Immigration and Invention:

Evidence from the Quota Acts

Kirk Doran* Chungeun Yoon†

2022

Abstract

How does mass immigration affect invention? Mass immigration may increase invention by making inventions both easier to produce and more profitable (Marshall, 1890), but it may decrease invention because more plentiful labor makes labor-saving inventions less useful (Hicks, 1932). We examine the end of America's period of mass immigration in the 1920s to explore this question. We find that the decrease in mass immigration caused by the Quota Acts decreased the rate of invention. This positive relationship between mass immigration and invention can be explained by immigration's effect on the labor supply and the potential scale of production, as well as externalities on existing innovators. Mass immigration was an essential input to the era of great American invention; a low-skilled labor force whose scale was unprecedented in human history helped shape the content of America's inventive portfolio.

JEL: J24, N32, O31

Keywords: Immigration, Invention, Economies of Scale

*University of Notre Dame

†KDI School of Public Policy and Management

We thank Pierre Azoulay, George Borjas, Brian Cadena, Prithwiraj Choudhury, Bill Evans, Ina Ganguli, Britta Glennon, Walker Hanlon, Exequiel Hernandez, Daniel Hungerman, Ben Jones, Joseph Kaboski, Shu Kahn, Bill Kerr, Ethan Lieber, Francesco Lissoni, Megan MacGarvie, Petra Moser, Gabriele Pellegrino, Sergio Petralia, Marco Tabellini, Reinhilde Veugelers, Fabian Waldinger, Bruce Weinberg, Heidi Williams, Abigail Wozniak, and seminar participants at the NBER 2019 Summer Institute and GEOINNO2018. We thank Grace Enright and Cody Kankel for excellent research assistance. All errors are our own.

Mass immigration is conceptually distinct from other types of immigration, and raises policy questions that are particularly relevant today (Abramitzky et al., 2014; Abramitzky and Boustan, 2017; Hatton and Williamson, 1998; Stuart and Taylor, 2021; Tabellini, 2020). Some economists and policy makers have called for another period of mass immigration to North America like the one that ended one hundred years ago, with the hope that it would produce large and immediate gains to human welfare (Ager and Brückner, 2013; Azoulay et al., 2022; Sequeira et al., 2020). However, the long-run effects will depend on how mass immigration affects innovation, a question that economists have only just begun to address.

On the one hand, mass immigration may increase innovation through a combination of effects that are both internal and external to firms. Larger populations allow for larger manufacturing plants and enterprises, making innovation less risky and more profitable. Larger populations also increase the potential for division of labor (Marshall, 1895) and have a host of well-known external effects, all of which tend to increase the profitability of inventions while decreasing their costs. On the other hand, since Hicks (1932), economists have noted that when labor is more plentiful, labor-saving inventions become less useful; this could cause mass immigration to change the direction of inventive activity and possibly decrease its rate.

We use data from the 1920s, which marked the end of America's period of mass immigration, to explore this question. Before 1921, the United States had few restrictions on immigration from European countries. After 1921, the United States imposed quotas that almost eliminated immigration from Southern and Eastern Europe, while allowing immigration from Northern and Western Europe to continue. Given history dependence in terms of which counties immigrants from specific source countries tended to move to, and which industries immigrants from specific source countries tended to work in, this immigration restriction produced, as a side effect, significant variation in the flow of immigrants into both counties and industries.

Studies that rely on variation in immigration often face endogeneity challenges because many shocks to immigration were intended (by policy makers, the immigrants themselves, or those demanding their labor) to change immigration to certain regions more than others (Cadena and Kovak, 2016). Crucially, in the case of the quota laws, the variation in immigration that they induced across counties and industries was an unintended side effect. The aim of the quotas was to reduce immigration from certain source countries and not others; as a side effect, they ended up reducing immigration to some cells (counties and industries) and not others.

It is plausible that this reduction in immigration affected innovation through several channels, including labor supply, consumer demand, and economies of scale both internal and external to firms. The first question we ask is: what is the overall effect of mass immigration on invention, through all channels? The best way to address this question is to examine immigration rates and invention rates at the county-year level. The quotas induced sudden drops in immigration to counties that had traditionally attracted immigrants from Southern and Eastern Europe, while similar counties that had traditionally attracted immigrants from Northern and Western Europe lost many fewer immigrants. By comparing invention rates in these counties over time, we can learn how mass immigration affected invention. We find that a 10% decrease in immigration to a county caused a 0.7% decrease in patent applications from inventors located in that county. Incumbent inventors located in these counties at the time of the shock were particularly affected, experiencing a 0.6% decrease in patent applications as a result.[1]

What mechanisms were primarily responsible for this effect? It is clear that the inventions of immigrants themselves could, at best, play a limited role. The quotas were designed to affect low-skilled immigration. Professors, members of "learned professions," and immigrants intending to become students were specifically exempted from the quotas (Parker, 1924), so it is ex ante unlikely that the effects at the county level were driven by the inventions of immigrants themselves. Furthermore, we verify that the immigration of individual Europeans whom we identify as having already filed patents before the quotas did not differentially change in quota-exposed counties versus other counties after the quotas.

This leaves mechanisms relating to immigrants' effects on the supply of labor (and various economies of scale on the supply side) and consumer demand (and various economies of scale on the demand side). Immigrants' effect on the supply side is particularly intriguing. On the

[1]Indeed, the magnitudes of the effects imply that incumbent inventors account for all of the decline in patents at the county level.

one hand, increasing the labor supply could disincentivize inventions that economize labor, as Hicks (1932) pointed out. The famous Habakkuk hypothesis (Habakkuk, 1962) applied this argument to the first Industrial Revolution, positing that relatively scarce labor in early nineteenth century America incentivized invention. On the other hand, these classic and intuitive arguments are incomplete.[2]

We can determine the effect of the loss of the immigrant labor supply on invention at the industry-year level. Some industries, such as apparel, steel works, and coal mining, relied on workers from Southern and Eastern Europe, while other industries, such as electrical machinery and medicines, relied on workers from Northern and Western Europe. Since the immigrants who worked in an industry were unlikely to disproportionately purchase the products of that industry, variation in the labor shock across industries is explained by labor supply, not derived labor demand.

We find that a 10% decrease in immigrant workers in a given industry caused a 1.2% decrease in patent applications associated with that industry. Furthermore, this effect was concentrated in more labor intensive industries, i.e., those with a large fraction of low-skilled workers or low labor costs. In other words, the decline in the labor supply itself was an important channel for the decrease in invention.

Thus, the intuitive argument of Hicks/Habakkuk (that a decline in labor supply encourages invention) is overturned here. Because the inventions characteristic of the era were all "labor-saving," in the sense that they were designed to provide more value for less labor, it can be difficult to imagine how this could be. A specific example can help shed light. Consider the dual clusters of inventions of the automated assembly line and mass-producible automobiles. These inventions were characteristic of the second industrial revolution, in that they used electric-powered machinery and interchangeable parts (the so-called "American system of manufacturing") to provide a new product that required very few hours of labor per unit of output. In a *casual* sense, therefore, these were labor-saving inventions, as were most of the famous inventions of the second industrial revolution in America. But the usefulness of these inventions was not unrelated to scale. The new product and method

[2]For example, the theoretical results in Acemoglu (2010) show that, contrary to Hicks and Habakkuk, plentiful labor supply encourages invention whenever new technology increases the marginal product of labor, and indeed, this is how technology is conceptualized in all canonical macroeconomic models.

of production made Henry Ford's automobile factory, by necessity, the largest production facility in the world. In his factory, 3,000 parts were combined in a total of 7,882 tasks. Given this number of unique tasks, in order to take full advantage of the division of labor, the new assembly line required 14,000 local employees.[3] Thus, it is possible that the inventions characteristic of America's second industrial revolution were only worth producing in the context of the plentiful local labor supply. The era of mass migration may have provided necessary fuel for the era of great American invention.[4]

The remainder of this paper is organized as follows. Section I describes the Immigration Acts and relevant historical background. Section II introduces the data set. In Section III, we describe our empirical strategy. We then present the effects of immigration restrictions at the county level in Section IV. Section V investigates how the end of mass immigration affected industries. Section VI explores the impact on individuals and possible mechanisms. Section VII concludes.

I Historical Context

A challenge in identifying the causal effects of immigration on the economy is that shocks to immigration are endogenous: immigration shocks are typically intended (by policy makers, the immigrants themselves, or those demanding their labor) to change immigration to certain regions more than others (Cadena and Kovak, 2016). In this section, we illustrate that the quotas were not intended to end immigration *to* some regions but not others; rather, they were intended to end immigration *from* some sources but not others. We also address the similarities and differences between the variation we exploit here and that used in the existing literature on the quotas, as well as the overall historical context.

Between 1850 and 1920, over 30 million Europeans migrated to the United States (Abramitzky, Boustan, and Eriksson, 2014). As Figure 1a shows, at its peak, the annual inflow was over 1.5% of the pre-existing U.S. population. This migration was unprecedented in its size,

[3] Furthermore, the work was so repetitive (and thus turnover so rampant), that the actual number of employees required in a year was considerably higher than 14,000. See: Beniger (1997); Meyer (1981)

[4] Indeed, this conclusion is consistent with the literature that relates the era of mass migration to changes in manufacturing and productivity during the second industrial revolution. Immigration during this era may have encouraged mass production (Hirschman and Mogford, 2009), complemented assembly-line machinery (Lafortune, Tessada, and Lewis, 2018), and allowed for larger, more productive firms (Kim, 2007).

and numerous economists and historians have analyzed its correlates and circumstances. As Figure 1b shows, Southern and Eastern Europeans comprised an increasing portion of the immigrants as the century progressed (Abramitzky and Boustan, 2017; Spitzer and Zimran, 2018).

American concerns about the effects of immigration grew in proportion to the increased prevalence of Southern and Eastern European immigrants seen in Figure 1b. Figure 1a demonstrates that World War I temporarily reduced immigration rates, but it was federal government policy that effectively brought mass immigration to an end. A literacy requirement established in 1917 via an override of President Woodrow Wilson's veto was ineffective, but the 1921 Emergency Quota Act and the 1924 Immigration Act effectively reduced immigration to considerably lower rates for the next four decades.

Remarkably, these quotas were precisely calibrated to allow immigration from Northern and Western European countries to continue nearly as before, while almost eliminating immigration from much of Southern and Eastern Europe. The precise calibration of the 1921 and 1924 Quotas is apparent when pre-quota immigration rates from Scandinavia and Italy are compared with the quotas for Scandinavia and Italy. The 1921 law set an annual quota of new immigrants from a given nationality at 2% of the number of foreign-born persons of such nationality residing in the U.S. in 1910. The 1924 law set an annual quota of each nationality at 3% of the number of foreign-born persons of such nationality residing in the U.S. in 1890. The results of these calculations were startling. The 1921 Scandinavian immigration flow totaled 22,854. The post-1921 Scandinavian quota was 41,412. The 1921 Italian immigration flow totaled 222,260. The post-1921 Italian quota was 40,294. Thus, at the 1921 quota levels, immigration from Italy would still be twice the immigration from all of Scandinavia combined, because the Scandinavian quota was underutilized. It is not surprising, therefore, that the 1924 Quota used new calculations, arriving at a Scandinavian quota of 18,665, and an Italian quota of only 3,845. The final 1924 quotas appear to have been carefully calibrated to keep immigration from some nations roughly constant, while nearly eliminating immigration from other nations. Table 1 reports the average quotas over the course of the period, comparing them with actual immigration numbers from Willcox (1929) and U.S. Department of Commerce (1924, 1929, 1931).

In Figures 2a and 2b, we present a slightly modified replication of the results in Ager and Hansen (2018), which starkly demonstrates the effects of the quotas' careful calibration. We follow them in regressing actual immigration inflows from 1900 through 1913 on a simple quadratic in time and projecting forward; in our case performing the analysis twice, once for Southern and Eastern Europe and once for Northern and Western Europe. It is apparent that the actual quotas were strictly binding for Southern and Eastern Europe as a whole, and barely binding for Northern and Western Europe. The quotas resulted in a massive number of "missing" immigrants, nearly all of them from Southern and Eastern Europe.[5]

All of the papers in the recent literature on the quotas use identification strategies that take advantage of the fact that this variation in quotas across source countries induced variation across U.S. locations. Following Abramitzky and Boustan (2017), in Figure 3a, we map the share of the 1920 population in each U.S. county who were from Northern and Western Europe; and in Figure 3c, the share from Southern and Eastern Europe. Clearly, there is variation between and within regions of the United States, in terms of where immigrants from these different source countries tended to settle. Due to history dependence in where immigrants tend to settle (Card, 2001; Moretti, 1999), these pre-quota patterns in the source countries of immigration across U.S. locations induced variation in post-quota impacts across U.S. cities. This provides the first source of variation we use in this paper. We also expand on the existing literature by demonstrating similar history dependence in terms of which industries immigrants tended to work in before the quotas, which is the second source of variation we use. The identification strategy thus consists of comparing cities and industries that had experienced substantial inflows from Southern and Eastern Europe with otherwise similar locations and industries that had experienced substantial inflows from Northern and Western Europe, both before and after the quotas.

