particular, the reduction in immigration was more pronounced for European regions that were more directly involved in the War and which did not belong to the Allies (Figure 3.3, Panel A). Moreover, during the 1920s, quotas were set so as to limit the inflow of immigrants from new sending regions, while favoring that from old sources such as the UK, Germany, and Scandinavia (Figure 3.3, Panel B). Since immigrants tend to cluster along ethnic lines (Card, 2001), the post-1915 events generated substantial variation in the number as well as in the mix of immigrants received by US cities over time (Figure A3.3): this is the variation we exploit in our empirical analysis.

3.2.2 Immigration, Natives' Marriage, and Fertility

During the Age of Mass Migration, many prominent scholars expressed concerns over the effects of immigration on natives' fertility and marriage. As discussed in Leonard (2005), Edward Ross was among the first to propose the theory of "race suicide". According to this theory, not all immigrants were the same, and members of new, inferior races (i.e. immigrants from new sending regions) would eventually outbreed the "superior national stock" (i.e. natives and immigrants from old source countries) because industrial capitalism was conducive to the survival of the unfit (Leonard, 2005). More specifically, Francis A. Walker argued that "the native element failed to maintain its previous rate of increase because the foreigners came in such swarms", and natives were unwilling not only to engage in competition with "these new elements of the population", but also, they did not want "to brings sons and daughters into the world to enter that competition" (Walker, 1899, p. 424).

In contrast with these predictions, the inflow of immigrants might have increased marriage rates and fertility of native women by altering sex ratios (i.e. the relative number of men and women). At the time, more than 60% of immigrants entering the United States were young men between 20 and 35 (Figure A3.1). Since in the early twentieth century the median age at first marriage was around 21 for women and 25 for men (Figure 3.4), even though marriage markets were highly segmented along ethnic lives and native women loss their U.S. citizenship after marrying a foreign born,[6] immigration might have made it easier for native

[5]With the 1924 National Origins Act, the total number of immigrants that could be admitted in a given year was capped at 150,000. In 1921, quotas
were specified reflecting the 1910 composition of immigrants. However, they were rapidly changed to 1890 to limit immigration from new sending countries even further (Goldin, 1994).

[6]The loss of citizenship after the marriage with an "alien" men was established by Section 3 of the

women to find a mate and to have kids (Angrist, 2002).[7]

Yet another possibility is that immigration affected natives' marriage, fertility, and trends in family formation by altering employment and occupational standing of native men. Historical accounts tend to view immigrants as one of the key determinants of American industrialization and economic development during the Age of Mass Migration. When describing the economic impact of European immigrants, historian Maldwyn Jones wrote that "The realization of America's vast economic potential has...been due in significant measure to the efforts of immigrants. They supplied much of the labor and technical skill needed to tap the underdeveloped resources of a virgin continent" (Jones, 1992, pp. 309-310). Similarly, John F. Kennedy argued that "every aspect of the American economy has profited from the contribution of immigrants" (Kennedy, 1964, p. 88).

During the Age of Mass Migration, the US economy had large potentials for growth. In this context, immigrants provided a cheap and unskilled supply of labor which could not only be absorbed, but that may have even allowed industries to expand (Foerster, 1924), in turn creating new job opportunities for native workers (Tabellini, 2017). It is thus possible that, by increasing the supply of "marriageable men", immigration raised fertility and marriage rates not only of native women, but also of native men. Moreover, if native men could find a stable job earlier in their working life, they might have been able to leave their parental house and set up their own household earlier. Somewhat ironically, then, immigration might have had exactly the opposite effect relative to what was argued by advocates of the theory of "race suicide".

3.3 Data

Our analysis is based on a balanced panel of the 180 US cities with at least 30,000 residents in each of the three census years from 1910 to 1930, and where at least some Europeans were living in 1900 (see Figure A3.4 and Table A3.2 for the complete list of cities). The dataset used in this paper was assembled using the decennial US Census of Population, made available by IPUMS (Ruggles *et al.*, 2015).[8] From this source, we collected data on city population, on the number of immigrants by country of origin at the city and at the national level, and on most of the outcomes considered in our analysis, including marital

Expatriation Act of 1907. The law was aimed at avoiding cases of multiple nationality among women.

[7] In particular, Figure 3.4 plots the distribution of the age at first marriage for native men and women in 1930 (the first year in which this question was asked in the US Census). See also estimates reported at https://www.thespruce.com/estimated-median-age-marriage-2303878.

[8] For 1900, we used the 5% sample, while for 1910, 1920, and 1930, we relied on the full count census datasets.

status, relationship to the household head, and the number of children.[9] To investigate the mechanisms, we also collected data on employment, labor force participation, and occupation of native men of age 20 to 65 from the US Census, and on several measures of economic activity and industrialization from the 1904 to 1929 quinquennial Census of Manufactures (Tabellini, 2017).[10]

Table 3.1 reports the summary statistics for the main variables used in our analysis. City population ranges from more than 6.9 million (New York City in 1930) to as little as 30,200 (Pasadena in 1910). There is also wide variation in the fraction of immigrants across cities and over time, which was higher in the northeastern states of Connecticut, Massachusetts, New Jersey, and New York, and lower in the US South. As already discussed in Section 3.2, WWI and the Immigration Acts drastically reduced immigration: in 1910, the fraction of immigrants over city population was, on average, 0.18, but this number fell to 0.12 in 1930. The decline in the fraction of foreign born that entered the United States in the previous decade was even starker: for the average city, this number was 0.08 in 1910, but fell to 0.02 in 1930.

In Panel B of Table 3.1, we report the summary statistics of the main outcomes of this paper, i.e. marriage rates, fertility, and the propensity to leave the parental house for young native men and women. By the age of 33 for women and 35 for men, 65\% of the native population was married. As shown in Table 3.2 for 1910, among native women of native parentage, 73% were married to a native husband with both native parents, 20% to a husband with one or both foreign born parents, and only 8% to a foreign born husband. Interestingly, most of the foreign born husbands arrived to the US more than ten years before.[11] Instead, the probability of being married with a foreign born husband was as high as 24% for second generation women.[12]

Between 1910 and 1930, among women aged 18-33, the average children to women ratio was 0.65: 34% of native women had at least one child, while those who were mothers had on average almost 2 children each. Table 3.1 also suggests that the decision of leaving parents' home was strongly correlated with financial independence and with the choice of getting married: the proportion of men and women who were household head or spouse was close to marriage rates (45% and 43% for women and men respectively).

Finally, Panel C presents the summary statistics for the key labor market outcomes

[9] See Table A3.1 for the list of European countries used in our work. To classify individuals based on their country of origin, we followed the classification made by IPUMS (Ruggles *et al.*, 2015).

[10] In 1920, the US Census did not report employment status, but rather only an indicator for holding any gainful occupation. For this year, we imputed values from the latter to proxy for employment.

[11] We analyze the impact of immigration in the previous decade on natives' marriage rates. As shown in Table 3.2, very few native women were getting married to the immigrants just arrived to the US.

[12] In our sample, second generation women accounted for roughly one fourth of all native women.

considered below. In 1910, the average employment to population ratio for native men aged 20-35 in our sample was 91%, and then fell to 84% in 1930, with the onset of the Great Depression. Average labor force participation for native women was 42%, with an increasing trend over time which was slowed down by the economic downturn in 1930.[13]

Immigration data are available for all the 540 city-year observations in our sample. However, for 1920, Sacramento (CA) and New Bedford (MA) had unreasonably low values for marriage, fertility, and the other demographic outcomes considered in our work, probably reflecting mis-reporting in the original documents. For this reason, in our baseline specification, we drop 1920 data for these two cities, but our results remain unchanged when all 540 city-year observations are included.[14]

3.4 Empirical Strategy

In this section, we present the baseline estimating equation (Section 3.4.1), construct the instrument for immigration (Section 3.4.2), and report first stage results (Section 3.4.3). To deal with the potential endogeneity of immigrants' location decision, we instrument the actual number of immigrants by interacting 1900 settlements of different ethnic groups with subsequent migration flows from each sending region, leaving out immigrants that eventually settled in the city's MSA.

3.4.1 Baseline Estimating Equation

To investigate the effects of immigration on natives' marriage, probability of having children, and family structure across US cities, we stack the data for the three Census years between 1910 and 1930, and estimate

$$y_{cst} = \gamma_c + \delta_{st} + \beta Imm_{cst} + u_{cst} \tag{14}$$

where y_{cst} is the outcome for city c in state s in Census year t,and Imm_{cst} refers to the fraction of immigrants received by city c in the previous decade, over city population. γ_c and δ_{st} are city and state by year fixed effects, implying that β is estimated from changes in the fraction of immigrants within the same city over time, compared to other cities in the same state in a given year. Since city population could itself be an outcome of immigration, the number of immigrants is scaled by predicted (rather than actual) city population, constructed

[13]Until 1930, the US Census classified individuals as participating in the labor force if they were holding any gainful occupation.

[14]Also, data from the Census of Manufactures were not available for Superior (WI), Washington DC in 1909 and 1919, and for Flint (MI), Galveston (TX), Huntington (WV), Lexington (KY), McKeesport (PA), Pueblo (CO), Quincy (IL), and Roanoke (VA) in 1929.

by multiplying 1900 population by average urban growth in the US, excluding that of the Census division where the city is located. Below, we also report results obtained when scaling immigration by 1900 population. Standard errors are clustered at the MSA level, and MSA boundaries are fixed to 1940 in order to keep geography constant.[15]

3.4.2 Instrument for Immigration

A priori, we may expect immigrants to be attracted to cities with better employment opportunities. Alternatively, immigrants might settle in otherwise declining cities, where house prices are lower. In either case, OLS estimates of equation (14) will likely be biased. To deal with this endogeneity problem, we construct a "leave-out" version of the shift-share instrument (Card, 2001). The instrument predicts the number of immigrants received by US cities over time by interacting 1900 settlements of different ethnic groups with subsequent migration flows from each sending region, excluding individuals that eventually settled in a given city's MSA. Formally, Imm_{cst} in (14) is instrumented with

$$Z_{cst} = \frac{1}{\hat{P}_{cst}} \sum_j \alpha_{jc} O_{jt}^{-M} \tag{15}$$

where $\hat{P}_{cst}$ is predicted city population; α_{jc} is the share of individuals of ethnic group j living in city c in 1900; and O_{jt}^{-M} is the number of immigrants from country j that entered the US between t and $t-1$, net of those that eventually settled in city c's MSA.[16]

The instrument constructed in equation (15) exploits two sources of variation: first, cross-sectional variation in the share of individuals from each ethnic group living in different US cities in 1900 (α_{jc}); second, time-series variation induced by changes in the total number of immigrants from any sending region entering the United States in a given decade $\left(O_{jt}^{-M}\right)$. Section 1.4.2.1 above presents a simple example to illustrate graphically how the instrument combines them.

3.4.2.1 Geographic Variation in Immigrants' Settlements

The cross-sectional variation underlying the instrument in equation (15) is based on the idea that immigrants cluster geographically and when newcomers arrive, they tend to move

[15] In our baseline specification, we restrict attention to European immigrants that entered the United States during the previous decade, but results are robust to using immigrants' stock or considering immigrants from all sources.

[16] A similar "leave-out" strategy is used in Burchardi *et al.* (2016). Results are also robust to using a specification where the endogenous regressor, Imm_{cst}, is constructed by scaling the number of immigrants by actual (rather than predicted) city population, and is instrumented with Z_{cst} in (15), i.e. the predicted number of immigrants over predicted city population.

where their ethnic community is larger because of social networks and family ties, and not because of local pull factors (Card, 2001; Stuart and Taylor, 2016). As documented in Sequeira *et al.* (2017), the gradual expansion of railroads during the nineteenth century is a strong predictor of the geographic distribution of immigrants in the US: places that gained access to the railroad just before an immigration boom received more immigrants in the following decade. Moreover, upon arrival, early settlers tended to locate in places that were relatively more attractive at that time. Since the timing of outmigration varied widely across European countries, depending on local political and economic conditions (Hatton and Williamson, 1998), different US regions were populated by different ethnic groups before 1900. Early settlers then acted as a catalyst for subsequent migrants from the same ethnic group (Lafortune and Tessada, 2014).

To visually display the degree of geographic concentration of different ethnic groups, Figure 3.5 plots the share of individuals from different European regions living in selected US cities in 1900.[17] For example, while Italian communities were present in Boston, Philadelphia, and San Francisco, they were practically non-existent in Minneapolis. On the other hand, while almost 4% of Swedes living in the US in 1900 were settled in Minneapolis, less than 1% of them were located in north-eastern cities like Philadelphia or Boston. Finally, in 1900, more than 8% of Eastern Europeans were living in Cleveland, but their share in the other cities displayed in Figure 3.5 was well below 1%. Presenting a similar example for Ohio, Figure A3.5 shows that differences in immigrants' settlements existed not only across, but also within states. This is important since our empirical strategy exploits only within state variation in immigration.

3.4.2.2 Identifying Assumptions and Instrument Validity

The key identifying assumption behind the instrument is that cities receiving more immigrants (from each sending area) before 1900 must not be on different trajectories for the evolution of economic and social conditions in subsequent decades. Said differently, outmigration from European regions must be independent of cross-city pull factors systematically related to 1900 settlers' country of origin. For example, between 1910 and 1920, immigration to the US was higher from Poland than from Norway. The exclusion restriction would be violated if this happened because cities that in 1900 had attracted more Poles were growing more than cities where more Norwegians had moved to in 1900.

Another threat to the validity of the identifying assumption is that the characteristics of cities that attracted early immigrants might have time-varying, confounding effects on

[17] See also Abramitzky and Boustan (2017) for a discussion of the geographic concentration of Europeans in the United States during the Age of Mass Migration.

migration patterns as well as on changes in the outcomes of interest. It is possible, for instance, that larger urban centers attracted more immigrants in the nineteenth century, and that these cities kept growing more also in subsequent decades. In turn, more sustained economic growth may have increased marriage and fertility of natives, invalidating the instrument constructed in equation (15). To deal with these and similar issues, we perform several robustness checks, which we describe below when presenting our main results.

3.4.3 First Stage Results

First stage results for the relationship between actual and predicted immigration are reported in Table 3.3, after partialling out city and state by year fixed effects. In column 1, the dependent variable is the fraction of immigrants over actual city population, and the regressor of interest is the baseline instrument constructed in equation (15). Columns 2 and 3 replicate column 1 by scaling both the actual and the predicted number of immigrants by, respectively, 1900 and predicted population. In all cases, the F-stat is very high, and there is a strong and significant relationship between the endogenous regressor and the instrument.

Figure 3.6 plots the residual scatterplot of the regression reported in column 3. As it appears, the city of Passaic (NJ) experienced a large drop in immigration between 1910 and 1930, and one may be concerned that, for this reason, it influences the strength of the first stage. However, omitting this city barely affects the slope of the regression line (see red dashed line in Figure 3.6 and additional results in Tabellini, 2017). From column 3 onwards, Table 3.3 presents estimates for specifications where both the actual and the predicted number of immigrants are scaled by predicted city population, and explores the stability of the baseline specification to the inclusion of interactions between year dummies and 1900 city characteristics.

First, we augment the specification reported in column 3 by interacting the 1900 (log of) city and immigrants' population (column 4).[18] Next, in columns 5 and 6, we include interactions between year dummies and, respectively, the 1904 (log of) value added by manufacture and the marriage rate of native women in 1900. Even though the F-stat falls relative to column 1, it remains well above conventional levels. Also, and importantly, neither the economic nor the statistical significance of coefficients is affected.

Overall, Table 3.3 suggests that there is a strong relationship between actual and predicted immigration, which is robust to the use of different specifications and alternative ways of constructing the instrument.

[18] This check is important since the instrument mechanically predicts higher immigration to cities that had a larger 1900 fraction of immigrants, and, at the same time, larger ethnic enclaves might have direct and time-varying effects on economic and social conditions.

3.5 Main Results

In this section, we present three sets of results. First, immigration had a positive and large effect on marriage rates of both native women and native men (Section 3.5.1). Second, the inflow of immigrants raised the probability of having children for natives by increasing the share of young women with at least one child (Section 3.5.2). Third, immigration induced native young men and women to anticipate the age at which they chose to leave their parental house (Section 3.5.3).

3.5.1 Immigration and Marriage Rates of Natives

In Table 3.4, we study the impact of immigration on natives' marriage focusing on the age groups with the highest marriage rates, i.e. women aged 18-33 and men aged 20-35.[19] In Panel A (resp. Panel B), the dependent variable is the fraction of native women (resp. men) who were married. OLS results of equation (14) are presented in column 1, while column 2 reports 2SLS estimates for the baseline specification, where we instrument the fraction of immigrants over predicted population using the "leave-out" shift-share instrument described in Section 3.4.2. Throughout the paper, we always report the mean of the dependent variable at baseline and the F-stat associated with first stage results shown in Table 3.3.

Starting from Panel A, both OLS and 2SLS estimates suggest that immigration increased marriage rates for native women aged 18-33.[20] These effects are not only statistically significant but also economically relevant: the coefficient in column 2 implies that one standard deviation increase in the fraction of immigrants raised marriage rates of native women aged 18-33 by 2.2% relative to the 1910 mean (see Figure 3.7). Panel B documents a similar pattern (both qualitative and quantitative) for native men aged 20-35: a five percentage point increase in immigration (equivalent to a one standard deviation) raised men's marriage rates by 2.1% relative to their baseline mean.[21] How do these estimates compare to the existing literature? Our findings are quantitatively close to those obtained in Autor *et al.* (2017), who document that, over the last thirty years, a one percentage point increase in import

[19] As discussed in Section 3.2.2, the median age at first marriage was around 21 for women and 25 for men (Figure 3.4).

[20] Both the OLS and 2SLS coefficients reported in columns 1 and 2 respectively are positive and significant, with the latter being only slightly smaller than the former.

[21] OLS estimates are sensitive to the inclusion of three cities (Duluth, Superior, and Tacoma) for which in 1910 marriage rates were very low. The mean value of marriage rates of men aged 20-35 in 1910 is 24 percentage points lower compared to the mean value of the same cities in 1920 and 23 percentage points lower compared to other US cities in our sample in 1910. The latter effect corresponds to 4.6 lower standard deviations in the marriage rates of men in these cities compared to the rest of our sample. In the Appendix Table A3.3, we present estimates of OLS and 2SLS results with and without these three cities. Once we restrict the sample, OLS and 2SLS are closer in magnitude.

competition from China lowered female marriage rates by 1.8%.

Subsequent columns of Table 3.4 explore the robustness of our baseline results. First, in column 3, we test for pre-trends by regressing the 1900 to 1910 change in marriage rates against the 1910 to 1920 instrumented change in immigration. Reassuringly, in both Panel A and Panel B, the coefficient on immigration is statistically indistinguishable from zero and different from that reported in column 2. Next, in columns 4 and 5, we augment our baseline specification by interacting year dummies with the (log of) 1900 city and immigrants' population and the 1900 marriage rates, respectively. This exercise is performed to check that results in column 2 are not due to city-specific characteristics that may have simultaneously attracted more immigrants before 1900 and affected the evolution of natives' marriage rates in subsequent decades. In all cases, the point estimate remains statistically significant and quantitatively close to that estimated in the baseline specification. Finally, in column 6 we provide evidence that results are robust to scaling both the actual and the predicted number of immigrants by 1900, rather than predicted, population.

Up to now, we reported results for the "marriage-relevant" age groups by gender. In Figure 3.7, we separately document the effect of immigration on marriage rates of native men and women for different age groups. All of the effect estimated in Table 3.4 comes from the youngest cohorts: one standard deviation increase in the fraction of immigrants raised marriage rates of native women aged 18-25 and men aged 20-27 by 3.4% and 4.0%, respectively, relative to their baseline means. Instead, the effect of immigration is not statistically significant for older cohorts. The point estimates and standard errors related to this figure are reported in Appendix Table A3.4, where we also show the probability of being never married for the oldest cohorts. While immigration had no effect on the probability of being never married for men, it lowered the likelihood that women aged 34-65 remained unmarried.

3.5.2 Immigration and Natives' Probability of Having Children

In Table 3.5, we study how exposure to immigration affected the probability of having children for native women in our sample of 180 cities. The first two columns focus on the children to women ratio, while in subsequent columns we separately analyze the effect of immigration on the extensive and the intensive margin. We define the former as the share of women with at least one child, and the latter as the children to mothers ratio. In odd (resp. even) columns, the dependent variable is the total number of children in the household (resp. children below the age of 5). Since full-count data allow to match mothers with children only if they are living in the same household, we restrict the sample to women aged 18-33, whose children are likely to live with their parents.

Both OLS and 2SLS results, reported in Panels A and B respectively, document a positive and significant relationship between immigration and the probability of having children of native women. The point estimate in column 1 of Panel B implies that a one standard deviation increase in immigration raised the children to women ratio by 3.3% relative to its 1910 mean. When decomposing this effect along the extensive and the intensive margin, we note that immigration increased the number of women with children by 2.4%, and raised the average number of children per woman by 1%. Said differently, for every ten new babies born from native women, seven were due to the extensive margin, while three to the intensive margin. The magnitude of the effect is similar when we restrict our attention to children below the age of 5.[22]

Between the late nineteenth century and the 1930s, the US went through a demographic transition, with a reversal of the positive relationship between income and economic growth (Galor and Weil, 2000). The fertility rate of the total white population declined substantially, with the birth rate moving from almost 50 per thousand population in 1850 to 20 per thousand in 1930 (Zelnik, 1959). However, in our analysis, the inclusion of state by year fixed effects takes care of these national trends, since the effect of immigration is estimated from changes in the fraction of immigrants within the same city over time, as compared to other cities in the same state in a given year. Moreover, as noted by Easterlin (1961), the decline in fertility was driven by rural areas; instead, fertility of the urban native (white) population remained stable in this time period.[23]

In Appendix Table A3.5, we separately report the effect of immigration on fertility of native women by age groups. As for marriage, the effect is driven mainly by native women aged 18-25, especially on the extensive margin: one standard deviation increase in immigration raised the number of women in the younger age cohort with at least one child by 3.1%.

3.5.3 Household Formation

In Table 3.6, we provide evidence that immigration anticipated the choice of natives to leave their parental house, and set up their own independent family unit. In the first two columns, we focus on women aged 18-33, while in subsequent columns we report the effects of immigration on men aged 20-35.[24]

[22]As before, our estimates are quantitatively in line with those from Autor *et al.* (2017), who, for the more recent period, find that a 1 percentage point increase in import competition from China reduced fertility by 2.8%.

[23]Guinnane *et al.* (2006) find that fertility of immigrants in the late nineteenth and early twentieth century was higher than that of natives, but converged to US standards for second generation immigrants.

[24]Both OLS and 2SLS results, reported in Panels A and B respectively, are statistically significant and close in magnitude for women. As for marriage rates, for men, OLS estimates are instead sensible to the inclusion of three cities (Duluth, Superior, and Tacoma). In these cities, only 20% of men were household

Specifically, the coefficients in Table 3.6 imply that one standard deviation increase in immigration raised the probability of living in an independent family unit by 2.4% for women and 2.2% for men, relative to the mean in 1910. This effect is quantitatively close to that estimate for marriage rates, suggesting that the decisions of getting married and of leaving the parental house were both part of a unique lifetime plan. Interestingly, focusing on the contemporaneous period, Autor *et al.* (2017) find that one percentage point increase in import competition from China, not only decreased marriage rates and fertility, but also lowered the probability of living with the spouse by 1.6%.

In Figure A3.6, we provide evidence that the effect of immigration on the probability of leaving the parental house was driven by women aged 18-25 and men aged 20-27: for these age groups, a five percentage point increase in immigration raised the probability of setting up their own household by more than 3%. Incidentally, these cohorts also experienced the largest increase in marriage and fertility because of immigration.

Stitching together the three sets of results presented in this section, our estimates paint a coherent picture of how immigration affected family formation, marriage rates, and fertility of native men and women in the urban early twentieth century US. The inflow of immigrants induced natives to get married more (and, possibly, earlier);[25] this decision was accompanied by the choice of leaving the parental house and set up an independent family unit. In a period in which oral contraception was not yet available (Bailey, 2006), higher fertility was probably mechanically related to marriage and family formation decisions.

3.6 Mechanisms

In this section, we explore the mechanisms behind the results presented above. In Section 3.6.1, we start by documenting that immigration raised employment and occupational standing of native men, and then argue that such higher supply of "marriageable men" was the key driver of the positive effect of immigration on natives' marriage and fertility shown in Section 3.5. Next, we provide evidence that changes in sex ratios (Section 3.6.2), natives' cultural responses (Section 3.6.3), and direct effects of immigration on native female labor force participation (Section 3.6.4) cannot account for a quantitatively relevant fraction of our main findings.

---

head in 1910, as compared to 42% in 1920 or 39% in the other cities of our sample in 1910. 2SLS estimates are instead unaffected by the inclusion of these three cities.

[25] Using information from the 1940 US census for the same cohorts of individuals of our main specification, we check whether the increase in marriage rates is merely due to the anticipation of the timing of marriage. Although there are some potential concerns related to migration and data quality, we find that individuals in cities exposed to more immigration have overall higher marriage rates during the entire life and slightly lower age at first marriage. The table is available upon request to the authors.

3.6.1 Natives' Employment and the Supply of Marriageable Men

In two important contributions, Wilson (1987, 1996) argues that the decline in marriage and the rise in the share of single-mother households in the US during the last forty years have been, at least in part, due to deteriorating employment opportunities in manufacturing. Along these lines, exploiting exogenous variation in exposure to import competition from China across US local labor markets, Autor *et al.* (2017) find that job losses in manufacturing caused a steep decline in marriage rates and a significant increase in the proportion of single-mother households. In this section, we investigate the possibility that a similar mechanism, with the opposite sign, was at play in our context. Specifically, we advance and empirically test the hypothesis that immigration had a positive effect on natives' marriage, fertility, and patterns of family formation by increasing employment and occupational standing of native men, in turn raising the supply of "marriageable men".

In Table 3.7, we study the effects of immigration on natives' employment to population ratio, focusing on men in the "marriageable relevant" age range, i.e. 20-35 (see Section 3.5.1). As for Table 3.4, columns 1 and 2 estimate the baseline specification (see equation (14)) with OLS and 2SLS respectively. In both cases, there is a strong and positive relationship between immigration and natives' employment. The coefficient in column 2, which is quantitatively very close to OLS results reported in column 1, implies that a five percentage points increase in immigration (equivalent to one standard deviation) raised natives' employment to population ratio by 0.9% relative to its 1910 mean. Said differently, for every ten new immigrants, one and a half more jobs were created for native men aged 20 to 35.

As documented in Figure A3.7, the effect of immigration is slightly larger for men in the age range 20-27, but remains positive and statistically significant also for those aged 28-35. The point estimate is positive and quantitatively very similar, albeit not statistically significant, also for older natives, i.e. those in the age range 36-65.[26] As we did in Table 3.4, we next test the robustness of our baseline specification in subsequent columns of Table 3.7. First, as for marriage rates, there is no evidence of pre-trends (column 3). Second, results are robust to interacting year dummies with the 1900 log of city and immigrants population (column 4) and log of value added by manufacture (column 5). Third, our estimates are unchanged when scaling both the actual and the predicted number of immigrants by 1900, rather than predicted, city population (column 6).[27]

The positive effects of immigration on natives' employment estimated in Table 3.7 are

[26] Very similar results are obtained in Tabellini (2017), who studies the effects of immigration on natives' employment for natives in the age range 15-65 in the same sample of cities. Tabellini (2017) also shows that immigration had a positive and large effect on natives' occupational standing.

[27] For many other robustness checks, see Tabellini (2017).

in contrast with some of the results from the contemporary immigration literature such as Tabellini (2017), Borjas and Katz (2007), and Dustmann *et al.* (2017) who find a negative effect of immigration on natives' labor market outcomes. Our findings are also somewhat different from those of some contemporaneous cross-city design studies that estimate a zero effect of immigration on natives' wages (Card, 2001, 2005). However, they are consistent with a recent body of the literature which documents a positive impact of immigrants on natives' occupational mobility (Foged and Peri, 2016), and more specifically for this historical period with Tabellini (2017).

In line with the latter works, in Figure 3.8, we show that immigration induced natives to leave occupations that were more exposed to immigrants' competition and to take up jobs where immigrants were prevented from entering, because of skill and language mismatch or because of discrimination. Specifically, Figure 3.8 plots the implied effect (expressed in percent change) of a one standard deviation increase in immigration on the fraction of native men aged 20-35 employed in specific occupations (see also Table A3.6 in the appendix).

The first three (orange) bars starting from the left refer to occupations that were highly exposed to immigrants' competition:[28] as it appears, there is a negative effect of immigration on the share of young natives working in these occupations. This effect is statistically significant and economically large especially for manufacturing laborers - one of the jobs with the highest exposure to immigrants' competition and with the lowest skill requirements. Moving rightward along the graph, the three (blue) bars on the right show that immigration increased the fraction of natives employed in more skilled and less exposed occupations such as manufacturing foremen, electricians, and engineers. The pattern displayed in Figure 3.8 can be effectively summarized using the words of the economist and statistician Isaac Hourwich who, in 1912, noted that "the effect of immigration upon the occupational distribution of industrial wage earners has been the elevation of the English-speaking workmen to the status of an aristocracy of labor, while the immigrants have been employed to perform the rough work of all industries" (Meyer, 1981).

For natives' employment to increase, immigration must have also stimulated economic activity, inducing firms to create new jobs. Otherwise, absent changes in labor demand, it would be hard to reconcile the labor supply shock induced by immigration with the positive employment effects estimated above. Consistent with this idea, in Table A3.7, we show that immigration had a positive and large effect on value added and the value of products per establishment (columns 1 and 2), establishment size (column 3), capital accumulation (column 4), and total factor productivity (column 5).[29] Specifically, the coefficients in Panel

[28] We proxy for the degree of immigrants' competition using the ratio of the probability that natives and immigrants held a given occupation in 1910 (see also Table A3.6 and Tabellini (2017)).

[29] Panel A and Panel B report, respectively, OLS and 2SLS results. We proxy for capital utilitazion

B of Table A3.7 imply that a one standard deviation increase in immigration increased industrial production by roughly 10% relative to its 1910 level. Such sizeable effects are not only consistent with the historical literature reviewed in Section 3.2, but they can also explain the positive employment effects estimated in Table 3.7. On the one hand, industrial expansion allowed the economy to absorb the large supply shock by creating new jobs for both high and low skilled workers. Second, it provided natives with opportunities for skill upgrading (see also Tabellini, 2017 for a more extensive discussion).

Overall, this section documents that immigration boosted natives' employment and induced men to take up better jobs. We argue that, in turn, the larger pool of "marriageable men" was responsible for the positive effects of immigration on natives' marriage rates, fertility, and propensity to leave the parental household earlier. In Table A3.8, we provide an additional piece of evidence consistent with this interpretation by showing that immigration lowered the share of children below the age of 10 born from native parents living in a household where the father was unskilled (column 1).[30] Similarly, even if the coefficient is not statistically significant at conventional levels, there is a positive relationship between immigration and the share of children of native parentage whose father is employed. These results suggest that, because of immigration, children of native parentage were likely to grow up in a better environment at home.

Consistent with the latter observation, as it appears from Table A3.9, immigration increased the fraction of sons of native parentage aged 6-14 who were enrolled in school (column 1). Somewhat interestingly, though, we do not find a similar effect for daughters (column 4), even if the 1910 average enrollment was very similar for boys and girls. One possible explanation for this pattern is that families were credit constrained and, as more resources became available, parents chose to first invest them in sons rather than in daughters. Especially in an urban context, higher employment opportunities brought about by immigration might have increased the opportunity cost of schooling, in turn inducing some boys to opt out of high school. Indeed, column 3 of Table shows that immigration had a negative and significant effect on enrollment of sons of native parentage aged 15-18.

3.6.2 Changes in Sex Ratios

The literature has documented that sex ratios, i.e. the relative number of men and women, can be an important determinant of marriage and family formation decisions (Angrist, 2002;

using the log of horsepower (column 4), and estimate the effects of immigration on productivity (column 5), assuming a Cobb-Douglas production function with two factors of production, capital and (homogeneous) labor.

[30] OLS and 2SLS results are reported respectively in Panel A and Panel B. Very similar results are obtained when focusing on the share of families rather than on the share of children (see columns 3 and 4).

Abramitzky *et al.*, 2011). Since more than 60% of immigrants entering the United States at the beginning of the twentieth century were young men (Figure A3.1), immigration likely altered sex ratios, possibly increasing the availability of potential mates for native women. However, in this section, we argue that this channel cannot explain a relevant fraction of our main results.

First, while changes in the relative number of men and women might have contributed to the increase in marriage rates and fertility of native women documented above, they cannot explain why immigration also raised native males' marriage rates.[31] Second, as we show in Tables 3.8 and A3.10, only 4% of native women had a foreign born husband and 3% of native men had a foreign born wife as of 1910.[32] Also, the increase in marriage rates for men and women was quantitatively similar (see Table 3.4), suggesting that natives, in most cases, were marrying with each other. Said differently, marriage markets at the time were highly segmented along ethnic lines (Angrist, 2002).

Focusing on results reported in Table 3.8, in Panel A, we find that one standard deviation increase in immigration raised the probability of getting married with a husband of native parentage by around 6% for all native women, irrespective of their parentage (columns 2 and 5). Instead, while the effect of immigration on the probability of having a foreign born spouse for native women was indistinguishable from zero (column 3), it was positive and significant for second generation women (column 6).[33] Yet, focusing on the relevant age group (i.e. 18-33), since second generation women who had a foreign born husband represented less than 2.5% of all native women, the implied effect of immigration on the overall marriage rates of native women was negligible.[34] Finally, Panel B documents that these effects were mirrored by a corresponding increase in fertility precisely for couples with higher marriage rates, in turn supporting the idea that immigration raised natives' fertility by fostering marriage in an era when oral contraception was not yet available (Bailey, 2006).

Having established that most of the effects of immigration were not driven by native women marrying foreign born husbands, in the last part of this section, we study how the

[31] Indirectly, higher competition in the marriage market may have induced men to increase their investment in education and on-the-job training and their earnings, as suggested by Becker (1981) in his notion of male "efficiency" (see also (Angrist, 2002)). However, even in this case, changes in sex ratios should have had a stronger impact on women as compared to men.

[32] In Tables 3.8 and A3.10, we explore the characteristics of partners of native women aged 18-33 and men 20-35 respectively.

[33] Interestingly, Table A3.10 shows a similar impact of immigration on marriage rates for all native and second generation men. While the effect of immigration on the probability of having a foreign born spouse is indistinguishable from zero for native men (column 3), the impact is positive and significant for male second generation immigrant (column 6).

[34] In the age group 18-33, second generation women were 25% of native females, and their probability of marrying with a foreign born was 10% at baseline (see the last column of Table 3.8).

inflow of immigrants affected marriage prospects of second generation men and women, via changes in sex ratios. Sex ratios can have important implications for the marriage market of second generation immigrants, both directly and indirectly through the allocation of bargaining power within the couple. For example, in the same historical context of our paper, Angrist (2002) finds that a higher relative number of men in their own ethnic group improved marriage prospects of second generation females. Figure 3.9 documents a pattern in line with this idea: because of immigration, marriage rates of second generation women aged 18-25 increased twice as much as those of women of native parentage. Similarly, while immigration had a positive and large effect on marriage rates for men of native parentage, it did not have any significant impact for second generation men. This finding is consistent with the idea that immigrants increased competition in the marriage market for second generation men. In Table A3.11, we separately report the effect of immigration on marriage rates of native men and women for different age groups and parentage, and document that all of the effect comes from the youngest cohorts represented in Figure 3.9 (that is, women aged 18-25 and men aged 20-27).

To sum up, even though sex ratios were affected by immigration, they can hardly explain the increase in marriage rates of natives with native parentage, a group for which the relative number of men and women in the reference population was not significantly affected. Since natives of native parentage were by far the largest group among US born individuals, their decisions disproportionately affected natives' overall marriage and fertility.