One concern with this identification strategy is whether the laws were passed with precisely these differential effects across American cities and industries in mind. For example, a scenario in which senators and representatives from some US locations or industries sought to decrease the economic potential of competing US locations and industries by cutting off

[5]See Figures 1 and 2 in Ager and Hansen (2018), page 31 for the original figures that we replicate as Figures 2a and 2b in this paper.

their supply of low skilled labor while preserving their own would pose a problem for the exclusion restriction implicit in the identification strategy. In this scenario, the identification strategy would confuse the effects of the quotas with the effects of a host of correlated political acts designed by powerful senators and representatives in the early 1920s to help some U.S. locations and industries and harm others. tion of the immigrants as the century progressed (Abramitzky and Boustan, 2017; Spitzer and Zimran, 2018).

American concerns about the effects of immigration grew in proportion to the increased prevalence of Southern and Eastern European immigrants shown in Figure 1b. Figure 1a demonstrates that World War I temporarily reduced immigration rates, but it took federal government policy to nearly end it. A literacy requirement established in 1917 over President Woodrow Wilson's veto was ineffective, but it was the 1921 Emergency Quota Act and the 1924 Immigration Act that effectively reduced immigration to considerably lower rates for the next four decades.

Remarkably, these quotas were precisely calibrated to leave immigration from Northern and Western European countries nearly constant, while nearly ending immigration from much of Southern and Eastern Europe. The precise calibration of the 1921 and 1924 Quotas is apparent through comparing pre-quota immigration from Scandinavia and Italy with the quotas for Scandinavia and Italy. The 1921 law set an annual quota of new immigrants from each nationality at two percent of the number of foreign-born persons of such nationality resident in the US in 1910. The 1924 law set an annual quota of each nationality at three percent of the number of foreign-born persons of such nationality resident in the US in 1890. The results of these calculations were startling. The 1921 Scandinavian immigration flow was 22,854. The post-1921 Scandinavian quota was 41,412. The 1921 Italian immigration flow was 222,260. The post-1921 Italian quota was 40,294. Thus, at the 1921 quota levels, immigration from Italy would still be twice the immigration from all of Scandinavia combined, because the Scandinavian quota was underutilized. It is not surprising, therefore, that the 1924 Quota used new calculations, to arrive at a Scandinavian quota of 18,665, and an Italian quota of only 3,845. The final 1924 quotas appear to have been carefully calibrated to keep immigration from some nations roughly constant, while nearly eliminating immigration from other nations. Table 1 reports the average quotas throughout the period, comparing them with actual immigration numbers from Willcox (1929) and U.S. Department of Commerce (1924, 1929, 1931).

In Figures 2a and 2b, we present a slightly modified replication of the results in Ager and Hansen (2018) which starkly demonstrates the effects of the quotas' careful calibration. We follow them in regressing actual immigration inflows from 1900 through 1913 on a simple quadratic in time and projecting forward; in our case performing the analysis twice, once for Southern and Eastern Europe and once for Northern and Western Europe. It is apparent that the actual quotas were strictly binding for Southern and Eastern Europe as a whole, and barely binding for Northern and Western Europe. The quotas resulted in a massive number of "missing" immigrants, nearly all of them from Southern and Eastern Europe.[6]

All of the papers in the recent literature on the quotas use identification strategies that take advantage of the fact that this variation in quotas across source countries induced variation across US locations. Following Abramitzky and Boustan (2017), we map in Figure 3a the share of population in each U.S. county in 1920 from Northern and Western Europe; in Figure 3c from Southern and Eastern Europe. Clearly, there is variation between and within regions of the United States in where immigrants from these different sources tended to settle. Due to history dependence in where immigrants tend to settle (Card, 2001; Moretti, 1999), these pre-quota patterns in immigrant source countries across U.S. locations induced variation in post-quota impacts across U.S. cities, providing the first source of identifying variation we use in this paper. We also expand on the existing literature by demonstrating similar history dependence in which industries immigrants tended to work in before the quotas, thus providing a second source of identifying variation. The identification strategy thus consists in comparing cities and industries that had experienced substantial inflows from Southern and Eastern Europe with otherwise similar locations and industries that had experienced substantial inflows from Northern and Western Europe, before and after the quotas.

One concern with this identification strategy would be if the laws were passed with precisely these induced effects across American cities and industries in mind. A problem for

[6]See Figures 1 and 2 of Ager and Hansen (2018), page 31 for the original Figures that we replicate in Figures 2a and 2b of this paper.

the exclusion restriction implicit in the identification strategy would be a scenario in which senators and representatives from some U.S. locations and industries sought to decrease the economic potential of competing U.S. locations and industries by cutting off their supply of low skilled labor while preserving their own. Under this scenario, the identification strategy would confuse the effects of the quotas with the effects of a host of correlated political acts designed by powerful Senators and Representatives to help some U.S. locations and industries and harm others during the early 1920s.

Goldin (1994) has examined the motives behind immigration restrictions by running regressions on vote counts. She argues that early and unsuccessful attempts to limit immigration in 1915 and 1917 were based in part on economic concerns about immigration in general. Goldin's argument is that these economic concerns were national, and that so many voters supported immigration restrictions precisely because native rural voters (who did not live near immigrants) were concerned about the perceived plight of native workers in cities. Her work does not address the successful attempts to curtail immigration in the 1921 and 1924 quotas, the votes for which were nearly unanimous (89% of votes cast in the House were in favor of the 1921 restrictions, along with 99% of votes cast in the Senate). However, it can be argued that the 1921 and 1924 quotas likely also represented national concerns that affected natives everywhere, and were not part of an organized campaign to promote the economic well-being of one location over another.

Finally, we can learn what these national concerns were by examining the historical record of the debate leading up to the passage of these laws. During the discussions on the 1924 restrictions, senators and representatives from around the country repeatedly expressed concerns about the ethnic heritage of people from Southern and Eastern Europe, as well as their religious affiliations (i.e., Catholicism or Judaism). At the same time, they extolled the ethnic heritage of people from "Nordic" countries as well as people of Protestant background. For example, Representative Ira Hersey of Maine complained, "We have thrown open wide our gates and through them have come other alien races, of alien blood, from Asia and southern Europe...with their strange and pagan rites, their babble of tongues." Senator Earl Michener of Michigan explained, "The Nordic People laid the foundations of society in America. They have built this Republic, and nothing would be more unfair to them and their descendants than to turn over this Government and this land to those who had so little part in making us what we are." Senator Reed of Pennsylvania stated that his goal was to "maintain the racial preponderance of the basic strain on our people and thereby to stabilize the ethnic composition of the population." Representative William Vaile of Colorado stated, "What we do claim is that the Northern Europeans, and particularly Anglo-Saxons, made this country."[7] As historian Robert Fleegler recounts, "during the 1924 congressional debate over immigration restriction...the supporters of restriction espoused a conception of American identity that excluded eastern and southern European migrants. Only a small minority disagreed," (Fleegler, 2013).

Thus, far from being local efforts to curtail immigration to some locations but not others, these laws were national efforts to curtail immigration from some sources but not others. This national effort did exempt certain categories of immigrants, however. The categories of immigrants granted blanket exceptions to the quotas (who could therefore continue to immigrate without restriction) included: "professors;" "lecturers;" people belonging to "any recognized learned profession;" "an immigrant who is a bona fide student at least 15 years of age and who seeks to enter the United States solely for the purpose of study;" domestic servants (from 1921–1924); and singers and actors.[8] These exemptions imply which mechanisms are most likely at work in the results we report below.

In the next section, we explain how we construct the data necessary to investigate the quotas' effect on immigration and innovation.

II Data and Matching

Census Data

Administrative data from Willcox (1929) and U.S. Department of Commerce (1924, 1929, 1931) give us exact immigration counts by source country and year. IPUMS full count Census data (Ruggles et al., 2017) provide comprehensive information on immigrants' locations, industry and year of arrival in the United States via reports of the total population, foreign-

[7] Quotes are from "Ellis Island Nation: Immigration Policy and American Identity in the 20th Century" by Fleegler (2013).

[8] See Sec. 2(d) of Pub. L. No. 67-8 and Sec. 4 of Pub. L. No. 68-190.

born population, Southern and Eastern European foreign-born population, and Northern and Western European foreign-born population in 1900, 1910, 1920, and 1930. Complete count Census data with names from 1920 tell us: full names, genders, birth years, birthplaces, arrival years, locations, and occupations of everyone living in the United States in 1919 (the year the 1920 Census took place).

The identification strategy depends on variation across locations, industries, and years. Thus, it is helpful to observe immigration inflows into locations and industries on a yearly basis if possible. The 1910, 1920, and 1930 United States Censuses report the nativity status, birth country, and year of arrival for every person living in the United States in 1909, 1919, and 1929, respectively. Thus, we can use these three censuses to determine the exact initial location choice and industry choice of immigrants who arrived in 1909, 1919, and 1929.

This would provide us with two pre-quota years of data on immigration inflows across locations and industries, and one post-quota year for immigration inflows across locations and industries. A difference-in-differences strategy relies on the assumption that treated and comparison groups have similar levels and trends of relevant variables before the treatment begins. While two pre-quota years (1909 and 1919) are useful for establishing pre-quota trends, it would be helpful to have richer data, and to establish the exact year when the trends diverge post-quota. To establish this richer data, we develop a proxy for the initial locations and industries of immigrants who arrived in the years between censuses. Our proxy uses information from the 1910, 1920, and 1930 censuses, assigning immigrants who arrived in year t to the county and industry they report living and working in in the census closest to year t. Thus, for immigrants who arrived in 1919, the proxy corresponds with the true (contemporaneously observed) observations of their initial locations and industries gleaned from the 1920 census. For immigrants who arrived in 1925, the proxy corresponds with the circa-1929 locations and industries they reported in the 1930 Census.

While the proxy corresponds with the truth during Census years themselves, it will diverge from the truth in the intervening years in two ways: via immigrants' movement within the United States between year t and the next Census year, and via return migration. Fortunately, we can test the accuracy of the proxy by comparing the proxy vectors of the number of 1919-arrival immigrants across locations and industries reported in the 1930 Census with the true (contemporaneously observed) vectors of the number of 1919-arrival immigrants across locations and industries reported in the 1920 Census. We find that the location proxy and the industry proxy have correlations of approximately 0.9 with their respective true vectors. We can also perform all of the analyses below ignoring the proxy and relying only on the three observations of newly arrived immigrants in 1909, 1919, and 1929, as in the existing literature on the effects of the quotas.

Patent Data

We create three types of patent data: county-level, industry-level, and inventor-level. The European Patent Office's PATSTAT database (European Patent Office, 2017) provides details on each patent application to the United States Patent Office from 1899 to the present, including: the inventor's full name, year of application, International Patent Classification (IPC), and number of citations. To construct the county-level patent data, we use the HistPat database (Petralia et al., 2016), which provides the locations of patent applicants based on digitized records of their original applications to the United States Patent and Trademark Office (USPTO) from 1790 to 1975. We then identify the year of patent applications in the PATSTAT database and link this with the HistPat data using the publication number for each patent. The county-level patent data thus presents the number of patent applications per county per year.

We construct the industry-level data by matching patent classifications with industry classifications. Specifically, we use an IPC (International Patent Classification) to NAICS (the North American Industry Classification System) concordance (Lybbert and Zolas, 2014) and then link patents to the industry classifications used in historical censuses (using the IPUMS variable IND1950, which ensures that industry classifications are comparable across time periods). In the 1910 and 1920 U.S. Censuses, there are a total of 55 industries represented among the industry-patent matches. The industry-level patent data reports the number of patent applications per industry per year.

Finally, we consider inventor-level patent data. To determine the effect of the quotas on inventors who lost geographically-close immigrant labor, we need to identify where inventors were living just before the quotas were instituted. The immigration quota treatment applies to inventors who lived in a county with a large fraction of Southern and Eastern European immigrants in 1919, just before the quotas. Patent applications report the inventors' locations at the moment the patent application was filed. But the median number of patent applications conditional on ever having filed a patent is one (Bell et al., 2019), so the vast majority of incumbent inventors living in any given city in 1919 would be unlikely to happen to apply for a patent (and thereby reveal their current location) in 1919. This means that using location data based on 1919 patent applications would cause us to substantially underestimate the number of inventors living in each location.

Therefore, we merge patent data into census data at the level of individual inventors. We can then determine where all inventors subject to the matching criteria were living in 1919, regardless of whether they applied for a patent that year. Furthermore, we can also control for demographic characteristics, which are proven determinants of the probability of invention (Bell et al., 2019), thus improving the precision of our estimates.