3.6.3 Preservation of "Natives"

Opposition to immigration was widespread during the Age of Mass Migration, with a heated aversion towards individuals coming from non Anglo-Saxon and non English-speaking countries (Abramitzky and Boustan, 2017; Leonard, 2016). Since immigrants from Southern and Eastern Europe were linguistically and culturally far from natives (Hatton and Williamson, 2006), it is possible that natives reacted to immigration by marrying more and having more kids, in order to preserve their own race and culture (see Section 3.2.2).

The role of culture in affecting marriage and fertility decisions has been stressed, among others, by Bisin and Verdier (2000) and Fernández and Fogli (2006), who study the transmission of cultural norms among second generation immigrants in the US.[35] More broadly, social interactions can influence the diffusion of cultural norms and have historically contributed to the convergence of fertility rates, both within and across countries (Spolaore and Wacziarg,

[35]Interestingly, in our context, we find that native men and women married to a foreign born were almost exclusively matching in the marriage market with immigrants from old sending regions who were culturally more similar to US born.

2016). For instance, Daudin *et al.* (2016) find that the demographic transition at the end of the nineteenth century in France was affected by the diffusion of low-fertility norms through internal migration.

To test if native men and women changed their family formation decisions to preserve their own culture, we analyze whether the effect on marriage rates and fertility was stronger when natives were exposed to linguistically farther individuals (which we take as a proxy for cultural distance). Specifically, we construct an index of immigrants' linguistic distance from English, $LD_{ct} = \sum_j sh^j_{ct} L^j$, where sh^j_{ct} is the share of ethnic group j among the foreign born population of city c in year t, and L^j is the linguistic distance from English of country j, computed in Chiswick and Miller (2005).[36] To ease the interpretation of results, which are reported in Table 3.9, we standardize our measure of linguistic distance by subtracting its mean and dividing it through its standard deviation. Differently from what we would have expected if this mechanism was driving our results, marriage rates were not differentially affected by immigrants with different linguistic distance from English. These results thus suggest that cultural considerations were unlikely to explain our key findings.

3.6.4 Increased Labor Market Competition for Women

From the end of the nineteenth century to the 1920s, female workers were mainly young, unmarried, and from low-income households (Goldin, 2006). Most women were employed as piece workers in manufacturing, as private household workers or laundresses, or in clerical jobs. Upon getting married, women typically quit their jobs because of the stigma attached to wives working outside their home (Cherlin, 2014). Goldin (1990) estimates that more than 80% of all married women exited the labor force at marriage, before 1940 (see Goldin, 1990, page 7). As shown in Table A3.12, in our sample of cities, the 1910 average labor force participation of native women aged 18-25 was 0.49, but was substantially lower for older women (0.33 and 0.25 for women aged 26-33 and 34-65, respectively).[37]

Studying the link between immigration, female labor force participation, and fertility, Furtado (2016) shows that the availability of lower cost childcare opportunities brought about by immigration induced native women to have more kids and work longer hours. In contrast with these results, at the beginning of the twentieth century, immigration may have increased competition in the labor market for women, in turn inducing them to first leave their job and then, as a consequence, to get married and have more children (Angrist and Evans, 1998). While possible, this interpretation seems to be inconsistent with the historical context studied

[36] We instrument the actual ethnic shares, sh_{ct}, using the same logic of the instrument constructed in equation (Z).

[37] Goldin (2006) notes that labor force participation of married women may be underestimated before 1940 because they were often reluctant to report that they had a job.

in our paper: at that time, as already discussed above, women most frequently took care of their own children, and used to quit their job upon marriage. Moreover, even though immigrants provided a cheap and unskilled supply of labor, which in principle might have displaced women, during the Age of Mass Migration, the US economy had large potential for economic expansion (Higgs, 1971). Thus, the displacement of female workers due to immigration seems unlikely, even more so as immigrants were more closely substitutes for men than for women, and we showed above that immigration increased natives' employment (see Section 3.6.1).

In line with this discussion, in Table A3.12, we document that immigration decreased labor force participation only for native women in the age group that experienced a significant increase in marriage rates (i.e., women aged 18-25). The impact is instead indistinguishable from zero for all older age cohorts, including women between 26 and 33 years old, among which one third was in the labor force.[38] In Figure 3.10, we report the implied coefficients for the effect of a one standard deviation increase in immigration, and show that female labor force participation in the age group 18-25 fell by 1.6% relative to its 1910 mean. Incidentally, this effect is only slightly smaller (in absolute value) than the increase in marriage induced by immigration for women in the same age group (see Figure 3.7). Our interpretation of these results is that immigration first induced native women to marry and have children, and then, as a consequence of the latter two decisions, to leave the labor force.

3.7 Conclusions

Today, immigration is at the forefront of the political debate, and there are increasing concerns over its economic and social consequences. If we look at American history, however, this is not the first time that immigration is such a relevant and controversial issue. In fact, at the beginning of the twentieth century, following the inflow of more than 30 million Europeans, the share of foreign born in the US population was even higher than it is today, and opposition towards immigration was widespread.

In this paper, we exploit plausibly exogenous variation in the number of European immigrants to US cities between 1910 and 1930 induced by WWI and the Immigration Acts to study the impact of immigration on marriage rates, the probability of having children, and the propensity to leave the parental house for young native men and women. We find that, by promoting industrial expansion and economic activity, immigration increased the supply of native "marriageable" men who, because of their better employment prospects and occupational standing, became more attractive spouses. This, in turn, fostered natives'

[38] Furthermore, women aged 26-33 were likely to work in the same sectors and occupations as women aged 18-25.

marriage rates for both men and women, and induced young adults to leave their parents' house earlier in their life. Higher marriage rates, in a period when oral contraception was not yet available, raised natives' probability of having children, mainly by increasing the number of women with at least one child (extensive margin).

In our context, the inflow of immigrants was largely beneficial to natives' economic and social outcomes. However, this does not imply that immigration will always promote fertility and marriage among young natives. In fact, if immigrants increase labor market competition, they may deteriorate, rather than promote, family stability as well as the environment where children grow up. Moreover, while we showed that in the early twentieth century, immigration to US cities affected marriage rates and fertility of natives mostly through (positive) income shocks, other channels may be at play in other settings. These observations suggest that one needs to be careful when extrapolating our results to other contexts.

Findings in this paper provide motivation for future work in at least two directions. First, in this study, we have not explored how changes in the supply of "marriageable men" affected the quality of the match between husbands and wives. If higher marriage rates were associated with worse matching between partners, this might have increased divorce rates and family instability, in turn lowering children's well-being (Stevenson and Wolfers, 2007; Lundberg *et al.*, 2016). Second, this setting seems ideal to study the dynamics of cultural assimilation - between immigrants and natives as well as between different ethnic groups - using intermarriage as a proxy for the latter.

References

Ran Abramitzky and Leah Platt Boustan. Immigration in american economic history. *Journal of Economic Literature*, 2017.

Ran Abramitzky, Adeline Delavande, and Luis Vasconcelos. Marrying up: the role of sex ratio in assortative matching. *American Economic Journal: Applied Economics*, 3(3):124–157, 2011.

Joseph G Altonji and David Card. The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, pages 201–234. University of Chicago Press, 1991.

Josh Angrist and William Evans. Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, 1998.

Josh Angrist. How do sex ratios affect marriage and labor markets? evidence from america's second generation. *The Quarterly Journal of Economics*, 117(3):997–1038, 2002.

David Autor, David Dorn, Gordon Hanson, et al. When work disappears: manufacturing decline and the falling marriage-market value of men. *NBER Working Paper*, 23173, 2017.

Martha J Bailey. More power to the pill: the impact of contraceptive freedom on women's life cycle labor supply. *The Quarterly Journal of Economics*, 121(1):289–320, 2006.

Gary Becker. *A Treatise on the Family.* Harvard university press, 1981.

Alberto Bisin and Thierry Verdier. Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics*, 115(3):955–988, 2000.

George J Borjas and Lawrence F Katz. The evolution of the mexican-born workforce in the united states. In *Mexican immigration to the United States*, pages 13–56. University of Chicago Press, 2007.

George J Borjas. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, 118(4):1335–1374, 2003.

Konrad B Burchardi, Thomas Chaney, and Tarek A Hassan. Migrants, ancestors, and investments. *National Bureau of Economic Research*, 2016.

David Card. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64, 2001.

David Card. Is the new immigration really so bad? *The Economic Journal*, 115(507), 2005.

Andrew J Cherlin. *Labor's love lost: the rise and fall of the working-class family in America.* Russell Sage Foundation, 2014.

Barry R Chiswick and Paul W Miller. Linguistic distance: A quantitative measure of the distance between english and other languages. *Journal of Multilingual and Multicultural Development*, 26(1):1–11, 2005.

Guillaume Daudin, Raphaël Franck, and Hillel Rapoport. The cultural diffusion of the fertility transition: Evidence from internal migration in 19th century france. *CESifo Working Paper*, 2016.

Christian Dustmann, Uta Schonberg, and Jan Stuhler. Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, 132(1):435, 2017.

RA Easterlin. The american baby boom in historical perspective. *The American Economic Review*, 51(5):869–911, 1961.

Raquel Fernández and Alessandra Fogli. Fertility: The role of culture and family experience. *Journal of the European Economic Association*, 4(2-3):552–561, 2006.

Robert Franz Foerster. *The Italian emigration of our times*, volume 20. Harvard University Press, 1924.

Mette Foged and Giovanni Peri. Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, 8(2):1–34, 2016.

Delia Furtado and Heinrich Hock. Low skilled immigration and work-fertility tradeoffs among high skilled us natives. *The American Economic Review*, 100(2):224–228, 2010.

Delia Furtado. Fertility responses of high-skilled native women to immigrant inflows. *Demography*, 53(1):27–53, 2016.

Oded Galor and David N Weil. Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *The American Economic Review*, pages 806–828, 2000.

Claudia Goldin. *The gender gap: An economic history of American women.* New York: Cambridge University Press, 1990.

Claudia Goldin. The political economy of immigration restriction in the united states, 1890 to 1921. In *The regulated economy: A historical approach to political economy*, pages 223–258. University of Chicago Press, 1994.

Claudia Goldin. The quiet revolution that transformed women's employment, education, and family. *The American Economic Review*, 2006.

Timothy W Guinnane, Carolyn M Moehling, and Cormac Ó Gráda. The fertility of the irish in the united states in 1910. *Explorations in Economic History*, 43(3):465–485, 2006.

Martin Halla, Alexander F Wagner, and Josef Zweimüller. Immigration and voting for the far right. *Journal of the European Economic Association*, page jvw003, 2017.

Timothy J Hatton and Jeffrey G Williamson. *The age of mass migration: Causes and economic impact.* Oxford University Press on Demand, 1998.

Timothy J Hatton and Jeffrey G Williamson. International migration in the long run: Positive selection, negative selection, and policy. In *Labor Mobility and the World Economy*, pages 1–31. Springer, 2006.

R. Higgs. *The transformation of the American economy, 1865-1914: An essay in interpretation.* Wiley New York., 1971.

John Higham. *Strangers in the land: Patterns of American nativism, 1860-1925.* Rutgers University Press, 1955.

Maldwyn Allen Jones. *American immigration.* University of Chicago Press, 1992.

Melissa S Kearney and Riley Wilson. Male earnings, marriageable men, and nonmarital fertility: Evidence from the fracking boom. *National Bureau of Economic Research*, 2017.

Drew Keeling. The transportation revolution and transatlantic migration, 1850-1914. *Research in Economic History*, 19:39–74, 1999.

President John F Kennedy. *A nation of immigrants.* New York: Harper & Row Publishers, 1964.

Jeanne Lafortune and José Tessada. Smooth (er) landing? the role of networks in the location and occupational choice of immigrants. *Mimeo, Instituto de Economía, Pontificia Universidad Católica de Chile September 2013*, 2014.

Thomas C. Leonard. Retrospectives: Eugenics and economics in the progressive era. *The Journal of Economic Perspectives*, 19(4):207–224, 2005.

Thomas C Leonard. *Illiberal reformers: race, eugenics, and American economics in the Progressive Era.* Princeton University Press, 2016.

Shelly Lundberg, Robert A Pollak, and Jenna Stearns. Family inequality: Diverging patterns in marriage, cohabitation, and childbearing. *The Journal of Economic Perspectives*, 30(2):79–101, 2016.

Anna Maria Mayda, Giovanni Peri, and Walter Steingress. Immigration to the us: A problem for the republicans or the democrats? *National Bureau of Economic Research Working Papers*, 2016.

Stephen Meyer. *The five dollar day: Labor management and social control in the Ford Motor Company, 1908-1921.* State University of New York Press, 1981.

E. Porter. Can immigration hurt the economy? an old prejudice returns. The New York Times, February 14, 2017, 2017.

Steven Ruggles, Katie Genadek, Ronald Goeken, Joasiah Grover, and Matthew Sobek. Integrated public use microdata series: Version 6.0 [dataset]. *Minneapolis: University of Minnesota*, 2015.

Sandra Sequeira, Nathan Nunn, and Nancy Qian. Migrants and the making of america: The short-and long-run effects of immigration during the age of mass migration. *National Bureau of Economic Research*, 2017.

Enrico Spolaore and Romain Wacziarg. Fertility and modernity. *UCLA CCPR Population Working Papers*, 2016.

Betsey Stevenson and Justin Wolfers. Marriage and divorce: Changes and their driving forces. *The Journal of Economic Perspectives*, 21(2):27–52, 2007.

Bryan A Stuart and Evan J Taylor. Social interactions and location decisions: Evidence from us mass migration. *Unpublished working paper*, 2016.

Marco Tabellini. Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *Working Paper*, 2017.

Francis Amasa Walker. *Discussions in economics and statistics*, volume 2. H. Holt, 1899.

WJ Wilson. The truly disadvantaged: The inner city, the underclass, and public policy. chicago: Univ. *Press, Chicago*, 1987.

Julius William Wilson. When work disappears: The world of the new urban poor. *New York: Alfred A Knopf*, 1996.

Melvin Zelnik. *Estimates of Annual Births and Birth Rates for the White Population of the United States from 1855 to 1934.* PhD thesis, Princeton University, 1959.

Table 3.1 Summary Statistics

Table 1: Summary Statistics

	Count	Mean	SD	Min	Max
Panel A: City Demographics					
Fr. Immigrant	538	0.04	0.05	0.00	0.44
City Population (thousand)	538	190	511	30	6930
Panel B: Key Outcomes					
Marriage Rates of Women					
Aged 18-33	538	0.49	0.08	0.28	0.67
Aged 18-25	538	0.35	0.08	0.12	0.58
Aged 26-33	538	0.65	0.07	0.46	0.81
Marriage Rates of Men					
Aged 20-35	538	0.47	0.07	0.18	0.65
Aged 20-27	538	0.31	0.07	0.11	0.49
Aged 28-35	538	0.65	0.07	0.26	0.81
Fertility of Women 18-33					
Children to Women Ratio	538	0.65	0.12	0.40	1.00
Mothers to Women Ratio	538	0.34	0.05	0.21	0.49
Children to Mothers Ratio	538	1.90	0.11	1.59	2.27
Living with parents					
Women Aged 18-33	538	0.36	0.09	0.17	0.58
Men Aged 20-35	538	0.33	0.09	0.12	0.55
Living in own household					
Women Aged 18-33	538	0.45	0.08	0.25	0.67
Men Aged 20-35	538	0.43	0.06	0.18	0.60
Panel C: Labor Market					
Employment Men 20-35	538	0.90	0.05	0.71	0.98
Labor Force Participation Women 18-33	538	0.42	0.09	0.20	0.67

Note: The Table shows the summary statistics of the main variables used in this paper for the 180 US cities with at least 30,000 residents in each Census year report. Source: Authors' calculations using IPUMS data.

Table 3.2 Characteristics of Husbands of Women Aged 18-33

Husband	**Native**			**Immigrant**	
	Native Parents	*Mix Parents*	*Foreign Parents*	> 10 *years*	≤ 10 *years*
Native Wife					
Native Parents	0.73	0.07	0.13	0.06	0.02
Mix Parents	0.50	0.12	0.24	0.11	0.03
Foreign Parents	0.35	0.09	0.32	0.18	0.06
Immigrant Wife					
> 10 *years*	0.18	0.04	0.14	0.50	0.13
≤ 10 *years*	0.07	0.01	0.05	0.20	0.67

Note: The Table shows the probability of marriage with husband of different parentage for women aged 18-33 of the 180 US cities in 1910 with at least 30,000 residents in each Census year report. Source: Authors' calculations using IPUMS data.

Table 3.3 First Stage

	Dep. Variable: *Fraction of Immigrants*					
	(1)	(2)	(3)	(4)	(5)	(6)
Z	0.830***	0.944***	0.990***	0.905***	0.889***	0.986***
	(0.053)	(0.071)	(0.063)	(0.090)	(0.086)	(0.066)
Immigrants over	*Actual pop.*	*1900 pop.*	*Predicted pop.*	*Predicted pop.*	*Predicted pop.*	*Predicted pop.*
Year interacted with 1900				*Immigrants and city population*	*Value added by manufacture*	*Fr. native women married*
F-stat	249.3	175.3	251.3	100.2	107.5	224.5
Observations	538	538	538	538	526	538

Note: the sample includes a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. In Col 1 the actual number of immigrants is scaled by actual population, and the instrument is the leave-out version of the shift-share IV in equation (2) (Section 4.2). Cols 2 and 3 replicate Col 1 by scaling the actual and predicted number of immigrants by, respectively, 1900 and predicted population. From Col 3 onwards, Table X presents results from specifications where both the predicted and the actual number of immigrants are scaled by predicted population. Cols 4 to 6 include the interaction between year dummies and, respectively: the (log of) 1900 city and immigrants population; the (log of) 1904 value added by manufacture per establishment; and the marriage rate of native women in 1900. F-stat refers to the K-P F-stat for weak instrument. All regressions partial out city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table 3.4 Immigration and Natives' Marriage

	(1) OLS	(2) 2SLS	(3) Pre-trends	(4) 2SLS	(5) 2SLS	(6) 2SLS
Panel A: Dep. Var. Marriage Rates of Women (Age 18-33)						
Fr. Immigrant	0.238***	0.209***	0.128	0.329***	0.197***	0.154***
	(0.057)	(0.044)	(0.204)	(0.058)	(0.053)	(0.027)
F-stat		251.3	318.4	100.2	107.5	175.3
Mean dep. var. in 1910	0.45	0.45	0.45	0.45	0.45	0.45
Obs.	538	538	178	538	538	538
Panel B: Dep. Var. Marriage Rates of Men (Age 20-35)						
Fr. Immigrant	-0.006	0.190***	0.078	0.181***	0.217***	0.121***
	(0.135)	(0.054)	(0.092)	(0.059)	(0.061)	(0.038)
F-stat		251.3	318.4	100.2	107.5	175.3
Mean dep. var. in 1910	0.42	0.42	0.42	0.42	0.42	0.42
Obs.	538	538	178	538	538	538
Pre-period			*Yes*			
Year by 1900 city and imm. pop				*Yes*		
Year by 1900 fr married					*Yes*	
Imm over 1900 pop						*Yes*

Note: this Table presents results of OLS and 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is the fraction of women married in the age range 18-33 in Panel A and the fraction of men married in the age range 20-35 in Panel B. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table 3.5 Immigration and Fertility of Native Women

	Dep. Variable: Fertility of Native Women (aged 18-33)					
	Children to Women Ratio		Mothers to Women Ratio		Children to Mothers Ratio	
	All Children	Children<5	All Children	Children<5	All Children	Children<5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Fr. Immigrant	0.291**	0.200***	0.132**	0.123***	0.149	0.209**
	(0.131)	(0.056)	(0.052)	(0.036)	(0.159)	(0.085)
Panel B: 2SLS						
Fr. Immigrant	0.431***	0.194***	0.165***	0.105***	0.342***	0.111**
	(0.072)	(0.037)	(0.027)	(0.026)	(0.103)	(0.052)
F-stat	251.3	251.3	251.3	251.3	251.3	251.3
Mean Dep.Var.	0.650	0.340	0.340	0.250	1.900	1.010
Obs.	538	538	538	538	538	538

Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is: in column 1 (column 2), the total number of children (toddlers) with native mother in the age range 18-33 over the total number of women in the age range 18-33, in column 3 (column 4) the fraction of women in the age range 18-33 who have children (toddlers) and in column 5 (column 6) the average number of children (toddlers) per mother in the age range 18-33. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table 3.6 Immigration and Living Choices of Natives

	(1)	(2)	(3)	(4)
Dep. Var.	**Living with Parents**	**Living in Own House**	**Living with Parents**	**Living in Own House**
	Women 18-33		**Men 20-35**	
Panel A: OLS				
Fr. Immigrant	-0.383***	0.231***	-0.493***	-0.004
	(0.086)	(0.057)	(0.131)	(0.137)
Panel B: 2SLS				
Fr. Immigrant	-0.285***	0.204***	-0.316***	0.171***
	(0.043)	(0.040)	(0.045)	(0.056)
F-stat	251.3	251.3	251.3	251.3
Mean dep. var.	0.370	0.418	0.317	0.387
Obs.	538	538	538	538

Note: this Table presents results of OLS and 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variables are described on the top part of the Table. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table 3.7 Immigration and Employment of Native Men

Dep. Var. : Natives' Employment to Population Ratio (Men, Age 20-35)						
	(1) OLS	(2) 2SLS	(3) Pre-trends	(4) 2SLS	(5) 2SLS	(6) 2SLS
Fr. Immigrant	0.151***	0.152***	-0.071	0.094**	0.130**	0.113***
	(0.043)	(0.044)	(0.124)	(0.041)	(0.053)	(0.033)
F-stat		251.3	318.4	100.2	107.5	175.3
Mean dep. var. in 1910	0.911	0.911	0.911	0.911	0.911	0.911
Obs.	538	538	180	538	538	538
Pre-period			*Yes*			
Year by 1900 city and imm. pop				*Yes*		
Year by 1900 value added manuf.					*Yes*	
Imm over 1900 pop						*Yes*

Note: this Table presents results of OLS and 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is the natives' employment to population ratio in the age range 20-35 for men. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0.10, **p<0.05, ***p<0.01.

Table 3.8 Immigration, Marriage Rates, and Fertility of Native Women aged 18-33 (2SLS Results)

	(1)	(2)	(3)	(4)	(5)	(6)
	All Native Women			Second Generation Women		
Husband	**All**	**Native Parentage**	**Immigrant**	**All**	**Native Parentage**	**Immigrant**
Panel A: Marriage rates						
Fr. Immigrant	0.209***	0.309***	0.001	0.193***	0.169**	0.178***
	(0.044)	(0.046)	(0.020)	(0.071)	(0.066)	(0.046)
Mean dep. var.	0.47	0.27	0.04	0.45	0.14	0.10
Panel B: Fertility (Children to Women Ratio)						
Fr. Immigrant	0.431***	0.443***	-0.005	0.359**	0.177*	0.259**
	(0.072)	(0.087)	(0.053)	(0.162)	(0.103)	(0.127)
Mean dep. var.	0.65	0.35	0.07	0.58	0.19	0.17
F-stat	251.3	251.3	251.3	251.3	251.3	251.3
Obs.	538	538	538	538	538	538

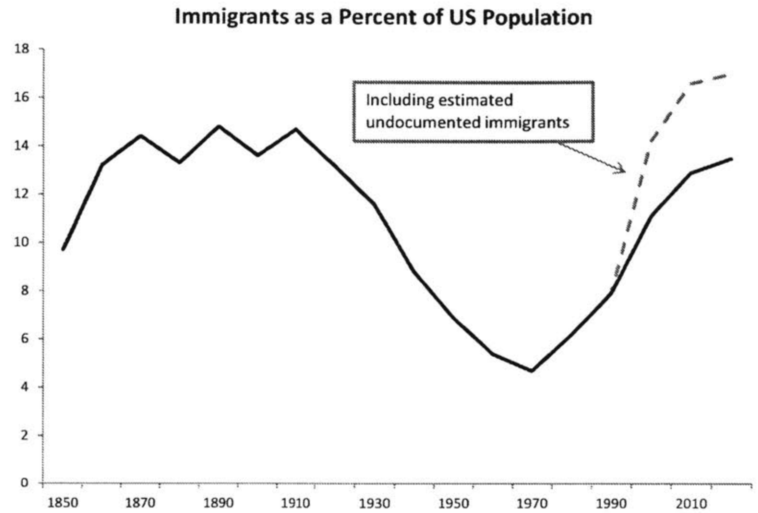
Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. In panel A, the dependent variable is the marriage rates of women aged 18-33 by husband parentage. In panel B, the dependent variable is the children to women ratio by father parentage. We consider only children of women aged 18-33. For example, in column 2 of Panel B, the dependent variable is the number of children with native mother aged 18-33 and father with a native parentage over the number of native women aged 18-33. Columns 4-6 focus on women who are second generation immigrants. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0.10, **p<0.05, ***p<0.01.

Table 3.9. Immigration, Linguistic Distance, Employment and Marriage Rates of Natives (2SLS Results)

Dep Var.:	**Employment Men 20-35**	**Marriage Rate Women 18-33**			
Own Parents		**All**	**Native**	**Mix**	**Immigrants**
	(1)	(2)	(3)	(4)	(5)
Fr. Immigrant	0.136**	0.207***	0.126**	0.282***	0.274***
	(0.066)	(0.060)	(0.064)	(0.091)	(0.102)
Ling. Distance	0.001	0.000	0.001	-0.004	-0.007
	(0.004)	(0.003)	(0.004)	(0.005)	(0.006)
Mean Dep. Var.	0.340	0.257	0.277	0.642	0.603
Obs.	538	538	538	538	538

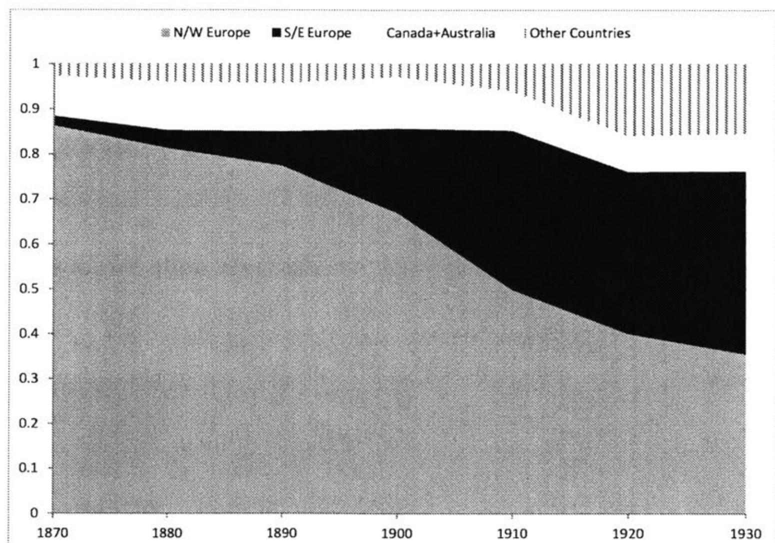
Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is the marriage rate of the groups described in each panel. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0.10, **p<0.05, ***p<0.01.

Figure 3.1 Immigrants as a Percent of US Population



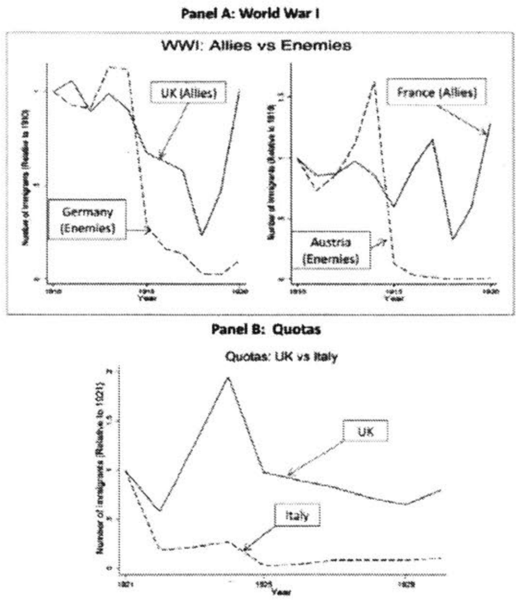
Note: The solid line shows the number of legal immigrants as a percent of US population. The dashed line includes also the estimated number of illegal immigrants, available from 2000 onwards. Source: the number of legal immigrants comes from the Migration Policy Institute, while the number of illegal immigrants was taken from the Pew Research Center tabulations.

Figure 3.2 Share of Foreign Born in the US



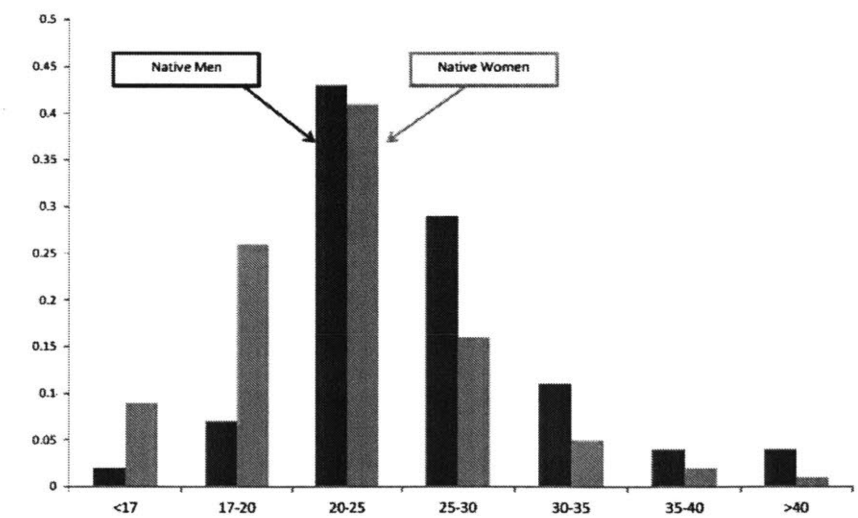
Note: share of immigrant stock living in the US, by sending region and decade. Authors' calculations from IPUMS sample of US Census (Ruggles et al., 2015).

Figure 3.3 Impact of Quotas and WWI on Share of Immigrants in the US



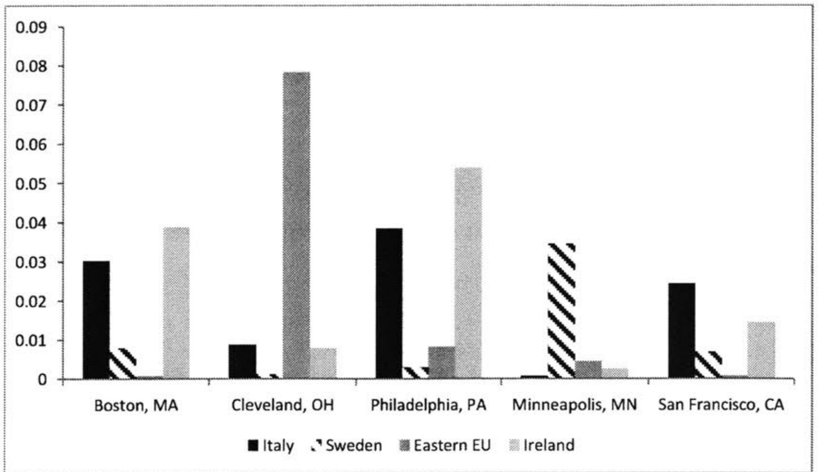
Note: the figure plots the number immigrants from each European country that entered the United States in each year, scaled by the number of immigrants from that country in 1910 (Panel A) and 1921 (Panel B). Source: adapted from Tabellini (2017).

Figure 3.4 Marriage Rates by Age and Gender



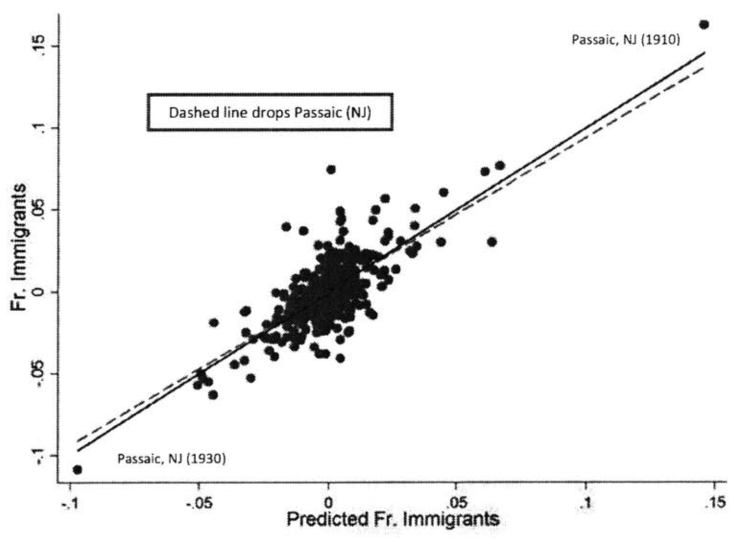
Note: Authors' calculations using IPUMS data.

Figure 3.5 Share of Immigrants from Selected Regions in US Cities, 1900



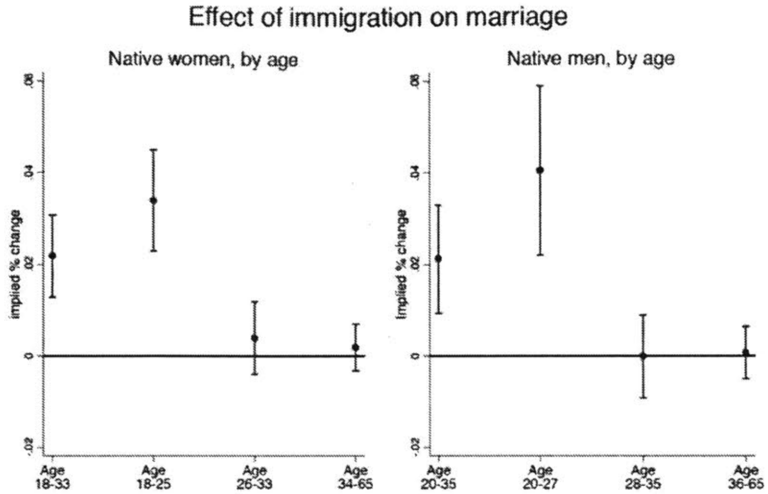
Note: share of individuals of European ancestry living in US cities in 1900, for selected ethnic groups. Source: Authors' calculations using IPUMS data.

Figure 3.6 First Stage



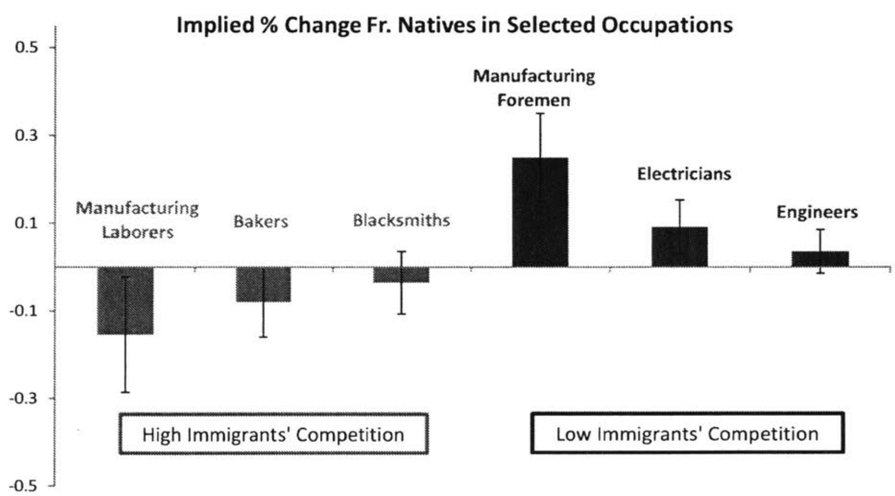
Note: the y-axis (resp. x-axis) reports the actual (resp. predicted) number of immigrants over predicted city population in each of the three Census years, 1910, 1920, and 1930. Each point in the scatter diagram represents the residual change in a city's actual and predicted fraction of immigrants after partialling out city and year by state fixed effects. The predicted number of immigrants is constructed as discussed in Section 4.2 in the text. Predicted city population is obtained by multiplying 1900 city population with average urban growth, excluding that of the Census division where a city is located. The solid line shows the regression coefficient for the full sample (coefficient=0.990, standard error=0.063). The dotted (red) line shows the regression coefficient obtained when dropping the city of Passaic, NJ (coefficient=0.940, standard error=0.068).

Figure 3.7 The Impact of Immigration on Marriage Rates by Gender and Age



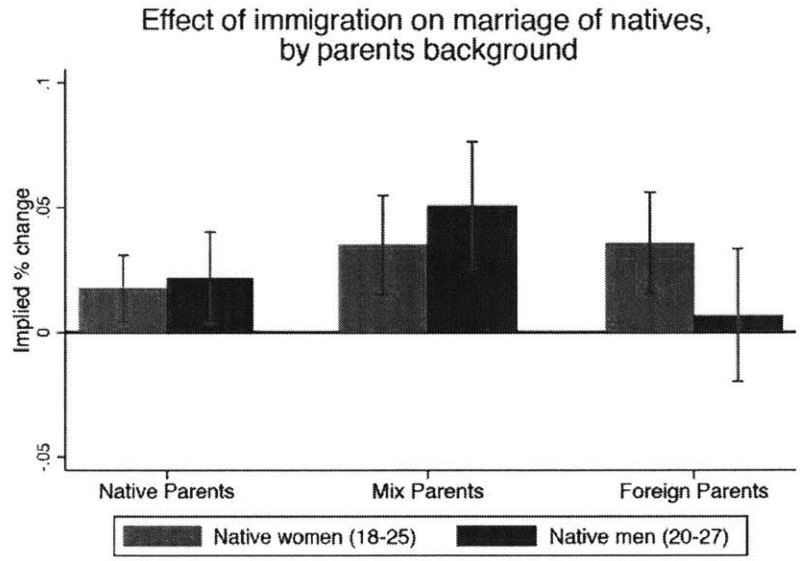
Notes: This graph shows the impact of one standard deviation increase of the fraction of immigrants on the increase in marriage rates with respect to the mean value in 1910. We report the standardized coefficients by age group and for men and women separately.

Figure 3.8 Natives' Occupation Mobility



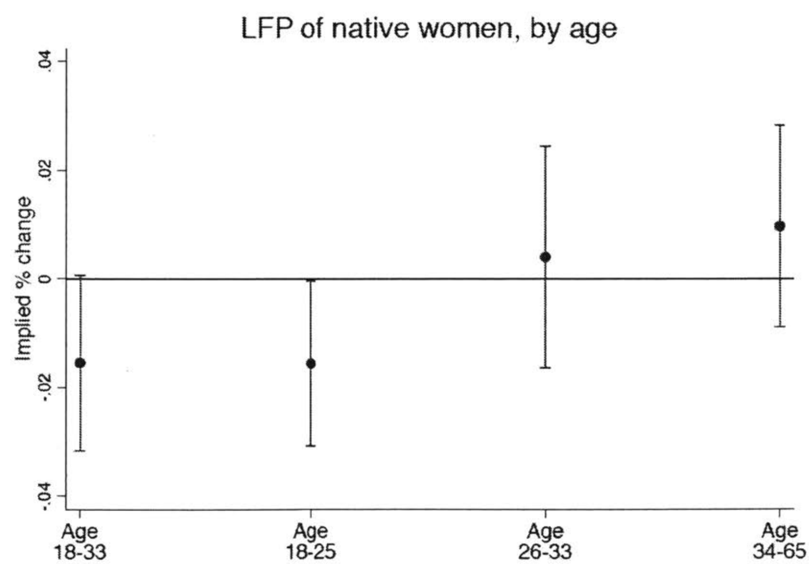
Note: the figure plots the percent change in the fraction of natives aged 20-35 in each occupation (relative to the 1910 mean) implied by a one standard deviation increase in immigration, according to 2SLS estimates (with corresponding 95% confidence intervals).

Figure 3.9 Impact of Immigration on Marriage Rates by Parentage



Note: This graph shows the impact of one standard deviation increase of the fraction of immigrants on the marriage rates of men and women by parentage with respect to the mean value in 1910. We report the standardized coefficients.

Figure 3.10 Impact of Immigration on Female Labor Force Participation by Age



Note: This graph shows the impact of one standard deviation increase of the fraction of immigrants on the decrease in labor force participation of women with respect to the mean value in 1910. We report the standardized coefficients separately by age group.

A1.1 Supplementary Tables and Figures to Chapter 1

Table A1.1. European Regions

UK	Russia
Ireland	Eastern Europe (Yugoslavia, Czechoslovakia, etc.)
Denmark	Austria-Hungary
Finland	Switzerland
Norway	France
Sweden	Belgium-Netherlands
Germany	Greece-Portugal-Spain
Poland	Italy

Note: this table lists the European sending regions used to construct the instrument for immigration.

Table A1.2. City List

Akron, OH	Elizabeth, NJ	McKeesport, PA	Saint Joseph, MO
Albany, NY	Elmira, NY	Memphis, TN	Saint Louis, MO
Allentown, PA	Erie, PA	Milwaukee, WI	Saint Paul, MN
Altoona, PA	Evansville, IN	Minneapolis, MN	Salem, MA
Amsterdam, NY	Everett, MA	Mobile, AL	San Antonio, TX
Atlanta, GA	Fall River, MA	Montgomery, AL	San Diego, CA
Atlantic City, NJ	Fitchburg, MA	Mount Vernon, NY	San Francisco, CA
Auburn, NY	Flint, MI	Nashville, TN	Savannah, GA
Augusta, GA	Fort Wayne, IN	New Bedford, MA	Schenectedy, NY
Baltimore, MD	Fort Worth, TX	New Britain, CT	Scranton, PA
Bay City, MI	Galveston, TX	New Castle, PA	Seattle, WA
Bayonne, NJ	Grand Rapids, MI	New Haven, CT	Sioux City, IA
Berkeley, CA	Hamilton, OH	New Orleans, LA	Somerville, MA
Binghamton, NY	Harrisburg, PA	New York, NY	South Bend, IN
Birmingham, AL	Hartford, CT	Newark, NJ	Spokane, WA
Boston, MA	Haverhill, MA	Newton, MA	Springfield, IL
Bridgeport, CT	Hoboken, NJ	Niagara Falls, NY	Springfield, MA
Brockton, MA	Holyoke, MA	Norfolk, VA	Springfield, MO
Buffalo, NY	Houston, TX	Oakland, CA	Springfield, OH
Butte, MT	Huntington, WV	Oklahoma City, OK	Superior, WI
Cambridge, MA	Indianapolis, IN	Omaha, NE	Syracuse, NY
Camden, NJ	Jackson, MI	Oshkosh, WI	Tacoma, WA
Canton, OH	Jacksonville, FL	Pasadena, CA	Tampa, FL
Cedar Rapids, IA	Jamestown , NY	Passaic, NJ	Taunton, MA
Charleston, SC	Jersey City, NJ	Paterson, NJ	Terre Haute, IN
Charlotte, NC	Johnstown, PA	Pawtucket, RI	Toledo, OH
Chattanooga, TN	Joliet, IL	Peoria, IL	Topeka, KS
Chelsea, MA	Kalamazoo, MI	Perth Amboy, NJ	Trenton, NJ
Chester, PA	Kansas City, KS	Philadelphia, PA	Troy, NY
Chicago, IL	Kansas City, MO	Pittsburgh, PA	Utica, NY
Cincinnati, OH	Knoxville, TN	Pittsfield, MA	Washington, DC
Cleveland, OH	La Crosse, WI	Portland, ME	Waterbury, CT
Columbus, OH	Lancaster, PA	Portland, OR	Wheeling, WV
Covington, KY	Lansing, MI	Portsmouth, VA	Wichita, KS
Dallas, TX	Lawrence, MA	Providence, RI	Wilkes-Barre, PA
Davenport, IA	Lexington, KY	Pueblo, CO	Williamsport, PA
Dayton, OH	Lima, OH	Quincy, IL	Wilmington, DE
Decatur, IL	Lincoln, NE	Quincy, MA	Woonsocket, RI
Denver, CO	Little Rock, AR	Racine, WI	Worcester, MA
Des Moines, IA	Los Angeles, CA	Reading, PA	Yonkers, NY
Detroit, MI	Louisville, KY	Richmond, VA	York, PA
Dubuque, IA	Lowell, MA	Roanoke, VA	Youngstown, OH
Duluth, MN	Lynn, MA	Rochester, NY	
East Orange, NJ	Macon, GA	Rockford, IL	
East St. Louis, IL	Malden, MA	Sacramento, CA	
El Paso, TX	Manchester, NH	Saginaw, MI	

Table A1.3. Labor Market Characteristics of Immigrants and Natives

	Natives	Immigrants	Ratio (Natives to Immigrants)
Panel A: Industries			
Manufacturing	0.216	0.437	0.494
Construction	0.089	0.107	0.832
Trade	0.182	0.169	1.077
Services (excluding personal)	0.098	0.037	2.649
Public Sector	0.034	0.005	6.800
Panel B: Skills and Broad Occupational Groups			
High Skilled	0.345	0.126	2.738
Unskilled	0.347	0.614	0.565
Clerical and Sales	0.198	0.065	3.046
Laborers	0.110	0.311	0.354
Panel C: Narrowly Defined Occupations			
Manuf. Laborers	0.038	0.150	0.253
Waiters	0.007	0.012	0.583
Blacksmiths	0.006	0.008	0.750
Manuf. Supervisors	0.007	0.002	3.500
Electricians	0.010	0.003	3.667
Engineers	0.021	0.005	4.200

Note: this table presents the fraction of natives and of immigrants in selected industries (Panel A), skill categories (Panel B), and narrowly defined occupations (Panel C) in 1910. For both natives and immigrants, the sample is restricted to males in working age living in the 180 cities in my sample. The last column on the right shows the ratio of the fraction of natives over the fraction of immigrants in a given industry/skill category/occupation.

Table A1.4. Additional Results and Placebo Checks

	Natives Only					Natives and Immigrants	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Var:	In Labor Force	High-Low Skill Ratio	Employed Illiterate	Employed Blacks	Employed Labor manuf	Log workers manuf	Log avg. wage manuf
Panel A: OLS							
Fr. Immigrants	0.205***	-0.030	-0.147	-0.108	0.037	1.671***	-0.091
	(0.050)	(0.034)	(0.217)	(0.273)	(0.098)	(0.557)	(0.237)
Panel B: 2SLS							
Fr. Immigrants	0.204***	0.061*	-0.109	-0.107	0.078	1.471***	-0.186
	(0.065)	(0.036)	(0.332)	(0.269)	(0.114)	(0.527)	(0.291)
F-stat	251.3	251.3	251.3	251.3	251.3	270.5	270.5
Mean dep var	0.954	0.978	0.745	0.750	0.941	9.063	6.275
Observations	538	538	538	538	538	525	525

Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930 (Cols 1-5), and for which data were reported in the Census of Manufacture between 1909 and 1929 (Cols 6-7). Variables in Cols 1 to 5 refer to native men in the age range 15 to 65 who were not enrolled in schools. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is: an indicator for holding any gainful occupation (*In Labor Force*) in Col 1; the log of high skill natives over the log of low skill natives in Col 2; and the employment rate for illiterate natives, for African Americans, and for natives working as manufacturing laborers in Cols 3 to 5 respectively. Variables in Cols 6-7 refer to the whole labor force in the manufacturing sector (from the Census of Manufacture), and include both immigrant and native workers. The dependent variable is (the log of) the number of workers employed in manufacturing in Col 6; and (the log of) the average wage in manufacturing in Col 7. To classify individuals across skill categories, I use the classification made by Katz and Margo (2013). *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. The mean of each dependent variable at baseline is shown at the bottom of the Table. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.5. Additional Results for Economic Activity

VARIABLES	(1) Log value added per capita	(2) Log value of products per establishment	(3) Log value of products per capita	(4) Log horsepower	(5) TFP
Panel A: OLS					
Fr. Immigrants	0.785	2.264***	0.992*	1.267***	0.295
	(0.580)	(0.704)	(0.556)	(0.475)	(0.358)
Panel B: 2SLS					
Fr. Immigrants	1.404**	3.549***	2.065**	1.906***	1.013*
	(0.586)	(1.214)	(0.845)	(0.705)	(0.540)
F-stat	270.5	270.5	270.5	270.5	270.5
Cities	178	178	178	178	178
Observations	525	525	525	525	525

Note: this Table presents results for a balanced panel of the 178 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930, and for which data were reported in the Census of Manufacture between 1909 and 1929. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is: the log of value added per capita in Col 1; the log of value of products per establishment (per capita) in Col 2 (Col 3); the log of horsepower in Col 4; and total factor productivity (TFP) in Col 5. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.6. Tax Revenues and Property Values

Dep. Var.	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property values PC	(4) Property values over 1910 pop	(5) Business Taxes PC
Panel A: OLS					
Fr. Immigrants	-8.525	-8.060	372.4	240.3	0.268
	(6.490)	(5.515)	(740.6)	(562.1)	(1.677)
Panel B: 2SLS					
Fr. Immigrants	-11.15	-11.08*	294.6	518.3	1.843
	(6.982)	(6.467)	(915.3)	(740.9)	(1.604)
F-stat	288.3	288.3	288.3	288.3	288.3
Mean of dep var	12.53	12.04	715.9	715.9	0.889
Cities	180	180	180	180	180
Observations	540	540	540	540	540

Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is total (resp. property) tax revenues per capita in Col 1 (resp. Col 2); property values per capita (resp. over 1910 population) in Col 3 (resp. Col 4); and business taxes per capita in Col 5. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.7. Public Spending Per Capita, by Category

VARIABLES	(1) Education	(2) Police	(3) Fire	(4) Charities and hospitals	(5) Sanitation
Panel A: OLS					
Fr. Immigrants	-7.453***	0.227	-0.369	0.486	-0.537
	(2.332)	(0.560)	(0.552)	(0.747)	(0.696)
Panel B: 2SLS					
Fr. Immigrants	-6.170***	-0.345	-0.213	-1.258	-1.318*
	(2.146)	(0.663)	(0.680)	(1.897)	(0.717)
F-stat	248.6	288.3	288.3	220.3	288.3
Mean dep var	4.250	1.338	1.485	0.635	1.129
Cities	180	180	180	175	180
Observations	534	540	540	516	540

Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable, in per capita terms, is displayed at the top of each column. Sanitation (Col 5) includes garbage collection, sewerage, and other spending on sanitation. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.8. Additional Electoral Outcomes

VARIABLES	(1) Republicans' vote share	(2) Other parties' vote share	(3) Democrats-Republicans Margin	(4) Turnout
Panel A: OLS				
Fr. Immigrants	0.337**	0.191	-0.866***	-1.033***
	(0.133)	(0.127)	(0.219)	(0.233)
Panel B: 2SLS				
Fr. Immigrants	0.169	0.235**	-0.573**	-1.422***
	(0.149)	(0.101)	(0.272)	(0.183)
F-stat	83.14	83.14	83.14	83.52
Mean dep var	0.310	0.200	0.181	0.504
MSAs	126	126	126	125
Observations	378	378	378	375

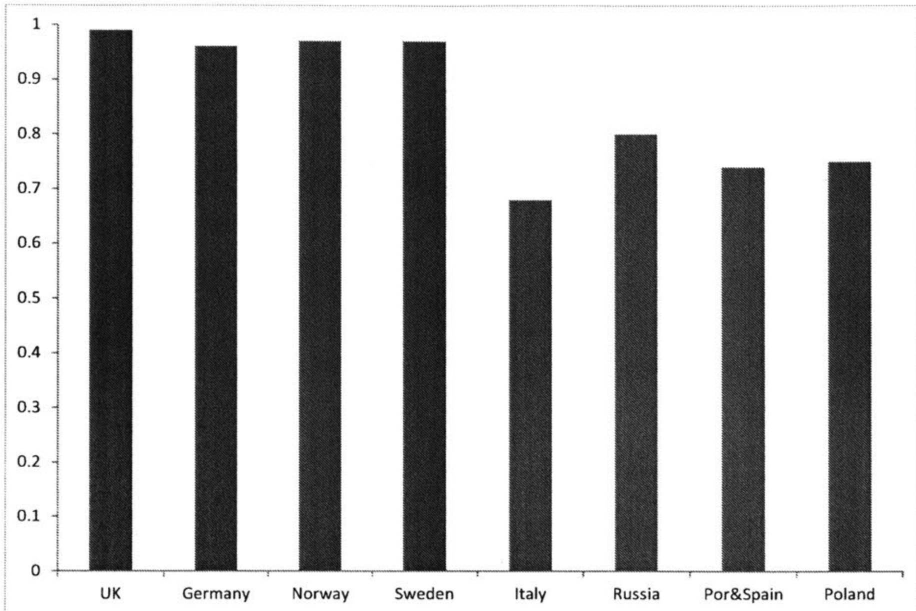
Note: this Table presents results for a balanced panel of the 126 metropolitan statistical areas (MSAs) including at least one of the 180 cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is reported at the top of each column, and refers to Presidential elections. All electoral outcomes were aggregated from the county to the MSA level, using the 1940 MSAs' definitions, and were computed as the average between the closest two elections after each Census year. Results are unchanged when taking the average from the two closest election years (see the online appendix). *Other parties' vote share refers* to the vote share of all parties other than Democrats and Republicans. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include MSA and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.9. Linguistic Distance vs Literacy

Dep. Var.	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Public spending PC	(5) Education	(6) Police	(7) Charities and Hospitals	(8) Sanitation
Panel A: OLS								
Ling. Distance	-0.292	-0.260	-0.997	-0.183	-0.062	-0.020	-0.044	-0.028
	(0.185)	(0.180)	(0.701)	(0.151)	(0.054)	(0.019)	(0.036)	(0.033)
Literacy	0.058	0.160	0.404	0.093	0.099	0.026	-0.060	-0.028
	(0.181)	(0.169)	(0.327)	(0.132)	(0.063)	(0.020)	(0.041)	(0.026)
Panel B: 2SLS								
Ling. Distance	-0.946**	-0.861*	-2.340	-0.575*	-0.177	0.001	-0.131	-0.065
	(0.458)	(0.450)	(1.553)	(0.314)	(0.128)	(0.046)	(0.092)	(0.054)
Literacy	-0.294	-0.217	-0.129	-0.234	0.096	0.062	-0.091	-0.054
	(0.327)	(0.303)	(0.801)	(0.266)	(0.099)	(0.039)	(0.097)	(0.051)
KP F-stat	14.30	14.30	14.57	14.30	14.45	14.30	10.89	14.30
F-stat (Imm.)	101.7	101.7	102.1	101.7	87.48	101.7	83.47	101.7
F-stat (Ling.)	36.48	36.48	37.87	36.48	34.74	36.48	26.10	36.48
F-stat (Lit.)	21.77	21.77	21.68	21.77	21.70	21.77	21.27	21.77
Mean of dep var	12.76	12.10	19.75	12.16	4.250	1.338	0.635	1.129
Observations	540	540	539	540	534	540	516	540

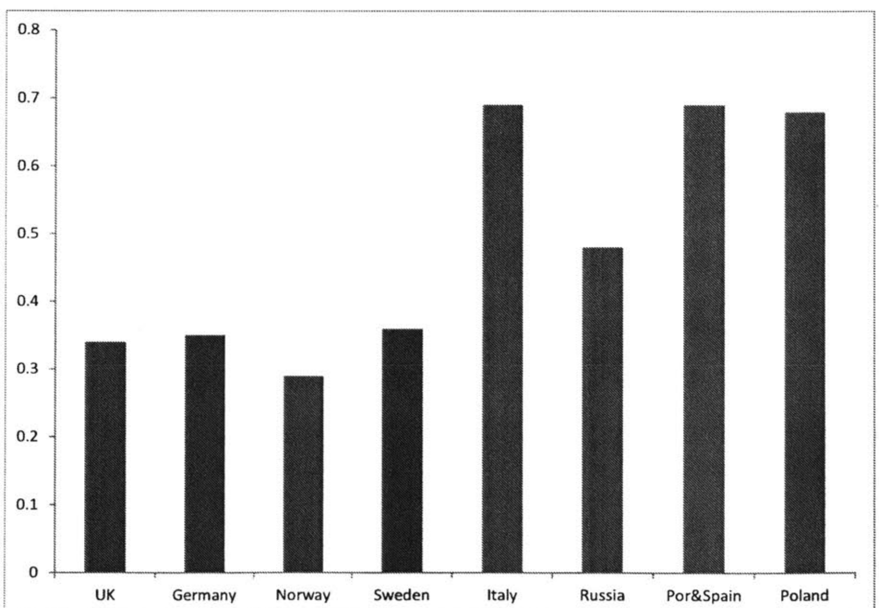
Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is displayed at the top of each column. In Cols 5 to 8, the dependent variable is spending per capita on the category listed at the top of the column. The main regressors of interest are the (standardized) weighted average linguistic distance and literacy index, and are instrumented using predicted shares of immigrants from each sending region obtained from (2) in Section 4.2. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. F-stat (Imm.), F-stat (Ling.), and F-stat (Lit.) refer to the partial F-stats for joint significance of the instruments in the three separate first-stage regressions. All regressions include the (instrumented) fraction of immigrants, and control for city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Figure A1.1. Literacy Rates, for Selected Sending Regions (1910)



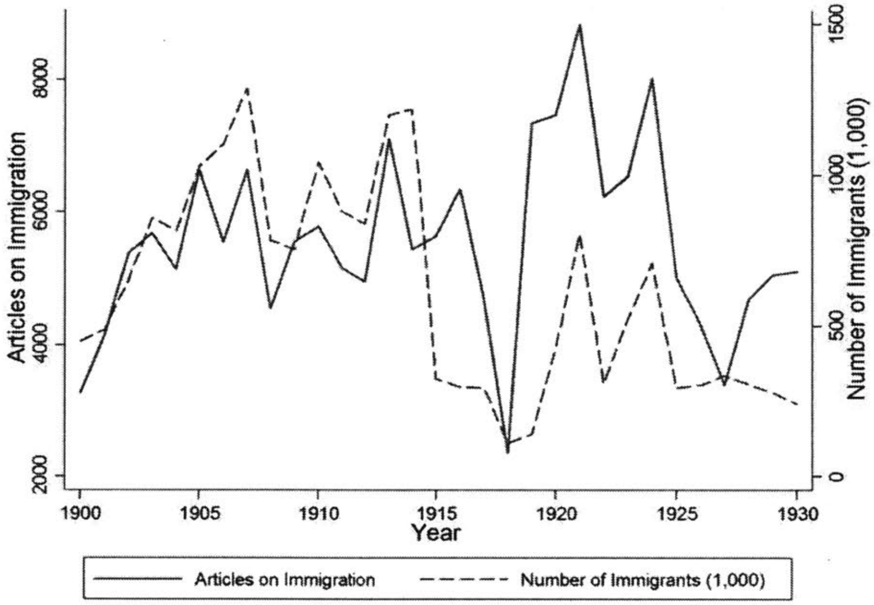
Note: this Figure reports the literacy rate for men in the age range (15-65) for selected immigrants' groups in 1910. Source: Author's calculations using IPUMS data.

Figure A1.2. Fraction Unskilled, for Selected Sending Regions (1910)



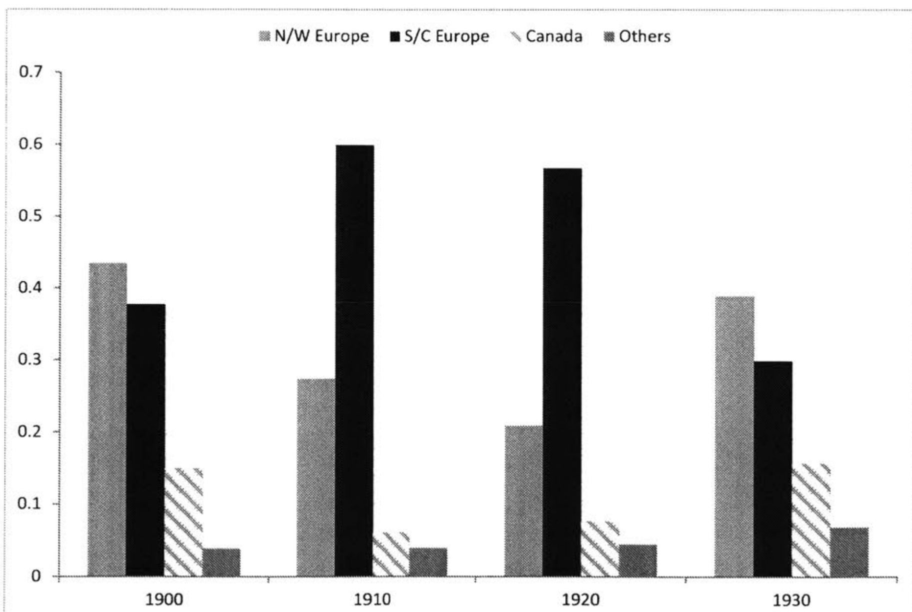
Note: this Figure reports the fraction of men in the age range (15-65) in unskilled occupations for selected immigrants' groups in 1910. Source: Author's calculations using IPUMS data.

Figure A1.3. Immigration and Newspapers' Coverage



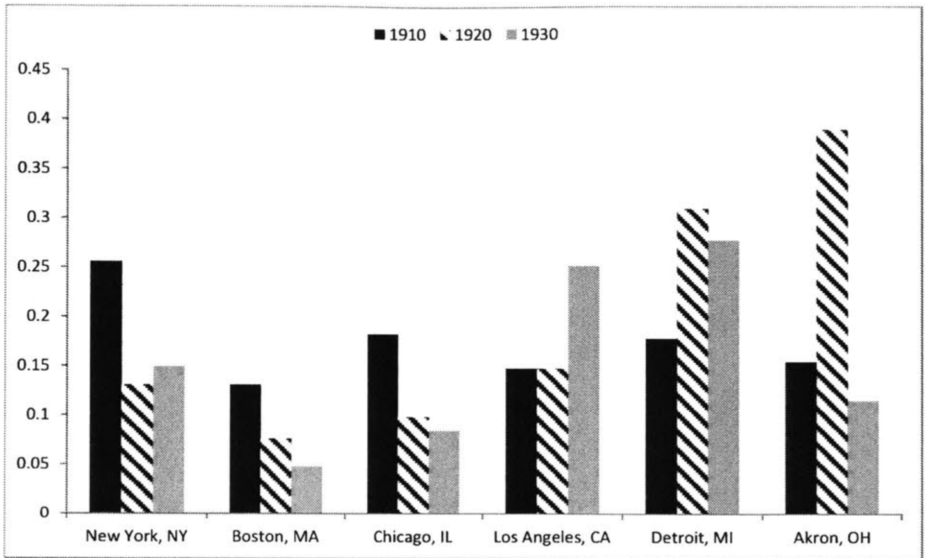
Note: the Figure plots the annual number of immigrants in thousands (dashed blue line, right-axis) and the number of times the words "immigration" and "immigrants" appeared in local newspapers for all cities with at least 30,000 residents and for which data were available in the database of Newspapersarchive (solid red line, left-axis). Source: author's calculation using data from Newspapersarchive.

Figure A1.4. Share of Recent Immigrants, by Region and Decade



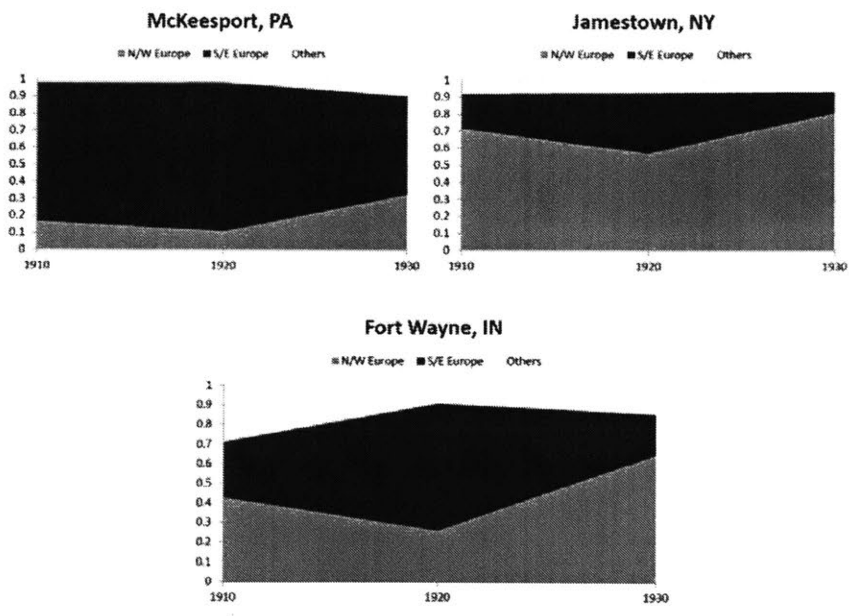
Note: Share of immigrant entering the United States in the previous ten years, by sending region and by decade. Source: Author's calculations from IPUMS sample of US Census (Ruggles et al., 2015).

Figure A1.5. Recent Immigrants Over 1900 City Population, by Decade



Note: Number of European immigrants that arrived in the United States in the last decade over 1900 city population, for selected cities and by decade. Source: Author's calculations from IPUMS sample of US Census (Ruggles et al., 2015).

Figure A1.6. Changing Composition of Immigrants in Selected Cities

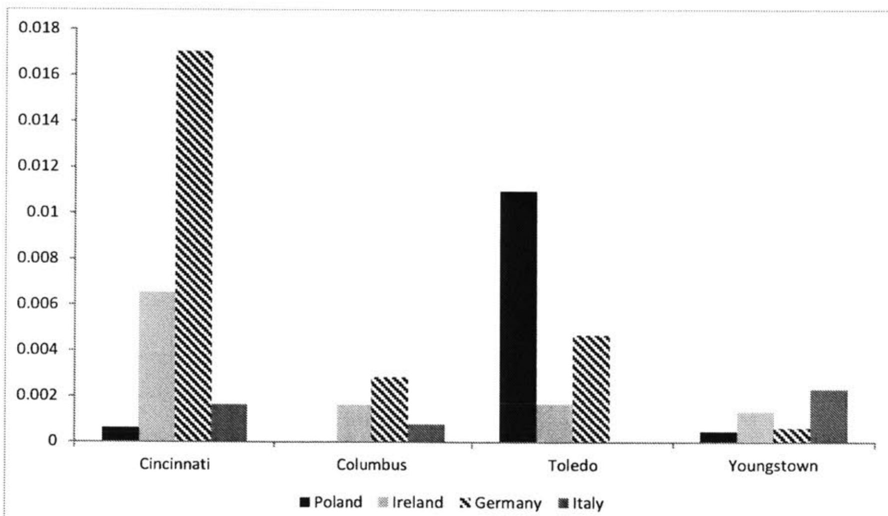


Note: Share of immigrants entering the US in the previous decade from different regions living in selected cities. Source: Author's calculations from IPUMS sample of US Census (Ruggles et al., 2015).

Figure A1.7. Map of Cities

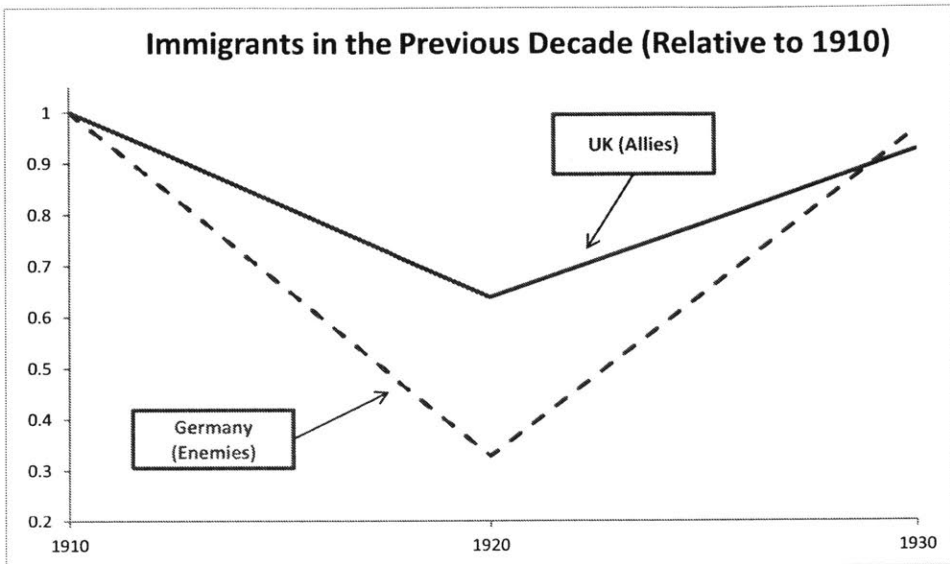
Note: The map plots the 180 cities with at least 30,000 residents in each of the three Census years 1910, 1920, and 1930.

Figure A1.8. Share of European Immigrants in Ohio, 1900



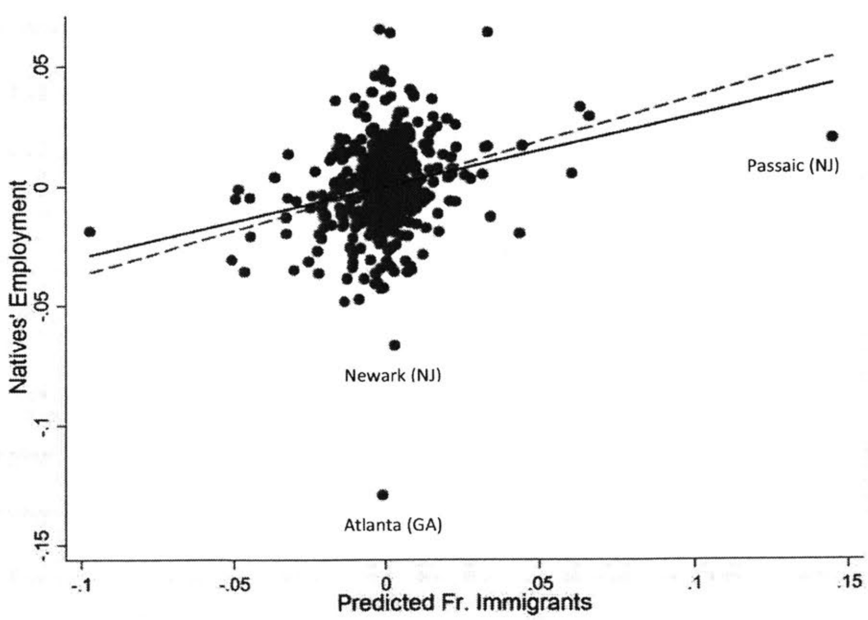
Note: share of individuals of European ancestry living in selected cities of Ohio in 1900, for selected ethnic groups. Source: Author's calculations using IPUMS data.

Figure A1.9. The Effect of WWI on Immigration from Allies and Enemies



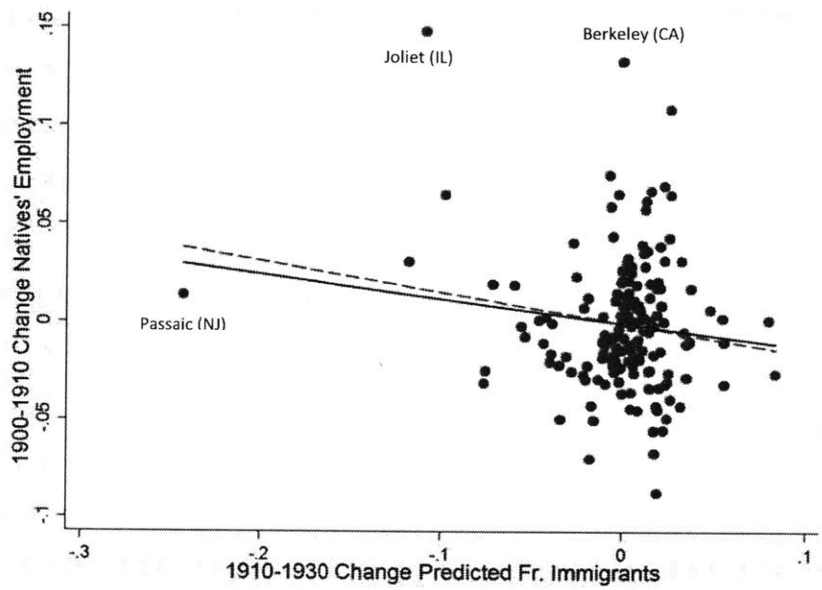
Note: the figure plots the number immigrants from Germany (blue, dashed line) and from the UK (red line) that entered the United States during the previous decade, normalizing them to 1 relative to 1910. Source: author's calculation using IPUMS data.