We use a match between the PATSTAT database and the complete count 1920 U.S. Census with names. A fuzzy matching procedure merges patents and publications at the individual name level into the 1900, 1910, 1920, 1930, and 1940 complete count U.S. Censuses with names. Each such Census can tell us how many people living in the U.S. at the time of that Census had any given first name/middle name/last name combination. In any given Census, almost half of the population is made up of people who are the only person in the country with their first name/middle name/last name combination. For example, in the 1920 U.S. Census, 43% of the U.S. population was made up of people with unique names. The fuzzy matching procedure accounts for common misspellings and assigns each patent to the person or persons with a matching name in the Census. We impose three restrictions to increase the probability of correct matches. First, and most importantly, we only consider the 43% of the population with a unique name. Second, we only consider matching patent applications when the inventor's implied age at the time of the application is between 18 and 80. Finally, in most regressions, we restrict our attention to patent applications matched between the years 1919 and 1929.

Given these restrictions, it is very likely that the resulting matched patents are correct. Given a person with a unique name in the 1920 Census (observed in 1919), we know that any patent applications between 1919 and 1929 with that unique name must be either from that person, or from someone with an identical name who immigrated to the United States during those years. They could not be from someone born after 1919 with the same unique name, because any such person would be younger than 10 years old. They could not be from someone with the same name who was born before 1919 and had died by 1919, because such a person would be dead and unable to file a patent. Thus, for the 43% of people that comprise our restricted sample, and for the eleven years in our primary regressions, the matched patents should only be incorrect if there are transcription errors in the names recorded in the raw data or if a new immigrant arrived with the exact same full name and filed a patent application shortly after arrival.

In the next section, we describe the empirical strategy that we use to analyze the impacts of these quota-related declines in mass immigration on American innovation.

III Empirical Strategy

Our empirical strategy is to compare otherwise similar counties, industries, and individuals over time, some that were more exposed to the quotas and some that were less exposed. We operationalize this strategy by defining cells at the county-year level, the industry-year level, or the individual inventor–year level.

Importantly, our identifying strategy relies on the assumption of parallel trends, in which the outcomes of groups highly affected by exposure to the quotas and the outcomes of groups little affected by exposure to the quotas would not change differently in the absence of the quotas. To be specific, among counties with a large proportion of immigrants before the quotas, some counties with many Southern and Eastern European immigrants should have similar characteristics to other counties with few Southern and Eastern European immigrants. Table 2 reports summary statistics of baseline county characteristics and outcomes before the quotas, such as immigration inflows and patent applications. In column 5, we find no significant differences in these variables between high–quota exposure counties and low–quota exposure counties among the counties with a large fraction of foreign-born populations, with the exception of number of low-skilled workers. When the fraction of low-skilled workers is used as the outcome in the empirical analysis, there are no pre-treatment differences.

Furthermore, we find consistent results at the industry level as illustrated in Table 6. None of the differences in baseline characteristics and outcomes before the quotas is statistically significant at the industry level.

Our difference-in-differences strategy will thus involve regressing equations of the following general form:

$$Y_{jt} = \alpha + \beta(QuotaExposure_j \times PostTreatment_t) + \gamma_j + \tau_t + \epsilon_{jt} \tag{1}$$

where j is a specific county or a specific industry and t is a calendar year. The outcome variable will be either a measure of population levels or flows (e.g., new immigrants arriving each year) or a measure of invention (e.g., new patent applications). In order to control for differences in the level of the outcome variable Y across cells, we include fixed effects at both levels (i.e., county/industry and calendar year).

The treatment variable $QuotaExposure_i$ measures local exposure to the immigration quotas given a group's proportion of Southern and Eastern Europeans. Following the method in Abramitzky et al. (2019), we calculate the quota exposure through the following equation:

$$QuotaExposure_j = \sum_c \frac{FB_{cj1900}}{Pop_{j1900}} \times 1\{Restricted_c\} \tag{2}$$

where FB_{cj1900} is the size of the foreign-born population living in county j or working in industry j and born in country c in 1900, and Pop_{j1900} is the total population of county j or industry j in 1900. $1\{Restricted_c\}$ indicates that a country c is restricted by the immigration quotas, thus it takes a value of one for Southern and Eastern European countries and zero otherwise.

We construct other measures of quota exposure following Ager and Hansen (2018) and Tabellini (2020). We follow the "missing immigrants" method in Ager and Hansen (2018) to assign the missing immigrants to different locations or industries over time. For each location or industry, we calculate the quota exposure through the following equation:

$$QuotaExposure_j = \frac{100}{P_{j,1920}} \sum_{c=1}^{C} \left(\widehat{Immig}_{c,22-30} - Quota_{c,22-30}\right) \frac{FB_{cj,1920}}{FB_{c,1920}} \tag{3}$$

where $\widehat{Immig}_{c,22-30}$ is the estimated average immigration inflow that would have taken place per year from country c during the post-quota years from 1922 and 1930 had the quota acts not been enacted.[9] The variable $QuotaExposure_j$ represents the average annual number of "missing" immigrants per 100 inhabitants of county j or industry j due to quotas.

We construct an alternative measure of quota exposure based on Tabellini (2020). It is a modified version of the standard shift-share instrument that "predicts the number of immigrants received by U.S. 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."

For individual inventors, we modify the equation above as follows:

$$Y_{ijt} = \alpha + \beta(QuotaExposure_j \times PostTreatment_t) + \theta X_{it} + \gamma_i + \tau_t + \epsilon_{ijt} \tag{4}$$

where a specific inventor i living in county j was exposed to $QuotaExposure_j$. We include a control variable X_{it}, the quartic of the age of person i in year t.

Challenges to identification in this context could arise from two sources: (1) differential pre-trends across more and less exposed counties, industries, or individuals; and (2) other shocks which were roughly contemporaneous with the quotas and which affected the more quota-exposed counties, industries, or individuals differently than they affected those that were less exposed.

In order to address the first set of concerns, we estimate the difference-in-differences analysis with time-varying coefficients in the following form:

$$Y_{jt} = \alpha + \sum_t \beta_t(QuotaExposure_j \times YearDummy_t) + \gamma_j + \tau_t + \epsilon_{jt} \tag{5}$$

By plotting the β_t coefficients over time, we can determine whether differential pre-trends affect the analysis. We modify the equation for individual inventors as follows:

$$Y_{ijt} = \alpha + \sum_t \beta_t(QuotaExposure_j \times YearDummy_t) + \theta X_{it} + \gamma_i + \tau_t + \epsilon_{ijt} \tag{6}$$

where inventor i was living in county j.

In order to address the second set of concerns, we rely first on the historical context

[9] The estimates are predicted using the pre-WWI annual immigration flows from 1900 to 1914 based on the following regression model: $Immig_{ct} = \beta_1 lnt + \beta_2(lnt)^2 + \epsilon_{ct}$.

introduced above: namely, that the classic endogeneity problem in immigration studies (that many shocks to immigration were intended (by policy makers, the immigrants themselves, or those demanding their labor) to change immigration to some regions more than others (Cadena and Kovak, 2016)), is unlikely to be of concern here. Crucially, the variation in immigration that the quotas induced across counties and industries was an unintended side effect of the laws. The aim of the quotas was to reduce immigration from some sources and not others; an unintended side effect is that they ended up reducing immigration to some cells (counties and industries) and not others.

The second way of addressing the concern about contemporaneous shocks is by controlling for state-year fixed effects. We compare counties and individual inventors with their counterparts in the same state, capturing contemporaneous shocks and economic trends using the state-year. Another method we consider is adjusting for World War I, which affected immigration and innovation from 1914 through 1918. We address World War I in two ways: first, by running many of our regressions from 1919 through 1929, and, second, by dropping patents in categories that may have been directly affected by the war effort itself.

We can also ask whether mass immigration changed the *direction* of inventive activity at the county or industry level. In other words, did mass immigration alter the kind of inventions that inventors pursued? If so, in which direction did invention change? In particular, we measure each county or industry's baseline "location" in the space of ideas, the subject area in which it specialized before the quotas. We can then measure how the quotas changed this location by assessing the frequency of patent applications in categories not covered by this specialty.

In the next section, we investigate the impacts of the immigration quotas on innovation within counties.

IV Effect of the End of Mass Immigration on Innovation within Counties

Immigration Rates and the Labor Force

We begin our analysis by verifying that the quotas decreased immigration rates, the labor force, and the population size in quota-exposed locations. In Table 3, we report the results when the outcome variable is newly arrived immigrant inflows in a given location in a given year, proxied for the years between censuses using the technique described in the section above. It is apparent that regardless of the date range included in the sample, the cutoff year chosen for the beginning of the quotas, or the sample chosen based on demographic characteristics before the quotas, the quotas resulted in substantial reductions of immigration inflows relative to the pre-quota means. In particular, the number of newly arrived low-skilled immigrants significantly decreased in response to the quotas restricting immigration to the U.S. from Southern and Eastern European countries. We find similar results using the alternative measures of quota exposure that control for state-year fixed effects. These results are available upon request.

In Figures 4, 5, 6, and 7, we depict the results using the characteristics of locations during the 1910, 1920, and 1930 Censuses. We take first differences within locations before the quotas and first differences within locations after the quotas, and report the results separately. It is apparent that there were substantial declines in the total population and in the total number of workers in counties more exposed to quotas between 1920 and 1930. In particular, the number of low-skilled workers substantially declined, but the number of high-skilled workers did not change. It is also clear that this relative decline was not the result of differential pre-quota trends because none of the changes between 1910 and 1920 were statistically significant.

It is important to note that this evidence shows that counties more exposed to the quotas experienced a decline in total population and total workers after the quotas, and that any internal migration did *not* erase this decline (such as, for example, the movement of African-Americans from rural areas in the South to urban areas in the North). Furthermore, the Great Migration began in the 1910s, before the quotas (Collins, 1997); greater quota exposure is only associated with declines after the quotas, not before.

Innovation

We determine how the quotas, which reduced the population and the labor force in affected counties, affected innovation as measured in terms of the number of patents and patent citations. We report the results from equation (1) in Table 4. Clearly, regardless of the sample restrictions, years covered, or the cutoff year for the post-quota period, we find large declines in the number of patent applications per year in quota-exposed locations. A 1 percentage point increase in quota exposure decreased patent applications per year by 1.2% in column 2.[10] According to the results in Table 3, the equivalent increase in quota exposure decreases immigration inflows by 18%.[11] Thus, we find that for every 18% decrease in immigration, patent applications decreased by 1.2%. For brevity, a 10% decrease in immigration inflows per year in a county exposed to the quotas led to a 0.7% reduction in patent applications per year per county.

It is possible that the marginal inventions were not useful ones; perhaps the inventors would have invented the most useful inventions anyway, regardless of the quota-induced shock. To examine this possibility, we estimate the results in panel B using citation-weighted patents as the outcome variable. It is evident that the results are very similar in sign, significance, and magnitude. Specifically, as seen in column 2 of Table 4, a 1 percentage point increase in quota exposure decreased citation-weighted patents by 1.4%, thus implying that a 10% decrease in immigration inflows decreased citation-weighted patents by 0.8%.

We consider the case in which counties with greater exposure to the quotas had different demographic characteristics in the pre-quota years. To address this issue, we restrict our sample to counties with large populations of foreign born workers in 1910 where there were no pre-quota differences. It is clear in columns 3 and 4 that counties with greater exposure to the quotas saw significant drops in patenting activities.

We also report the results from equation (5) in Figure 8. The variable quota exposure is interacted with a set of indicator variables corresponding to each particular year after

[10] Given a dependent variable mean of 0.689, $\frac{-0.814}{0.689} = 1.18\%$.

[11] In Table 3, column 2 in panel A, $\frac{-0.073}{0.004} = 18.25\%$.

the base year 1921 is omitted. Figure 8 shows the quotas decreased inventions, and that this decrease in inventions is not an artifact of pre-trends. This provides evidence for our identifying assumption that the number of inventions in treatment and control counties would not differ in the absence of the quota exposure.

Another possibility is that the observed effects are mere artifacts of the increased demand for military-related inventions during World War I, and that after the war was over, this demand decreased. We therefore create a modified dependent variable in which patents related to military applications are removed. Sorting by IPC codes, we remove patent applications relevant to the arms industry, such as weapons, ammunition, and explosives. [12] We find that neither the size nor the precision of the estimates is affected by this change. We also find consistent results using two other measures of quota exposure and controlling for state-year fixed effects. The results are available upon request.

Addressing Additional Concerns

In the previous subsection, we provide evidence that there was a decrease in inventions in the areas exposed to the quotas. While we showed that there were no differential trends in patents before the quotas, we further verify the parallel pre-trend assumption using the following equation:

$$Y_{jt} = \alpha + \beta(QuotaExposure_j \times YearTrend_t) + \gamma_j + \tau_t + \epsilon_{jt} \tag{7}$$

The sample is restricted to either the period from 1919 to 1921 or the period from 1919 to 1923. The coefficient β measures differential trends in patent activities before the quotas. Table A.1 shows that none of these coefficients are statistically significant. This supports the parallel pre-trends assumption underlying our difference-in-differences strategy.