Figure A1.10. Natives' Employment and Immigration: Reduced Form



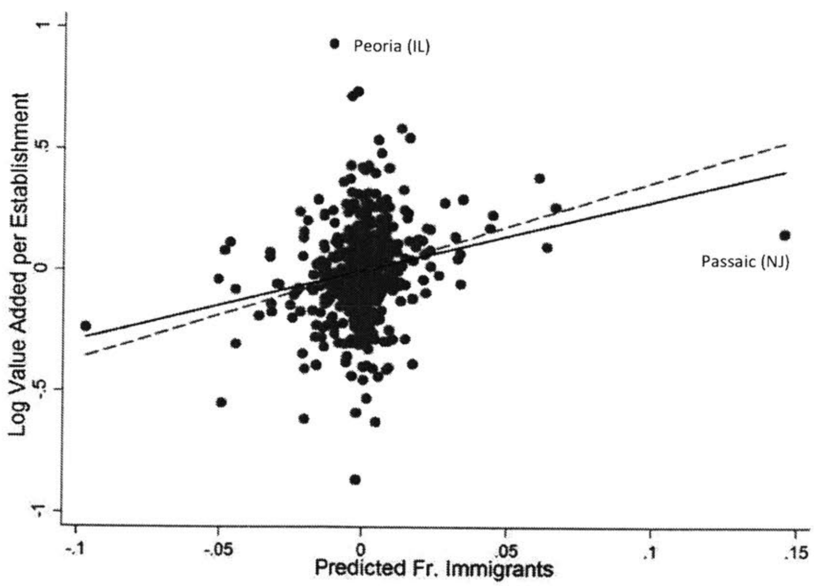
Note: the y-axis and the x-axis report, respectively, the employment to population ratio for native males in working age who were not in school and predicted fraction of immigrants over predicted city population in each of the three Census years, 1910, 1920, and 1930. Each point in the scatter diagram represents the residual change in each of the two variables after partialling out city and state by year fixed effects. The solid line shows the regression coefficient for the full sample (coefficient=0.296, standard error=0.054). The dotted (red) line shows the regression coefficient obtained when dropping the city of Passaic, NJ (coefficient=0.371, standard error=0.065).

Figure A1.11. Natives' Employment and Immigration: Placebo Check



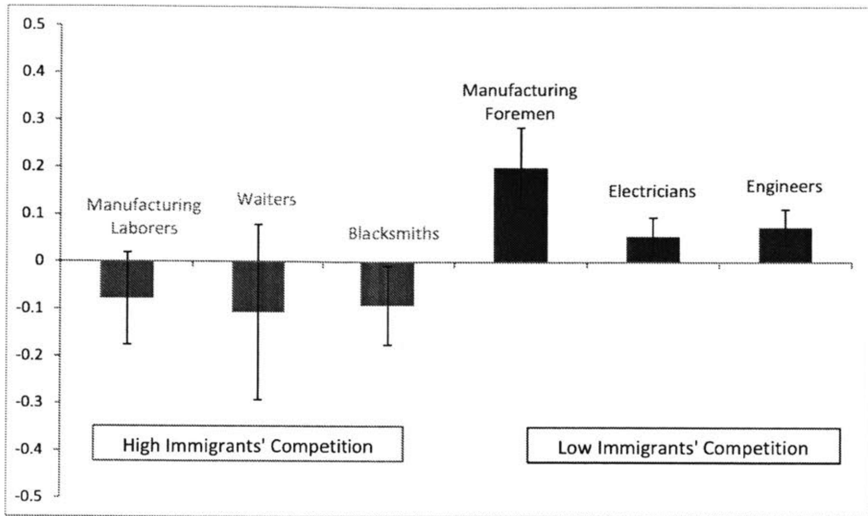
Note: this figure shows the residual plot of the 1900-1910 change in employment to population ratio (y-axis) against the 1910-1930 change in the predicted fraction of immigrants over predicted city population (x-axis) after partialling out city and state by year fixed effects. The solid line shows the regression coefficient for the full sample (coefficient= -0.119, standard error=0.110). The dotted (red) line shows the regression coefficient obtained when dropping the city of Passaic, NJ (coefficient= -0.151, standard error=0.155).

Figure A1.12. Value Added and Immigration: Reduced Form



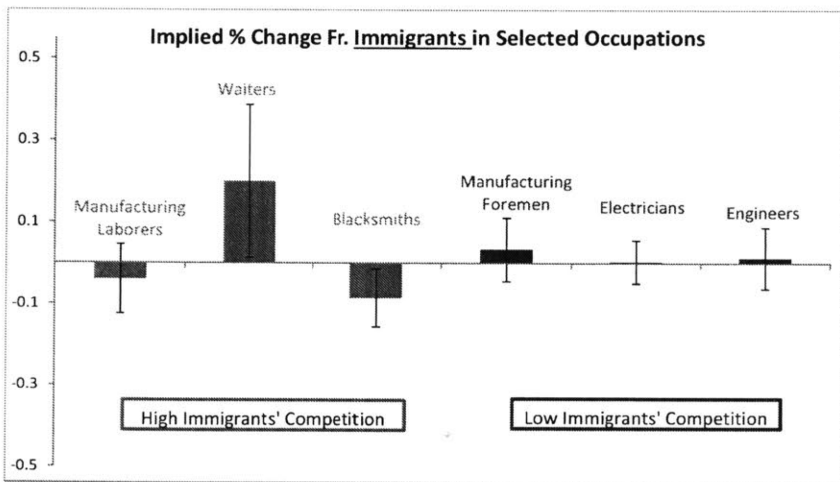
Note: the y-axis and the x-axis report, respectively, the log of value added per establishment and predicted fraction of immigrants over predicted city population in each of the three Census years, 1910, 1920, and 1930. Each point in the scatter diagram represents the residual change in each of the two variables after partialling out city and state by year fixed effects. The solid line shows the regression coefficient for the full sample (coefficient=2.874, standard error=0.868). The dotted (red) line shows the regression coefficient obtained when dropping the city of Passaic, NJ (coefficient=3.685, standard error=0.825).

Figure A1.13. Percent Change in Fraction of Natives in Selected Occupations



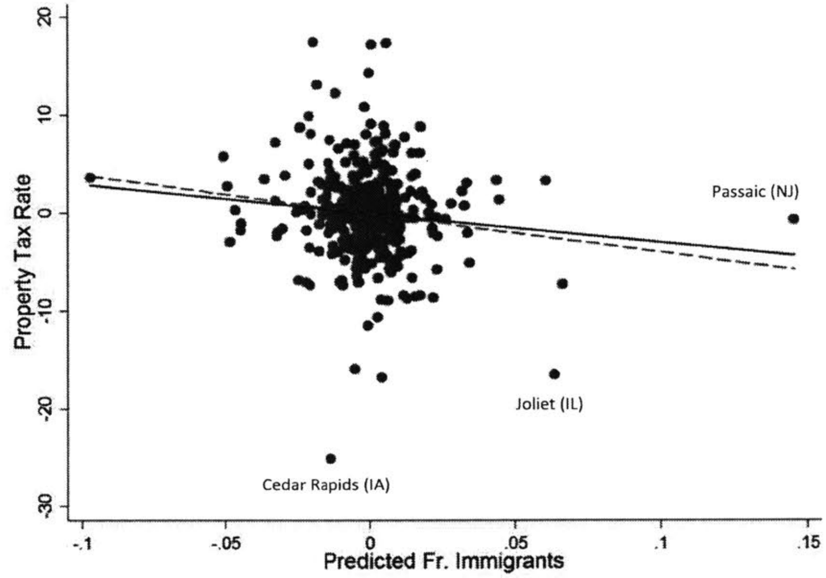
Note: the figure plots the percent change in the fraction of natives in each occupation (relative to its 1910 mean) implied by a one standard deviation increase in immigration, according to 2SLS estimates (with corresponding 95% confidence intervals) reported in Table 4.

Figure A1.14. Percent Change in Fraction of Immigrants Across Occupations



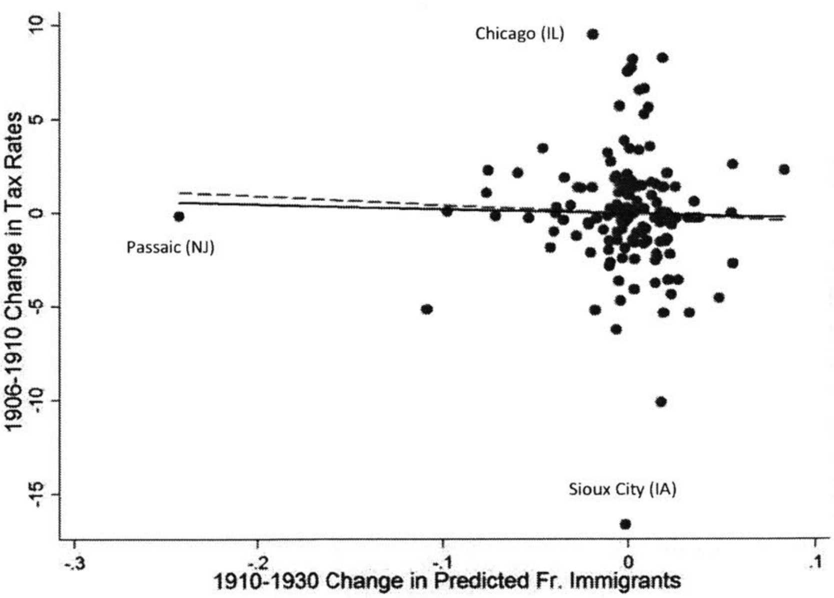
Note: the figure replicates Figure 9 by plotting the percent change in the fraction of immigrants arrived at least 10 year before in each occupation (relative to its 1910 mean) implied by a one standard deviation increase in immigration, according to 2SLS estimates (with corresponding 95% confidence intervals).

Figure A1.15. Tax Rates and Immigration: Reduced Form



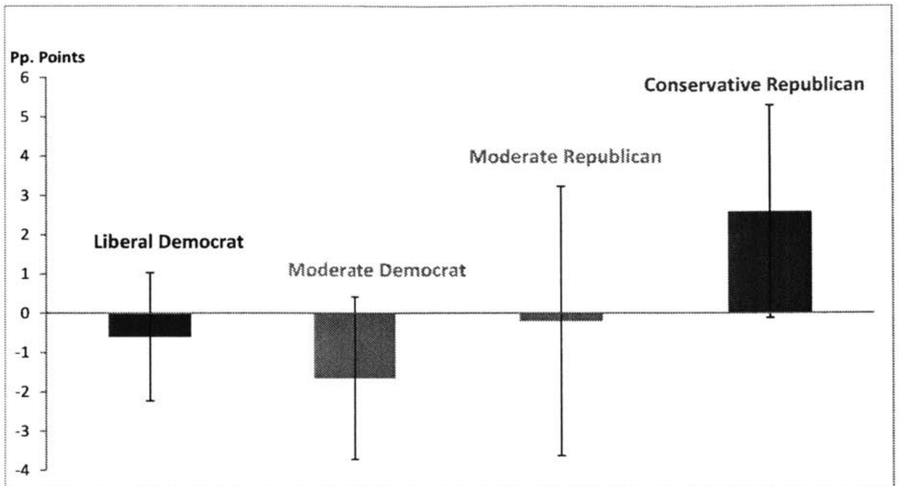
Note: the y-axis and the x-axis report, respectively, the property tax rate and predicted fraction of immigrants over predicted city population in each of the three Census years, 1910, 1920, and 1930. Each point in the scatter diagram represents the residual change in each of the two variables after partialling out city and state by year fixed effects. The solid line shows the regression coefficient for the full sample (coefficient= -29.45, standard error=16.03). The dotted (red) line shows the regression coefficient obtained when dropping the city of Passaic, NJ (coefficient= -39.37, standard error=22.39).

Figure A16. Tax Rates and Immigration: Placebo Check



Note: this figure shows the residual plot of the 1900-1910 change in the property tax rate (y-axis) against the 1910-1930 change in the predicted fraction of immigrants over predicted city population (x-axis) after partialling out city and state by year fixed effects. The solid line shows the regression coefficient for the full sample (coefficient= -2.279, standard error=6.869). The dotted (red) line shows the regression coefficient obtained when dropping the city of Passaic, NJ (coefficient= -4.518, standard error=8.810).

Figure A1.17. Probability that Winner Has Given Political Orientation



Note: the figure plots 2SLS estimates (with corresponding 95% confidence intervals) reported in columns 3 to 6 of Table 8 (Panel B) for the probability that the member of the House of Representatives elected has a given political orientation. Liberal (resp. moderate) Democrats are defined as legislators with a Nominate score below (resp. above) the median score for Democrats in the 61[st] Congress. A Republican legislator is classified as moderate (resp. conservative) if his Nominate score is below (resp. above) the median score for Republicans in the 61[st] Congress.

A1.2 Predicting Migration Using Weather Shocks

A number of works have documented the link between agricultural output and weather conditions in Europe during the Age of Mass Migration. For instance, Solomou and Wu (1999) find that more than half of the variation in agricultural production in France, Germany, and the UK between 1850 and 1915 can be explained by temperature and precipitation shocks. At the same time, in a seminal contribution, Hatton and Williamson (1995) argue that agricultural conditions were strongly related to outmigration rates in Europe in this historical period. A similar pattern is found also for contemporary international migration in less developed countries (e.g. Feng et al. (2010)).

Motivated by this evidence, I exploit variation in weather shocks across European countries over time to predict migration flows that are independent of US economic or political conditions. As in Sequeira et al. (2017), I make use of historical precipitation and temperature data from, respectively, Pauling et al. (2006) and Luterbacher et al. (2004). The data are measured at annual frequency for each season of the year, and are available at a 0.5 degree spatial resolution. Since out-migration is available at the country-level, I averaged weather variables over all grid cells for each country.

To estimate yearly outmigration rates, I digitized data from the *Commissioner General of Immigration* between 1900 and 1930.[1] I use migration flows classified by race rather than by country of origin to deal with the non-trivial problem that the boundaries of several European countries changed significantly between 1900 and 1930. Ethnic groups were then mapped to the country of birth reported in the US Census of Population to match migration flows with 1900 immigrants' shares, α_{jc} in equation (2) in the main text.[2]

For each year between 1900 and 1930 and for each country, I estimated

$$\ln Outmig_{jy} = \alpha + \sum_{s=1}^{4} \sum_{m \in M} \beta_{jsm} I_{jy-1}^{s,m} + \varepsilon_{jy-1} \tag{A1}$$

where $\ln Outmig_{jy}$ is the log of migrants from European country j in year y; and $I_{jy-1}^{s,m}$ is a dummy equal to 1 if the average precipitation (or temperature) in season s falls in the range m. In my baseline specification, I consider precipitation shocks, but results are unchanged when using temperature.

Following Sequeira et al. (2017), I consider the following six categories $m \in M$: more

[1] The US Census of Population records migration only at decadal frequency, and so cannot be used to perform this exercise.

[2] This exercise was relatively straightforward, except when matching individuals of Hebrew race to the corresponding country of origin. I experimented with several alternatives and, reassuringly, results remained always very similar.

than 3 standard deviations below the mean; between 2 and 3 standard deviations below the mean; between 1 and 2 standard deviations below the mean; between 1 and 2 standard deviations above the mean; between 2 and 3 standard deviations above the mean; and more than 3 standard deviations above the mean. That is, I omit the category for precipitations (or temperatures) that are within one standard deviation below or above the mean.

After separately estimating $(A1)$ for each country in my sample, I predict log migrant flows (for each country in each year), $\ln \widehat{Outmig}_{jy}$, using the β_{jsm}'s estimated from $(A1)$. Figure A1.18 plots the relationship between actual and predicted (log) migration flows, and shows that the two are strongly correlated. Next, I aggregate predicted flows at the decadal frequency to get

$$\hat{O}_{jt} = \sum_{y} \exp\left(\ln \widehat{Outmig}_{jy}\right) \tag{A2}$$

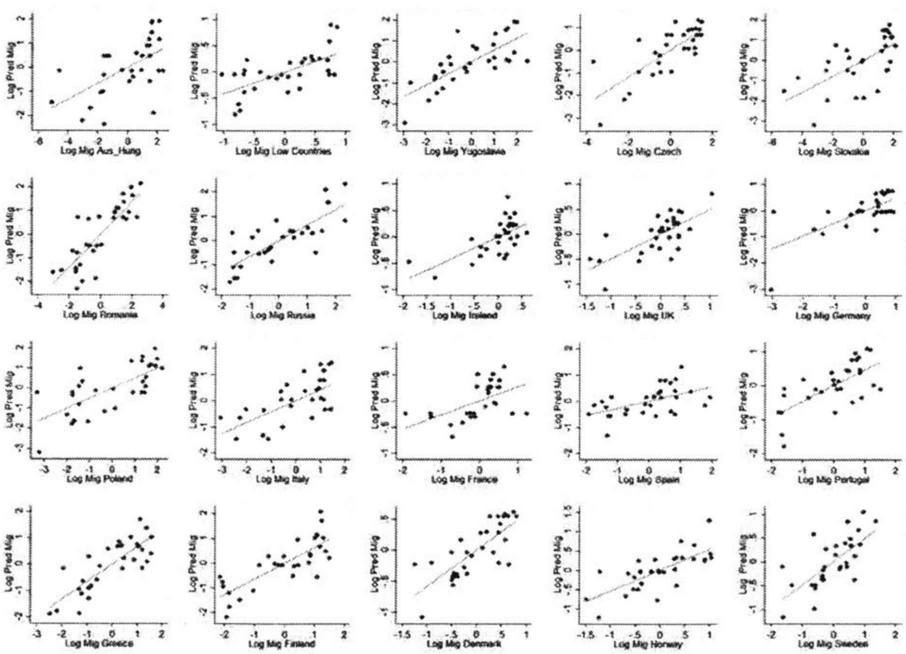
Below, I assess the robustness of my results using an alternative version of the shift-share instrument, obtained by replacing O_{jt}^{-M} with $\hat{O}_{jt}$ in equation (2) in the main text. In Table A1.10, I report estimates for the relationship between actual immigration and the instrument constructed in the previous paragraph. To ease comparisons, columns 1 and 2 show the baseline specification estimated in the main text, and reported in columns 1 and 2 of Table 1.2. Next, columns 3 and 4 replace actual migration flows from each sending country with those predicted using only weather shocks at origin (see $(A2)$). Column 3 scales the number of immigrants by predicted population, while column 4 divides it by 1900 population. Both the coefficients and the F-stat in columns 3 and 4 fall, but the relationship between actual and predicted immigration remains positive and highly significant even when using this alternative instrument.

Table A1.10. First Stage for Weather Shocks Instrument

	Dep. Variable: *Fraction of Immigrants*			
	(1)	(2)	(3)	(4)
Predicted Fr. Immigrants	0.999***	0.968***	0.725***	0.738***
	(0.059)	(0.064)	(0.168)	(0.155)
Immigrants over	Predicted Pop.	1900 Pop.	Predicted Pop.	1900 Pop.
Weather Shocks			X	X
F-stat	288.3	226.7	18.70	22.65
Observations	540	540	540	540

Note: the sample includes a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. Col 1 presents results for the baseline specification, where the number of immigrants is scaled by predicted population, and the instrument is the leave-out version of the shift-share IV in equation (2) (Section 4.2). Col 2 replicates Col 1 for a long differences specification, regressing the 1910 to 1930 change in the fraction of immigrants against the corresponding change in the instrument. Cols 3 and 4 replicate Cols 1 and 2 replacing actual aggregate flows (by country of origin) with those predicted exploiting only weather shocks at origin (see (A2) in the online appendix). All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Figure A1.18. Actual Versus Predicted Migration Using Precipitation Shocks



A1.3 Additional Robustness Checks

In this section, I conduct a number of robustness checks, in addition to those reported in the main text.

A1.3.1 Outliers

First, I check that results are not driven by the city of Passaic (NJ), which often appeared as a potential outlier in the scatterplots presented in the main paper, or by other cities with extreme values (either high or low) of immigration. In Table A1.11, I start by replicating each of the specifications for the first stage reported in Table 1.2 in the main text by dropping Passaic (NJ). Reassuringly, coefficients always remain highly significant and quantitatively close to those presented in Table 1.2 and, in all cases, the F-stat is above conventional levels. Figure A1.19 replicates Figure 1.6 in the paper, and reports the graphical analogue of column 3, plotting the relationship between the fraction of immigrants and the instrument, after partialling out city and state by year fixed effects. Then, in Table A1.12, I replicate Table 1.2 by excluding cities in the 1^{st} (Flint, MI, and Pasadena, CA) and 99^{th} (Passaic, NJ, and Perth Amboy, NJ) percentiles of the 1910 to 1930 change in immigration. As for Table A1.11, all results remain in line with those reported in the main text.[3]

Next, in Table A1.13, I assess the robustness of second stage estimates for the key economic (natives' employment and occupational scores; value added per establishment and establishment size) and political (property tax rates; public spending per capita; Democrats' vote share; DW Nominate scores) outcomes. Panel A reports the baseline specification, Panel B drops the city of Passaic, and Panel C omits cities in the 1^{st} and 99^{th} percentiles of the 1910 to 1930 change in immigration. As in Tables A1.11 and A1.12, reassuringly, all results remain in line with those obtained for the full sample, becoming, if anything, somewhat larger (in absolute value).[4]

A1.3.2 Full Sample and Cities with Data for the 1904 Census of Manufacture

As noted in Section 1.6.1 of the paper, when interacting year dummies with the 1904 value added by manufacture in Panel B of Table 1.6 (column 8), the coefficient for public spending becomes twice as large (in absolute value) as that from the baseline specification (column

[3]Very similar results are obtained when trimming the top and bottom percentiles of the 1900 fraction of immigrants.

[4]Results, not reported for brevity, also remain very similar to those from my baseline specification when using the Li (1985) procedure to downweight outliers.

2). Yet, as documented in Table A1.14, this is due to the slightly different sample of cities for which 1904 Census of Manufacture data are available.[5] Specifically, in Table A1.14, I estimate the baseline specification for the full sample (Panel A) and for the sample of cities for which data were reported in the 1904 Census of Manufacture. As it appears, coefficients on public spending per capita (column 6) in Panel B are twice as large as those for the full sample. Also, and importantly, the coefficient in column 6 of Panel B is quantitatively very close to that reported in column 8 of Table 1.6 (Panel B).

A1.3.3 Cross-Sectional Regressions for Legislators' Ideology

In Section 1.6.3.2 of the paper, I investigate the relationship between the 1910 to 1920 change in immigration and votes of members of the House of Representatives on the 1924 National Origins Act. Since I examine voting behavior at a specific point in time, my analysis for this outcome relies on a cross-sectional regression, implying that city and state by year fixed effects cannot be included. To indirectly assess the size and the direction of the bias that this may generate, in Table A1.15, I replicate columns 1 and 3 to 6 of Table 1.8 using cross-sectional regressions. To mirror as close as possible the specification reported in columns 7 and 8 of Table 1.8, 1920 DW Nominate scores are regressed against the 1910 to 1920 (instrumented) change in the fraction of immigrants and on state fixed effects.

Results from this exercise, reported in Panel B of Table A1.15, are similar to those from the baseline specification, which, to ease comparisons, are presented in Panel A. While only suggestive, the estimates in Table A1.15 indicate that failing to include city and state by year fixed effects, at least for this set of outcomes, does not seem to introduce substantial bias in 2SLS estimates.

A1.3.4 Pre-Migration Characteristics and Differential Trends

One important condition for the validity of the empirical strategy is that the characteristics of cities that attracted early immigrants did not have a confounding effect on migration patterns as well as on changes in the outcomes of interest. In the main text, I already addressed this concern by replicating the main economic and political results interacting year dummies with the (log of) 1900 city and immigrants population and with the (log of) 1904 value added by manufacture. I now extend this exercise to the inclusion of additional city characteristics that may have simultaneously attracted more immigrants in 1900 and affected the evolution of economic and political variables in subsequent decades.

[5] Cities for which 1904 industrial data are missing are: Pasadena (CA), Perth Amboy (NJ), Superior (WI), and Washington DC.

First, Table A1.16 interacts year dummies with the 1900 ratio of high to low skill workers (Panel B) and the fraction of black population (Panel C).[6] Reassuringly, all results remain in line with those from my baseline specification, which are reported in Panel A to ease comparisons. Next, in Table A1.17, I check whether differences in early industrialization, which may have influenced the distribution of immigrants before 1900, differentially affected changes in the outcomes of interest in subsequent decades. I do so by interacting year dummies with, respectively, the log of production per establishment (Panel A) and the employment share in manufacturing (Panel B) in 1904.

As for Table A1.16, results remain close to those estimated in the baseline specification. When comparing results in column 6 with those reported in Panel A of Table A1.16, one may be worried that the magnitude of coefficients is now somewhat larger. As shown in Table A1.14 and discussed above, though, this is due to the slightly different sample for which 1904 industrial data are available.

Finally, in Panel B of Table A1.18, I perform an even more stringent robustness check by augmenting the stacked first difference specification (equation (5) in the main text) with the interaction between year dummies and (the log of) 1900 city and immigrants population. This amounts to comparing cities that in 1900 had the same fraction of immigrants, and that experienced changes in immigration only because of variation in sending countries induced by World War I and by the Immigration Acts. Not surprisingly, the precision of the estimates deteriorates. However, their magnitude remains very close to that from the baseline stacked first difference specification (see Panel A).[7]

A1.3.5 Alternative Specifications

A1.3.5.1 Immigrants' Stock and Immigrants from Any Source

As discussed in Section 1.4.1 of the paper, in my baseline specification, I only consider European immigrants that entered the United States in the previous decade. However, one may be worried that the effect of immigrants' stock differs from that of immigrants' flows. For this reason, in Panel A of Table A1.19, I repeat the analysis considering the fraction of all foreign born individuals, and not only those arrived in the previous ten years. To instrument for immigrants' stock, I adopt a strategy very similar to that in Burchardi et al. (2016): at any point in time, the number of foreign born in a given city is predicted by

[6] Interactions between year dummies and the fraction of blacks are included to address the specific concern that during the first wave of the Great Migration (1915-1930), cities experiencing sharper drops in European immigration received larger inflows of southern born African Americans Collins (1997).

[7] As noted in the main text (see Table 1.7), the only case in which results are not robust to the use of the stacked first difference specification is for the Democrats' vote share.

interacting 1900 shares (i.e. α_{jc} in equation (2) in the main text) with both current and lagged aggregate migration flows (from each sending region). This strategy is also akin to the "double instrumentation" procedure suggested by Ruist et al. (2017) to isolate the component of immigration uncorrelated with both current demand and past supply shocks.[8] Reassuringly, results remain very similar to those obtained using only immigrants' flows. As a further check, Panel B of Table A1.19 considers (recently arrived) immigrants from all source countries, and not only from Europe. As expected, results are barely affected.

A1.3.5.2 Actual Population to Construct the Fraction of Immigrants

Since actual city population is likely to be an outcome of immigration, in my baseline specification I present results where the actual and the predicted number of immigrants are both scaled by predicted city population, constructed by multiplying 1900 population with average urban growth in the US, excluding the Census division where the city is located. In the main text, I also present results obtained scaling the (actual and predicted) number of immigrants by 1900 city population. As an additional robustness check, in Panel A of Table A1.20, I show that all results are unchanged (and, if anything, become stronger and more precise) when the endogenous regressor is scaled by actual city population, instrumented with predicted immigration constructed by scaling the predicted number of immigrants over predicted city population, i.e. Z_{cst} in (2) in the main text.

A1.3.5.3 Weather Shocks Instrument

As noted in Section 1.4.2 of the paper, one additional concern is that aggregate migration flows (by country of origin) may be endogenous to city-specific pull factors. This issue was addressed in the paper, where I showed that results are qualitatively unchanged when using the alternative instruments which exploits variation solely induced by WWI and the quotas. The only exception, discussed in the paper, is that the estimates for the effect of immigration on the Democrats' vote share are not robust to using the stacked first difference specification. As a further robustness check, in Panel B of Table A1.20, I replicate the analysis replacing actual aggregate migration flows, O_{jt}^{-M} in (2), using the measure of predicted immigration that only exploits temperature and precipitation shocks in origin countries ($\hat{O}_{jt}$ in $(A2)$ above). In all cases, results are consistent with those obtained in the baseline specification, even if they are an order of magnitude larger, especially for employment (column 1) and tax rates (column 5).

[8]As noted in Ruist et al. (2017), for this procedure to work, significant innovation in aggregate flows by country of origin is needed. As discussed in Section 1.2, World War I and the Immigration Acts provide large variation in the composition of immigrants entering the United States between 1910 and 1930.

A1.3.5.4 Presidential Elections

Finally, in Table A1.21, I replicate the electoral results presented in Table 1.7 (Section 1.6.2) and Table A1.8 above by computing vote shares and turnout by taking the average between the two closest election years rather than between the two elections after each Census year (Panel A), and excluding MSAs in the US South (Panel B). In both cases, results remain in line with those presented in the main text.

Table A1.11. First Stage Omitting Passaic (NJ)

	Dep. Variable: *Fraction of Immigrants*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Z	0.884***	0.900***	0.940***	0.948***			0.795***	0.786***
	(0.059)	(0.077)	(0.068)	(0.105)			(0.092)	(0.078)
(Z_WWI)*1920					0.730***	0.783***		
					(0.146)	(0.121)		
(Z_Quotas)*1930					0.774**	1.246***		
					(0.349)	(0.189)		
(Z_WWI)*1930					0.053			
					(0.092)			
(Z_Quotas)*1920					0.472			
					(0.427)			
1900 population		X						
Predicted population			X					
MSA analysis				X				
WWI-Quotas IV					First Diff.	Long Diff.		
Year by 1900 Log							City and imm pop	Value added manuf.
F-stat	226.4	137.0	193.2	81.88	36.31	64.47	75.38	102.0
P-value Overid. Test					0.431	0.603		
Cities	179	179	179	127	179	179	179	175
Observations	537	537	537	379	358	179	537	525

Note: this table presents the full replica of Table 2 in the main text excluding the city of Passaic (NJ). In Col 1 the actual number of immigrants is scaled by actual population, and the instrument is the leave-out version of the shift-share IV in equation (2) (Section 4.2). Cols 2 and 3 replicate Col 1 by scaling the actual and predicted number of immigrants by, respectively, 1900 and predicted population. From Col 3 onwards, Table A11 presents results from specifications where both the predicted and the actual number of immigrants are scaled by predicted population. Col 4 replicates the analysis aggregating the unit of analysis at the MSA level. Cols 5 and 6 estimate stacked first differences equation (5) and long differences equation (7) replacing the standard shift-share instrument with those constructed exploiting World War I and the quotas (equations (3) and (4) in Section 4.3). Cols 7 and 8 include the interaction between year dummies and, respectively, the (log of) 1900 city and immigrants population, and the (log of) 1904 value added by manufacture per establishment. F-stat refers to the K-P F-stat for weak instrument. Cols 5-6 report the p-value for the test of overidentifying restrictions. All regressions partial out city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.12. First Stage Trimming 1^{st} and 99^{th} Percentiles of Immigration

	Dep. Variable: *Fraction of Immigrants*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Z	0.889***	0.845***	0.897***	0.945***			0.758***	0.785***
	(0.067)	(0.097)	(0.079)	(0.105)			(0.099)	(0.078)
(Z_WWI)*1920					0.759***	0.790***		
					(0.162)	(0.163)		
(Z_Quotas)*1930					0.787**	1.209***		
					(0.347)	(0.188)		
(Z_WWI)*1930					0.031			
					(0.099)			
(Z_Quotas)*1920					0.422			
					(0.412)			
1900 population		X						
Predicted population			X					
MSA analysis				X				
WWI-Quotas IV					First Diff.	Long Diff.		
Year by 1900 Log							City and imm pop	Value added manuf.
F-stat	176.4	75.71	128.6	80.74	23.76	42.48	58.36	101.5
P-value Overid. Test					0.456	0.557		
Cities	176	176	176	127	176	176	176	173
Observations	528	528	528	379	352	176	528	519

Note: this table presents the full replica of Table 2 in the main text by dropping cities in the 1^{st} and 99^{th} percentiles of the 1910-1930 change in immigration. Cities in the top 99^{th} percentile of the change in immigration are Perth Amboy (NJ) and Passaic (NJ), while those in the bottom 1^{st} percentile are Flint (MI) and Pasadena (CA). F-stat refers to the K-P F-stat for weak instrument. Cols 5-6 report the p-value for the test of overidentifying restrictions. All regressions partial out city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.13. Main Results Omitting Potential Outliers

VARIABLES	(1) Employed	(2) Log Occ. Scores	(3) Log Value Added per Establ.	(4) Log Establ. Size	(5) Prop. Tax Rate	(6) Public Spending PC	(7) Democrats' Vote Share	(8) DW Nominate Scores
Panel A: Baseline Specification								
Fr. Immigrants	0.299***	0.097***	2.889***	2.532***	-29.44*	-8.699*	-0.404***	1.658**
	(0.064)	(0.036)	(0.954)	(0.815)	(16.95)	(4.453)	(0.141)	(0.808)
F-stat	251.3	251.3	270.5	270.5	288.3	288.3	83.14	23.11
Observations	538	538	525	525	540	540	378	460
Panel B: Drop Passaic (NJ)								
Fr. Immigrants	0.401***	0.103*	3.947***	3.339***	-41.79*	-9.035	-0.404***	2.032**
	(0.078)	(0.055)	(0.902)	(0.855)	(25.17)	(6.655)	(0.141)	(0.838)
F-stat	178.2	178.2	181.9	181.9	193.3	193.2	82.37	24.86
Observations	535	535	522	522	536	537	378	457
Panel C: Trim 1st and 99th Percentiles of Immigration								
Fr. Immigrants	0.418***	0.121*	4.421***	3.978***	-28.83	-18.25***	-0.434***	2.205*
	(0.098)	(0.069)	(0.960)	(0.953)	(30.96)	(5.557)	(0.138)	(1.298)
F-stat	123.0	123.0	121.0	121.0	128.0	128.6	81.22	42.23
Observations	526	526	513	513	527	528	375	448

Note: this table replicates the main results for the full sample (reported in Panel A) by dropping Passaic, NJ (Panel B), by excluding cities in the 1st and 99th percentiles of the 1910-1930 change in immigration (Panel C). Cities in the top 99th percentile of the change in immigration are Perth Amboy (NJ) and Passaic (NJ), while those in the bottom 1st percentile are Flint (MI) and Pasadena (CA). The dependent variable is displayed at the top of each column. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. In addition to the set of interactions described above, all regressions include city (MSA) and state by year fixed effects in Cols 1 to 6 (Col 7), and state by year and congressional district to city fixed effects in Col 8. Robust standard errors, clustered at the MSA (resp. at the congressional district) level, in Cols 1 to 7 (resp. Col 8) are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.14. Full Sample vs Cities with 1904 Industrial Data

VARIABLES	(1) Employed	(2) Log Occ. Score	(3) Log Value Added per Establ.	(4) Log Establ. Size	(5) Property Tax Rate	(6) Public Spending PC	(7) Democrats' Vote Share	(8) DW Nominate Scores
Panel A: Full Sample								
Fr. Immigrants	0.299***	0.097***	2.889***	2.532***	-29.44*	-8.699*	-0.404***	1.658**
	(0.064)	(0.036)	(0.954)	(0.815)	(16.94)	(4.453)	(0.141)	(0.808)
F-stat	251.3	251.3	270.5	270.5	292.7	288.3	83.14	23.11
Cities	180	180	178	178	180	180	126	157
Observations	538	538	525	525	539	540	378	460
Panel B: Drop Cities without 1904 Industrial Data								
Fr. Immigrants	0.287***	0.109***	2.964***	2.715***	-20.85	-13.38***	-0.415***	2.159*
	(0.074)	(0.038)	(1.038)	(0.838)	(18.55)	(3.840)	(0.140)	(1.291)
F-stat	216.6	216.6	240	240	251.3	247.9	83.75	42.17
Cities	176	176	176	176	176	176	126	154
Observations	526	526	519	519	527	528	378	451

Note: this table replicates the baseline specification comparing the full sample (Panel A) with the sample of cities for which 1904 industrial data were reported in the 1904 Census of Manufacture (Panel B). The 4 cities for which industrial data is not available are: Pasadena (CA), Perth Amboy (NJ), Superior (WI), and Washington D.C. The dependent variable is displayed at the top of each column. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. In addition to the set of interactions described above, all regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.15. Cross-Sectional Estimates of Legislators' Ideology

Dep. Variable:	*DW Nominate Scores*	*Pr. that Winner has Given Political Orientation*			
	(1)	(2)	(3)	(4)	(5)
Panel A: Baseline					
Fr. Immigrants	1.658**	-0.601	-1.655	-0.198	2.592*
	(0.808)	(0.817)	(1.039)	(1.717)	(1.354)
F-stat	23.11	23.11	23.11	23.11	23.11
Observations	460	460	460	460	460
Panel B: Cross-Sectional					
Fr. Immigrants	2.112**	0.198	-1.898	-0.359	2.974**
	(1.060)	(0.351)	(1.522)	(1.454)	(1.226)
F-stat	15.24	15.24	15.24	15.24	15.24
Observations	146	146	146	146	146
Political Orientation		Liberal Democrat	Moderate Democrat	Moderate Republican	Conservative Republican

Note: this table replicates results for the ideology of members of the House of Representatives reported in Table 8 in the main text using a cross-sectional specification. Panel A reports 2SLS estimates shown in columns 1 and 3 to 6 of Table 8. In Panel B, the dependent variable refers to 1920, and the regressor of interest (*Fr. Immigrants*) is the 1910 to 1920 change in the fraction of immigrants arrived during the previous decade, instrumented with the corresponding change in the instrument constructed in Section 4.2 of the main text (equation (2)). F-stat refers to the K-P F-stat for weak instrument. Panel A includes city by congressional district and state by year fixed effects, while Panel B controls for state fixed effects. Robust standard errors, clustered at the congressional district level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.16. Differential Trends/1

VARIABLES	(1) Employed	(2) Log Occ. Score	(3) Log Value Added per Establ.	(4) Log Establ. Size	(5) Property Tax Rate	(6) Public Spending PC	(7) Democrats' Vote Share	(8) DW Nominate Scores
Panel A: Baseline Specification								
Fr. Immigrants	0.299***	0.097***	2.889***	2.532***	-29.44*	-8.699*	-0.404***	1.658**
	(0.064)	(0.036)	(0.954)	(0.815)	(16.94)	(4.453)	(0.141)	(0.808)
F-stat	251.3	251.3	270.5	270.5	292.7	288.3	83.14	23.11
Panel B: Skill Ratios								
Fr. Immigrants	0.294***	0.101***	2.879***	2.520***	-29.52*	-8.257*	-0.393***	1.614**
	(0.065)	(0.035)	(0.948)	(0.816)	(17.00)	(4.686)	(0.142)	(0.821)
F-stat	223.8	223.8	247.1	247.1	260.3	258.8	83.60	25.33
Panel C: Fraction of Blacks								
Fr. Immigrants	0.286***	0.092**	3.009***	2.825***	-27.25	-9.968**	-0.384***	1.756**
	(0.063)	(0.036)	(1.006)	(0.920)	(16.94)	(4.480)	(0.146)	(0.794)
F-stat	235.1	235.1	249.9	249.9	273.8	269.7	76.74	24.06
Observations	538	538	525	525	539	540	378	460

Note: this table replicates the baseline specification (reported in Panel A), for each of the outcomes reported at the top of each column, including the interaction between year dummies and the ratio of high to low skilled natives (Panel B); and the fraction of blacks (Panel C) in 1900. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. In addition to the set of interactions described above, all regressions include city (MSA) and state by year fixed effects in Cols 1 to 6 (Col 7), and state by year and congressional district to city fixed effects in Col 8. Robust standard errors, clustered at the MSA (resp. at the congressional district) level, in Cols 1 to 7 (resp. Col 8) are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.17. Differential Trends/2

VARIABLES	(1) Employed	(2) Log Occ. Score	(3) Log Value Added per Establ.	(4) Log Establ. Size	(5) Property Tax Rate	(6) Public Spending PC	(7) Democrats' Vote Share	(8) DW Nominate Scores
Panel A: Value of Products								
Fr. Immigrants	0.265***	0.108***	2.642**	2.690**	-19.63	-17.19***	-0.226	2.302
	(0.075)	(0.038)	(1.189)	(1.054)	(20.15)	(4.752)	(0.179)	(1.550)
F-stat	112.9	112.9	128.6	128.6	130.7	130.6	34.89	37.73
Panel B: Employment Share Manufacture								
Fr. Immigrants	0.284***	0.127***	3.220***	2.860***	-20.47	-15.45***	-0.407***	2.310*
	(0.092)	(0.044)	(1.118)	(0.941)	(22.37)	(4.455)	(0.146)	(1.247)
F-stat	204.7	204.7	222.3	222.3	232.0	230.5	60.27	34.45
Observations	526	526	519	519	527	528	378	451

Note: this table replicates the baseline specification, for each of the outcomes reported at the top of each column, including the interaction between year dummies and the log of 1904 value of products per establishment (Panel A); and the 1904 employment share in manufacturing (Panel B). *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. In addition to the set of interactions described above, all regressions include city (MSA) and state by year fixed effects in Cols 1 to 6 (Col 7), and state by year and congressional district to city fixed effects in Col 8. Robust standard errors, clustered at the MSA (resp. at the congressional district) level, in Cols 1 to 7 (resp. Col 8) are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.18. Stacked First Difference Specification and 1900 Immigration

VARIABLES	(1) Employed	(2) Log Occ. Scores	(3) Log Value Added Establ.	(4) Log Establ. Size	(5) Prop. Tax Rate	(6) Public Spending PC	(7) Democrats' Vote Share	(8) DW Nominate Scores
			Panel A: Baseline Stacked 1^{st} Differences					
Fr. Immigrants	0.213***	0.082**	1.778***	1.983***	-24.29	-5.739*	0.048	1.939**
	(0.043)	(0.033)	(0.665)	(0.596)	(19.35)	(2.970)	(0.162)	(0.773)
F-stat	102.2	102.2	106.0	106.0	106.2	106.8	23.43	8.693
			Panel B: Stacked 1^{st} Differences AND Year by 1900 City and Immigrants Pop.					
Fr. Immigrants	0.205***	0.087*	1.515**	1.526**	-20.45	-7.790	-0.069	1.735*
	(0.055)	(0.047)	(0.724)	(0.764)	(23.84)	(5.080)	(0.197)	(1.046)
F-stat	80.24	80.24	91.18	91.18	87.26	86.99	15.28	5.945
Observations	356	356	347	347	359	360	252	303

Note: Panel A reports baseline estimates obtained from the stacked first difference specification (equation (7) in the main text), where immigration is instrumented with the interaction of year dummies and the World War I and the quota instruments constructed in Section 4.2.4 in the main text (see equations (5) and (6)). Panel B augments this specification including the interaction between year dummies and (the log of) 1900 city and immigrants population. The dependent variable is displayed at the top of each column. F-stat refers to the K-P F-stat for weak instrument. All regressions include city (MSA) and state by year fixed effects in Cols 1 to 6 (Col 7), and state by year and congressional district to city fixed effects in Col 8. Robust standard errors, clustered at the MSA (resp. at the congressional district) level, in Cols 1 to 7 (resp. Col 8) are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.19. Alternative Specifications/1

VARIABLES	(1) Employed	(2) Log Occ. Score	(3) Log Value Added per Establ.	(4) Log Establ. Size	(5) Property Tax Rate	(6) Public Spending PC	(7) Democrats' Vote Share	(8) DW Nominate Scores
Panel A: Immigrants' Stock								
Fr. Immigrants	0.335***	0.109***	3.323***	2.911***	-33.29*	-9.843*	-0.489***	2.089**
	(0.071)	(0.041)	(1.047)	(0.884)	(18.42)	(5.156)	(0.181)	(1.001)
F-stat	81.20	81.20	76.66	76.66	80.90	81.14	19.49	38.10
Panel B: Immigrants from All Sources								
Fr. Immigrants	0.189***	0.095***	1.945***	1.605**	-17.78	-7.107**	-0.256***	0.957**
	(0.059)	(0.027)	(0.627)	(0.634)	(11.67)	(2.819)	(0.075)	(0.465)
F-stat	85.19	85.19	89.38	89.38	86.32	86.59	32.02	16.71
Observations	538	538	525	525	539	540	378	460

Note: Panel A replicates the baseline specification measuring immigration (*Fr. Immigrants*) as the fraction of all foreign born individuals over predicted city population, instrumenting it with both current and lagged migration flows interacted with the share of immigrants (from each country of origin) living in the city in 1900. Panel B replicates the baseline specification considering immigrants from all sending countries: *Fr. Immigrants* refers to the fraction of immigrants arrived during the previous decade over predicted city population, and is instrumented with the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). The dependent variable is reported at the top of each column. F-stat refers to the K-P F-stat for weak instrument. Cols 1 to 6 (Col 7) include city (MSA) and state by year fixed effects, while Col 8 controls for congressional district and state by year fixed effects. Robust standard errors, clustered at the MSA (congressional district) level in Cols 1 to 7 (Col 8), in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.20. Alternative Specifications/2

VARIABLES	(1) Employed	(2) Log Occ. Score	(3) Log Value Added per Establ.	(4) Log Establ. Size	(5) Property Tax Rate	(6) Public Spending PC	(7) Democrats' Vote Share	(8) DW Nominate Scores
Panel A: Immigrants Over Actual Pop.								
Fr. Immigrants	0.357***	0.116***	3.456***	3.029***	-35.06*	-10.34**	-0.408***	1.916**
	(0.056)	(0.040)	(0.926)	(0.810)	(18.96)	(4.870)	(0.005)	(0.885)
F-stat	249.3	249.3	241.8	241.8	223.5	225.1	145.3	168.6
Panel B: Weather Shocks Instrument								
Fr. Immigrants	0.480***	0.141**	2.171**	1.924*	-86.07*	-15.88**	-0.387*	2.205**
	(0.113)	(0.060)	(0.969)	(0.983)	(48.12)	(7.848)	(0.230)	(1.042)
F-stat	15.68	15.68	18.28	18.28	18.72	18.70	28.17	14.39
Observations	538	538	525	525	539	540	378	460

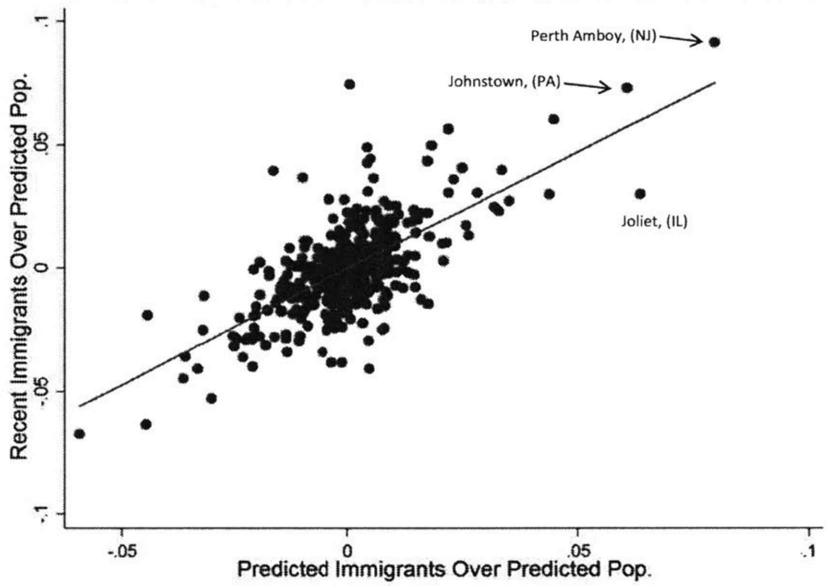
Note: Panel A replicates the baseline specification measuring immigration (*Fr. Immigrants*) as the fraction of immigrants over actual (rather than predicted) city population, instrumented with the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). Panel B estimates the baseline specification, constructing the instrument with aggregate migration flows predicted with temperature and precipitation shocks in the country of origin (see (A2) in the online appendix). The dependent variable is reported at the top of each column. F-stat refers to the K-P F-stat for weak instrument. Cols 1 to 6 (Col 7) include city (MSA) and state by year fixed effects, while Col 8 controls for congressional district and state by year fixed effects. Robust standard errors, clustered at the MSA (congressional district) level in Cols 1 to 7 (Col 8), in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.21. Presidential Elections: Alternative Definitions and Samples

VARIABLES	(1) Democrats' vote share	(2) Republicans' vote share	(3) Other parties' vote share	(4) Democrats-Republicans Margin	(5) Turnout
Panel A: Average Outcomes Between 2 Closest Elections					
Fr. Immigrants	-0.743***	0.431***	0.312***	-1.174***	-1.588***
	(0.154)	(0.145)	(0.119)	(0.275)	(0.157)
F-stat	83.14	83.14	83.14	83.14	82.20
Mean dep var	0.455	0.341	0.204	0.114	0.525
MSAs	126	126	126	126	125
Observations	378	378	378	378	373
Panel B: Exclude Southern MSAs					
Fr. Immigrants	-0.396**	0.197	0.199*	-0.593**	-1.532***
	(0.154)	(0.158)	(0.106)	(0.293)	(0.180)
F-stat	71.55	71.55	71.55	71.55	71.55
Mean dep var	0.423	0.351	0.225	0.073	0.570
MSAs	94	94	94	94	94
Observations	282	282	282	282	282

Note: this table replicates results reported in Table 7 and in Table A8 computing vote shares and turnout by taking the average between the two closest election years rather than between the two elections after each Census year (Panel A), and excluding southern MSAs (Panel B). The dependent variable is reported at the top of each column, and refers to Presidential elections. All electoral outcomes were aggregated from the county to the MSA level, using 1940 MSAs' definitions. *Other parties' vote share refers* to the vote share of all parties other than Democrats and Republicans. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include MSA and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Figure A1.19. First Stage Omitting Passaic (NJ)



Note: this figure replicates Figure 7 in the main text omitting the city of Passaic (NJ). The y-axis (resp. x-axis) reports the actual (resp. predicted) number of immigrants over predicted city population in each of the three Census years, 1910, 1920, and 1930. Each point in the scatter diagram represents the residual change in a city's actual and predicted fraction of immigrants after partialling out city and year by state fixed effects. The predicted number of immigrants is constructed as discussed in Section 4.2 in the text (see (2)). Predicted city population is obtained by multiplying 1900 city population with average urban growth, excluding that of the Census division where a city is located. The solid line shows the regression coefficient (coefficient=0.940, standard error=0.068).

A1.4 Additional Results

A1.4.1 Immigration and Electricity in Manufacturing

As discussed in the main text, one possible explanation for the positive effect of immigration on firms' productivity is that the inflow of immigrants encouraged the adoption of new technologies that made intensive use of electricity, e.g. the assembly line, in turn raising the demand for managers and supervisors, and for high skilled workers such as electricians (Goldin and Katz, 2009; Katz and Margo, 2014). Lack of systematic data on electricity use at the city level before 1940 prevents me from investigating this idea directly. However, I digitized data on the share of horsepower coming from electricity reported in the 1929 Census of Manufactures for selected US counties. Aggregating the data to the MSA level, and running cross-sectional regressions, I find that MSAs that received more immigrants between 1910 and 1930 had a larger share of power coming from electricity in 1930 (Table A1.22). Because of the cross-sectional nature of this exercise, the evidence in Table A1.22 should be interpreted as only suggestive.[9] Nevertheless, it is consistent with the idea that immigration may have induced a faster adoption of electricity and of related technologies.

A1.4.2 Immigration and Internal Migration

If immigrants increased labor market opportunities for natives and made cities economically more attractive, immigration may have also encouraged internal in-migration. Since prior to 1940 statistics on internal migration in the US do not exist, I proxy for the number of internal movers by looking at the fraction of males in working age that were born outside the state of their city of residence (see also Bandiera et al., 2017). As I show in Table A1.23, immigration had a positive and significant effect on the fraction of internal movers (column 1). Reassuringly, the 1900 to 1910 change in the fraction of natives born in another state does not predict the (instrumented) change in immigration in subsequent decades (column 2).

Somewhat surprisingly, the positive effect of immigration on internal migrants is not driven by growing cities (column 3), but, instead, by cities whose 1910-1930 population growth rate was below the median (column 4). This result suggests that the inflow of immigrants may have acted as a lift for otherwise declining cities. These findings are also consistent with the idea that at least part of the employment responses to immigration are due to the behavior of "outsiders", i.e. natives not originally living in the local labor markets

[9] All regressions control for state fixed effects. Columns 2 and 4 also include a number of 1900 controls such as city and immigrants population, skill ratios, and measures of industrial production.

exposed to the immigration shock (see also Dustmann et al., 2017).

A1.4.3 Natives-Immigrants Complementarity

In Section 1.5.1 of the paper, I show that immigration boosted natives' employment and induced natives to move away from more exposed and less skill-intensive occupations and towards jobs that had higher skill intensity and where immigrants' competition was lower. One possibility is that this happened because of gains from specialization and complementarity between natives and immigrants. A direct implication of this idea is that natives' occupational upgrading should be stronger when immigrants are less skilled and have lower English proficiency. To investigate this conjecture, I classify immigrants as coming from linguistically close and far countries using the measure constructed by Chiswick and Miller (2005), which is based on the difficulty that Americans have in learning other languages. I define a country as linguistically far (resp. close) if its linguistic distance from English is above (resp. below) the median distance.

Relying on this admittedly crude measure of linguistic distance, in Table A1.24, I re-estimate equation (1) in the main text allowing immigrants from linguistically far and close countries to have differential effects on natives' employment and occupational standing. To ease comparisons across groups, I standardize both regressors by subtracting their means and dividing them by their standard deviations. In this way, coefficients in Table A1.24 can be interpreted as the effect of a one standard deviation change in the fraction of linguistically far and close immigrants respectively.[10]

Results, reported in Panel A, are consistent with the mechanism discussed above and in the main text: while the employment effects of immigration are positive regardless of "immigrants' type", they are significant only when immigrants came from linguistically far countries (columns 1 and 2). More importantly, there is evidence of natives' occupational upgrading (columns 3 and 4) only when immigrants were linguistically farther from natives. Similar results are obtained when splitting immigrants between new (Russia, Eastern, and Southern Europe) and old (British Isles, Western Europe, Scandinavia, and Germany) source countries (see Panel B of Table A1.24), exploiting the fact that immigrants from new sending regions were more likely to be illiterate and unskilled (see Figures A1.1 and A1.2 above).[11]

[10]Figure A1.20 plots the implied percent change in employment and occupational scores due to a one standard deviation change in immigrants from linguistically far (orange bar) and linguistically close (blue bar) countries.

[11]See also Biavaschi et al. (2017) and Greenwood and Ward (2015).

A1.4.4 Effects of Immigration on Employment of Previously Arrived Immigrants

In this section, I perform a placebo exercise to check that immigration did not increase employment of previously arrived immigrants. Because the dependent variable in this analysis refers to the foreign born, it seems particularly important to directly control for the size of immigrant groups in 1900. Thus, I re-estimate a version of (1) in the paper where, in addition to instrumenting the fraction of immigrants with the standard instrument from equation (2) in the paper and to partialling out city and state by year fixed effects, I include the interaction between year dummies and the 1900 fraction of immigrants.[12] Results are reported in Figure A1.21, where I plot the coefficient (and corresponding 95% confidence intervals) for the effect of immigration on employment of different groups of immigrants.

Starting from the left, Figure A1.21 first presents the coefficient for the effect of immigration on natives' employment (which is positive and statistically significant). Next, moving from the left to the right of the Figure on the x-axis, one can see that there is no positive effect of immigration on employment of immigrants from either Southern and Eastern (second group from the left) or Northern and Western (third group from the left) Europe. In fact, the coefficient is negative, although not statistically significant. Consistent with the idea that European immigrants gradually assimilated over time (Abramitzky et al. (2016); Abramitzky and Boustan (2016)), the point estimate is negative for relatively recent immigrants (i.e. those arrived between 10 and 20 years before), while it is zero for foreign born that had spent at least 20 years in the United States.

A1.4.5 Immigration and Natives' Rents

To directly assess the causal effect of immigration on rents, ideally, one would want to exploit data that vary both over time and across neighborhoods within the same city. Unfortunately, such data are not consistently available for the historical period studied in my paper.

Instead, to indirectly investigate the possibility that higher rents fueled natives' discontent, in Figure A1.22, I plot the relationship between the 1910 to 1930 instrumented change in immigration (x-axis) and 1930 natives' average rents (y-axis), after partialling out state fixed effects.[13] Because of the cross-sectional nature of this regression, results in Figure A1.22 should be interpreted with some caution, but they suggest that immigration was not correlated with rents paid by natives. This, in turn, weighs against the possibility that natives'

[12] Equivalent results are obtained when interacting year dummies with the 1900 log of city and of immigrants population.

[13] 1930 is the first year in which data on rents were collected by the US Census.

backlash was triggered (mainly) by higher rents. As discussed in the main text of the paper, one potential explanation for why, despite its positive effect on productivity, immigration did not increase rents is that immigrants represented a *production amenity*, but were perceived as a *consumption disamenity*, as documented for the contemporaneous period by a number of papers in both Europe and the US (e.g. Card et al., 2012; Saiz and Wachter, 2011).

A1.4.6 A Horse-Race Between Religion and Linguistic Distance

Tables 1.9 and 1.10 in the paper document that political discontent: *i*) took place only when immigrants came from non-Protestant countries; and *ii*) was increasing in the linguistic distance between immigrants and natives. In this section I investigate if either of the two measures of cultural diversity dominates over the other by replicating the analysis in Table 1.9 including simultaneously religion and linguistic distance. To ease the interpretation of results, the fraction of immigrants from Protestant and non-Protestant countries are both standardized by subtracting their mean and dividing through their standard deviation. Thus, as for the index of linguistic distance, the coefficient on immigration from each religious group should be interpreted as the effect of a one standard deviation increase in the fraction of immigrants from Protestant and non-Protestant countries.

2SLS results are reported in Table A1.25. As it appears, when focusing on taxes and spending (columns 1 to 4), only the index of linguistic distance is statistically significant. Instead, even if the coefficient on immigration from non-Protestant countries is negative, it is quantitatively small and imprecisely estimated. However, for electoral outcomes (columns 5 to 8), the opposite pattern emerges: only immigration from non-Protestant countries is associated with a significant reduction (resp. increase) in support for Democrats (resp. in legislators' ideology and support for the 1924 Immigration Act).[14] One possible (tentative) interpretation is that the salience of different cultural attributes might differ across political issues.

A1.4.7 Ethnic Diversity

A large literature has shown that ethnic diversity is associated with lower public goods provision and with more limited redistribution (e.g. Alesina et al., 1999; Beach and Jones, 2017; Luttmer, 2001). The argument advanced in these works is that both altruism and the utility from public goods' consumption are lower when they involve inter-ethnic interactions. It follows that, if immigration reduced natives' demand for public goods by increasing ethnic

[14] Somewhat surprisingly, the index of linguistic distance seems to have a positive effect on the support for Democrats.

diversity, this effect should be stronger when the ethnic composition of foreign born was more heterogeneous. Also, a more diverse foreign born population may reduce immigrants' ability to act as a unified political group, in turn reinforcing the effectiveness of natives' actions.[15] To test these conjectures, I interact immigration, Imm_{cst}, with an index of ethnic diversity (Alesina et al., 1999) of the foreign born population, $ED_{cst} = 1 - \sum_j \left(sh^j_{cst}\right)^2$, where sh^j_{cst} is the share of ethnic group j among the foreign born population introduced in the previous section. I then estimate

$$y_{cst} = \gamma_c + \delta_{st} + \beta_1 Imm_{cst} + \beta_2 Imm_{cst} * ED_{cst} + \beta_3 ED_{cst} + u_{cst} \quad \text{(A3)}$$

As in the main text (Section 1.7.2), to ease the interpretation of results, which are reported in Table A1.26, I standardize ED_{cst} by subtracting its mean and dividing it by its standard deviation. The coefficient on the interaction between immigration and ethnic diversity, β_2, can thus be interpreted as the additional effect of immigration for a city with ethnic diversity one standard deviation above the sample mean. When estimating $(A3)$, the interaction term, $Imm_{cst} * ED_{cst}$, is instrumented with the interaction between ED_{cst} and predicted immigration, i.e. Z_{cst} in equation (2) in the main text.

The (negative) effect of immigration on tax revenues per capita is larger when ethnic diversity among foreign born is higher (columns 1 and 2). Somewhat surprisingly, though, when looking at tax rates (column 3), the coefficient on the interaction between immigration and ethnic diversity is not statistically significant, even if it is negative. Next, in line with columns 1 and 2, column 4 shows that the effects of immigration on public spending are larger (i.e. more negative) when ethnic diversity is higher. This result is consistent with the existing literature (e.g. Alesina et al., 1999), and corroborates the interpretation advanced in Section 6 that immigrants lowered natives' utility from consumption of public goods.

[15] An alternative view is discussed in Borjas (2016), who suggests that higher diversity could make immigration less salient, in turn reducing natives' backlash.

Table A1.22. Share of Electric Power in Manufacture (1930)

	Dep. Variable: *Share of Horsepower from Purchased Electricity*			
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
Fr. Immigrants	2.449***	1.799**	2.520***	1.867**
	(0.557)	(0.774)	(0.522)	(0.744)
F-stat			61.14	27.23
Mean Dep. Var.	0.617	0.617	0.617	0.617
Additional Controls		X		X
MSAs	101	101	101	101

Note: the sample is restricted to the 101 MSAs spanning counties for which data on purchased electricity used in production was reported in the 1929 Census of Manufacture, and that include at least one of the 180 cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. MSA boundaries are fixed to 1940. Cols 1 and 2 (resp. 3 and 4) present OLS (resp. 2SLS) results. The dependent variable is the share of horsepower coming from purchased electricity in 1930. *Fr. Immigrants* is the 1910 to 1930 change in the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include state fixed effects. Cols 2 and 4 also control for the fraction of immigrants and the fraction of blacks in 1900, and the log of value added per establishment in 1904. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.23. Immigration and Internal Migration

	Dep. Variable: *Fr. Natives Born Outside the State*			
	(1)	(2)	(3)	(4)
Panel A: OLS				
Fr. Immigrants	0.290***	0.090	0.244	0.307***
	(0.097)	(0.126)	(0.224)	(0.099)
Panel B: 2SLS				
Fr. Immigrants	0.296***	0.044	-0.169	0.377***
	(0.096)	(0.115)	(0.190)	(0.113)
F-stat	288.3	313.0	116.0	144.2
Mean dep var	0.350	0.350	0.391	0.264
Cities	180	180	90	90
Observations	540	180	270	270
Sample	Full	Full	High growth	Low growth
Pre-period		X		

Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is the fraction of native males in working age that were born outside the state of their city of residence. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). Col 2 reports results for a regression of the 1900-1910 change in the dependent variable against the 1910 to 1930 change in the fraction of immigrants. Col 3 (resp. 4) restricts the sample to the 90 cities with population growth between 1910 and 1930 above (resp. below) median. F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. The mean of each dependent variable at baseline is shown at the bottom of the Table. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.24. Immigrants' Characteristics and Natives' Occupational Upgrading

Dependent Variable:	Natives' Employment		Natives' Log Occupational Scores	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Panel A: Linguistically Far vs Close				
Fr. Imm. Far	0.012***	0.011***	0.003*	0.005***
	(0.002)	(0.003)	(0.002)	(0.002)
Fr. Imm. Close	0.003	0.006	-0.008***	-0.007*
	(0.003)	(0.004)	(0.003)	(0.004)
KP F-stat		22.20		22.20
F-stat (Far)		86.31		86.31
F-stat (Close)		27.11		27.11
Panel B: New vs Old Sending Regions				
Fr. Imm. New	0.011***	0.011***	0.003**	0.006***
	(0.002)	(0.003)	(0.002)	(0.002)
Fr. Imm. Old	0.003	0.007	-0.008**	-0.008*
	(0.003)	(0.004)	(0.003)	(0.004)
KP F-stat		20.91		20.91
F-stat (New)		88.52		88.52
F-stat (Old)		29.44		29.44
Observations	538	538	538	538

Note: this table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. The sample is restricted to native men in the age range 15 to 65 who are not enrolled in schools. Panels A and B classify immigrants as coming from linguistically far vs close and new vs old sending countries. The dependent variable is displayed at the top of each column. *Fr. Imm. Far* (resp. *Close*) refers to the fraction of immigrants arrived in the previous decade that come from linguistically far (resp. close) countries, over predicted city population. *Fr. Imm_New* (resp. *Close*) refers to the fraction of immigrants arrived in the previous decade that come from new (resp. old) source countries, over predicted city population. Each endogenous regressor is instrumented with the predicted fraction of immigrants (see (2) in Section 4.2), obtained by summing (predicted) immigration across linguistically far and close (Panel A) and new and old (Panel B) sending countries. F-stat (Far), F-stat (Close), F-stat (New), and F-stat (Old) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.25. A Horse-Race Between Religion and Linguistic Distance (2SLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Total tax revenues PC	Property tax revenues PC	Property tax rate	Public spending PC	Dem-Rep. margin	Smith's pct. votes	DW Nominate Scores	1[Restrict Immigration]
Ling. Distance	-0.860*	-0.802*	-2.309	-0.516*	0.036**	0.065*	0.028	-0.033
	(0.474)	(0.463)	(1.593)	(0.303)	(0.015)	(0.033)	(0.030)	(0.061)
Fr. Non-Prot	-0.108	-0.104	-0.122	-0.122	-0.041***	-0.148***	0.049	0.157***
	(0.417)	(0.382)	(0.879)	(0.257)	(0.013)	(0.031)	(0.037)	(0.060)
Fr. Prot	0.213	0.086	-0.051	0.005	-0.009	0.072	0.006	-0.071
	(0.411)	(0.375)	(1.029)	(0.267)	(0.013)	(0.044)	(0.030)	(0.081)
KP F-stat	14.72	14.72	15.01	14.72	8.111	5.841	13.57	9.314
F-stat (Ling.)	37.91	37.91	40.05	37.91	30.54	30.33	25.25	19.45
F-stat (Non-Prot)	65.52	65.52	66.81	65.52	39.39	29.81	57.43	46.80
F-stat (Prot)	20.91	20.91	20.66	20.91	27.37	23.31	23.01	16.79
Mean of dep var	12.76	12.10	19.75	12.16	0.180	0.398	0.165	0.676
Observations	540	540	539	540	378	126	460	155

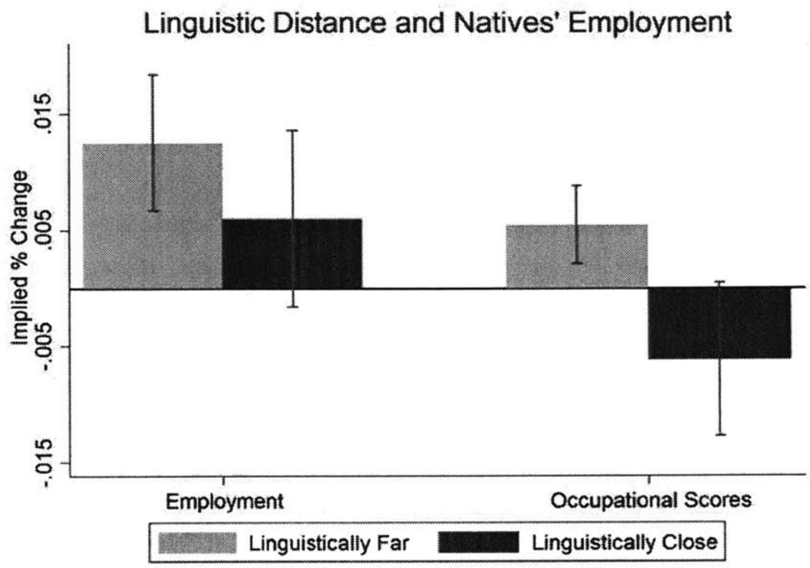
Note: this table replicates results reported in Table 9 in the main text, including simultaneously immigration from different religious groups and the index of linguistic distance introduced in Section 7.2 in the paper. The dependent variable is displayed at the top of each column. To ease the interpretation of results, the fraction of immigrants from Protestant (*Fr. Prot*) and from non-Protestant (*Fr. Non-Prot*) countries are both standardized by subtracting their mean and dividing them by their standard deviation. All regressors are instrumented using the instruments constructed in the main text. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. F-stat (Ling.), F-stat (Non-Prot), and F-stat(Prot) refer to the partial F-stats for joint significance of the instruments in the three separate first-stage regressions. Cols 1 to 4 (resp. 5) include city (resp. MSA) and state by year fixed effects, while Col 7 includes congressional district by city and state by year fixed effects. Cols 6 and 8 present results from a cross-sectional regression and control for state dummies. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A1.26. Immigration and Ethnic Diversity

Dep. Var.	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Public spending PC	(5) Education	(6) Police	(7) Charities and Hospitals	(8) Sanitation
Panel A: OLS								
Fr. Immigrants	-7.092	-6.817	-28.35***	-4.803	-7.178***	0.263	0.828	-0.433
	(6.030)	(5.055)	(10.82)	(3.705)	(2.248)	(0.586)	(0.701)	(0.667)
(Fr.Imm.)*ED	-9.749**	-9.390*	0.626	-6.107**	-2.882**	-0.760*	-0.480	-1.614**
	(4.647)	(4.749)	(7.772)	(2.969)	(1.253)	(0.423)	(0.740)	(0.672)
Panel B: 2SLS								
Fr. Immigrants	-9.885	-10.133*	-30.31*	-7.564*	-5.933***	-0.305	-0.759	-1.211*
	(6.477)	(5.934)	(17.709)	(4.125)	(2.097)	(0.680)	(1.703)	(0.716)
(Fr. Imm.)*ED	-15.43***	-15.28***	-13.71	-10.69***	-1.903	-0.223	-0.800	-0.897
	(4.587)	(4.458)	(11.26)	(3.665)	(1.414)	(0.648)	(0.802)	(0.562)
KP F-stat	21.39	21.39	21.37	21.39	20.80	21.39	15.80	21.39
F-stat (Imm.)	146.4	146.4	148.4	146.4	130.3	146.4	114.4	146.4
F-stat (Imm_ED)	18.31	18.31	18.30	18.31	30.06	18.31	16.00	18.31
Mean of dep var	12.76	12.10	19.75	12.16	4.250	1.338	0.635	1.129
Observations	540	540	539	540	534	540	516	540

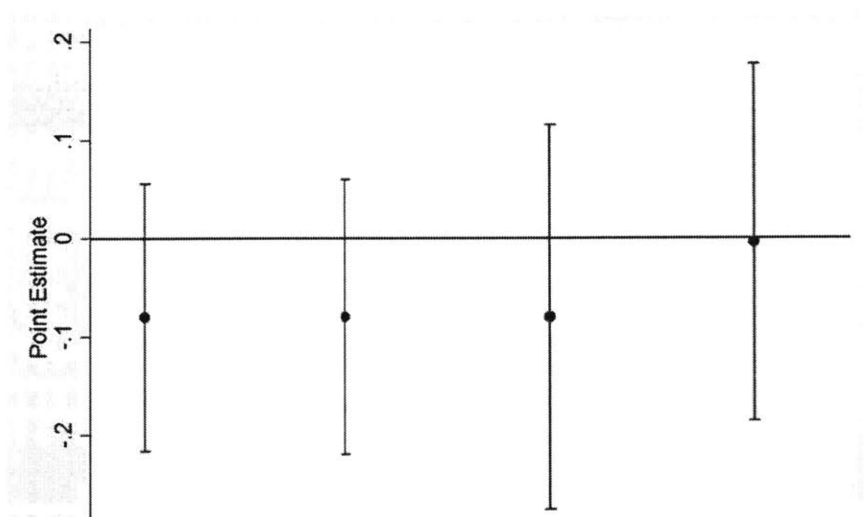
Note: this Table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is displayed at the top of each column. In Cols 5 to 8, the dependent variable is spending per capita on the category listed at the top of the column. *Fr. Immigrants* refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). *(Fr. Imm.)*ED* is the interaction between the fraction of immigrants and the (standardized) index of ethnic diversity of the foreign born population constructed in online appendix A Section A4.7. It is instrumented with the interaction between predicted immigration and the index of ethnic diversity. F-stat (Imm.) and F-stat (Imm_ED) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include the main effect of the index of ethnic diversity, and control for city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Figure A1.20. Linguistic Distance and Natives' Labor Market Outcomes



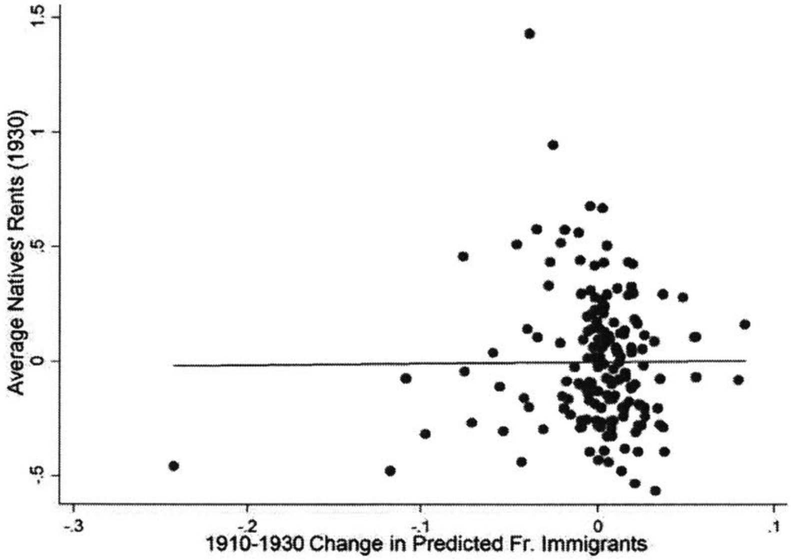
Note: this figure plots the 2SLS estimates for the percent change in employment and occupational scores for native men of working age implied by one standard deviation change (0.05) in the fraction of immigrants with associated 95% confidence intervals, for immigrants coming from linguistically far and linguistically close countries. Countries are classified as linguistically far (resp. close) if they are above (resp. below) the median linguistic distance from English as computed by Chiswick and Miller (2004). See online appendix A for a detailed description of the measure of linguistic distance from English and for the Table displaying both OLS and 2SLS results (Table B13).