Additional concern should be taken when we interpret the decrease in inventions caused by the quotas. The threat to this interpretation is that our results might reflect the different characteristics of areas exposed to the quotas. To alleviate this concern, we provide results in which our sample was restricted to counties with large populations of foreign-born workers

[12] The following IPC codes are relevant to the arms industry: F41A, F41B, F41C, F41F, F41G, F41H, F41J, F42B, F42C, F42D, B63G, C06B, and G21J.

where there were no pre-quota differences across many attributes. We address this threat further by conducting a placebo test to examine the effect of the quotas on patent activity in counties with fewer foreign-born workers.

Table A.2 reports the estimated coefficients using the sample of counties with fewer foreign-born workers. The results of the placebo test show that none of these coefficients are statistically significant, and quota exposure is not negatively associated with inventions. The placebo test suggests that it was the quotas that were responsible for the decrease in inventions in areas where many immigrants were living prior to the quotas, which supports the validity of our specification.

Mechanism

We explore a possible mechanism for the decline in innovation after the decrease in mass immigration. While our work above analyzes the effect of the extant labor market on the *rate* of inventive activity, a fuller analysis must address possible effects of the labor market on the *direction* of inventive activity within this space of ideas (Lerner and Stern, 2012). To investigate the direction of innovation, we create two measures, an index of dissimilarity, and a measure of patent originality.

The index of dissimilarity is calculated following Borjas and Doran (2012) and Cutler and Glaeser (1997). The index takes the value of one when patent classifications (4-digit IPC codes) in a given county and year do not match any preexisting patent classifications in that county between 1900 and 1919, before the quotas. If the patent classifications perfectly match previous patent classifications in the same county, then the index equals zero. This allows us to identify whether counties continue to produce the same types of patents or produce new types of patents. We report the results in Table 5 where the dependent variable is the index of dissimilarity. We find that quota exposure increased the index of dissimilarity, meaning it took inventive activity in a new direction.

Furthermore, we create another measure of new inventive activity using text analysis. Patent originality is measured by counting the patents with new words in their title. We collect preexisting words in the titles of patent applications filed between 1900 and 1919, prior to the quotas, and identify original patents that use new words, weighting these by citations. In panel B of Table 5, we find that counties more exposed to the quotas saw an increase in patent originality. Taken together, exposure to the quotas decreased the rate of inventive activity, but changed its direction. An overall decrease in patents and a reduced labor force due to immigration restrictions spurred new types of inventions.

In the next section, we report the results at the industry level.

V Effect of the End of Mass Immigration on Innovation within Industries

Immigration Rates and the Labor Force

We first determine whether some industries were more exposed to the quotas than others and experienced a decline in immigration inflows. Table 6 describes the summary statistics of the industries included in our analysis. We restrict industries in the censuses (the variable IND1950) to 55 industries that can be matched to patent classifications using an IPC code to NAICS concordance (Lybbert and Zolas, 2014). In column 5, there were no statistically significant pre-quota differences in baseline characteristics and outcomes between industries more exposed to the quotas and other industries less exposed to the quotas. We find that industries with and without exposure to the quota are comparable.

To investigate the effect of the quotas on the rate of immigration inflow, we estimate equation (1) and report the results in Table 7. Industries with greater exposure to the quotas experienced a large decline in immigration inflows. Specifically, a 1 percentage point increase in quota exposure decreased the number of immigrants by 2.5% per year per industry in column 1. Consistent with the results at the county level, we also find a substantial decrease in low-skilled immigrants among industries more exposed to the quotas.

Innovation

We investigate whether the industries that experienced a decline in immigrant workers saw a decrease in the number of associated patent applications. Table 8 shows that in industries more exposed to the quotas, patents did decrease. In particular, in column 1, a 1 percentage point increase in quota exposure decreased patents by 0.3% per year per industry. We find that a 10% decrease in immigration inflows per year per industry exposed to the quotas decreased patent applications by 1.2%.

The results are robust to specifications that restrict the sample to industries with a large fraction of foreign workers before the quotas, with no pre-quota differences in characteristics and outcomes. In columns 3 and 4 of Table 7 and Table 8, industries where low-skilled immigrant workers decreased due to the quotas saw a decrease in inventive activity as well. The results support the hypothesis that these quota-exposed industries had less demand for inventions after the quotas than they had before. This implies that it was difficult to increase the scale of production due to the decrease in the inflow of immigrant workers.

To provide evidence for our identifying assumption, we complement our empirical analysis with a difference-in-differences specification relative to the base year prior to the quotas in equation (5). We find the decrease in inventions with a lag following immigration restrictions in Figure 9. The lack of difference before the quotas provides evidence that the results are not an artifact of pre-quota differential trends.

Mechanism

Mass immigration was an essential input to American invention in the beginning of the last century. Acemoglu (2010) conceptually shows that a plentiful labor supply encourages innovation when technology increases the marginal product of labor, thus supporting the proposition that inventions and the labor supply are complementary rather than substitutes. We investigate one possible mechanism through which mass immigration could affect the incentives to invent: a simple case of increased scale available for production. To test this hypothesis, we ask: for industries that depended on immigrants to maintain their production scale, which types of inventions decreased after the quotas? Industry level analysis allows us to investigate how different types of inventions depended on the extant labor market conditions.

We first measure the effect of the quotas on the direction of inventions at the county level in Table 5, which employs the index of dissimilarity and patent originality. Table 9 shows the positive effect of immigration restrictions on those outcomes. We find that industries decreased their overall patent applications in response to the decreased workforce following immigration restrictions, but invested in new types of patents. The quotas took inventions in a new direction.

To investigate in which direction inventive activity went, we explore possible mechanisms through which some inventions complement labor while others are strong labor substitutes. We report the results in Table A.6, broken down by the skill levels of the workforce in the industries. It is apparent that nearly all of the reduction in patent applications reported in Table 8 was due to a reduction in patent applications in industries with a large fraction of low-skilled workers. Patent applications in high-skilled industries did not change significantly. The results suggest that the end of mass migration caused a significant decrease in low-skilled immigrant workers from Southern and Eastern European countries, and substantially decreased inventions relevant to low-skilled industries.

Finally, we identify the labor intensity of industries using labor costs from workers' occupational income scores. We report the results in Table A.7. These results suggest that what declined substantially after the quotas was the invention of technology relevant to labor-intensive industries that lost workers due to the quotas. In these industries, labor became scarce, and this discouraged invention. In the language of Acemoglu (2010), this suggests that much of the invention at the time was labor-complementary.

In the next section, we determine how the quotas affected individual inventors and explore the mechanisms through which labor scarcity from immigration restrictions affects innovation.

VI Inventors, Cognitive Mobility, and Other Mechanisms

Individual Inventors

We investigate how individual inventors coped with the decrease in the labor force following immigration restrictions. We first consider incumbent inventors who were active before the quotas. Table 10 reports the results from equation (4). We include the quartic of age of person i in year t, the individual fixed effects, and the year fixed effects. We find a significant decrease in the number of patents applied for per year by incumbent inventors living in quota-exposed locations. Specifically, a 1 percentage point increase in quota exposure decreased patent applications per year by approximately 1%, as seen in column 2. Regardless of the sample restrictions in columns 3 and 4, the cutoff year for the post-quota period, or the citation-weighted patents in panel B, we find large declines in the number of patent applications from incumbent inventors. Furthermore, we estimate the results from equation (6) in Figure 10. This shows no difference in patents prior to the quotas and a decrease in patents after the decreased immigration inflows due to the quotas.

Cognitive Mobility in the Space of Possible Inventions

If inventors, like mathematicians, face "Cognitive Mobility Costs" for changing their position in the space of ideas in response to a new set of opportunities and incentives (Borjas and Doran, 2015), then this movement in the space of ideas may explain much of the decline in the rate of patent applications among incumbent inventors reported above. In contrast, new inventors and young inventors would have little invested in an ongoing invention program at the time of the quotas, thus their cognitive mobility costs would be lower or non-existent. If it were these cognitive mobility costs that led to the decline in inventions among pre-existing inventors, then the new and young inventors may not have experienced the same decline in the rate of their patent applications.

We test this possibility by estimating equation (4) using a sample of all individuals aged 18 to 80 at the time of the quotas, including those who had never submitted a patent application before the quotas were enacted. We report the results in Table A.3. In Panel A, we show that across all seventy million adults living in the United States at the time the quotas were enacted, the total patents per year declined among quota-exposed people. The magnitude of the effect relative to the pre-quota dependent variable mean is a 2% decline, about half the size of that reported in the main results for incumbent inventors in Table 10. In Panel B, we show similar results when the dependent variable is a binary indicator for any patent applications at all in a given year. In Panels C and D, we consider the 99.8% subset of the seventy million adult Americans who had never held a patent before the quotas were enacted. We find that after the quotas began, the quota-exposed subset of these individuals were more likely to file a first patent application, and to complete more patents than otherwise similar non-quota-exposed individuals. Thus, all of the negative effects reported in Panels A and B are the result of the 0.2% subset of the American population who were incumbent inventors at the time of the quotas. New inventors increased their likelihood and rate of patenting when they were exposed to the quotas.

Likewise, in Table A.4, we estimate the same specifications above separated by age group. Once again, we see that all of the decline in the rate of patent applications reported in Panel A of Table A.3 is due to a decline among individuals aged 31–80 at the time of the shock. Younger individuals, aged 18–30 at the time of the shock, actually increased their rate of patenting.

Taken together, this evidence is consistent with: (a) the incentives and opportunities for different types of inventions depending on extant labor market conditions; (b) inventors choosing where in the space of ideas to specialize in response to these extant labor market conditions; (c) inventors who had already specialized at the time of the change in labor market conditions incurring substantial cognitive mobility costs for changing their location in the idea space; and (d) new and young inventors who had been on the margins of invention during earlier labor market conditions taking advantage of the change to increase their inventive output.

Other Mechanisms

The quotas decreased immigration and decreased patenting. We argue that one mechanism at work is a scale effect due to the overall decline in mass migration. This scale effect can numerically account for nearly all of the decline in patents, as shown above. But it is possible that other mechanisms could be at work as well. With such a large immigration shock, there are many small subsets among the overall set of missing immigrants that could be particularly important for the patenting effects. We consider two such subsets of immigrants.

One such subset is immigrant inventors. The quotas were not intended to limit exceptionally skilled individuals, as explained above. But what if the quotas inadvertently prevented prolific scientists and inventors from migrating? The effects could then be due in part to lost knowledge spillovers due to a reduction in knowledge transfer from Europe.

We therefore construct a database of all inventor migrants to the United States who both: (a) appeared as foreign-born individuals in the 1930 U.S. Census; and (b) made at least one patent application between the years 1900 and 1940. We run a difference-in-differences analysis with specifications analogous to those reported above, at the location-year level, where the dependent variable is the number of inventor migrants to that location (as observed in the 1930 U.S. Census), rescaled by the number of inventors who were already there (as observed in the 1920 U.S. Census). In this alternative mechanism, we are concerned that European knowledge transfer may have been reduced by the quotas, so we consider all inventor migrants who had applied for at least one patent *before* their year of migration (observed in the 1930 U.S. Census), and therefore had some knowledge from their time in Europe that they could have transferred.

The results, in Table A.5, show that the quotas had no effect on the migration of European inventors to quota-exposed locations after the quota. The results are similar when we rescale by the pre-existing population in each location.

One potential alternative mechanism for our results is that the quotas may have affected invention through an (unintended) effect on highly skilled immigration. Moser and San (2019) have used biographical data on American scientists to investigate these effects and estimate that more than 1,000 Eastern and Southern European–born scientists were missing from U.S. science up to 1956 as a result of the quotas. The loss of these scientists, however, did not reduce American invention in quota-affected *topic areas* until the 1930s, almost ten years after our estimates indicate a decline in American invention in quota-affected *cities*. Given the substantial differences in timing and incidence, we view the Moser and San results as distinct from (and complementary to) our own. Indeed, given the timing of this change, any unintended effects working through high-skilled scientists cannot explain our results.

Another alternative mechanism is that, before the quotas, immigrants may have disproportionately taken occupations that freed up natives' time for invention instead. After the quotas, natives would have to spend time doing tasks immigrants had formerly done, when they otherwise would have spent such time on invention. Indeed, there is evidence that in general, low-skilled immigrants free high-skilled women's time, although the evidence does not address invention in particular (Cortes and Tessada, 2011; Cortes and Pan, 2013). To consider this hypothesis, we examine the fraction of each occupation in the 1920 Census held by foreign-born workers, as well as specifically Southern and Eastern European foreign-born workers. We do not find that occupations related to household help were especially taken up by either group.

VII Conclusion

In this paper, we provide the first causal evidence for the effect of mass immigration on US inventors. We do so by examining the end of the largest international migration in history, which took place at the tail end of America's second Industrial Revolution. Our results suggest that a 10% reduction in mostly low-skilled immigration results in a 0.5% reduction in the number of patent applications by incumbent U.S. inventors. The results are not an artifact of a changing pool of inventors, differential pre-quota trends, or a loss of uncited patent applications.