Figure A1.21. Effects of Immigration on Previously Arrived Immigrants



Note: the figure plots the coefficient (with corresponding 95% confidence intervals) from a regression of immigration on employment of different groups of foreign born men of working age (15-65). The fraction of immigrants is instrumented with the instrument constructed in equation (2) of Section 4.2. All regressions control for city and state by year fixed effects, and include interactions between the 1900 fraction of immigrants and year dummies. *Immigrants S/E* (resp. *N/W*) refers to immigrants from Eastern and Southern (resp. Northern and Western) Europe. *Immigrants 10-20Y* (resp. *+20Y*) refers to immigrants that spent between 10 and 20 (resp. more than 20) years in the United States.

Figure A1.22. 1910-1930 Immigration and 1930 Natives' Rents



Note: this figure plots the relationship between the log of 1930 average rents paid by natives (y-axis) and the 1910 to 1930 predicted change in immigration (x-axis) after partialling out state fixed effects.

References

Abramitzky, R. and L. P. Boustan (2016). Immigration in american economic history. Technical report, Journal of Economic Literature.

Abramitzky, R., L. P. Boustan, and K. Eriksson (2016, July). Cultural assimilation during the age of mass migration. Working Paper 22381, National Bureau of Economic Research.

Alesina, A., R. Baqir, and W. Easterly (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics 114*(4), 1243–1284.

Bandiera, O., M. Mohnen, I. Rasul, and M. Viarengo (2017). Nation-building through compulsory schooling during the age of mass migration. Working Paper.

Beach, B. and D. B. Jones (2017). Gridlock: Ethnic diversity in government and the provision of public goods. *American Economic Journal: Economic Policy 9*(1), 112–36.

Biavaschi, C., C. Giulietti, and Z. Siddique (2017). The economic payoff of name americanization. *Journal of Labor Economics Forthcoming.*

Borjas, G. J. (2016). *We Wanted Workers: Unraveling the Immigration Narrative.* WW Norton & Company.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2016). Migrants, ancestors, and investments. Technical report, National Bureau of Economic Research.

Card, D., C. Dustmann, and I. Preston (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association 10*(1), 78–119.

Chiswick, B. R. and P. W. Miller (2005). Linguistic distance: A quantitative measure of the distance between english and other languages. *Journal of Multilingual and Multicultural Development 26*(1), 1–11.

Collins, W. J. (1997). When the tide turned: Immigration and the delay of the great black migration. *The Journal of Economic History 57*(03), 607–632.

Dustmann, C., U. Schonberg, and J. Stuhler (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics 132*(1), 435.

Feng, S., A. B. Krueger, and M. Oppenheimer (2010). Linkages among climate change, crop yields and mexico–us cross-border migration. *Proceedings of the National Academy of Sciences 107*(32), 14257–14262.

Goldin, C. D. and L. F. Katz (2009). *The race between education and technology.* Harvard University Press.

Greenwood, M. J. and Z. Ward (2015). Immigration quotas, world war i, and emigrant flows from the united states in the early 20th century. *Explorations in Economic History 55*, 76–96.

Katz, L. F. and R. A. Margo (2014). Technical change and the relative demand for skilled labor: The united states in historical perspective. In *Human capital in history: The American record*, pp. 15–57. University of Chicago Press.

Li, G. (1985). Robust regression. *Exploring data tables, trends, and shapes 281*, U340.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner (2004). European seasonal and annual temperature variability, trends, and extremes since 1500. *Science 303*(5663), 1499–1503.

Luttmer, E. F. (2001). Group loyalty and the taste for redistribution. *Journal of political Economy 109*(3), 500–528.

Pauling, A., J. Luterbacher, C. Casty, and H. Wanner (2006). Five hundred years of gridded high-resolution precipitation reconstructions over europe and the connection to large-scale circulation. *Climate Dynamics 26*(4), 387–405.

Ruist, J., J. Stuhler, and D. A. Jaeger (2017). Shift-share instruments and the impact of immigration.

Saiz, A. and S. Wachter (2011). Immigration and the neighborhood. *American Economic Journal: Economic Policy 3*(2), 169–188.

Sequeira, S., N. Nunn, and N. Qian (2017). Migrants and the making of america: The short- and long-run effects of immigration during the age of mass migration. Technical report, National Bureau of Economic Research.

Solomou, S. and W. Wu (1999). Weather effects on european agricultural output, 1850–1913. *European Review of Economic History 3*(3), 351–373.

B1. Theoretical Framework for Chapter 1

B.1 Overview

In what follows, I present a simple model to explain the three key findings of Section 1.5 in the paper, namely that immigration

1. Increases natives' employment, without generating negative effects even for workers in highly exposed occupations

2. Boosts economic activity, capital utilization, and productivity

3. Increases (reduces) the fraction of natives employed in high (low) occupations, and promotes natives' occupational upgrading

I build on a model of biased technical change (Acemoglu, 2002), where a final good is produced combining two intermediate inputs. One of the two intermediate inputs is produced using only non-production (proxy for high skilled) workers, while the other uses both laborers (proxy for low skilled workers) and capital.[1] Capital is, in turn, endogenously supplied by a continuum of manufacturing establishments, each producing a different variety. In this standard set-up, I formally show under what conditions an immigration shock in the unskilled sector can benefit high skilled natives without harming workers in the more exposed sector. As in the more general model of Acemoglu (2002), the key intuition is that, by increasing the supply of unskilled labor, immigration can induce an endogenous response from the production side (i.e., the entry of new plants), which can partly (or even completely) accommodate the inflow of immigrants.

Next, I present two extensions of the model. First, I assume that immigrants and native laborers are imperfect substitutes, and show that the degree of capital adjustment needed to absorb the immigration shock is lower than in the baseline version of the model. This is intuitive: on the one hand, the negative (competition) effect induced by immigration is lower, since immigrants are only imperfect substitutes for unskilled natives; on the other, the complementarity between the skills of natives and those of immigrants makes firms' investment even more profitable than before. Second, I endogeneize natives' sectoral choice, assuming that natives can work in both the skilled and the unskilled sector, while immigrants are barred from non-production occupations. Following the inflow of immigrants, natives reallocate their labor away from the unskilled (and more exposed) sector and towards more skilled occupations. In this case, immigration is absorbed by two distinct channels: first,

[1]See Goldin and Katz (2009) for the relationship between production and non-production workers and education or skills in the early twentieth century.

through an increase in firms' investment, as before; second, via occupational mobility of natives who tend to take up jobs where they have a comparative advantage relative to immigrants.[2]

B.2 Set-Up

B.2.1 Demand Side

I consider a general equilibrium model with two types of workers, skilled and unskilled, who have the same utility function over consumption of the final good

$$U\left(C\left(t\right)\right) = \int_0^{\infty} \exp\left(-\rho t\right) \frac{C^{1-\theta}\left(t\right)}{1-\theta} dt$$

where ρ is the discount rate and θ is the intertemporal elasticity of substitution (or, equivalently, the coefficient of relative risk aversion). To ease notation, whenever possible, I drop the time index. The budget constraint is given by

$$C + I + Z \leq Y$$

where I and Z denote respectively investment and expenditures to enter the manufacturing sector and produce capital supplies (introduced below).[3]

B.2.2 Supply Side

The final good (Y) is produced combining two intermediate inputs, Y_H and Y_L, according to a CES production function

$$Y = [Y_H^{\gamma} + Y_L^{\gamma}]^{\frac{1}{\gamma}} \tag{B1}$$

where $\gamma \leq 1$ governs the elasticity of substitution between the two intermediate goods.[4] The price of the final good is normalized to 1, and both Y_H and Y_L are produced by a large number of perfectly competitive firms. Since I am interested in evaluating the effects of a change in the supply of unskilled labor (induced by an immigration shock), to simplify the analysis, I assume that Y_H is produced using only high skilled workers, while both unskilled

[2] Peri and Sparber (2009) is the first paper that formally shows empirically and theoretically this mechanism. However, the forces highlighted in my model are rather different from those originally proposed in Peri and Sparber (2009).

[3] I assume that the standard no Ponzi condition holds, so that the lifetime budget constraint is satisfied.

[4] The elasticity of substitution betwen Y_H and Y_L is given by $\varepsilon = \frac{1}{1-\gamma}$. When $\gamma = 1$, i.e. $\varepsilon \to \infty$, the two intermediate goods are perfect substitutes; when $\gamma \to 0$, i.e. $\varepsilon \to 1$, Y is produced according to a Cobb-Douglas; when $\gamma \to -\infty$, i.e. $\varepsilon \to 0$, Y_H and Y_L are perfect complements.

labor and capital are used in the production of Y_L:[5]

$$Y_H = H$$

and

$$Y_L = KL^{\beta} \tag{B2}$$

Capital is, in turn, the aggregate of inputs (that I refer to as machines) supplied by a continuum of manufacturing plants, each producing a different variety, $k_L(v)$

$$K = \frac{1}{1-\beta}\int_0^{N_L} k_L^{1-\beta}(v)\,dv$$

where N_L is the number of manufacturing plants (and thus of varieties).

B.2.3 Production of Machines

As in Acemoglu (2002), machines are assumed to fully depreciate after use, and are supplied by monopolists at price $p_L^k(v)$ for all $v \in [0, N_L]$. Once a specific machine is invented, the monopolist has full property rights over that variety, and can produce it at marginal cost $\lambda \equiv 1-\beta$. Finally, I assume that one unit of the final good used in the development of machines directed towards Y_L generates η_L new varieties of L-complementary machines. That is,

$$\frac{dN_L(t)}{dt} = \eta_L Z(t) \tag{B3}$$

B.3 Equilibrium

An equilibrium is defined as a set of prices of machines, p_L^k, that maximizes monopolists' profits, demand for machines, x_L, that maximizes profits of producers of intermediate good Y_L, factor and product prices, w_L, w_H, p_L, and p_H, such that markets clear, and number of machine varieties, N_L, that satisfies the free entry condition.

First, because of perfect competition, prices of Y_H and Y_L, p_H and p_L, are equal to their marginal products:

$$p_H = Y_H^{\gamma-1}\left[Y_H^{\gamma} + Y_L^{\gamma}\right]^{\frac{1}{\gamma}-1} \tag{B4}$$

and

$$p_L = Y_L^{\gamma-1}\left[Y_H^{\gamma} + Y_L^{\gamma}\right]^{\frac{1}{\gamma}-1} \tag{B5}$$

[5] I assume that the labor markets are competitive and clear at every instant. For now, I also assume that skill supplies are given, but below I endogeneize native workers' occupational choice (see Section B.5.2).

The price ratio is thus[6]

$$p \equiv \frac{p_H}{p_L} = \left(\frac{H}{Y_L}\right)^{\gamma-1} \tag{B7}$$

Since $Y_H = H$, it follows directly that

$$w_H = p_H \tag{B8}$$

Next, from the maximization problem of producers of good Y_L, it is possible to derive the demand for machines:

$$k_L(v) = \left(\frac{p_L}{p_L^k(v)}\right)^{\frac{1}{\beta}} L \ \forall v \tag{B9}$$

The profit maximization of monopolists, in turn, implies that the price of each variety is given by

$$p_L^k(v) = 1 \ \forall v \tag{B10}$$

so that

$$k_L(v) = p_L^{\frac{1}{\beta}} L \ \forall v \tag{B11}$$

Using $(B11)$ and $(B10)$, monopolists' profits are then

$$\pi_L = \beta p_L^{\frac{1}{\beta}} L \tag{B12}$$

implying that the net present discounted value of profits for a monopolist is

$$V_L = \frac{\beta p_L^{\frac{1}{\beta}} L}{r} \tag{B13}$$

where r is the interest rate. Even though, in principle, the interest rate can be time-varying, I focus on a balanced growth path (BGP), where r is constant and equal to $(\theta g + \rho)$, where g is the steady state growth rate of output (see below).

Replacing $(B11)$ in $(B2)$, we get

$$Y_L = \frac{N_L L}{1-\beta} p_L^{\frac{1-\beta}{\beta}} \tag{B14}$$

Using $(B14)$, and solving the maximization problem of intermediate producers in sector L,

[6] It should be noted that normalizing the price of the final good to 1 is equivalent to write

$$\left[p_H^{\frac{\gamma}{\gamma-1}} + p_L^{\frac{\gamma}{\gamma-1}}\right]^{\frac{\gamma-1}{\gamma}} = 1 \tag{B6}$$

one can derive the unskilled wage, given by

$$w_L = \frac{N_L}{1-\beta} p_L^{\frac{1}{\beta}} \tag{B15}$$

Finally, the free entry condition in the machine-producing market implies that

$$V_L \eta_L = 1$$

Or,

$$\eta_L \beta p_L^{\frac{1}{\beta}} L = r$$

The previous expression pins down the price of Y_L as a function of r, η_L, β, and L:[7]

$$p_L = \left(\frac{r}{\eta_L \beta L} \right)^{\beta} \tag{B16}$$

In online appendix C, I show that, using $(B16)$ in $(B14)$ and combining the resulting expression with $(B5)$ and $(B6)$, it is possible to derive an equation that characterizes the relationship between the equilibrium number of plants, N_L, and the supply of both high and low skilled workers (H and L):

$$N_L = \frac{H(1-\beta) L^{\frac{\beta\gamma}{1-\gamma}}}{\psi^{(1-\beta)} \left[\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}} \right]^{\frac{1}{\gamma}}} \tag{B17}$$

where $\psi \equiv \frac{r}{\eta_L \beta}$.

The last step to fully characterize the steady state equilibrium of the economy is to determine the BGP growth rate, g. As noted above, along the BGP, $r = \theta g + \rho$. Using the free entry condition into the monopolist sector, it can be shown that (see also Acemoglu, 2002)[8]

$$g = \frac{1}{\theta} \left[\beta \eta_L L - \rho \right] \tag{B18}$$

Before turning to the comparative statics exercise of the next section, where I study the effects of immigration on the economy, let me highlight three important results, which will be used extensively below. Direct inspection of $(B16)$ and of $(B17)$ shows that

$$\frac{\partial p_L}{\partial L} < 0 \tag{B19}$$

[7]Note that, once we have p_L, it is immediate to get p_H from $(B6)$: $p_H = \left(1 - \left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{\frac{\gamma-1}{\gamma}}$.

[8]Note that, from the No Ponzi condition it directly follows that $\rho > g(1-\theta)$.

$$\frac{\partial N_L}{\partial H} > 0 \ \forall \gamma \tag{B20}$$

and, most importantly,

$$\gamma > 0 \implies \frac{\partial N_L}{\partial L} > 0 \tag{B21}$$

The three results, $(B19)$, $(B20)$, and $(B21)$, are standard in the biased technical change literature (Acemoglu, 1998, 2002). However, especially $(B21)$ will be very important when studying the effects of immigration in the next section, so it is worth briefly discussing the intuition behind it. Specifically, incentives to enter the manufacturing sector depend on two forces - a price and a market size effect. When the former dominates, an increase in the supply of a given factor reduces incentives to introduce technologies complementary to that factor. When the latter prevails, instead, higher supplies of a factor will make it more profitable to develop technologies biased towards that factor. As stated in $(B21)$ (see the proof in online appendix C), if $\gamma > 0$, i.e. when the degree of complementarity between high and low skilled workers is not too high, the market size effect will be stronger, and an increase in the supply of unskilled labor will induce capital accumulation in the unskilled sector, by increasing the number of plants producing technologies that are unskill-biased.

B.4 Evaluating the Effects of Immigration

In this section, I study how an exogenous increase in immigration affects the economy. To mirror the empirical setting considered in my paper, I assume that immigrants can only be employed in the unskilled sector, and do not have access to high skilled jobs (see Table A4), either because of skill mismatch or because of discrimination. For the moment, I assume that unskilled natives and immigrants are perfect substitutes, and that natives' labor supply in each sector is fixed. Below, I relax both these assumptions. Before turning to the analysis, note the followings. First, it is trivial to see that an increase in N_L mechanically favors capital accumulation. Second, from $(B15)$ it is immediate to verify that the unskilled wage is increasing in N_L and decreasing in L. Third, from $(B6)$, it follows directly that an increase in p_L will lower p_H, so that higher (lower) p_L will depress (increase) the high skilled wage.

Now, assume that the economy experiences an exogenous inflow of immigrants, which increases L. What happens to capital, wages, and the skill premium?

Capital Accumulation. First, from $(B21)$, we know that if

$$\gamma > 0 \tag{B22}$$

N_L is increasing in L. Hence, the first result is that, if $(B22)$ holds, immigration favors capital accumulation in the unskilled sector.[9]

High Skilled Wages. Second, it is immediate to see from $(B16)$ that higher immigration will reduce the price of Y_L, p_L, and, in turn raise p_H and w_H (see $(B6)$). Thus, immigration has a positive and unambiguous effect on high skilled wages.

Unskilled Wages. Turning to the impact of immigration on wages of unskilled workers, there are two countervailing forces. First, immigration has a negative effect on unskilled wages - the standard *substitution effect* that takes place as the economy moves along the (downward sloping) demand curve. Second, if $\gamma > 0$, there is a *directed technology effect* (Acemoglu, 1998): the increase in skill supplies (induced by immigration) increases incentives to open new plants and develop skill-complementary technologies, in turn exerting positive pressure on w_L. Remember that

$$w_L = \frac{\psi N_L}{L(1-\beta)} \tag{B23}$$

Then, from the previous expression, it is immediate to see how the two channels (the substitution effect and the capital response) just described affect the unskilled wage. Online appendix C provides an expression showing for which parameter values the directed technology effect prevails over the substitution effect. In line with Acemoglu (2002), this happens when γ is sufficiently large.[10] The main take-away from this discussion is that, when technology is allowed to be directed and as long as $\gamma > 0$, the standard (substitution) negative effect of immigration on earnings of unskilled natives will be partly (or even completely) offset by the endogenous technology response.

Skill Premium. Finally, I evaluate the effects of immigration on the skill premium, $\omega \equiv \frac{w_H}{w_L}$. Using the equilibrium conditions derived above, the skill premium can be written as

$$\omega = \left(\frac{1-\beta}{\psi}\right) \frac{\left(1 - \left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{\frac{\gamma-1}{\gamma}}}{N_L(L)} L \tag{B24}$$

where I am emphasizing the fact that, in equilibrium, N_L is a function of L (see $(B17)$). From $(B24)$, it is clear that an increase in L (induced by immigration) has two separate effects on the skill premium. First, higher L reduces w_L because of substitution and increases w_H because of complementarity (at least as long as $\gamma < 1$). Second, there is an indirect effect, operating through changes in N_L. Whenever $\gamma > 0$, the latter will tend to offset (and, if γ is sufficiently high even reverse) the positive effect of immigration on the skill

[9]This result follows directly from the fact that, in equilbrium, $K = \frac{N_L}{1-\beta}\psi$.

[10]In particular, a sufficient (but not a necessary) condition for the total effect of immigration on the unskilled wage to be positive is that $\gamma > \frac{1}{1+\beta}$. This condition can be equivalently expressed in terms of the derived elasticity of substitution, $\sigma \equiv \left(\frac{1}{1-\gamma} - 1\right)\beta + 1$, as $\sigma > 2$ (Acemoglu, 2002).

premium. In online appendix C, I explicitly derive expressions for each of the two forces, and provide a sufficient condition (in terms of γ and β) under which immigration reduces the skill premium.[11]

To summarize, when technology is endogenous and ($B22$) holds, an exogenous shock to immigration:

1. Increases capital accumulation in the unskilled sector
2. Raises the high skilled wage
3. Has ambiguous effects on both the unskilled wage and the skill premium. If the degree of substitutability between factors (i.e. γ) is sufficiently high, immigration can even be beneficial to unskilled natives.

Of course, one should not conclude that immigration is necessarily beneficial to *all* natives. In fact, the previous analysis makes it clear that, for immigration to benefit (or at least not to harm) natives in the more exposed sector, specific conditions - in particular, scope for capital accumulation and technological upgrading - must be satisfied.

B.5 Extensions

Thus far, I have neglected two potentially important mechanisms that, in addition to the capital response highlighted above, can help natives in more exposed occupations to cope with a sudden increase in immigration. First, I assumed that immigrants and unskilled natives are perfect substitutes in production; second, I fixed natives' labor supply in each sector. Yet, a large body of the literature has documented that neither condition is likely to hold in practice (Card, 2005; Peri and Sparber, 2009; Ottaviano and Peri, 2012; Foged and Peri, 2016). For this reason, and to more thoroughly analyze the channels through which immigration affects natives' labor market outcomes, I now relax each of the two assumptions.

B.5.1 Imperfect Substitutability Between Immigrants and Natives

I start by relaxing the assumption that immigrants and unskilled natives are perfect substitutes. In particular, I specify the total supply of unskilled labor as

$$L = [I^{\alpha} + U^{\alpha}]^{\frac{1}{\alpha}} \tag{B25}$$

[11]As in Acemoglu (2002), a sufficient condition for ω to fall with L is that $\gamma > \frac{1}{1+\beta}$.

where I and U refer, respectively, to immigrants and unskilled natives, and $\alpha \leq 1$ governs the elasticity of substitution between the two. When $\alpha \rightarrow 1$, we are in the limit case of perfect substitutability considered above. Since immigrants and unskilled natives are likely to display at least some degree of substitutability, I assume that $\alpha > 0$, but do not restrict this parameter any further.

When $\alpha \in (0,1)$, an increase in immigration will raise the unskilled labor aggregate in $(B25)$ more than one for one. To see this, note that

$$\frac{\partial L}{\partial I} = \left[1 + \left(\frac{U}{I}\right)^{\alpha}\right]^{\frac{1-\alpha}{\alpha}} \tag{B26}$$

As long as $\alpha \in (0,1)$, the term inside the square brackets is strictly greater than 1, and elevating this to $\left(\frac{1-\alpha}{\alpha}\right)$ will never yield a number below 1 (in the limit case of $\alpha = 1$, the increase in I will imply a one for one increase in L). It follows that

$$\frac{\partial L}{\partial I} \geq \frac{\partial L}{\partial L} = 1 \tag{B27}$$

with a strict inequality whenever $\alpha \in (0,1)$. The result in $(B27)$ is going to be important for some of the comparative static exercises below.

From now onwards, let us consider only the (empirically relevant) case in which $0 < \alpha < 1$. As before, I now study the effects of an exogenous increase in immigration on capital, wages, and on the skill premium.

Capital Accumulation. Remember from above that as long as $\gamma > 0$, $\frac{\partial N_L}{\partial L} > 0$. Hence, $(B27)$ immediately implies that

$$\frac{\partial N_L}{\partial I} > \frac{\partial N_L}{\partial L} > 0 \tag{B28}$$

In words, once we allow for immigrants and unskilled natives to be imperfect substitutes (i.e. $\alpha \in (0,1)$), if $\gamma > 0$, not only immigration has a positive effect on the number of plants producing machines complementary to unskilled workers, but also, this effect is going to be larger than in the baseline case of perfect substitutability.

High Skilled Wages. Since

$$w_H = \left(1 - \left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{\frac{\gamma-1}{\gamma}}$$

it follows that $\frac{\partial w_H}{\partial L} > 0$. From ($B27$) we know that $\frac{\partial L}{\partial I} > \frac{\partial L}{\partial L}$, and so

$$\frac{\partial w_H}{\partial I} = \frac{\partial w_H}{\partial L}\frac{\partial L}{\partial I} > \frac{\partial w_H}{\partial L}\frac{\partial L}{\partial L} > 0 \tag{B29}$$

That is, as for capital accumulation, also the high skilled wage increases more in response to immigration when immigrants are imperfect (and not perfect) substitutes for unskilled natives.

Unskilled (Natives) Wages. Differently from above, we now have to distinguish between wages of unskilled natives and those of immigrants. In particular, it can be shown that, in equilibrium,[12]

$$w_U = \frac{\psi N_L}{(1-\beta)}\frac{L^{-\alpha}}{U^{1-\alpha}} \tag{B30}$$

As in Section B.4, it is immediate to see how the two channels (the substitution effect and the capital response) affect the wage of unskilled natives: on the one hand, higher immigration increases competition for unskilled natives, thereby lowering their marginal product; on the other, when $\gamma > 0$, immigration favors the entry of establishments producing unskilled-complementary technologies, in turn exerting positive pressure on unskilled wages. By comparing ($B30$) to ($B23$), it is clear that, because of imperfect substitutability between immigrants and natives (i.e. $\alpha < 1$), the (negative) substitution effect is now smaller than in the baseline model presented above.

In online appendix C, I provide a sufficient condition for the directed technology effect to prevail over the substitution effect, and show that the range of values of γ for which immigration raises the wage of unskilled natives is larger than in the case of perfect substitutability between immigrants and natives.[13] More formally, defining $\tilde{\gamma}$ (resp. $\tilde{\gamma}'$) the threshold value of γ above which immigration increases earnings of unskilled natives when $\alpha = 1$ (resp. $\alpha < 1$), online appendix C shows that

$$\tilde{\gamma} > \tilde{\gamma}' \ \forall \alpha \in (0,1) \tag{B31}$$

This result is intuitive: when immigrants and natives are imperfect substitutes, the direct

[12]To see this, note that

$$\begin{aligned} w_U &= \frac{\partial (p_L Y_L)}{\partial U} = \frac{\partial (p_L Y_L)}{\partial L}\frac{\partial L}{\partial U} \\ &= w_L \left(\frac{L}{U}\right)^{1-\alpha} \end{aligned}$$

[13]In particular, a sufficient condition for the wage of unskilled natives to increase with immigration is that $\gamma > \frac{\alpha}{\alpha+\beta}$.

negative (competition) effect of immigration on natives' wages is counterbalanced by two distinct forces. First, as before, capital accumulation and the development of (unskilled) biased technologies. Second, complementarity between the skills of immigrants and natives and the resulting gains from diversity (e.g. Peri and Sparber, 2009; Foged and Peri, 2016, among others).

Skill Premium. The skill premium can be now expressed as

$$\omega = \frac{w_H}{w_U} = \left(\frac{1-\beta}{\psi}\right) \frac{\left(1-\left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{-\frac{1-\gamma}{\gamma}}}{N_L(L)} L^{\alpha}U^{1-\alpha} \tag{B32}$$

As before, it is possible to show that the direct effect of immigration on the skill premium is positive. This result is intuitive, and follows directly from the assumption that immigrants are closer substitutes for unskilled than for high skilled natives. Also, similar to Section B.4, the indirect effect of immigration mediated by capital deepening tends to lower the skill premium. The total effect of immigration is, as usual, given by

$$\left(\frac{\partial\omega}{\partial I}\right)^{TOT} = \left[\frac{\partial\omega}{\partial L} + \frac{\partial\omega}{\partial N_L}\frac{\partial N_L}{\partial L}\right]\frac{\partial L}{\partial I}$$

and, as already noted above, is ambiguous. In online appendix C, I derive an explicit condition that shows under which parameter values the skill premium falls with immigration.[14] As for the unskilled wage, also in this case, introducing the assumption of imperfect substitutability between immigrants and natives ($\alpha < 1$) increases the range of values of γ for which immigration can reduce income inequality, relative to the scenario of perfect substitution ($\alpha = 1$).

To conclude, assuming (consistent with the empirical evidence) that immigrants and unskilled natives are imperfect substitutes in the production of Y_L lowers the degree of capital adjustment needed for the economy to absorb an immigration shock . Even in this case, however, whether or not there is room for major technological change is probably a key condition for immigration to benefit native workers, without harming even those in more exposed jobs.

B.5.2 Endogeneizing Natives' Occupational Choice

In this sub-section I formalize the idea that, in response to immigration, natives might reallocate their labor away from occupations more exposed to immigrants' competition and

[14]Specifically, if $\gamma > \frac{\alpha}{\alpha+\beta}$, immigration will reduce income inequality among natives.

take up more skilled jobs. As argued in Peri and Sparber (2009) among others, such labor reallocation can take place because natives and immigrants differ in terms of skills, language proficiency, and education. As a result, natives may be induced to specialize in occupations where they have a comparative advantage relative to immigrants.

The structure of the model is as before, but I now assume that there are two types of domestic labor: first, native whites; second, African Americans and previously arrived immigrants. Native whites can be employed in both sectors, whereas African Americans and immigrants can only work in the unskilled sector, due to skill mismatch and discrimination. To simplify the analysis, I assume, as in the baseline model, that native whites working in the unskilled sector are perfect substitutes for immigrants and African Americans.[15]

Wages are allowed to differ across sectors, but all workers are paid the same within each sector. I denote native whites working in the high and low skilled sectors respectively with H and U, and, without loss of generality I normalize $H + U = 1$. The assumption of perfect substitutability between unskilled natives and immigrants implies that $L = U + I$, where I refers to immigrants and African Americans. It is straightforward to verify that native whites choose the sector paying the higher wage, and so, for them to work in both sectors, wages must be equalized, i.e.

$$\omega \equiv \frac{w_H}{w_L} = 1 \tag{B33}$$

Suppose that, before the immigration shock, $(B33)$ holds so that native whites are employed in both sectors. Combining $(B33)$ with $(B24)$, we get

$$1 = \left(\frac{1-\beta}{\psi}\right) \frac{\left(1 - \left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{\frac{\gamma-1}{\gamma}}}{N_L} L \tag{B34}$$

Replacing $(B17)$ in $(B34)$, it is possible to determine the equilibrium number of native whites working as laborers (before the immigration shock), which is given by[16]

$$U = \frac{\psi^{\frac{\gamma\beta}{\gamma(1+\beta)-1}}}{(1+I)^{\frac{1-\gamma}{\gamma(1+\beta)-1}}} - I \tag{B35}$$

Having determined U from $(B35)$, and noting that $H = 1 - U$, all other equations follow as in the baseline model of Section B.3, with the only difference that, now, skill supplies (of native whites) are endogenously determined according to $(B34)$.

In what follows, I investigate how an immigration shock affects capital, wages, and the

[15] Relaxing this assumption does not alter any of the results below.
[16] See online appendix C.

distribution of native workers across the two sectors. Two cases can arise. First, even after the immigration shock, wages are equalized across sectors, and native whites continue to work in both sectors.[17] Second, after the immigration shock ($B34$) no longer holds, and all native whites move to the high skilled sector. To keep the analysis close to my empirical results, I focus on the second scenario, and show that, in this framework, after the immigration shock: *i*) all native whites work in the high skilled sector and earn a higher wage (relative to the pre-migration equilibrium); *ii*) the number of manufacturing plants in the new equilibrium is higher; *iii*) it is possible even for wages of African Americans and previously arrived immigrants not to fall (or, to experience only a small decline).

Sector and Wages of Native Whites. First, by assumption, the new equilibrium entails $H = 1$, $U = 0$, and $\omega > 1$. Second, when the immigration shock is sufficiently large relative to the initial (native) labor force in the unskilled sector, it is possible for the high skilled wage to be higher after the immigration shock (relative to its pre-immigration level). Remembering that

$$w_H = \left(1 - \left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{-\frac{1-\gamma}{\gamma}}$$

and denoting with the subscript 1 (resp. 0) the equilibrium variables after (resp. before) the immigration shock, the condition $w_{1,H} > w_{0,H}$ can be written as

$$\left(1 - \left(\frac{I_0 + U_0}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{\frac{1-\gamma}{\gamma}} > \left(1 - \left(\frac{I_1}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{\frac{1-\gamma}{\gamma}}$$

Or, after a few rearrangements,[18]

$$I_1 - I_0 > U_0 \qquad \text{(B36)}$$

That is, for natives' wage to increase, the immigration shock must be sufficiently *large* (relative to the fraction of native whites initially working in the unskilled sector).[19]

Unskilled Wages. Next, using ($B23$), the new and the old equilibrium wages in the

[17] It is easy to check that, even in this case, the fraction of natives in the unskilled sector falls when γ is sufficiently high.

[18] Using ($B35$), ($B36$) can be equivalently written as

$$I_1 > \left(\frac{\psi^{\gamma\beta}}{(1 + I_0)^{1-\gamma}}\right)^{\frac{1}{\gamma(1+\beta)-1}}$$

[19] The intuition for this result is discussed below.

unskilled sector are given by

$$w_{1,L} = \frac{\psi N_{1,L}}{I_1 (1-\beta)} \tag{B37}$$

and

$$w_{0,L} = \frac{\psi N_{0,L}}{(I_0 + U_0)(1-\beta)} \tag{B38}$$

where $N_{0,L}$ and $N_{1,L}$ are the pre and post immigration number of manufacturing plants (determined below). For wages in the unskilled sector to be equal before and after the immigration shock, it must be that

$$\frac{N_{1,L}}{I_1 - I_0} = \frac{N_{0,L}}{U_0} \tag{B39}$$

From $(B36)$, it is clear that for both the high skilled wage to rise and the unskilled wage not to fall, the number of manufacturing plants must be higher in the post-immigration equilibrium, i.e. $N_{1,L} > N_{0,L}$. Moreover, the endogenous capital response needed to absorb the immigration shock is increasing in the term $\frac{I_1 - I_0}{U_0}$.