The results seem to be driven by inventors who had specialized in providing "strongly labor complementary" inventions for local industries (Acemoglu, 2010). Assigning each patent to its associated industry, we find that nearly all of the decline occurred among the subset of patents relevant to the industries whose workforces were most exposed to declining immigrant flows after the quotas.

Because inventions in general, and the inventions of the second industrial revolution in particular, are often designed to economize on labor, it is intuitive that making labor less plentiful should increase the incentive to invent. Since the work of Sir John Hicks (1932) and Sir John Habakkuk (1962), this intuition has suggested that America's early labor scarcity promoted its early technological development. Our paper suggests that at least during the golden age of American invention, it was plentiful labor that made the inventions characteristic of the era worthwhile.

From a historical perspective, therefore, it appears that it was not necessity that was the mother of invention, but rather opportunity.

References

Abramitzky, R., P. Ager, L. P. Boustan, E. Cohen, and C. W. Hansen (2019). The Effects of Immigration on the Economy: Lessons from the 1920s Border Closure. *NBER Working Paper Series 26536*.

Abramitzky, R. and L. Boustan (2017). Immigration in american economic history. *Journal of economic literature 55*(4), 1311–45.

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.

Acemoglu, D. (2010). When does labor scarcity encourage innovation? *Journal of Political Economy 118*(6), 1037–1078.

Ager, P. and M. Brückner (2013). Cultural diversity and economic growth: Evidence from the us during the age of mass migration. *European Economic Review 64*, 76–97.

Ager, P. and C. W. Hansen (2018). Closing heaven's door: Evidence from the 1920s us immigration quota acts.

Azoulay, P., B. F. Jones, J. D. Kim, and J. Miranda (2022). Immigration and entrepreneurship in the united states. *American Economic Review: Insights 4*(1), 71–88.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019). Who Becomes an Inventor in America? The Importance of Exposure to Innovation. *Quarterly Journal of Economics 134*(2), 647–713.

Beniger, J. R. (1997). *The Control Revolution: Technological and Economic Origins of the Information Society*. Harvard University Press.

Borjas, G. J. and K. B. Doran (2012). The collapse of the soviet union and the productivity of american mathematicians. *The Quarterly Journal of Economics 127*(3), 1143–1203.

Borjas, G. J. and K. B. Doran (2015). Cognitive mobility: Labor market responses to supply shocks in the space of ideas. *Journal of Labor Economics 33*(S1), S109–S145.

Cadena, B. C. and B. K. Kovak (2016). Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession. *American Economic Journal: Applied Economics 8*(1), 257–290.

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

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.

Cortes, P. and J. Pan (2013). Outsourcing household production: Foreign domestic workers and native labor supply in hong kong. *Journal of Labor Economics 31*(2), 327–371.

Cortes, P. and J. Tessada (2011). Low-skilled immigration and the labor supply of highly skilled women. *American Economic Journal: Applied Economics 3*(3), 88–123.

Cutler, D. M. and E. L. Glaeser (1997). Are Ghettos Good or Bad? *Quarterly Journal of Economics 112*(3), 827–872.

European Patent Office (2017). The EPO worldwide patent statistical database. PATSTAT: Version 5.09 [dataset]. 2017 Spring edition.

Fleegler, R. L. (2013). *Ellis Island nation: Immigration policy and American identity in the twentieth century*. University of Pennsylvania Press.

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.

Habakkuk, H. J. (1962). *American and British Technology in the Nineteenth Century: The search for labour saving inventions*. Cambridge University Press.

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

Hicks, J. R. (1932). *The Theory of Wages (2nd ed.)*. London: Macmillan.

Hirschman, C. and E. Mogford (2009). Immigration and the american industrial revolution from 1880 to 1920. *Social Science Research 38*(4), 897–920.

Kim, S. (2007). Immigration, industrial revolution and urban growth in the united states, 1820-1920: Factor endowments, technology and geography.

Lafortune, J., J. Tessada, and E. Lewis (2018). People and machines: A look at the evolving relationship between capital and skill in manufacturing 1860-1930 using immigration shocks.

Lerner, J. and S. Stern (2012). *The rate and direction of inventive activity revisited*. University of Chicago Press.

Lybbert, T. J. and N. J. Zolas (2014). Getting Patents and Economic Data to Speak to Each Other: An 'Algorithmic Links with Probabilities' Approach for Joint Analyses of Patenting and Economic Activity. *Research Policy 43*(3), 530–542.

Marshall, A. (1895). *Principles of Economics: Third Edition*. London: Macmillan and Co.

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.

Moretti, E. (1999). Social networks and migrations: Italy 1876-1913. *International Migration Review*, 640–657.

Moser, P. and S. San (2019). Immigration, Ethnicity, and American Science: Evidence from the Quota Acts.

Parker, A. W. (1924). The quota provisions of the immigration act of 1924. *American Journal of International Law 18*(4), 737–754.

Petralia, S., P.-A. Balland, and D. Rigby (2016). HistPat Dataset.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2017). Integrated public use microdata series: Version 7.0 [american community survey, 2016]. *Minneapolis: University of Minnesota*.

Sequeira, S., N. Nunn, and N. Qian (2020). Immigrants and the making of america. *The Review of Economic Studies 87*(1), 382–419.

Spitzer, Y. and A. Zimran (2018). Migrant self-selection: Anthropometric evidence from the mass migration of italians to the united states, 1907–1925. *Journal of Development Economics 134*, 226–247.

Stuart, B. A. and E. J. Taylor (2021). Migration networks and location decisions: Evidence from us mass migration. *American Economic Journal: Applied Economics 13*(3), 134–75.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *Review of Economic Studies 87*(1), 454–486.

U.S. Department of Commerce (1924). Statistical Abstract of the United States.

U.S. Department of Commerce (1929). Statistical Abstract of the United States.

U.S. Department of Commerce (1931). Statistical Abstract of the United States.

Willcox, W. F. (1929). International migrations, volume i: Statistics. *NBER Books*.

(a) Total immigration inflows per year

(b) Fraction of immigration from Southern and Eastern Europe

Figure 1. IMMIGRATION INFLOWS FROM ADMINISTRATIVE DATA

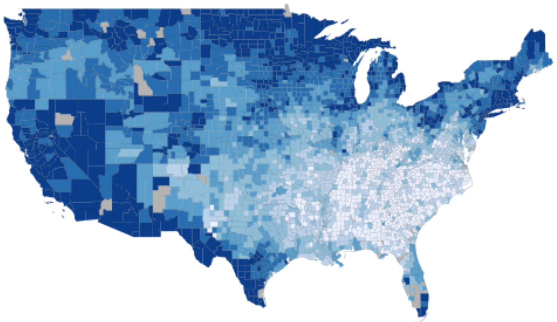
Notes: The administrative data come from Willcox (1929) and U.S. Department of Commerce (1924, 1929, 1931). Figure 1a shows the annual inflows and Figure 1b shows Southern and Eastern Europeans as a fraction of total immigrants.

(a) Southern and Eastern Europe

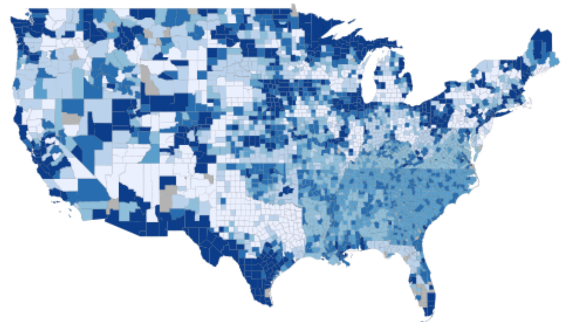
(b) Northern and Western Europe

Figure 2. IMMIGRATION INFLOWS UNDER QUOTAS BY REGION

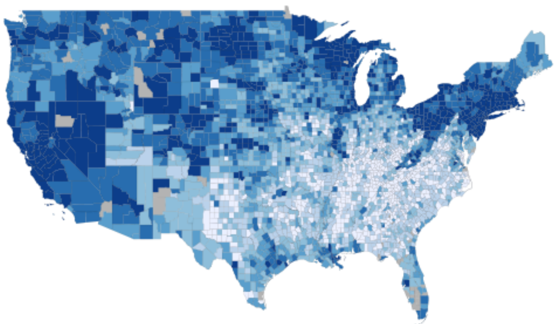
Notes: This figure is a replication of Figure 1 of Ager and Hanson (2018), modified through aggregating immigration into two groups: Southern and Eastern Europe, and Northern and Western Europe. The data from this replication come from Willcox (1929) and U.S. Department of Commerce (1924, 1929, 1931).



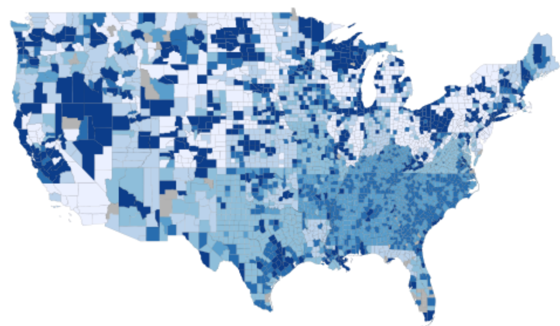
(a) Total foreign born as a fraction of total population



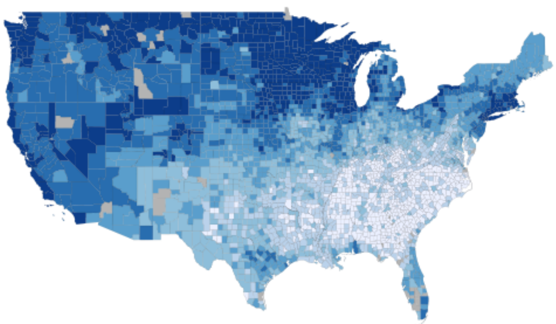
(b) State fixed effects



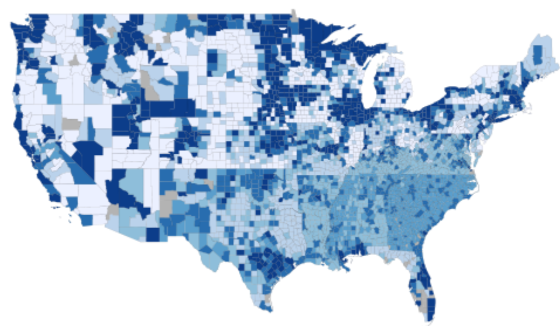
(c) Fraction of foreign born from Southern and Eastern Europe



(d) State fixed effects



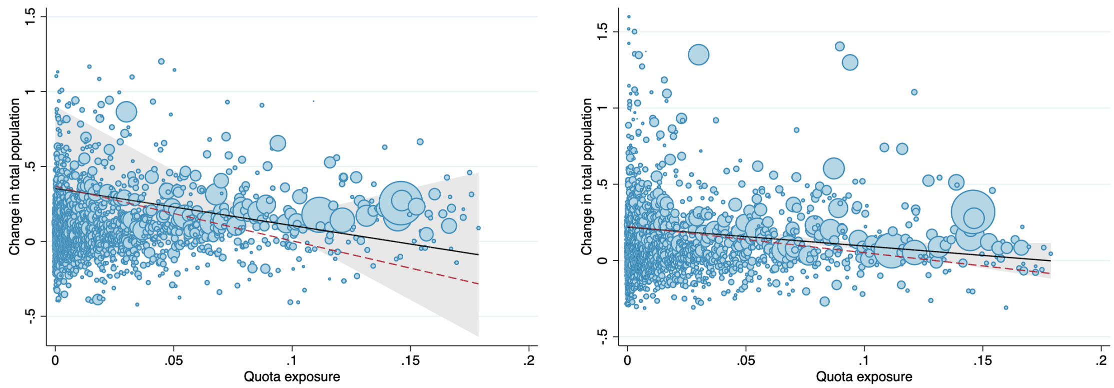
(e) Fraction of foreign born from Northern and Western Europe



(f) State fixed effects

Figure 3. GEOGRAPHIC DISTRIBUTION OF FOREIGN BORN

Notes: The figures show the share of foreign born population in each U.S. county in 1920; in Figure 4a total foreign born; in Figure 4c from Southern and Eastern Europe; in Figure 4e from Northern and Western Europe. Figure 4b, 4d, and 4f are the share of population with state fixed effects, respectively.

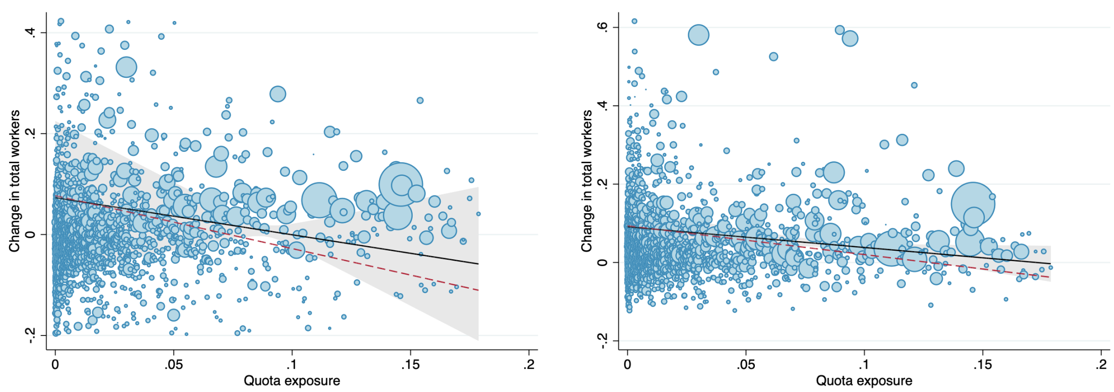


(a) Change between 1910 and 1920 as a fraction of 1910 population

(b) Change between 1920 and 1930 as a fraction of 1920 population

Figure 4. CHANGE IN TOTAL POPULATION

Notes: The figure shows the change in total population (change between 1910 and 1920 in Panel (a); change between 1920 and 1930 in Panel (b)) against the quota exposure. The solid line shows the coefficient from the regression of change on quota exposure and the red dot line shows the regression coefficient after the top 1% of highly quota exposed areas is excluded. Marker size represents the county population.

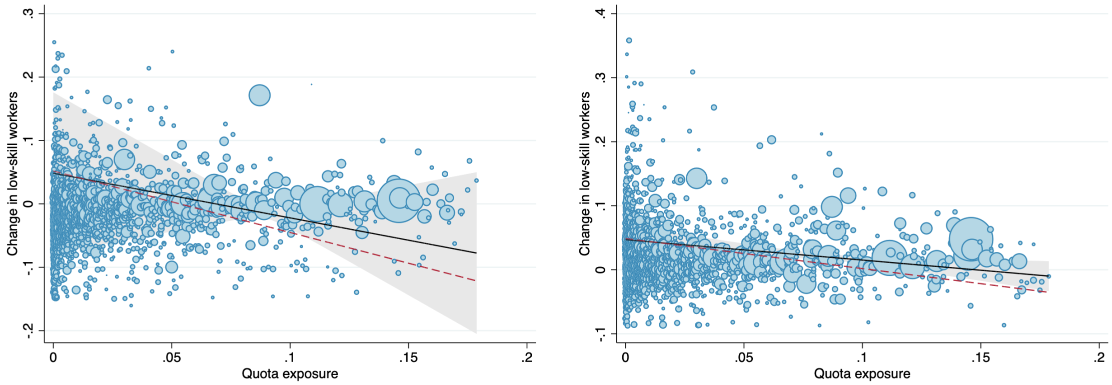


(a) Change between 1910 and 1920 as a fraction of 1910 population

(b) Change between 1920 and 1930 as a fraction of 1920 population

Figure 5. CHANGE IN TOTAL WORKER

Notes: The figure shows the change in total worker (change between 1910 and 1920 in Panel (a); change between 1920 and 1930 in Panel (b)) against the quota exposure. The solid line shows the coefficient from the regression of change on quota exposure and the red dot line shows the regression coefficient after the top 1% of highly quota exposed counties is excluded. Marker size represents the county population.

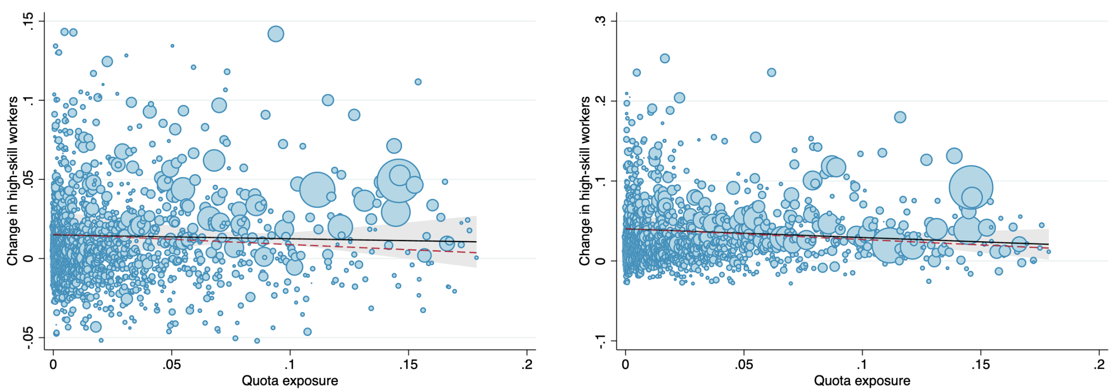


(a) Change between 1910 and 1920 as a fraction of 1910 population

(b) Change between 1920 and 1930 as a fraction of 1920 population

Figure 6. CHANGE IN LOW-SKILL WORKER

Notes: The figure shows the change in low-skill worker (change between 1910 and 1920 in Panel (a); change between 1920 and 1930 in Panel (b)) against the quota exposure. The solid line shows the coefficient from the regression of change on quota exposure and the red dot line shows the regression coefficient after the top 1% of highly quota exposed counties is excluded. Marker size represents the county population.



(a) Change between 1910 and 1920 as a fraction of 1910 population

(b) Change between 1920 and 1930 as a fraction of 1920 population

Figure 7. CHANGE IN HIGH-SKILL WORKER

Notes: The figure shows the change in high-skill worker (change between 1910 and 1920 in Panel (a); change between 1920 and 1930 in Panel (b)) against the quota exposure. The solid line shows the coefficient from the regression of change on quota exposure and the red dot line shows the regression coefficient after the top 1% of highly quota exposed counties is excluded. Marker size represents the county population.

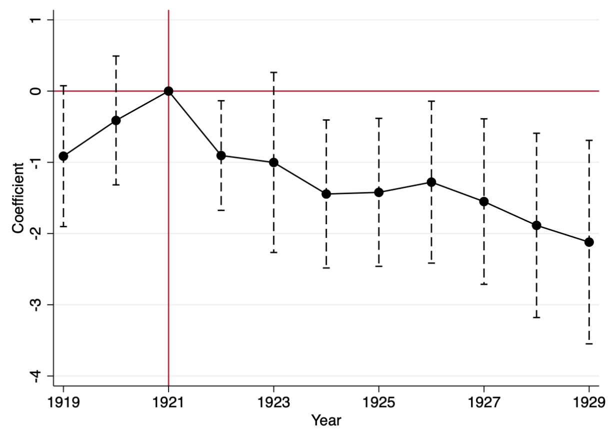


Figure 8. DIFFERENCE-IN-DIFFERENCES IN INVENTIONS AT THE COUNTY LEVEL

Notes: The figures show the estimated coefficients (with corresponding 95% confidence intervals) relative to the year 1921 from the difference-in-differences specification. The dependent variable is the number of patent application per year at the county level rescaled by the workers. Standard errors are clustered by counties.

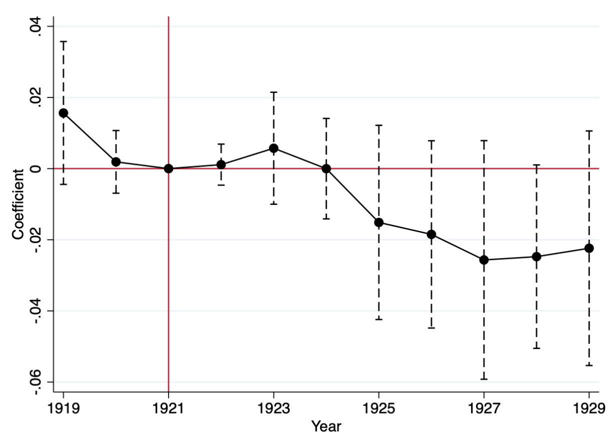


Figure 9. DIFFERENCE-IN-DIFFERENCES IN INVENTIONS AT THE INDUSTRY LEVEL

Notes: The figures show the estimated coefficients (with corresponding 95% confidence intervals) relative to the year 1921 from the difference-in-differences specification. The dependent variable is the number of patent applications per year in industry rescaled by the total number of workers in that industry. Standard errors are clustered by industries.

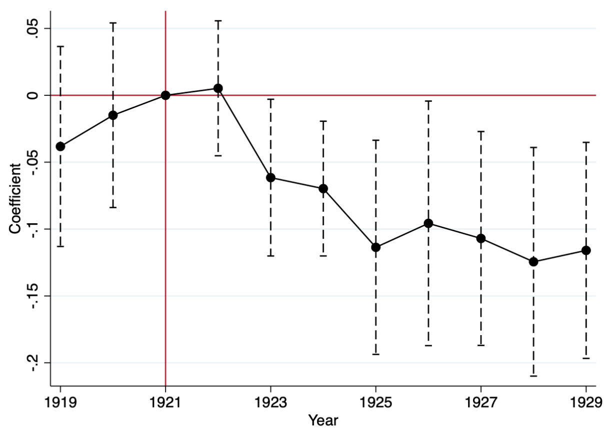


Figure 10. DIFFERENCE-IN-DIFFERENCES IN INVENTIONS AT THE INDIVIDUAL LEVEL

Notes: The figures show the estimated coefficients (with corresponding 95% confidence intervals) relative to the year 1921 from the difference-in-differences specification. The sample consists of incumbent inventors who had at least one patent in 1910. The dependent variable is the number of patent applications per year by inventors, winsorized at 10. Standard errors are clustered by counties.

Table 1. QUOTAS BY COUNTRY

Country	Quota	Actual immigrants	Missing immigrants	1920 Population in thousands	Fraction of missing immigrants
	(1)	(2)	(3)	(4)	(5)
A. Southern and Eastern Europe					
Austria	3,065	2,756	66,145	689	0.096
Bulgaria	167	160	7,600	10	0.781
Czechoslovakia	6,804	6,742	3,112	319	0.010
Greece	1,162	1,177	37,909	160	0.237
Hungary	2,251	2,279	67,420	407	0.166
Italy	16,800	16,655	187,287	1,609	0.116
Poland	13,820	13,594	129,258	1,135	0.114
Portugal	1,156	1,143	12,627	113	0.112
Romania	2,841	2,839	0	92	0.000
Russia	10,791	10,127	163,786	1,424	0.115
Spain	405	400	8,948	50	0.179
Turkey	714	760	47,282	27	1.767
Yugoslavia	2,609	2,598	31,160	128	0.244
Total	62,584	61,231	762,535	6,163	0.303
B. Northern and Western Europe					
Belgium	950	931	5,918	65	0.091
Denmark	3,562	3,155	433	186	0.002
Finland	1,632	1,532	3,067	151	0.020
France	4,449	4,084	4,502	155	0.029
Germany	53,929	45,165	0	1,633	0.000
Ireland	27,377	21,584	0	1,051	0.000
Netherlands	2,468	2,258	6,740	133	0.051
Norway	7,916	7,048	0	367	0.000
Sweden	12,361	10,758	0	631	0.000
Switzerland	2,596	2,500	255	121	0.002
UK	42,453	37,920	20,446	1,159	0.018
Total	159,695	136,934	41,361	5,651	0.019

Notes: This table shows information on quotas for countries restricted by quota limits. In columns (1), (2), and (3), the variable is calculated as the average number per year during the quotas, 1922-1930. Missing immigrants are estimated by the difference between average estimated immigrants per year without quotas based on immigration flows from 1900 and 1914 before the WWI and average actual quota limits per year. Column (5) reports the average missing immigrants as a fraction of 1920 population in that country.

Table 2. SUMMARY STATISTICS AT THE COUNTY LEVEL

	Means				Difference
	All	Low FB	High FB		
			Low quotas	High quotas	(4) - (3)
	(1)	(2)	(3)	(4)	(5)
A: Baseline characteristics					
Population	19,245	19,377	18,495	19,137	643
					(1237)
Workers	5,004	4,995	4,861	5,205	344
					(322)
Foreign born workers	625	282	1,730	1,741	11
					(107)
Low-skill workers	1,979	1,959	1,815	2,272	457
					(143)
Medium- and high-skill workers	3,024	3,035	3,046	2,933	-113
					(192)
B: Outcomes					
Immigration inflows per year	15	6	36	49	13
					(8)
Immigration inflows per year as a fraction of workers	0.002	0.001	0.006	0.005	-0.001
					(0.001)
Patent applications per year	5	3	8	12	5
					(2)
Patent applications per year as a fraction of workers	0.568	0.431	0.960	1.064	0.104
					(0.120)
Number of industries	2513	1921	296	296	592

Notes: This table describes baseline characteristics in the 1900 census and outcomes before the quota. Means of variables appear in columns (1), (2), (3), and (4). Workers are defined as people aged from 16 to 64 in the labor force. Skill-level of workers are defined as a variable of edscor50 that presents the percentage of people in the occupational category who had completed one or more years of college. The median of edscor50 of the entire workers in the 1900 census is used as a threshold for determining low-skill workers and median- and high-skill workers. Immigration inflows are the average of new immigrants per county and year before the Immigration Act of 1924 and after the end of WWI (between 1919 and 1923) and its fraction is divided by total workers in each county in the 1910 census. Patent applications per county and year are computed in the same way. We restrict data to counties that exist in all censuses from 1900 to 1930. Column (2) reports statistics for counties with less foreign born workers (below the 75th percentile of a fraction of foreign born workers in the 1900 census). Columns (3) and (4) use high-quota-exposure counties (above the median of quota exposure variable) and low-quota-exposure counties (below the median of quota exposure variable) in a subsample of counties that had more foreign born workers, respectively. Column (5) shows differences and standard errors (in parenthesis) between high-quota-exposure counties and low-quota-exposure counties in the subsample.