Capital Accumulation. The latter observation already anticipated that, in the new equilibrium, the number of manufacturing plants must be higher than before the immigration shock. Using $(B17)$, we know that

$$N_L = \frac{(1-U)(1-\beta)(I+U)^{\frac{\beta\gamma}{1-\gamma}}}{\psi^{(1-\beta)} \left[\psi^{\frac{\beta\gamma}{1-\gamma}} - (I+U)^{\frac{\beta\gamma}{1-\gamma}}\right]^{\frac{1}{\gamma}}}$$

Then,

$$N_{1,L} = \frac{(1-\beta) I_1^{\frac{\beta\gamma}{1-\gamma}}}{\psi^{(1-\beta)} \left[\psi^{\frac{\beta\gamma}{1-\gamma}} - I_1^{\frac{\beta\gamma}{1-\gamma}}\right]^{\frac{1}{\gamma}}}$$

and

$$N_{0,L} = \frac{(1-U_0)(1-\beta)(I_0+U_0)^{\frac{\beta\gamma}{1-\gamma}}}{\psi^{(1-\beta)} \left[\psi^{\frac{\beta\gamma}{1-\gamma}} - (I_0+U_0)^{\frac{\beta\gamma}{1-\gamma}}\right]^{\frac{1}{\gamma}}}$$

Combining the latter two expressions, $N_{1,L} > N_{0,L}$ whenever

$$\frac{I_1^{\frac{\beta\gamma}{1-\gamma}}}{\left[\psi^{\frac{\beta\gamma}{1-\gamma}} - I_1^{\frac{\beta\gamma}{1-\gamma}}\right]^{\frac{1}{\gamma}}} > \frac{(1-U_0)(I_0+U_0)^{\frac{\beta\gamma}{1-\gamma}}}{\left[\psi^{\frac{\beta\gamma}{1-\gamma}} - (I_0+U_0)^{\frac{\beta\gamma}{1-\gamma}}\right]^{\frac{1}{\gamma}}}$$

Taking logs on both sides and rearranging, we get

$$\frac{\beta\gamma}{1-\gamma}\log\left(\frac{I_1}{I_0+U_0}\right) > \log(1-U_0) + \frac{1}{\gamma}\log\left(\frac{\Phi_1}{\Phi_0}\right) \tag{B40}$$

where $\Phi_1 \equiv \psi^{\frac{\beta\gamma}{1-\gamma}} - I_1^{\frac{\beta\gamma}{1-\gamma}}$ and $\Phi_0 \equiv \psi^{\frac{\beta\gamma}{1-\gamma}} - (I_0+U_0)^{\frac{\beta\gamma}{1-\gamma}}$. Note that, from $(B36)$,

$$I_1 > I_0 + U_0$$

implying that $\log\left(\frac{I_1}{I_0+U_0}\right) > 0$. Similarly, $\Phi_1 < \Phi_0$, and so $\log\left(\frac{\Phi_1}{\Phi_0}\right) < 0$. Finally, since $U_0 \in (0,1)$, $\log(1-U_0) < 0$. But then, if $(B36)$ holds, $(B40)$ is always satisfied.

Discussion. The previous analysis showed that, if natives can reallocate their labor across sectors (but immigrants cannot), and if capital endogenously adjusts after the immigration shock, the followings can happen: *i*) all natives end up working in the high skilled sector; and *ii*) even workers that are prevented from entering the high skilled sector might experience only limited wage losses. Two mechanisms are responsible for (*i*) and (*ii*). First, natives' endogenous occupational choice allows them to move away from the sector most exposed to immigration and, potentially, take advantage of the complementarity between their skills and those of immigrants. Second, and crucially, capital endogenously adjusts to the inflow of immigrants - this is the capital response that was already operating in the previous versions of the model.

When the inflow of immigrants is *sufficiently large*, capital accumulation will not only boost wages in the skilled sector, but also, will partly or completely offset the direct, negative effect of immigration on earnings of workers in the unskilled sector. When analyzing these results from the lenses of a neoclassical framework, the latter observation might seem somewhat counterintuitive: the economy should be better able to cope with immigration when the latter is relatively contained. But, this line of reasoning misses the key point.

Specifically, the neoclassical framework fails to incorporate the endogenous (directed) technological response, which is key for the economy to absorb the immigration shock. By raising the supply of unskilled workers, immigration increases firms' incentives to invest. Capital accumulation, in turn, increases the marginal productivity of both high and low skilled workers, compensating (or reversing) the initial negative effect of immigration on wages.

B.6 Taking Stock

In this note, building on a standard model of biased technical change (Acemoglu, 2002), I presented a tractable framework to study the effects of immigration on natives' labor market outcomes, incorporating three important mechanisms. First, the degree to which firms can expand (or enter the market) and the scope for major capital adjustments. Second, complementarity in the skills, the language proficiency, and in education of immigrants and natives. Third, the potential decision of natives to reallocate their labor away from more exposed occupations, and into sectors where they have a comparative advantage relative to immigrants. I derived conditions under which the model is able to deliver the key findings documented in my paper, namely that immigration can: *i*) increase natives' employment, without harming any specific group; *ii*) promote capital accumulation and boost economic activity; and *iii*) favor natives' occupational mobility, by increasing (lowering) the fraction of natives in high (low) skilled occupations.

References

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. *The Quarterly Journal of Economics 113*(4), 1055–1089.

Acemoglu, D. (2002). Directed technical change. *The Review of Economic Studies 69*(4), 781–809.

Card, D. (2005). Is the new immigration really so bad? *The Economic Journal 115*(507).

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

Goldin, C. D. and L. F. Katz (2009). *The race between education and technology.* Harvard University Press.

Ottaviano, G. I. and G. Peri (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association 10*(1), 152–197.

Peri, G. and C. Sparber (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics 1*(3), 135–69.

C. Proofs for Theoretical Framework for Chapter 1

C1. Relationship Between Number of Machines, H, and L

In what follows, I first obtain an expression that relates N_L to H and L. Next, I show that: $i)$ $\frac{\partial N_L}{\partial H} > 0$ and; $ii)$ $\frac{\partial N_L}{\partial L} > 0$ if $\gamma > 0$. Using $(B16)$ in $(B14)$, we get

$$Y_L = \frac{N_L L^{\beta}}{1-\beta}\left(\frac{r}{\eta_L \beta}\right)^{1-\beta}$$

Plugging this back in $(B4)$, we get

$$\psi^{\beta} L^{-\beta} = \left[1 + \frac{H^{\gamma}(1-\beta)^{\gamma}}{N_L^{\gamma} L^{\beta\gamma} \psi^{\gamma(1-\beta)}}\right]^{\frac{1-\gamma}{\gamma}} \tag{1}$$

where $\psi \equiv \frac{r}{\eta_L \beta}$. Rearranging, we obtain

$$\psi^{\beta} = \left[L^{\frac{\beta\gamma}{1-\gamma}} + \frac{H^{\gamma}(1-\beta)^{\gamma}}{N_L^{\gamma}\psi^{\gamma(1-\beta)}} L^{\frac{\beta\gamma^2}{1-\gamma}}\right]^{\frac{1-\gamma}{\gamma}}$$

After some algebra, it is possible to get expression $(B17)$ in appendix B1

$$N_L = \frac{L^{\frac{\beta\gamma}{1-\gamma}} H (1-\beta)}{\psi^{(1-\beta)}\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)^{\frac{1}{\gamma}}} \tag{2}$$

Then, using (2), we can derive exact expressions for $\frac{\partial N_L}{\partial H}$ and $\frac{\partial N_L}{\partial L}$, and show that (i) and (ii) above hold. First,

$$\begin{aligned} \frac{\partial N_L}{\partial H} &= \frac{L^{\frac{\beta\gamma}{1-\gamma}}(1-\beta)}{\psi^{(1-\beta)}\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)^{\frac{1}{\gamma}}} \\ &= \frac{N_L}{H} > 0 \quad \forall\gamma \text{ and } \forall N_L > 0 \end{aligned}$$

Second,

$$\begin{aligned}\frac{\partial N_L}{\partial L} &= \frac{\left(\frac{\beta\gamma}{1-\gamma}\right) L^{\frac{\beta\gamma}{1-\gamma}-1} H (1-\beta)}{\psi^{(1-\beta)}\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)^{\frac{1}{\gamma}}} + \frac{\beta}{1-\gamma} L^{\frac{\beta\gamma}{1-\gamma}-1} \frac{L^{\frac{\beta\gamma}{1-\gamma}} H (1-\beta)}{\psi^{(1-\beta)}\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)^{\frac{1+\gamma}{\gamma}}} \\ &= \phi\left(\gamma + \frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)}\right) \qquad (3)\end{aligned}$$

where

$$\phi \equiv \frac{L^{\frac{\beta\gamma}{1-\gamma}-1} H (1-\beta)\beta}{\psi^{(1-\beta)}\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)^{\frac{1}{\gamma}}(1-\gamma)} > 0$$

The second line of the previous expression shows that, if $\gamma > 0$, $\frac{\partial N_L}{\partial L} > 0$. The intuition for this result is that, when Y_L and Y_H are gross substitutes in the production of the final good, then, the market size effect dominates over the price effect, inducing technological change that is biased towards Y_L. It is possible, instead, that if the two intermediate inputs are sufficiently complements, $\frac{\partial N_L}{\partial L} < 0$.[1]

C2. Quantifying the Effects of L on Unskilled Wages

C2.1 Perfect Substitutability Between Immigrants and Natives

Recall from $(B23)$ in appendix B1 that

$$w_L = \frac{\psi N_L}{L(1-\beta)}$$

Then,

$$\left(\frac{\partial w_L}{\partial L}\right)^{TOT} = \frac{\partial w_L}{\partial L} + \frac{\partial w_L}{\partial N_L}\frac{\partial N_L}{\partial L}$$

First, note that

$$\frac{\partial w_L}{\partial L} = -\frac{\psi N_L L^{-2}}{(1-\beta)} \qquad (4)$$

Next,

$$\frac{\partial w_L}{\partial N_L} = \frac{\psi}{L(1-\beta)}$$

[1] Note, however, that (3) is a sufficient, but not a necessary condition for N_L to be increasing in L.

and, from (3),

$$\frac{\partial N_L}{\partial L} = \phi\left(\gamma + \frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)}\right)$$

with $\phi \equiv \frac{L^{\frac{\beta\gamma}{1-\gamma}-1}H(1-\beta)\beta}{\psi^{(1-\beta)}\left(\psi^{\frac{\beta\gamma}{1-\gamma}}-L^{\frac{\beta\gamma}{1-\gamma}}\right)^{\frac{1}{\gamma}}(1-\gamma)}$. Thus,

$$\frac{\partial w_L}{\partial N_L}\frac{\partial N_L}{\partial L} = L^{-2}\tilde{\phi}\left[\gamma + \frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)}\right] \tag{5}$$

where

$$\tilde{\phi} \equiv \frac{\beta\psi}{(1-\gamma)(1-\beta)}N_L > 0$$

Then, combining (4) with (5), we get

$$\begin{aligned}\left(\frac{\partial w_L}{\partial L}\right)^{TOT} &= \left[-\frac{\psi N_L L^{-2}}{(1-\beta)} + L^{-2}\frac{\beta\psi}{(1-\gamma)(1-\beta)}N_L\left(\gamma + \frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)}\right)\right] \\ &= \frac{\psi N_L L^{-2}}{(1-\beta)}\left[-1 + \frac{\beta}{(1-\gamma)}\left(\gamma + \frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)}\right)\right]\end{aligned}$$

Then, $\left(\frac{\partial w_L}{\partial L}\right)^{TOT} > 0$ whenever

$$\frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)} > \frac{1-\gamma(1+\beta)}{\beta}$$

Since the left-hand side of the previous inequality is positive (when $N_L > 0$), it follows that

$$\gamma > \frac{1}{1+\beta} \Longrightarrow \left(\frac{\partial w_L}{\partial L}\right)^{TOT} > 0 \tag{6}$$

The latter inequality provides a sufficient (but not a necessary) condition for when immigration can raise the unskilled wage. Intuitively, when Y_L and Y_H are sufficiently substitutable in the production of the final good, the directed technology effect (Acemoglu, 1998) will prevail over the (negative) substitution effect - a result consistent with the more general case considered in Acemoglu (2002).

C2.2 Imperfect Substitutability Between Immigrants and Natives

When immigrants and unskilled natives are imperfect substitutes (see Section B.5.1), earnings of natives in the unskilled sector are given by[2]

$$w_U = \frac{\psi N_L}{(1-\beta)} \frac{L^{-\alpha}}{U^{1-\alpha}}$$

Then,

$$\begin{aligned} \left(\frac{\partial w_U}{\partial I}\right)^{TOT} &= \left[-\alpha L^{-\alpha-1} \frac{\psi N_L}{(1-\beta) U^{1-\alpha}} + \frac{\psi}{(1-\beta)} \frac{L^{-\alpha}}{U^{1-\alpha}} \frac{\partial N_L}{\partial L}\right] \frac{\partial L}{\partial I} \\ &= \chi \left[-\alpha N_L + \frac{L \partial N_L}{\partial L}\right] \frac{\partial L}{\partial I} \end{aligned} \quad (7)$$

with $\chi \equiv \frac{\psi L^{-\alpha-1}}{(1-\beta)U^{1-\alpha}} > 0$. Note that

$$\frac{\partial N_L}{\partial L} = \frac{\beta N_L}{(1-\gamma) L} \left(\gamma + \frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)}\right) \quad (8)$$

and so, (7) can be written as

$$\left(\frac{\partial w_U}{\partial I}\right)^{TOT} = \tilde{\chi} \left[-\alpha (1-\gamma) + \beta\gamma + \frac{\beta X}{(1-X)}\right] \frac{\partial L}{\partial I}$$

where $\tilde{\chi} \equiv \frac{\chi N_L}{1-\gamma}$ and $X \equiv \left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}$. Thus, $\left(\frac{\partial w_U}{\partial I}\right)^{TOT} > 0$ whenever

$$-\alpha (1-\gamma) + \beta\gamma + \frac{\beta X}{(1-X)} > 0$$

Or, whenever

$$X > \frac{(1-\gamma)\alpha - \beta\gamma}{(1-\gamma)\alpha - \beta\gamma + \beta\gamma} \quad (9)$$

And so, a sufficient condition for immigration to raise the wage of unskilled natives is that

$$\gamma > \frac{\alpha}{\alpha + \beta} \quad (10)$$

Note that when $\alpha = 1$, condition (10) coincides with (6). Moreover, for any $\alpha \in (0, 1)$, (10) is satisfied for values of γ lower than those needed to satisfy (6). This is intuitive. On the

[2] See $(B30)$ in appendix B.

one hand, if immigrants and natives are imperfect substitutes, the degree of competition induced by an immigration shock is lower than in the case of perfect substitutability. On the other, the capital response to immigration is larger the lower the degree of substitutability between immigrants and natives.

C3. Quantifying the Effects of L on the Skill Premium

C3.1 Perfect Substitutability Between Immigrants and Natives

In this paragraph I derive an explicit expression for the effects of changes in L on the skill premium. First, recall that

$$\omega = \left(\frac{1-\beta}{\psi}\right) \frac{\left(1-\psi^{\frac{\gamma\beta}{\gamma-1}} L^{-\frac{\gamma\beta}{\gamma-1}}\right)^{\frac{\gamma-1}{\gamma}}}{N_L(L)} L$$

As for w_L, also in this case, the effect of immigration on ω can be decomposed as

$$\left(\frac{\partial \omega}{\partial L}\right)^{TOT} = \frac{\partial \omega}{\partial L} + \frac{\partial \omega}{\partial N_L}\frac{\partial N_L}{\partial L}$$

Next,

$$\begin{aligned}
\frac{\partial \omega}{\partial L} &= \left(\frac{1-\beta}{\psi}\right) \frac{\left(1-\psi^{\frac{\gamma\beta}{\gamma-1}} L^{-\frac{\gamma\beta}{\gamma-1}}\right)^{\frac{\gamma-1}{\gamma}}}{N_L} + \left(\frac{1-\beta}{\psi}\right) \beta \frac{\psi^{\frac{\gamma\beta}{\gamma-1}} L^{-\frac{\gamma\beta}{\gamma-1}}}{N_L} \left(1-\psi^{\frac{\gamma\beta}{\gamma-1}} L^{-\frac{\gamma\beta}{\gamma-1}}\right)^{\frac{\gamma-1}{\gamma}-1} \\
&= \left(\frac{1-\beta}{\psi}\right) \frac{\left(1-\psi^{\frac{\gamma\beta}{\gamma-1}} L^{-\frac{\gamma\beta}{\gamma-1}}\right)^{\frac{\gamma-1}{\gamma}}}{N_L} \left[\frac{1-\psi^{\frac{\gamma\beta}{\gamma-1}} L^{-\frac{\gamma\beta}{\gamma-1}} - \beta\psi^{\frac{\gamma\beta}{\gamma-1}} L^{-\frac{\gamma\beta}{\gamma-1}}}{\left(1-\psi^{\frac{\gamma\beta}{\gamma-1}} L^{-\frac{\gamma\beta}{\gamma-1}}\right)}\right] \\
&= \frac{(1-\beta)\left(1-\left(\frac{L}{\psi}\right)^{\frac{\beta\gamma}{1-\gamma}}\right)^{-\frac{1}{\gamma}}}{\psi N_L} \left[1-(1-\beta)\left(\frac{L}{\psi}\right)^{\frac{\beta\gamma}{1-\gamma}}\right] \qquad (11)
\end{aligned}$$

It can be shown that, in a BGP with $N_L > 0$,

$$1 - \left(\frac{L}{\psi}\right)^{\frac{\beta\gamma}{1-\gamma}} > 0$$

and, since $\beta \in (0,1)$, $\frac{\partial \omega}{\partial L} > 0$. This is indeed consistent with the idea that an increase in immigration will lower w_L because of substitutability and increase w_H because of complementarity. Next, considering the indirect effect operating through changes in N_L, we have that

$$\frac{\partial \omega}{\partial N_L} = -\left(\frac{1-\beta}{\psi}\right)\frac{\left(1-\psi^{\frac{\gamma\beta}{\gamma-1}}L^{-\frac{\gamma\beta}{\gamma-1}}\right)^{\frac{\gamma-1}{\gamma}}}{N_L^2}L < 0$$

From (8) we know that

$$\frac{\partial N_L}{\partial L} = \frac{\beta N_L}{(1-\gamma)L}\left(\gamma + \frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)}\right)$$

and so, we get

$$\begin{aligned}\frac{\partial \omega}{\partial N_L}\frac{\partial N_L}{\partial L} &= -\frac{(1-\beta)\beta}{\psi(1-\gamma)N_L}(1-X)^{\frac{\gamma-1}{\gamma}}\left(\gamma+\frac{X}{1-X}\right) \\ &= -\frac{(1-\beta)\beta}{\psi(1-\gamma)N_L}(1-X)^{-\frac{1}{\gamma}}(\gamma + X(1-\gamma)) \qquad (12)\end{aligned}$$

with $X \equiv \left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}$, $X \in (0,1)$. Then, combining (11) with (12), we obtain

$$\begin{aligned}\left(\frac{\partial \omega}{\partial L}\right)^{TOT} &= \frac{(1-\beta)(1-X)^{-\frac{1}{\gamma}}}{\psi N_L}\left(1-(1-\beta)X - \frac{\beta}{(1-\gamma)}(\gamma + X(1-\gamma))\right) \\ &= \frac{(1-\beta)(1-X)^{-\frac{1}{\gamma}}}{\psi N_L}(1-\gamma-\gamma\beta-(1-\gamma)X)\end{aligned}$$

From the previous expression it then follows that $\left(\frac{\partial\omega}{\partial L}\right)^{TOT} > 0$ whenever

$$X < \frac{1-\gamma-\gamma\beta}{1-\gamma} \qquad (13)$$

Since $X \in (0,1)$, it is easy to show that a sufficient condition for the skill premium to fall with immigration is that

$$\gamma > \frac{1}{1+\beta} \qquad (14)$$

Note that, expressing (14) in terms of the derived elasticity of substitution, $\sigma \equiv (\varepsilon - 1)\beta + 1$, where $\varepsilon = \frac{1}{1-\gamma}$, we reach exactly the same condition as in Acemoglu (2002). That is, the skill premium falls following an increase in L, whenever $\sigma > 2$.

C3.2 Imperfect Substitutability Between Immigrants and Natives

Let us now consider the case in which immigrants and unskilled natives are imperfect substitutes. As discussed in Section B.5.1, the skill premium is given by

$$\omega = \frac{w_H}{w_U} = \left(\frac{1-\beta}{\psi}\right) \frac{\left(1-\left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{-\frac{1-\gamma}{\gamma}}}{N_L(L)} L^{\alpha} U^{1-\alpha}$$

We know from before that

$$\left(\frac{\partial \omega}{\partial L}\right)^{TOT} = \frac{\partial \omega}{\partial L} + \frac{\partial \omega}{\partial N_L}\frac{\partial N_L}{\partial L}$$

and so

$$\left(\frac{\partial \omega}{\partial I}\right)^{TOT} = \left[\frac{\partial \omega}{\partial L} + \frac{\partial \omega}{\partial N_L}\frac{\partial N_L}{\partial L}\right]\frac{\partial L}{\partial I}$$

where $\frac{\partial L}{\partial I} > 0$. Then,

$$\begin{aligned}\frac{\partial \omega}{\partial L} &= \lambda\left(\beta\frac{\left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}}{1-\left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}} + \alpha\right) \\ &= \lambda\left(\alpha + \frac{X\beta}{1-X}\right) > 0 \end{aligned} \tag{15}$$

where $\lambda \equiv \frac{1-\beta}{\psi N_L} L^{\alpha-1} U^{1-\alpha} (1-X)^{-\frac{1-\gamma}{\gamma}} > 0$, and $X \equiv \left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}$ as before. Next,

$$\begin{aligned}\frac{\partial \omega}{\partial N_L} &= -\left(\frac{1-\beta}{\psi}\right)\left(1-\left(\frac{L}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{-\frac{1-\gamma}{\gamma}} L^{\alpha}U^{1-\alpha}N_L^{-2} \\ &= -\left(\frac{1-\beta}{\psi}\right)(1-X)^{-\frac{1-\gamma}{\gamma}} L^{\alpha}U^{1-\alpha}N_L^{-2}\end{aligned}$$

and, as we already saw many times,

$$\begin{aligned}\frac{\partial N_L}{\partial L} &= \frac{\beta N_L}{(1-\gamma)L}\left(\gamma + \frac{L^{\frac{\beta\gamma}{1-\gamma}}}{\left(\psi^{\frac{\beta\gamma}{1-\gamma}} - L^{\frac{\beta\gamma}{1-\gamma}}\right)}\right) \\ &= \frac{\beta N_L}{(1-\gamma)L}\left(\gamma + \frac{X}{(1-X)}\right)\end{aligned}$$

The latter two expressions imply that

$$\frac{\partial \omega}{\partial N_L}\frac{\partial N_L}{\partial L} = -\frac{U^{1-\alpha}}{L^{1-\alpha}N_L}\frac{(1-\beta)\beta}{\psi(1-\gamma)}(1-X)^{-\frac{1-\gamma}{\gamma}-1}(\gamma(1-X)+X)$$
$$= -\frac{U^{1-\alpha}}{L^{1-\alpha}N_L}\frac{(1-\beta)\beta}{\psi(1-\gamma)}(1-X)^{-\frac{1}{\gamma}}(\gamma+X(1-\gamma)) \quad (16)$$

Finally, combining (15) and (16), we get

$$\left(\frac{\partial \omega}{\partial L}\right)^{TOT} = \frac{1-\beta}{\psi N_L}\frac{U^{1-\alpha}}{L^{1-\alpha}}(1-X)^{-\frac{1}{\gamma}}(\alpha(1-X)+X\beta) - \frac{U^{1-\alpha}}{L^{1-\alpha}N_L}\frac{(1-\beta)\beta}{\psi(1-\gamma)}(1-X)^{-\frac{1}{\gamma}}(\gamma+X(1-\gamma))$$
$$= \frac{1-\beta}{\psi N_L}\frac{U^{1-\alpha}}{L^{1-\alpha}}(1-X)^{-\frac{1}{\gamma}}\left[\alpha(1-X)+X\beta-\frac{\beta}{1-\gamma}(\gamma+X(1-\gamma))\right]$$
$$= \xi\left[\alpha(1-\gamma)-\beta\gamma-\alpha(1-\gamma)X\right]$$

where $\xi \equiv \frac{1-\beta}{\psi(1-\gamma)N_L}\frac{U^{1-\alpha}}{L^{1-\alpha}}(1-X)^{-\frac{1}{\gamma}} > 0$. Hence, it follows that $\left(\frac{\partial\omega}{\partial L}\right)^{TOT} > 0$ whenever

$$\alpha(1-\gamma)-\beta\gamma-\alpha(1-\gamma)X > 0$$

Or, whenever

$$X < \frac{\alpha(1-\gamma)-\beta\gamma}{\alpha(1-\gamma)} \quad (17)$$

Note that, when $\alpha = 1$, (17) coincides with (13) that we derived for the case of perfect substitutability between immigrants and natives. As we have done many times at this point, we can derive a sufficient condition, relating γ to α and β, such that immigration lowers the skill premium. In particular, if

$$\gamma > \frac{\alpha}{\alpha+\beta} \quad (18)$$

an inflow of (unskilled) immigrants will lower income inequality among natives. As for wages of unskilled natives, also in this case, the range of values of γ for which immigration compresses the income gap between high and low skilled workers is larger than when $\alpha = 1$.

C4 Natives' Occupational Choice

In this section, I derive the expression for the number of native whites in the unskilled sector before the immigration shock, i.e. ($B35$) in appendix B1. Start from ($B34$), and combine it with ($B17$). Remembering that $L = I + U$, and that $H = 1 - U$, ($B34$) becomes

$$\frac{\psi^{\beta}(1-U)(I+U)^{\frac{\beta\gamma}{1-\gamma}}}{\left[\psi^{\frac{\beta\gamma}{1-\gamma}}-(I+U)^{\frac{\beta\gamma}{1-\gamma}}\right]^{\frac{1}{\gamma}}}=\left(1-\left(\frac{I+U}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)^{\frac{\gamma-1}{\gamma}}(I+U)$$

$\Longrightarrow$

$$(1-U)\left(\frac{I+U}{\psi}\right)^{\frac{\beta\gamma}{1-\gamma}}=\left(1-\left(\frac{I+U}{\psi}\right)^{\frac{\gamma\beta}{1-\gamma}}\right)(I+U)$$

And, after some further rearrangements, it is possible to obtain

$$(I+U)^{\frac{\gamma(1+\beta)-1}{1-\gamma}}=\frac{\psi^{\frac{\gamma\beta}{1-\gamma}}}{1+I}\Longrightarrow U=\frac{\psi^{\frac{\gamma\beta}{\gamma(1+\beta)-1}}}{(1+I)^{\frac{1-\gamma}{\gamma(1+\beta)-1}}}-I$$

verifying $(B35)$.

A2. Supplementary Tables and Figures to Chapter 2

Table A2.1. List of Cities

Albany, NY	Evansville, IN	Providence, RI
Baltimore, MD	Hartford, CT	Rochester, NY
Boston, MA	Indianapolis, IN	St. Louis, MO
Bridgeport, CT	Kansas City, MO	Salt Lake City, UT
Buffalo, NY	Los Angeles, CA	San Diego, CA
Chicago, IL	Milwaukee, WI	San Francisco, CA
Cincinnati, OH	Minneapolis, MN	Scranton, PA
Cleveland, OH	New Haven, CT	Seattle, WA
Columbus, OH	New York, NY	Springfield, MA
Dayton, OH	Omaha, NE	Tacoma, WA
Denver, CO	Peoria, IL	Trenton, NJ
Des Moines, IA	Philadelphia, PA	Washington, DC
Detroit, MI	Pittsburgh, PA	Wichita, KS
Duluth, MN	Portland, OR	Youngstown, OH

Note: The sample includes the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration, as discussed in the main text.

Table A2.2. Variables' Sources and Definitions

Variable	Source
City population (total and by race)	US Census of Population
Number of dwellings	US Census of Population
Number of illiterate (total and by race)	US Census of Population
Pupils enrolled in public schools	US Census of Population
City population by age	US Census of Population
City population by nativity	US Census of Population
Net migration from southern states	Estimated using US Census data (ICPSR 2896)
Public Spending (total and by category)	Financial Statistics of Cities
Assessed valuation of property	Financial Statistics of Cities
Tax rate on $1,000 of assessed valuation	Financial Statistics of Cities
Tax revenues	Financial Statistics of Cities
Public debt and interest payments	Financial Statistics of Cities
Land area (in km)	Financial Statistics of Cities
Spending on poverty relief	Geddes (1937)
Number of highway rays planned and built	Baum-Snow (2007)
Number of special districts by county	US Census of Governments
Number of municipalities by county	US Census of Governments
Elevation data	USGS
Rivers and streams	USGS
Lakes and oceans	USGS and Saiz (2010)
Average January temperature	US climate data

Table A2.3. First Stage: Robustness Checks

VARIABLES	(1) Blacks	(2) Whites	(3) Blacks	(4) Whites	(5) Blacks	(6) Whites	(7) Blacks	(8) Whites
Z	1.711***	-0.193	1.732***	0.101	1.718***	-0.294	1.545**	-2.333
	(0.536)	(1.577)	(0.548)	(1.373)	(0.536)	(1.526)	(0.637)	(1.944)
Z_water	0.035**	0.360***	0.038**	0.369***	0.032*	0.334***	0.032*	0.400***
	(0.016)	(0.094)	(0.015)	(0.088)	(0.018)	(0.099)	(0.016)	(0.070)
Z_hills	-0.032	0.510***	-0.058	0.317	-0.044	0.479***	-0.039	0.527**
	(0.029)	(0.142)	(0.103)	(0.333)	(0.032)	(0.142)	(0.031)	(0.246)
Z_rivers	-0.299**	0.832	-0.302**	0.910	-0.206	1.408	-0.177	1.837*
	(0.140)	(0.865)	(0.129)	(0.843)	(0.150)	(1.022)	(0.194)	(1.090)
KP F-stat	12.47	12.47	24.84	24.84	14.59	14.59	5.112	5.112
AP F-stat	61.63	113.6	66.02	137.8	88.6	349.0	102.3	415.4
Drop 95pct	Water	Water	Hills	Hills	Rivers	Rivers	-	-
Radius	50	50	50	50	50	50	30	30
Cities	39	39	39	39	39	39	42	42
Observations	117	117	117	117	117	117	126	126

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. The dependent variable is the number of blacks (whites) in odd (even) columns. The regressors of interest are predicted black immigration (Z) and its interaction with the share of the area around the central city: i) occupied by lakes and oceans (*Z_water*); ii) with slope above 15% (*Z_hills*); iii) occupied by rivers and streams (*Z_rivers*). All regressions control for city and year by region fixed effects, and include interactions between year dummies and dummies for the presence of geographic features around the central city. Cols 1-2, 3-4, and 5-6 drop, respectively, cities with the share of lakes and oceans, land with slope above 15%, and rivers and stream above the 95^{th} percentile. Cols 7-8 replicate results in the main text (Cols 4-5, Table 3) defining each geographic variables using a 30 km radius. AP F-stat refers to the partial F-stat for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2.4. Spending on Education, Students Enrolled, and Literacy

VARIABLES	(1) Education spending over pop below 20	(2) Education spending per enrolled pupil	(3) Fr. population below 20 enrolled in school	(4) Fraction literate whites above 10
Panel A: OLS				
Blacks	-0.132**	-0.147**	-0.031	0.003
	(0.053)	(0.064)	(0.025)	(0.003)
Whites	0.018*	0.021*	0.005	-0.001
	(0.009)	(0.011)	(0.003)	(0.001)
Panel B: 2SLS				
Blacks	-0.148***	-0.172**	-0.032	0.001
	(0.051)	(0.068)	(0.024)	(0.004)
Whites	0.020**	0.024**	0.004	-0.000
	(0.008)	(0.011)	(0.004)	(0.001)
KP F-stat	11.44	11.44	11.93	11.93
AP (Blacks)	104.5	104.5	104.1	104.1
AP (Whites)	254.9	254.9	230.8	230.8
Mean of dep var	18.52	29.83	61.63	3.327
Observations	126	126	124	124

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. Panels A and B report, respectively OLS and 2SLS results. The dependent variable is: spending on education over population below 20 (over the number of students enrolled in public schools) in Col 1 (Col 2); the fraction of the population below the age of 20 enrolled in school in Col 3; and the fraction of whites above the age of 10 that cannot read or write in Col 4. The regressors of interest are the number of blacks (*Blacks*) and whites (*Whites*) in the central city, and are instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3). AP (blacks) and AP (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All columns report the mean of the dependent variable at baseline. All regressions include city and year by region fixed effects and interactions between year dummies and dummies for: i) 1900 fraction of blacks; ii) latitude and longitude; and iii) the presence of geographic features around the central city. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2.5. 1900-1910 Changes in European Immigration and in Economic Activity

	Dep. Variable: 1900-1910 Change in				
	(1) Fr. Immigrants	(2) Log value added per establishment	(3) Log value of products per establishment	(4) Log establishment size	(5) Log manufacturing wages
Blacks	-0.005	-0.893	-2.729	-0.543	-0.067
	(0.019)	(1.189)	(2.381)	(0.870)	(0.314)
Whites	0.001	-0.027	0.262	-0.059	-0.004
	(0.003)	(0.175)	(0.370)	(0.138)	(0.044)
KP F-stat	11.23	12.37	12.37	12.37	12.37
AP (Blacks)	63.94	69.96	69.96	69.96	69.96
AP (Whites)	207.7	257.4	257.4	257.4	257.4
Cities	41	41	41	41	41

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. Due to data availability, results in this Table do not include Washington DC. The dependent variable is the 1900-1910 change in outcome reported at the top of each column. Variables in Cols 2, 3, and 5 are expressed in 1910 dollars. The regressors of interest are the 1910-1930 change in black and white population in the central cities, and are instrumented with, respectively, the 1910-1930 change in predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3). AP (blacks) and AP (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions partial out trends for: region; 1900 fraction of immigrants; city coordinates; and geography. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2.6. Pre-Trends

VARIABLES	(1) Total Tax Revenues PC	(2) Property Tax Revenues PC	(3) Property Values PC	(4) Total Expenditures PC	(5) Expenditures on Education PC
Panel A: Baseline Specification					
Blacks	-0.111***	-0.110***	-5.796*	-0.041*	-0.041***
	(0.041)	(0.035)	(3.374)	(0.024)	(0.012)
Whites	0.012**	0.011**	0.766	0.003	0.005***
	(0.005)	(0.005)	(0.503)	(0.004)	(0.002)
Panel B: 1910-1930					
Blacks	-0.028	-0.029	-0.475	-0.001	-0.006
	(0.025)	(0.025)	(2.603)	(0.019)	(0.008)
Whites	0.008**	0.008**	0.088	0.001	0.001
	(0.003)	(0.003)	(0.311)	(0.003)	(0.001)
KP F-stat	11.01	11.01	11.01	11.01	11.01
AP (Blacks)	64.22	64.22	64.22	64.22	64.22
AP (Whites)	194.5	194.5	194.5	194.5	194.5
Cities	41	41	41	41	41

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. The outcome of interest, expressed in 1910 dollars, is displayed at the top of each column. Panel A reports results for the baseline specification (in the main text). In Panel B, the dependent variable is the 1906-1910 change in the variable at the top of each column, and the regressors of interest are the (instrumented) 1910-1930 change in black and white population, instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3 in the main text). AP (blacks) and AP (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions partial out trends for: region; 1900 fraction of immigrants; city coordinates; and geography. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2.7. Differential Trends and 1900 Characteristics

VARIABLES	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Property values over 1900 pop	(5) Total spending PC	(6) Spending on education PC
Panel A: Distance from the South						
Blacks	-0.113***	-0.109***	0.105	-16.28*	-0.039	-0.044***
	(0.042)	(0.036)	(0.100)	(9.507)	(0.024)	(0.015)
Whites	0.013**	0.011**	-0.022	2.074	0.003	0.006**
	(0.005)	(0.005)	(0.015)	(1.472)	(0.004)	(0.002)
KP F-stat	10.37	10.37	10.37	10.37	10.37	10.37
AP (Blacks)	104.0	104.0	104.0	104.0	104.0	104.0
AP (Whites)	255.3	255.3	255.3	255.3	255.3	255.3
Panel B: Fraction of immigrants						
Blacks	-0.115***	-0.115***	0.130*	-18.89**	-0.042*	-0.041***
	(0.043)	(0.037)	(0.075)	(8.828)	(0.024)	(0.012)
Whites	0.012**	0.010**	-0.016	1.980	0.002	0.005**
	(0.005)	(0.005)	(0.011)	(1.251)	(0.004)	(0.002)
KP F-stat	11.55	11.55	11.55	11.55	11.55	11.55
AP (Blacks)	69.10	69.10	69.10	69.10	69.10	69.10
AP (Whites)	142.1	142.1	142.1	142.1	142.1	142.1
Panel C: Skill Ratios						
Blacks	-0.120***	-0.115***	0.082	-19.55**	-0.049**	-0.042***
	(0.043)	(0.036)	(0.065)	(9.127)	(0.025)	(0.012)
Whites	0.012**	0.011**	-0.018*	1.973	0.002	0.005**
	(0.006)	(0.005)	(0.010)	(1.301)	(0.004)	(0.002)
KP F-stat	11.91	11.91	11.91	11.91	11.91	11.91
AP (Blacks)	57.07	57.07	57.07	57.07	57.07	57.07
AP (Whites)	76.01	76.01	76.01	76.01	76.01	76.01