Table 3. THE EFFECT OF THE QUOTAS ON IMMIGRATION INFLOWS AT THE COUNTY LEVEL

	All counties		Counties with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Immigration inflows in county				
Quota exposure × Post-treatment	-0.074***	-0.073***	-0.049***	-0.048***
	(0.004)	(0.004)	(0.005)	(0.005)
Dependent variable mean	0.004	0.004	0.009	0.009
Number of observations	83,250	83,250	20,970	20,970
Number of counties	2,775	2,775	699	699
R-squared	0.527	0.523	0.558	0.555
B. Dependent variable: Inflow of low-skill immigrants in county				
Quota exposure × Post-treatment	-0.061***	-0.058***	-0.041***	-0.039***
	(0.004)	(0.004)	(0.005)	(0.005)
Dependent variable mean	0.003	0.003	0.006	0.006
Number of observations	83,250	83,250	20,970	20,970
Number of counties	2,775	2,775	699	699
R-squared	0.482	0.476	0.539	0.535
C. Dependent variable: Inflow of high-skill immigrants in county				
Quota exposure × Post-treatment	-0.013***	-0.015***	-0.008***	-0.009***
	(0.001)	(0.001)	(0.002)	(0.002)
Dependent variable mean	0.001	0.001	0.003	0.003
Number of observations	83,250	83,250	20,970	20,970
Number of counties	2,775	2,775	699	699
R-squared	0.541	0.541	0.510	0.510

Notes: This table presents the coefficients from the difference-in-differences specification. The outcome variable is new immigrants per year and city as a fraction of 1910 workers. This variable is constructed by combining information from the 1910, 1920, and 1930 U.S. Census. Specifically, new immigrants per year between the years 1900 and 1909 are obtained from the 1910 Census data, those between 1910 and 1919 from the 1920 U.S. Census, etc. We restrict data to counties that exist in all three censuses to obtain a balanced panel. In columns (3) and (4), we use a subsample of counties that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). In panel B, we use the outcome of new arrived low-skill immigrants (below the median of edscor50 that presents the percentage of people in the occupational category who had completed one or more years of college) and panel C uses new arrived (medium- and) high-skill immigrants (above or equal to the median). Standard errors are clustered by counties.

Table 4. THE EFFECT OF THE QUOTAS ON INVENTIONS AT THE COUNTY LEVEL

	All counties		Counties with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Patent applications in county				
Quota exposure × Post-treatment	-0.781**	-0.814**	-0.925*	-0.974*
	(0.387)	(0.377)	(0.512)	(0.563)
Dependent variable mean	0.673	0.689	1.273	1.292
Number of observations	30,778	30,778	7,689	7,689
Number of counties	2,798	2,798	699	699
R-squared	0.766	0.766	0.784	0.784
B. Dependent variable: Citation-weighted patent applications in county				
Quota exposure × Post-treatment	-1.457	-2.032	-4.717**	-6.357***
	(1.416)	(1.499)	(1.896)	(2.146)
Dependent variable mean	1.409	1.499	2.756	2.915
Number of observations	30,778	30,778	7,689	7,689
Number of counties	2,798	2,798	699	699
R-squared	0.617	0.617	0.677	0.677

Notes: This table presents the coefficients from the difference-in-differences specification. The outcome variable is the number of patent applications per year at the county level rescaled by the 1910 workers in county. The data from the HistPat database (Petralia et al., 2016) is used and linked to PATSTAT database to identify the year of patent applications. In columns (3) and (4), we use a subsample of counties that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). In panel B, we use the outcome of citation-weighted patent applications per year in county. Standard errors are clustered by counties.

Table 5. THE EFFECT OF THE QUOTAS ON THE DIRECTION OF INVENTIONS AT THE COUNTY LEVEL

	All counties		Counties with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Index of dissimilarity in county				
Quota exposure × Post-treatment	0.222**	0.212***	0.358***	0.316***
	(0.087)	(0.079)	(0.114)	(0.104)
Dependent variable mean	0.674	0.674	0.741	0.736
Number of observations	29,458	29,458	7,370	7,370
Number of counties	2,678	2,678	670	670
R-squared	0.158	0.158	0.179	0.179
B. Dependent variable: Original patent applications in county				
Quota exposure × Post-treatment	2.642***	2.486**	2.070*	2.242**
	(0.920)	(0.967)	(1.080)	(1.091)
Dependent variable mean	0.924	0.955	1.777	1.851
Number of observations	30,778	30,778	7,689	7,689
Number of counties	2,798	2,798	699	699
R-squared	0.220	0.220	0.691	0.691

Notes: This table presents the coefficients from the difference-in-differences specification. The index of dissimilarity is based on Borjas and Doran (2012) and Cutler and Glaeser (1997). If the preexisting patent classifications between 1900 and 1919 and patent classifications per county and year never match, then the index will be one. The index takes a value of zero if they perfectly match. In panel B, original patent applications are defined as patents with new one/two/three word phrases in title that did not exist in preexisting patents in previous years, weighted by citations and workers in county. In columns (3) and (4), we use a subsample of counties that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). Standard errors are clustered by counties.

Table 6. SUMMARY STATISTICS AT THE INDUSTRY LEVEL

	Means				Difference
	All	Low FB	High FB		
			Low quotas	High quotas	(4) - (3)
	(1)	(2)	(3)	(4)	(5)
A: Baseline characteristics					
Workers	229,593	275,266	79,598	84,872	5,274
					(51,100)
Foreign born workers	41,027	42,376	34,626	39,055	4,429
					(22,276)
Low-skill workers	99,335	107,867	69,969	73,872	3,904
					(52,522)
Medium- and high-skill workers	130,257	167,399	9,629	11,000	1,371
					(7,116)
B: Outcomes					
Immigration inflows per year	1,526	1,474	1,941	1,404	-536
					(1,006)
Immigration inflows per year as a fraction of workers	0.013	0.013	0.012	0.012	0.001
					(0.004)
Patent applications per year	769	891	318	439	121
					(200)
Patent applications per year as a fraction of workers	0.046	0.059	0.005	0.006	0.001
					(0.004)
Number of industries	55	42	7	6	13

Notes: This table describes baseline characteristics of industries in the 1900 census and outcomes before the quota. Means of variables appear in columns (1), (2), (3), and (4). Workers are defined as people aged from 16 to 64 in the labor force. Skill-level of workers are defined as a variable of edscor50 that presents the percentage of people in the occupational category who had completed one or more years of college. The median of edscor50 of the entire workers in the 1900 census is used as a threshold for determining low-skill workers and median- and high-skill workers. Immigration inflows are the average of new immigrants per industry and year before the Immigration Act of 1924 and after the end of WWI (between 1919 and 1923) and its fraction is divided by total workers in each industry in the 1910 census. Patent applications per industry and year are computed in the same way. Column (2) reports statistics for industries with less foreign born workers (below the 75th percentile of a fraction of foreign born workers in the 1900 census). Columns (3) and (4) use high-quota-exposure industries (above the median of quota exposure variable) and low-quota-exposure industries (below the median of quota exposure variable) in a subsample of industries that had more foreign born workers, respectively. Column (5) shows differences and standard errors (in parenthesis) between high-quota-exposure industries and low-quota-exposure industries in the subsample.

Table 7. THE EFFECT OF THE QUOTAS ON IMMIGRATION INFLOWS AT THE INDUSTRY LEVEL

	All industries		Industries with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Immigration inflows in industry				
Quota exposure × Post-treatment	-0.027**	-0.015**	-0.017**	-0.016**
	(0.013)	(0.007)	(0.006)	(0.007)
Dependent variable mean	0.011	0.013	0.011	0.012
Number of observations	605	605	143	143
Number of industries	55	55	13	13
R-squared	0.779	0.774	0.806	0.807
B. Dependent variable: Inflow of low-skill immigrants in industry				
Quota exposure × Post-treatment	-0.017***	-0.011***	-0.014***	-0.014**
	(0.006)	(0.003)	(0.004)	(0.005)
Dependent variable mean	0.008	0.009	0.008	0.009
Number of observations	605	605	143	143
Number of industries	55	55	13	13
R-squared	0.807	0.803	0.811	0.814
C. Dependent variable: Inflow of high-skill immigrants in industry				
Quota exposure × Post-treatment	-0.010	-0.004	-0.003	-0.002
	(0.007)	(0.004)	(0.003)	(0.003)
Dependent variable mean	0.003	0.004	0.003	0.003
Number of observations	605	605	143	143
Number of industries	55	55	13	13
R-squared	0.721	0.717	0.803	0.802

Notes: This table presents the coefficients from the difference-in-differences specification. The dependent variable is the number of newly arrived immigrants per year in industry rescaled by the total number of workers in that industry in the 1910 census. This outcome variable is constructed by combining the 1910, 1920, and 1930 US censuses. Specifically, new immigrants per year between the years 1900 and 1909 are obtained from the 1910 census, those between 1910 and 1919 from the 1920 census, etc. We restrict data to industries that exist in all three censuses and have patenting applications. Columns (3) and (4), we use a subsample of industries that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). In panel B, we use the outcome of new arrived low-skill immigrants (below the median of edscor50 that presents the percentage of people in the occupational category who had completed one or more years of college) and panel C uses new arrived (medium- and) high-skill immigrants (above or equal to the median). The sample covers the years from 1919 to 1929. Standard errors are clustered by industries.

Table 8. THE EFFECT OF THE QUOTAS ON INVENTIONS AT THE INDUSTRY LEVEL

	All industries		Industries with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Patent applications in industry				
Quota exposure × Post-treatment	-0.014	-0.028	-0.005*	-0.006*
	(0.011)	(0.023)	(0.003)	(0.003)
Dependent variable mean	0.046	0.046	0.006	0.006
Number of observations	605	605	143	143
Number of industries	55	55	13	13
R-squared	0.983	0.983	0.975	0.977
B. Dependent variable: Citation-weighted patent applications in industry				
Quota exposure × Post-treatment	-0.310	-0.342	-0.028***	-0.035***
	(0.258)	(0.288)	(0.009)	(0.012)
Dependent variable mean	0.110	0.115	0.014	0.014
Number of observations	605	605	143	143
Number of industries	55	55	13	13
R-squared	0.945	0.945	0.941	0.947

Notes: This table presents the coefficients from the difference-in-differences specification. The dependent variable is the number of patent applications per year in industry rescaled by the total number of workers in that industry in the 1910 census. To match patent classifications with the industry classifications in the census (the variable of ind1950), we use an IPC (International Patent Classification) to NAICS (the North American Industry Classification System) concordance (Lybberta and Zolas, 2014). Columns (3) and (4), we use a subsample of industries that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). In panel B, we use the outcome of citation-weighted patent applications per year in industry. The sample covers the years from 1919 to 1929. Standard errors are clustered by industries.

Table 9. THE EFFECT OF THE QUOTAS ON THE DIRECTION OF INVENTIONS AT THE INDUSTRY LEVEL

	All industries		Industries with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Index of dissimilarity				
Quota exposure × Post-treatment	0.164**	0.186*	0.208*	0.387*
	(0.072)	(0.106)	(0.101)	(0.197)
Dependent variable mean	0.789	0.763	0.801	0.780
Number of observations	605	605	143	143
Number of industries	55	55	13	13
R-squared	0.695	0.698	0.472	0.509
B. Dependent variable: Original patent applications in industry				
Quota exposure × Post-treatment	0.026	0.015	0.002	0.001
	(0.019)	(0.010)	(0.001)	(0.001)
Dependent variable mean	0.005	0.004	0.001	0.001
Number of observations	605	605	143	143
Number of industries	55	55	13	13
R-squared	0.865	0.863	0.780	0.776

Notes: This table presents the coefficients from the difference-in-differences specification. The index of dissimilarity is based on Borjas and Doran (2012) and Cutler and Glaeser (1997). If the preexisting patent classifications before WWI (1900-1919) and patent classifications per industry and year never match, then the index will be one. The index takes a value of zero if they perfectly match. In panel B, original patent applications are defined as patents with new one/two/three word phrases in title that did not exist in preexisting patents in previous years. Columns (3) and (4), we use a subsample of industries that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). The sample covers the years from 1919 to 1929. Standard errors are clustered by industries.

Table 10. THE EFFECT OF THE QUOTAS ON INVENTIONS AT THE INDIVIDUAL LEVEL

	All counties		Counties with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Patent applications by inventors in county				
Quota exposure × Post-treatment	-0.067	-0.082	-0.044	-0.069
	(0.026)	(0.024)	(0.042)	(0.037)
Dependent variable mean	0.081	0.079	0.095	0.090
Number of observations	843,408	843,408	200,824	200,824
Number of inventors	78,055	78,055	18,480	18,480
Number of counties	2,766	2,766	699	699
R-squared	0.444	0.444	0.411	0.411
B. Dependent variable: Citation-weighted patent applications by inventors in county				
Quota exposure × Post-treatment	-0.124	-0.179	-0.197	-0.209
	(0.093)	(0.062)	(0.140)	(0.091)
Dependent variable mean	0.184	0.183	0.215	0.210
Number of observations	843,408	843,408	200,824	200,824
Number of inventors	78,055	78,055	18,480	18,480
Number of counties	2,766	2,766	699	699
R-squared	0.255	0.255	0.234	0.234

Notes: This table presents the coefficients from the difference-in-differences specification. The outcome variable is the number of patent applications per year by native-born incumbent inventors who already had at least one patent as of 1910, winsorized at 10. In columns (3) and (4), we use a subsample of inventors in counties that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). In panel B, we use the outcome of citation-weighted patent applications per year by inventors, winsorized at 20. Standard errors are clustered by counties.

Table A.1. PRE-TRENDS IN THE QUOTAS AND INVENTIONS

	All counties		Counties with high FB	
	Before the quota			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Patent applications in county				
Quota exposure × Year	0.005	0.006	-0.002	0.006
	(0.009)	(0.005)	(0.009)	(0.005)
Dependent variable mean	0.673	0.689	1.273	1.292
Number of observations	8,394	13,990	2,097	3,495
Number of counties	2,798	2,798	699	699
R-squared	0.797	0.782	0.826	0.810
B. Dependent variable: Citation-weighted patent applications in county				
Quota exposure × Year	-0.013	-0.005	-0.027	-0.010
	(0.020)	(0.010)	(0.023)	(0.012)
Dependent variable mean	1.409	1.499	2.756	2.915
Number of observations	8,394	13,990	2,097	3,495
Number of counties	2,798	2,798	699	699
R-squared	0.639	0.651	0.777	0.769

Notes: This table reports the estimated pre-trends in the relationship between the quota exposure and inventions before the quota using the sample period of 1919-1921 in columns 1 and 3, and 1919-1923 in columns 2 and 4. The dependent variable is the number of patent applications per year at the county level rescaled by the 1910 workers in county. In panel B, we use citation-weighted patent applications per year in county. In columns 3 and 4, we use a subsample of counties that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). Standard errors are clustered at the county level.

Table A.2. PLACEBO TEST USING COUNTIES WITH RELATIVELY FEW IMMIGRANTS BEFORE THE QUOTA

	Post-treatment year	
	1922	1924
	(1)	(2)
A. Dependent variable: Patent applications in county		
Quota exposure × Post-treatment	0.961	0.410
	(0.789)	(0.856)
Dependent variable mean	0.473	0.488
Number of observations	23,089	23,089
Number of counties	2,099	2,099
R-squared	0.675	0.675
B. Dependent variable: Citation-weighted patent applications in county		
Quota exposure × Post-treatment	4.085	2.791
	(2.935)	(2.927)
Dependent variable mean	0.960	1.027
Number of observations	23,089	23,089
Number of counties	2,099	2,099
R-squared	0.448	0.448

Notes: This table reports the estimated coefficients using a sample of counties that had relatively few foreign born workers (below the 75th percentile of a fraction of foreign born workers). The dependent variable is the number of patent applications per year at the county level rescaled by the 1910 workers in county. In panel B, we use citation-weighted patent applications per year in county. Standard errors are clustered at the county level.

Table A.3. THE EFFECT OF THE QUOTAS ON INVENTIONS BY ALL INDIVIDUALS (INCLUDING THOSE WHO HAD NEVER PATENTED BEFORE THE QUOTA)

	Post-treatment year	
	1922	1924
	(1)	(2)
A. Dependent variable: Patents, all individuals		
Quota exposure × Post-treatment	-0.00016	-0.00018
	(0.00006)	(0.00006)
Dependent variable mean	0.00039	0.00039
Number of observations	652,474,742	652,474,742
Number of individuals	70,746,414	70,746,414
Number of counties	3,065	3,065
R-squared	0.2994	0.2994
B. Dependent variable: Becoming an inventor, all individuals		
Quota exposure × Post-treatment	0.00146	0.00135
	(0.00033)	(0.00032)
Dependent variable mean	0.00298	0.00298
Number of observations	652,474,742	652,474,742
Number of individuals	70,746,414	70,746,414
Number of counties	3,065	3,065
R-squared	0.9275	0.9275
C. Dependent variable: Patents, all individuals who had no patent before the quota		
Quota exposure × Post-treatment	0.00052	0.00045
	(0.00012)	(0.00011)
Dependent variable mean	0.00015	0.00010
Number of observations	650,751,093	650,603,668
Number of individuals	70,585,845	70,571,886
Number of counties	3,065	3,065
R-squared	0.1264	0.1074
D. Dependent variable: Becoming an inventor, all individuals who had no patent before the quota		
Quota exposure × Post-treatment	0.00110	0.00084
	(0.00026)	(0.00021)
Dependent variable mean	0.00037	0.00021
Number of observations	650,751,093	650,603,668
Number of individuals	70,585,845	70,571,886
Number of counties	3,065	3,065
R-squared	0.5074	0.3823

Notes: This table presents the coefficients from the difference-in-differences specification. We use a sample of the population aged 18-80 from the complete count 1920 census. The outcome variable is the number of patent applications per year, winsorized at 10. Another outcome variable is the probability of becoming an inventor that takes a value of one if a person has at least one patent in or before year. In panels C and D, we use a sample of individuals who had no patent before the quota. Standard errors are clustered by counties.

Table A.4. HOW THE EFFECT OF THE QUOTAS ON INVENTIONS VARIES WITH AGE

	Post-treatment year 1922	Post-treatment year 1924
	(1)	(2)
A. Dependent variable: Patents, all individuals, aged 18-30		
Quota exposure × Post-treatment	0.00022	0.00027
	(0.00011)	(0.00011)
Dependent variable mean	0.00034	0.00032
Number of observations	225,946,801	224,708,146
Number of individuals	20,990,887	21,194,242
Number of counties	3,065	3,065
R-squared	0.2862	0.2800
B. Dependent variable: Becoming an inventor, all individuals, aged 18-30		
Quota exposure × Post-treatment	0.00122	0.00115
	(0.00034)	(0.00032)
Dependent variable mean	0.00158	0.00143
Number of observations	225,946,801	224,708,146
Number of individuals	20,990,887	21,194,242
Number of counties	3,065	3,065
R-squared	0.8499	0.8387
C. Dependent variable: Patents, all individuals, aged 31-80		
Quota exposure × Post-treatment	-0.00046	-0.00052
	(0.00011)	(0.00010)
Dependent variable mean	0.00046	0.00046
Number of observations	341,871,187	358,571,840
Number of individuals	31,974,023	33,462,680
Number of counties	3,065	3,065
R-squared	0.3172	0.3178
D. Dependent variable: Becoming an inventor, all individuals, aged 31-80		
Quota exposure × Post-treatment	0.00167	0.00156
	(0.00032)	(0.00033)
Dependent variable mean	0.00454	0.00445
Number of observations	341,871,187	358,571,840
Number of individuals	31,974,023	33,462,680
Number of counties	3,065	3,065
R-squared	0.9509	0.9497

Notes: This table presents the coefficients from the difference-in-differences specification. The sample is partitioned into two subsamples: those aged 18-30 and those aged 31-80. The outcome variable is the number of patent applications per year, winsorized at 10. Another outcome variable is the probability of becoming an inventor that takes a value of one if a person has at least one patent in or before year. Standard errors are clustered by counties.

Table A.5. THE EFFECT OF THE QUOTAS ON INFLOWS OF IMMIGRANT INVENTORS

	Year of immigration			
	1900-1929		1919-1929	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: New immigrant inventors as a fraction of 1920 inventors				
Quota exposure × Post-treatment	-0.0001	-0.0003	-0.0002	-0.0005
	(0.0003)	(0.0003)	(0.0004)	(0.0003)
Dependent variable mean	0.0007	0.0007	0.0007	0.0007
Number of observations	54,000	54,000	19,800	19,800
Number of counties	1,800	1,800	1,800	1,800
R-squared	0.0038	0.0038	0.0077	0.0079
B. New immigrant inventors from Northern and Western Europe as a fraction				
Quota exposure × Post-treatment	0.0001	-0.0000	-0.0002	-0.0003
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Dependent variable mean	0.0005	0.0005	0.0005	0.0005
Number of observations	54,000	54,000	19,800	19,800
Number of counties	1,800	1,800	1,800	1,800
R-squared	0.0041	0.0041	0.0048	0.0049
C. New immigrant inventors from Southern and Eastern Europe as a fraction				
Quota exposure × Post-treatment	-0.0001	-0.0001	-0.0000	-0.0001
	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Dependent variable mean	0.0001	0.0001	0.0001	0.0001
Number of observations	54,000	54,000	19,800	19,800
Number of counties	1,800	1,800	1,800	1,800
R-squared	0.0032	0.0032	0.0218	0.0220

Notes: This table presents the coefficients from the difference-in-differences specification. The sample consists of inventor migrants who had patented before they immigrated to the U.S. The outcome variable of new immigrant inventors is constructed by merging information from the 1930 U.S. Census. Standard errors are clustered by counties.

Table A.6. HOW THE QUOTAS AFFECT INVENTIONS BY SKILL LEVEL AT THE INDUSTRY LEVEL

	All industries		Industries with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Patent applications in low-skill industry				
Quota exposure × Post-treatment	-0.016	-0.037	-0.006	-0.007
	(0.017)	(0.039)	(0.004)	(0.005)
Dependent variable mean	0.057	0.056	0.003	0.003
Number of observations	308	308	88	88
Number of industries	28	28	8	8
R-squared	0.982	0.982	0.976	0.978
B. Dependent variable: Patent applications in high-skill industry				
Quota exposure × Post-treatment	-0.004	-0.009	0.010	0.013
	(0.018)	(0.022)	(0.014)	(0.015)
Dependent variable mean	0.035	0.036	0.010	0.010
Number of observations	297	297	55	55
Number of industries	27	27	5	5
R-squared	0.993	0.993	0.942	0.945

Notes: This table presents the coefficients from the difference-in-differences specification. The dependent variable is the number of patent applications per year in industry rescaled by the total number of workers in that industry in the 1910 census. To match patent classifications with the industry classifications in the census (the variable of ind1950), we use an IPC (International Patent Classification) to NAICS (the North American Industry Classification System) concordance (Lybberta and Zolas, 2014). Columns (3) and (4), we use a subsample of industries that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). In panel A, we restrict industries to those with a large fraction of low-skill workers (below the median of edscor50 that presents the percentage of people in the occupational category who had completed one or more years of college) and panel B uses industries with (medium- and) high-skill workers (above or equal to the median). The sample covers the years from 1919 to 1929. Standard errors are clustered by industries.

Table A.7. HOW THE QUOTAS AFFECT INVENTIONS BY LABOR INTENSITY AT THE INDUSTRY LEVEL

	All industries		Industries with high FB	
	Post-treatment year			
	1922	1924	1922	1924
	(1)	(2)	(3)	(4)
A. Dependent variable: Patent applications in more labor-intensive industry				
Quota exposure × Post-treatment	-0.029	-0.057	-0.007	-0.009
	(0.021)	(0.045)	(0.005)	(0.006)
Dependent variable mean	0.083	0.082	0.007	0.007
Number of observations	308	308	77	77
Number of industries	28	28	7	7
R-squared	0.983	0.983	0.974	0.976
B. Dependent variable: Patent applications in less labor-intensive industry				
Quota exposure × Post-treatment	0.002	0.003	-0.004	-0.005
	(0.004)	(0.005)	(0.003)	(0.004)
Dependent variable mean	0.008	0.008	0.005	0.005
Number of observations	297	297	66	66
Number of industries	27	27	6	6
R-squared	0.964	0.964	0.971	0.972

Notes: This table presents the coefficients from the difference-in-differences specification. The dependent variable is the number of patent applications per year in industry rescaled by the total number of workers in that industry in the 1910 census. To match patent classifications with the industry classifications in the census (the variable of ind1950), we use an IPC (International Patent Classification) to NAICS (the North American Industry Classification System) concordance (Lybberta and Zolas, 2014). Columns (3) and (4), we use a subsample of industries that had more foreign born workers (above the 75th percentile of a fraction of foreign born workers). In panel A, we restrict industries to those with a large fraction of labor cost (above the median of occscore that presents occupational income scores) and panel B uses industries with a small fraction of labor cost (below or equal to the median). The sample covers the years from 1919 to 1929. Standard errors are clustered by industries.