Table A2.7 (Continued). Differential Trends and 1900 Characteristics

VARIABLES	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Property values over 1900 pop	(5) Total spending PC	(6) Spending on education PC
Panel D: Value added by manufacture						
Blacks	-0.103**	-0.105***	0.082	-17.65*	-0.042	-0.043***
	(0.043)	(0.036)	(0.074)	(9.102)	(0.026)	(0.013)
Whites	0.011**	0.010**	-0.014	1.880	0.003	0.005***
	(0.005)	(0.005)	(0.011)	(1.296)	(0.004)	(0.002)
KP F-stat	11.58	11.58	11.58	11.58	11.58	11.58
AP (Blacks)	135.5	135.5	135.5	135.5	135.5	135.5
AP (Whites)	310.5	310.5	310.5	310.5	310.5	310.5
Panel E: Employment share manufacture						
Blacks	-0.109***	-0.109***	0.096	-17.18**	-0.047*	-0.043***
	(0.041)	(0.034)	(0.071)	(8.605)	(0.026)	(0.013)
Whites	0.012**	0.011**	-0.015	1.879	0.004	0.005***
	(0.005)	(0.005)	(0.011)	(1.225)	(0.004)	(0.002)
KP F-stat	13.34	13.34	13.34	13.34	13.34	13.34
AP (Blacks)	96.48	96.48	96.48	96.48	96.48	96.48
AP (Whites)	267.5	267.5	267.5	267.5	267.5	267.5
Observations	126	126	126	126	126	126

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. This Table replicates results in the main text interacting year dummies with distance from the South (rather than with latitude and longitude) in Panel A, and including interactions between year dummies and, respectively: the 1900 fraction of European immigrants (Panel B); the 1900 ratio of high to low skilled workers (Panel C); the 1904 value added by manufacture per establishment (Panel D); and the 1900 employment share in manufacturing (Panel E). The outcome of interest, expressed in 1910 dollars, is displayed at the top of each column. The regressors of interest are the number of blacks (*Blacks*) and whites (*Whites*) in the central city, and are instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3). AP (blacks) and AP (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and year by region fixed effects and interactions between year dummies and dummies for: i) 1900 fraction of blacks; ii) latitude and longitude (except for Panel A, where this is replaced with distance from the South); and iii) the presence of geographic features around the central city. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2.8. Trimming the Sample

VARIABLES	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Property values over 1900 pop	(5) Total spending PC	(6) Spending on education PC
Panel A: 1st-99th Percentiles						
Blacks	-0.076*	-0.077**	-0.019	-17.95	-0.065***	-0.046***
	(0.040)	(0.034)	(0.160)	(13.23)	(0.025)	(0.018)
Whites	-0.000	0.000	0.000	3.072	0.012**	0.009**
	(0.010)	(0.009)	(0.040)	(3.266)	(0.006)	(0.004)
KP F-stat	10.39	10.39	10.39	10.39	10.39	10.39
AP (Blacks)	20.20	20.20	20.20	20.20	20.20	20.20
AP (Whites)	16.50	16.50	16.50	16.50	16.50	16.50
Observations	120	120	120	120	120	120
Panel B: 5ht-95th Percentiles						
Blacks	-0.097***	-0.098***	-0.004	-30.09**	-0.088***	-0.054***
	(0.038)	(0.032)	(0.214)	(14.59)	(0.023)	(0.020)
Whites	0.004	0.004	0.000	5.645*	0.017***	0.011**
	(0.008)	(0.007)	(0.048)	(3.166)	(0.005)	(0.004)
KP F-stat	8.011	8.011	8.011	8.011	8.011	8.011
AP (Blacks)	35.28	35.28	35.28	35.28	35.28	35.28
AP (Whites)	20.18	20.18	20.18	20.18	20.18	20.18
Observations	108	108	108	108	108	108

Note: The sample includes, for Census years 1910, 1920, and 1930, the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration. This Table replicates results in the main text trimming the sample at the 1st and 99th (Panel A) and at the 5th and 95th (Panel B) percentiles of black population. In Panel A, the excluded cities are New York and Duluth, while in Panel B they are New York, Philadelphia, Washington DC, Duluth, San Diego, and Scranton. The outcome of interest, expressed in 1910 dollars, is displayed at the top of each column. The regressors of interest are the number of blacks (*Blacks*) and whites (*Whites*) in the central city, and are instrumented with, respectively, predicted black immigration and its interaction with the area around the city that: i) had slope above 15%; ii) is occupied by lakes and oceans; iii) is occupied by rivers and stream (see Sections 5.2 and 5.3). AP (blacks) and AP (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and year by region fixed effects and interactions between year dummies and dummies for the presence of geographic features around the central city. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2.9. Alternative Instrument for the White Flight

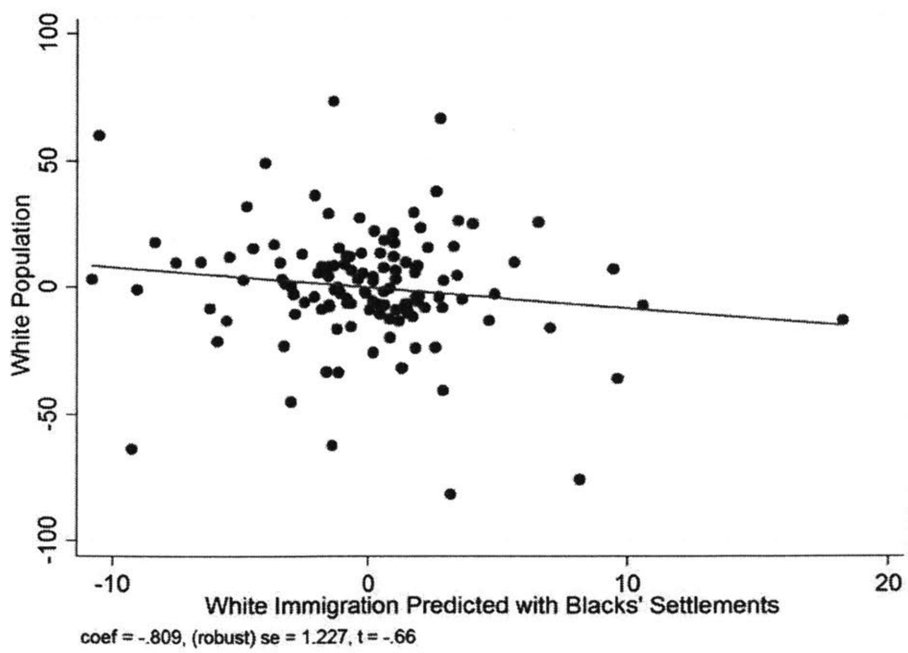
VARIABLES	(1) Total tax revenues PC	(2) Property tax revenues PC	(3) Property tax rate	(4) Property values over 1900 pop	(5) Total spending PC	(6) Spending on education PC
Blacks	-0.141***	-0.127***	0.015	-28.40**	-0.060*	-0.047**
	(0.047)	(0.042)	(0.237)	(11.63)	(0.036)	(0.018)
Whites	0.021***	0.017***	-0.016	4.605**	0.008	0.007**
	(0.007)	(0.006)	(0.032)	(2.023)	(0.006)	(0.003)
KP F-stat	10.69	10.69	10.69	10.69	10.69	10.69
AP (Blacks)	41.39	41.39	41.39	41.39	41.39	41.39
AP (Whites)	14.06	14.06	14.06	14.06	14.06	14.06
Mean of dep var	15.63	15.16	23.12	1,308	15.21	4.656
Observations	126	126	126	126	126	126

Note: This Table replicates the main results in Tables 4 and 5 instrumenting the number of whites in the central city with the interaction between predicted black immigration and average low January temperature. The dependent variable, expressed in 1910 dollars, is listed at the top of each column. AP (blacks) and AP (whites) refer to the partial F-stats for joint significance of the instruments in the two separate first-stage regressions. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All columns also report the mean of the dependent variable at baseline. All regressions include city and year by region fixed effects, and interactions between year dummies and dummies for: i) 1900 fraction of blacks; and ii) latitude and longitude. Robust standard errors, clustered at the MSA level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Figure A2.1. Map of Cities

Note: the map plots the 42 largest central cities that were anchored to a MSA and had at least 100,000 residents in 1930, and for which it was possible to construct the instrument for black migration.

Figure A2.2. White Immigration and 1900 Blacks' Settlements



Note: The y-axis reports the actual number of whites received by Northern cities in each decade between 1910 and 1930, and the x-axis shows the change in the predicted number of southern white migrants, constructed multiplying 1900 blacks' settlements with southern born white outmigration from each southern state. Each point in the scatter diagram represents the residual change in a city's actual and predicted number of whites after partialling out MSA population; city and year by region fixed effects; interactions between year dummies and 1900 fraction of blacks, city coordinates, and geography.

A3. Supplementary Tables and Figures to Chapter 3

Table A3.1 European Regions

UK	Russia
Ireland	Eastern Europe (Yugoslavia, Czechoslovakia, etc.)
Denmark	Austria-Hungary
Finland	Switzerland
Norway	France
Sweden	Belgium-Netherlands
Germany	Greece-Portugal-Spain
Poland	Italy

Note: this table lists the European sending regions used to construct the instrument for immigration.

Table A3.2 City List

Akron, OH	Elizabeth, NJ	McKeesport, PA	Saint Joseph, MO
Albany, NY	Elmira, NY	Memphis, TN	Saint Louis, MO
Allentown, PA	Erie, PA	Milwaukee, WI	Saint Paul, MN
Altoona, PA	Evansville, IN	Minneapolis, MN	Salem, MA
Amsterdam, NY	Everett, MA	Mobile, AL	San Antonio, TX
Atlanta, GA	Fall River, MA	Montgomery, AL	San Diego, CA
Atlantic City, NJ	Fitchburg, MA	Mount Vernon, NY	San Francisco, CA
Auburn, NY	Flint, MI	Nashville, TN	Savannah, GA
Augusta, GA	Fort Wayne, IN	New Bedford, MA	Schenectedy, NY
Baltimore, MD	Fort Worth, TX	New Britain, CT	Scranton, PA
Bay City, MI	Galveston, TX	New Castle, PA	Seattle, WA
Bayonne, NJ	Grand Rapids, MI	New Haven, CT	Sioux City, IA
Berkeley, CA	Hamilton, OH	New Orleans, LA	Somerville, MA
Binghamton, NY	Harrisburg, PA	New York, NY	South Bend, IN
Birmingham, AL	Hartford, CT	Newark, NJ	Spokane, WA
Boston, MA	Haverhill, MA	Newton, MA	Springfield, IL
Bridgeport, CT	Hoboken, NJ	Niagara Falls, NY	Springfield, MA
Brockton, MA	Holyoke, MA	Norfolk, VA	Springfield, MO
Buffalo, NY	Houston, TX	Oakland, CA	Springfield, OH
Butte, MT	Huntington, WV	Oklahoma City, OK	Superior, WI
Cambridge, MA	Indianapolis, IN	Omaha, NE	Syracuse, NY
Camden, NJ	Jackson, MI	Oshkosh, WI	Tacoma, WA
Canton, OH	Jacksonville, FL	Pasadena, CA	Tampa, FL
Cedar Rapids, IA	Jamestown , NY	Passaic, NJ	Taunton, MA
Charleston, SC	Jersey City, NJ	Paterson, NJ	Terre Haute, IN
Charlotte, NC	Johnstown, PA	Pawtucket, RI	Toledo, OH
Chattanooga, TN	Joliet, IL	Peoria, IL	Topeka, KS
Chelsea, MA	Kalamazoo, MI	Perth Amboy, NJ	Trenton, NJ
Chester, PA	Kansas City, KS	Philadelphia, PA	Troy, NY
Chicago, IL	Kansas City, MO	Pittsburgh, PA	Utica, NY
Cincinnati, OH	Knoxville, TN	Pittsfield, MA	Washington, DC
Cleveland, OH	La Crosse, WI	Portland, ME	Waterbury, CT
Columbus, OH	Lancaster, PA	Portland, OR	Wheeling, WV
Covington, KY	Lansing, MI	Portsmouth, VA	Wichita, KS
Dallas, TX	Lawrence, MA	Providence, RI	Wilkes-Barre, PA
Davenport, IA	Lexington, KY	Pueblo, CO	Williamsport, PA
Dayton, OH	Lima, OH	Quincy, IL	Wilmington, DE
Decatur, IL	Lincoln, NE	Quincy, MA	Woonsocket, RI
Denver, CO	Little Rock, AR	Racine, WI	Worcester, MA
Des Moines, IA	Los Angeles, CA	Reading, PA	Yonkers, NY
Detroit, MI	Louisville, KY	Richmond, VA	York, PA
Dubuque, IA	Lowell, MA	Roanoke, VA	Youngstown, OH
Duluth, MN	Lynn, MA	Rochester, NY	
East Orange, NJ	Macon, GA	Rockford, IL	
East St. Louis, IL	Malden, MA	Sacramento, CA	
El Paso, TX	Manchester, NH	Saginaw, MI	

Table A3.3 Immigration and Marriage of Native Men Aged 20-35

	(1)	(2)	(3)	(4)
	All Sample		**Restricted Sample**	
	OLS	2SLS	OLS	2SLS
Fr. Immigrant	-0.006	0.190***	0.077	0.147**
	(0.135)	(0.054)	(0.082)	(0.063)
F-stat		251.3		251.3
Mean dep. var. in 1910	0.42	0.42	0.43	0.43
Obs.	538	538	529	529

Note: this Table presents results of OLS and 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report in columns 1 and 2. In columns 3 and 4, we exclude three cities (Duluth, Superior, and Tacoma) with an extraordinary low level of marriage rates of men aged 20-35 in 1910. The dependent variable is the fraction of men married in the age range 20-35. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table A3.4 Immigration and Natives' Marriage (2SLS Results)

	(1)	(2)	(3)	(4)	(5)
Panel A: Women					
		Marriage rates			**Never Married**
Age Groups	**18-33**	**18-25**	**26-33**	**34-65**	**34-65**
Fr. Immigrant	0.209***	0.229***	0.053	0.025	-0.082***
	(0.044)	(0.038)	(0.054)	(0.035)	(0.019)
Mean dep. var.	0.47	0.34	0.65	0.63	0.15
Obs.	538	538	538	538	538
Panel B: Men					
		Marriage rates			**Never Married**
Age Groups	**20-35**	**20-27**	**28-35**	**36-65**	**36-65**
Fr. Immigrant	0.190***	0.236***	-0.001	0.011	-0.026
	(0.054)	(0.055)	(0.059)	(0.045)	(0.035)
Mean dep. var.	0.45	0.30	0.65	0.73	0.14
Obs.	538	538	538	538	538

Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is the fraction of women married in the different age range in Panel A and the fraction of men married in the different age range in Panel B. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table A3.5 Immigration and Fertility of Native Women (2SLS Results)

	Dep. Variable: Fertility of Native Women					
	Children to Women Ratio		Mothers to Women Ratio		Children to Mothers Ratio	
	Age 18-25	Age 26-33	Age 18-25	Age 26-33	Age 18-25	Age 26-33
	(1)	(2)	(3)	(4)	(5)	(6)
Fr. Immigrant	0.234***	0.357***	0.131***	0.071**	0.160**	0.391***
	(0.049)	(0.122)	(0.025)	(0.034)	(0.080)	(0.137)
F-stat	251.3	251.3	251.3	251.3	251.3	251.3
Mean Dep.Var.	0.330	1.050	0.210	0.500	1.530	2.090
Obs.	538	538	538	538	538	538

Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is: in column 1 and 2, the total number of children with native mother over the total number of women in the age range, in column 3 and 4 the fraction of women who have children and in column 5 and 6 the average number of children per mother. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table A3.6 Native Men (20-35) in Selected Occupations

	High Immigrants' Competition			Low Immigrants' Competition		
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction Natives:	Manuf. Laborers	Bakers	Blacksmiths	Manuf. Foremen	Engineers	Electricians
Panel A: OLS						
Fr. Immigrants	-0.079	-0.008*	-0.009	0.025***	0.017	0.007
	(0.058)	(0.004)	(0.005)	(0.004)	(0.014)	(0.006)
Panel B: 2SLS						
Fr. Immigrants	-0.117**	-0.008**	-0.005	0.030***	0.033***	0.010
	(0.050)	(0.004)	(0.005)	(0.006)	(0.011)	(0.007)
F-stat	251.3	251.3	251.3	251.3	251.3	251.3
Mean dep var	0.038	0.005	0.007	0.006	0.018	0.014
Observations	538	538	538	538	538	538

Note: this table presents results for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930 (see Table A? in the appendix). The dependent variable is the fraction of native males in age range (20-35) working in the occupation reported at the top of each column. Panels A and B report, respectively, OLS and 2SLS results. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. The mean of each dependent variable at baseline is shown at the bottom of the Table. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A3.7 Additional Results for Economic Activity

VARIABLES	(1) Log value added per capita	(2) Log value of products per establishment	(3) Log value of products per capita	(4) Log horsepower	(5) TFP
Panel A: OLS					
Fr. Immigrants	0.785	2.264***	0.992*	1.267***	0.295
	(0.580)	(0.704)	(0.556)	(0.475)	(0.358)
Panel B: 2SLS					
Fr. Immigrants	1.404**	3.549***	2.065**	1.906***	1.013*
	(0.586)	(1.214)	(0.845)	(0.705)	(0.540)
F-stat	270.5	270.5	270.5	270.5	270.5
Cities	178	178	178	178	178
Observations	525	525	525	525	525

Note: this Table presents results for a balanced panel of the 178 US cities with at least 30,000 residents in each Census year 1910, 1920, and 1930, and for which data were reported in the Census of Manufacture between 1909 and 1929. Panels A and B report, respectively, OLS and 2SLS results. The dependent variable is: the log of value added per capita in Col 1; the log of value of products per establishment (per capita) in Col 2 (Col 3); the log of horsepower in Col 4; and total factor productivity (TFP) in Col 5. *Fr. Immigrants* is the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 4.2 (see (2) in the main text). F-stat refers to the K-P F-stat for weak instrument. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *** p<0.01; ** p<0.05; * p<0.1.

Table A3.8 Immigration and Natives' Living Choices

	(1)	(2)	(3)	(4)
Dep. Var.	**Share of children < 10 (native parents)**		**Share of families (children < 10, native parents)**	
	Father employed	**Father unskilled**	**Father employed**	**Father unskilled**
Panel A: OLS				
Fr. Immigrant	0.052	-0.027	0.032	-0.035
	(0.049)	(0.075)	(0.045)	(0.076)
Panel B: 2SLS				
Fr. Immigrant	0.049	-0.138**	0.024	-0.171***
	(0.037)	(0.061)	(0.034)	(0.063)
F-stat	251.3	251.3	251.3	251.3
Mean dep. var.	0.908	0.332	0.901	0.318
Obs.	538	538	538	538

Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is reported at the top of the Table Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table A3.9 Immigration and Education of Native Children

	Dep. Var.: Fraction attending school					
	Sons of natives			Daughters of natives		
	(1)	(2)	(3)	(4)	(5)	(6)
Age group:	**Age 6-14**	**Age 15-18**	**Age 19-24**	**Age 6-14**	**Age 15-18**	**Age 19-24**
Panel A: OLS						
Fr. Immigrant	0.007	-0.081	0.010	-0.025	0.059	-0.006
	(0.040)	(0.080)	(0.019)	(0.041)	(0.076)	(0.027)
Panel B: 2SLS						
Fr. Immigrant	0.067***	-0.100**	0.011	0.017	0.044	-0.042*
	(0.025)	(0.049)	(0.017)	(0.027)	(0.059)	(0.023)
Dep. var:	.933	.241	.015	.936	.22	.013
Obs.	538	538	538	538	538	538

Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is reported at the top of the Table Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table A3.10 Immigration and Marriage Rates of Native Men aged 20-35 (2SLS Results)

	(1)	(2)	(3)	(4)	(5)	(6)
	All Native Men			Second Generation Men		
Wife	**All**	**Native Parentage**	**Immigrant**	**All**	**Native Parentage**	**Immigrant**
Dependent Variable: Marriage rates						
Fr. Immigrant	0.190***	0.296***	0.022	0.106*	0.215***	0.063*
	(0.054)	(0.064)	(0.016)	(0.063)	(0.075)	(0.036)
Mean dep. var.	0.45	0.27	0.03	0.42	0.16	0.04
F-stat	251.3	251.3	251.3	251.3	251.3	251.3
Obs.	538	538	538	538	538	538

Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. In panel A, the dependent variable is the marriage rates of men aged 20-35 by husband parentage. In panel B, the dependent variable is the children to women ratio by father parentage. We consider only children of women aged 18-33. For example, in column 2 of Panel B, the dependent variable is the number of children with native mother aged 18-33 and father with a native parentage over the number of native women aged 18-33. Columns 4-6 focus on women who are second generation immigrants. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table A3.11 Immigration and Marriage Rates of Natives by Parentage (2SLS Results)

	Dep. Variable: Marriage rates					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A:	**Women Age 18-25**			**Women Age 26-33**		
Own Parents	**Native**	**Mixed**	**Foreign**	**Native**	**Mixed**	**Foreign**
Fr. Immigrant	0.127***	0.192***	0.211***	0.068	0.163	0.117
	(0.049)	(0.055)	(0.061)	(0.042)	(0.113)	(0.112)
Mean Dep. Var.	0.340	0.257	0.277	0.642	0.603	0.587
Obs.	538	538	538	538	538	538
Panel B:	**Men Age 20-27**			**Men Age 28-35**		
Own Parents	**Native**	**Mixed**	**Foreign**	**Native**	**Mixed**	**Foreign**
Fr. Immigrant	0.139**	0.227***	0.035	-0.029	0.013	0.047
	(0.059)	(0.058)	(0.067)	(0.075)	(0.091)	(0.078)
Mean Dep. Var.	0.297	0.210	0.233	0.623	0.575	0.561
Obs.	538	538	538	538	538	538

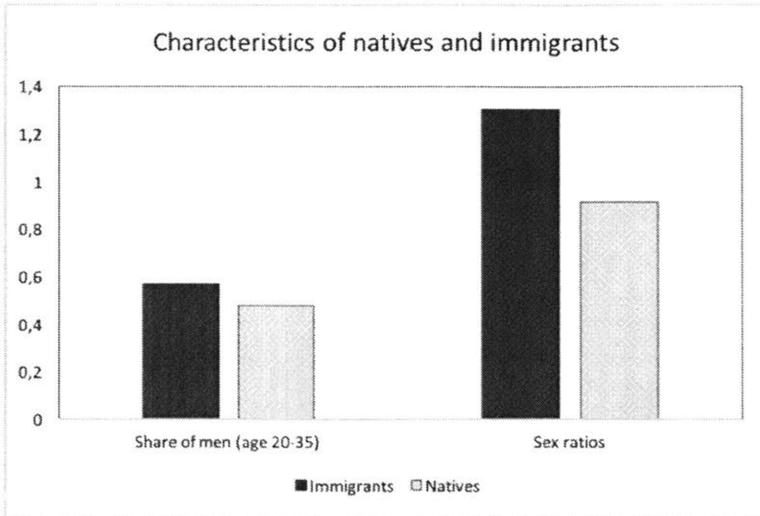
Note: this Table presents results of 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is the marriage rate of the groups described in each panel. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Table A3.12 Immigration and LFP of Native Women (2SLS Results)

	LFP of Native Women				
	(1)	(2)	(3)	(4)	(5)
	Age 18-33	Age 18-33	Age 18-25	Age 26-33	Age 34-65
	OLS	2SLS	2SLS	2SLS	2SLS
Fr. Immigrant	-0.084	-0.115*	-0.135**	0.023	0.042
	(0.083)	(0.061)	(0.067)	(0.061)	(0.041)
F-stat		251.3	251.3	251.3	251.3
Mean dep. var.	.42	.42	.49	.33	.25
Obs.	538	538	538	538	538

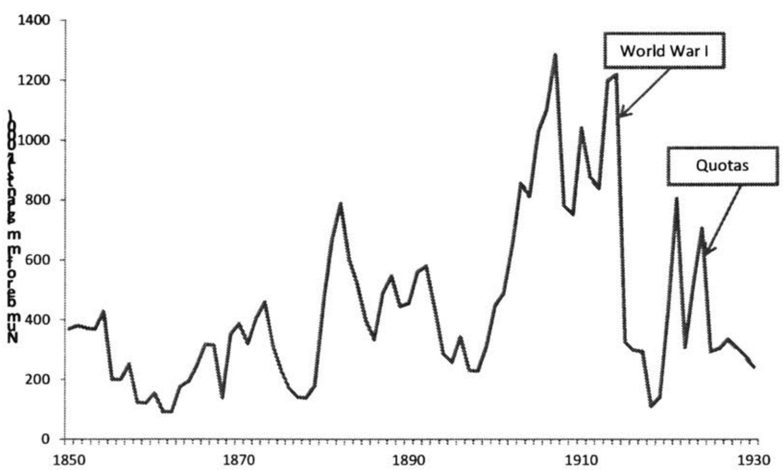
Note: this Table presents results of OLS and 2SLS for a balanced panel of the 180 US cities with at least 30,000 residents in each Census year report. The dependent variable is the labor force participation of women in the different age range. Fr. Immigrants refers to the fraction of immigrants arrived in the previous decade over predicted city population, and is instrumented using the baseline version of the instrument constructed in Section 3.4. The mean of dependent variables is shown at the bottom of the Table. KP F-stat is the Kleibergen-Paap F stat for joint significance of instruments. All regressions include city and state by year fixed effects. Robust standard errors, clustered at the MSA level, in parenthesis. *p<0:10, **p<0:05, ***p<0:01.

Figure A3.1 Share of Men and Sex Ratios (1910): Immigrants and Natives



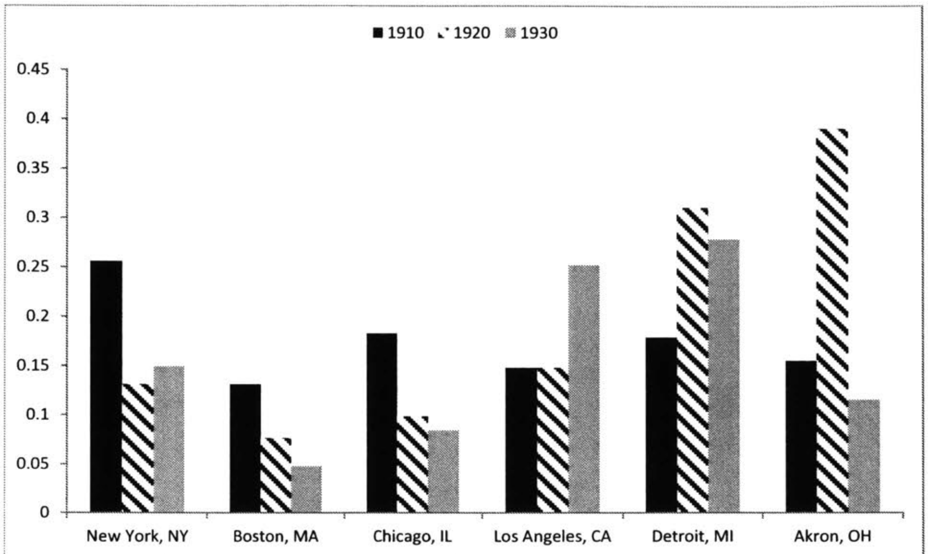
Note: Sex ratios are defined as the number of native men (resp. immigrant men) in the age group 20-35 over the number of native women in the age group 18-33 (resp. immigrant women). Source: Authors' calculations from IPUMS sample of 1910 US Census (Ruggles et al., 2015).

Figure A3.2 Total Number of Immigrants (in Thousands)



Note: Annual inflow of immigrants to the United States (1850-1930). Source: Migration Policy Institute.

Figure A3.3 Recent Immigrants Over 1900 City Population, by Decade

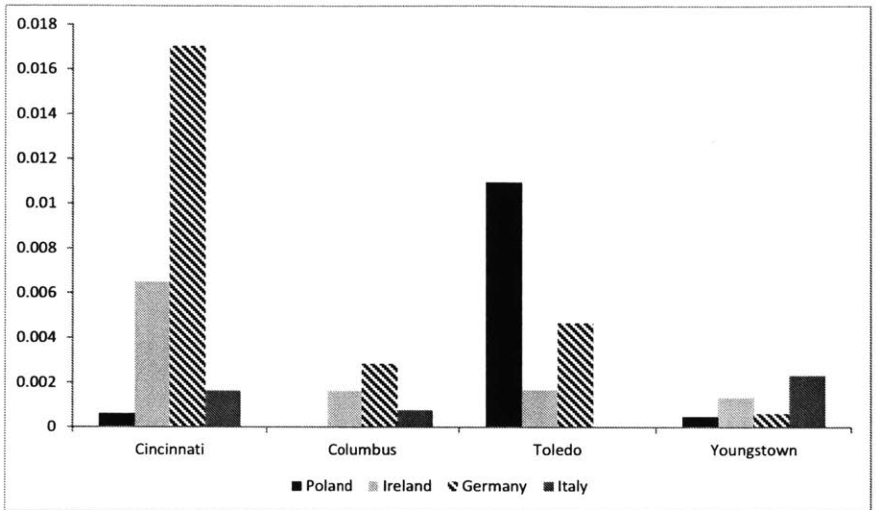


Note: Number of European immigrants that arrived in the United States in the last decade over 1900 city population, for selected cities and by decade. Source: Author's calculations from IPUMS sample of US Census (Ruggles et al., 2015).

Figure A3.4 Map of Cities

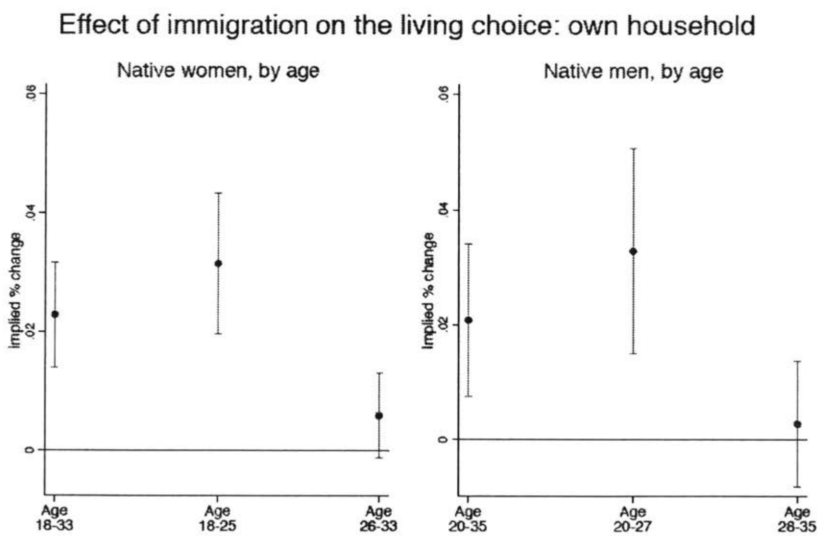
Note: The map plots the 180 cities with at least 30,000 residents in each of the three Census years 1910, 1920, and 1930.

Figure A3.5 Share of European Immigrants in Ohio, 1900



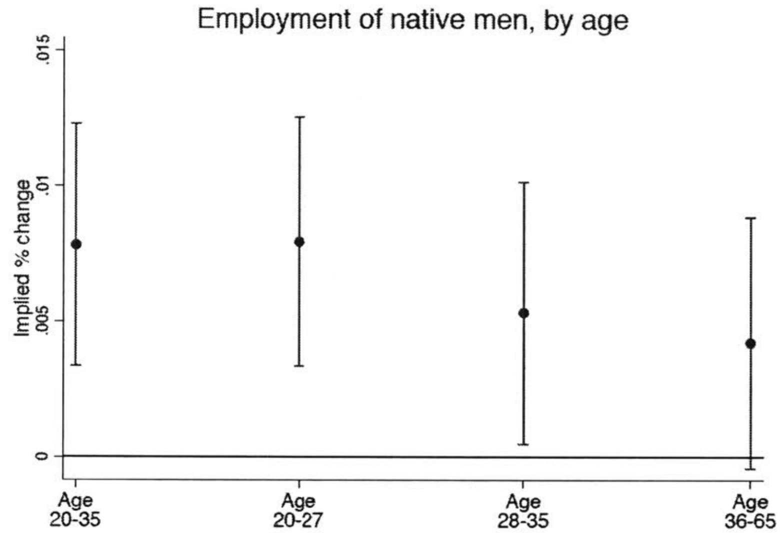
Note: share of individuals of European ancestry living in selected cities of Ohio in 1900, for selected ethnic groups. Source: Author's calculations using IPUMS data.

Figure A3.6 Impact of Immigration on the Creation of an Independent Family Unit



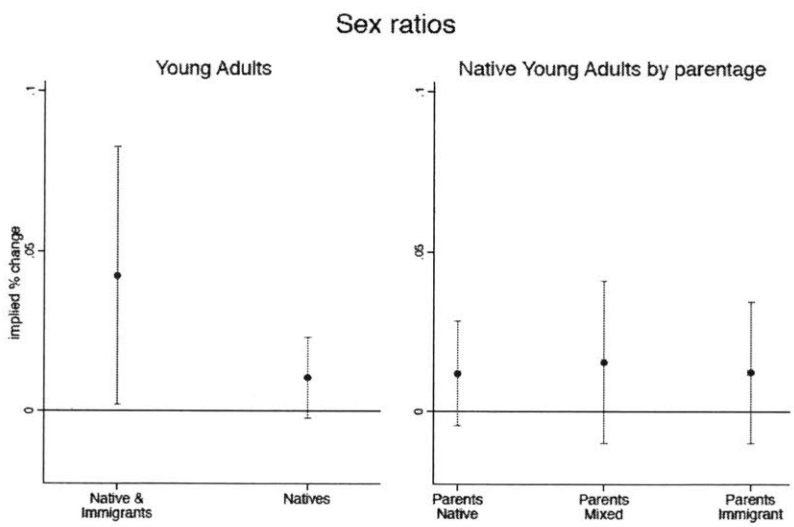
Note: This graph shows the impact of one standard deviation increase of the fraction of immigrants on the probability of being the household head or spouse by age and gender. 95% confidence intervals are reported.

Figure A3.7 Impact of Immigration on Natives' Employment (by Age)



Note: This graph shows the impact of one standard deviation increase of the fraction of immigrants on the employment to population ratio of native men, relative to the mean value in 1910.

Figure A3.8 Impact of Immigration on Sex Ratios



Note: this graph shows the impact of one standard deviation increase of the fraction of immigrants on the sex ratios or young adults, i.e. the number of men in the age group 20-35 over the number of women in the age group 18-33. The first bar shows the impact for the whole population (natives+ immigrants) living in the 180 US cities with at least 30,000 residents in each Census year. The following bars present the sex ratios for natives, divided by parentage.

Acknowledgements

I thank my advisors, Daron Acemoglu, Alberto Alesina, David Autor, and Heidi Williams for their constant and invaluable support throughout the years spent at MIT. I also thank Maristella Botticini, Nicola Gennaioli, and Francesco Giavazzi who constantly supported me from Italy, as well as Leah Boustan and Nico Voigtlaender who helped me from Princeton and UCLA both academically and personally.

Special thanks go to my family – to my parents and to my sister Anna in particular – and to Francesca. I simply would have not been able to accomplish this without them. I also thank my friends from "Mar Salada", Alessandro and Clara Cairo, Nicola Gaggero, Nicola Fontana, and Umberto Sagliaschi, my roommates – whom I consider "brothers and sisters" in Boston, Martina Uccioli and Enrico Innocenti – and my "little Italy" in the US, Giulia Brancaccio, Enrico Cantoni, Tommaso Denti, Andrea Di Miceli, Ludovica Gazzè, Giovanni Reggiani, and Irena Stojkov.

Finally, I am truly indebted to many fantastic people I met here in Cambridge: Andres and Malu, Juan, and, of course, all of my classmates.