Papers in Evolutionary Economic Geography

# 18.35

# Migration and Invention in the Age of Mass Migration

Dario Diodato, Andrea Morrison and Sergio Petralia

Utrecht University Human Geography and Planning

# Migration and Invention in the Age of Mass Migration

Dario Diodato*†, Andrea Morrison ‡§¶& Sergio Petralia ‖§

February 26, 2020

**Abstract**

More than 30 million people migrated to the US between the 1850s and 1920s and in order of thousands became inventors and patentees. Drawing on a novel dataset of immigrant inventors in the US, we assess the city-level impact of immigrants' patenting and their contribution to the technological specialization of the receiving US regions between 1870 and 1940. Our results show that native inventors benefited from the inventive activity of immigrants. We find that immigrant inventors imported knowledge from their home country, which generated positive local spill-overs. In addition, we show that the knowledge transferred by immigrants gave rise to new and previously not exiting technological fields in the US regions where immigrants moved to. Our findings are robust to several checks and the implementation of an instrumental variable strategy.

*CID Growth Lab, Harvard University, 79 JFK street, 02138 Cambridge (MA), USA

†Present address: European Commission, Joint Research Centre (JRC), Calle Inca Garcilaso 3, 41092 Seville, Spain

‡ICRIOS-Bocconi University, Department of Management and Technology, Via Roentgen 1, 20136, Milano, Italy

§Utrecht University, Department of Human Geography and Planning, Princetonlaan 8A, 3584 CB Utrecht, the Netherlands

¶corresponding author, email: a.morrison@unibocconi.it

‖Department of Economic Geography, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom

Keywords: immigration, innovation, knowledge spill-over, patent, age of mass migration, US.

JEL Classification: : F22, J61, O31, R3.

# 1 Introduction

Between 1850 and the mid-1920s more than thirty million people migrated to the US in search of a better life (Bandiera, Rasul, and Viarengo, 2013). The causes and economic impact of this mass migration have received already a good deal of attention in the literature (Hatton, Williamson, et al., 1998). More recently, also due to the backlash against immigration, this topic has regained popularity among scholars, who have initiated a new research line on the economic impact of historical migration in the US (Abramitzky and Boustan, 2017; Hatton and Ward, 2018; Rodriguez-Pose and Von Berlepsch, 2014; Sequeira, Nunn, and Qian, 2020; Tabellini, 2020). However, as noted by Abramitzky and Boustan (2017), very few of these works have focused on the link between migration and innovation. This link is an important one though, since many of today's largest US companies (e.g. General Electric) as well as several scientific and technological discoveries can be traced back to foreign born inventors and scientists who entered the US between the late nineteen century and 1940s (Hughes, 2004). Some recent evidence for this time period has indeed shown that inventor migrants greatly contributed to the rise of the US inventive activity in specific technological fields (Moser, Voena, and Waldinger, 2014) and in the long run for the US as whole (Akcigit, Grigsby, and Nicholas, 2017). Our work complements these studies by showing that the geographical distribution of immigrant inventors across US regions can explain their technological evolution.

We build a novel dataset of immigrant inventors to examine their impact on the US inventing activity between 1870 and 1940[1]. Did native inventors benefit from immigrants' inventive activity? Did immigrant inventors contribute to develop new technological activities in the regions they migrated to? While these questions have been somewhat addressed by

[1]The Age of Mass Migration usually ends in 1913, with the outburst of WWI, or in the 1920s with the introduction of the national-origins quotas. We extend the time span of our analysis until 1940 because other significant inflows of scientists occurred in this period (e.g. Jews escaping Nazism in Europe in the 1930s) (Moser, Voena, and Waldinger, 2014). Our data also show a robust (albeit declining) patenting activity by immigrants through the 1930s (see Figure II). The analysis is however robust to the exclusions of the years after 1930 (see tables S.10 and S.11)

the literature that analyses the effects of present-day immigration on innovation (Kerr, Kerr, Özden, and Parsons, 2016), there is less systematic evidence of these effects for historical migration in the US and in particular for the Age of Mass Migration.

Regarding contemporary studies, the literature has provided robust evidence showing that inventive activity as well as scientific outcomes of immigrant workers have been growing steadily in the US (Hunt, 2011; Kerr, 2007). Findings are instead mixed when it comes to measuring the impact of immigrants' inventive activity on natives (Kerr, Kerr, Özden, and Parsons, 2016). Some empirical works highlight the potential crowding-out effect of immigrant scientists (Borjas and Doran, 2012). Others instead show that inventor migrants have no negative effect (Kerr and Lincoln, 2010), or even strong positive effects on incumbents (Hunt and Gauthier-Loiselle, 2010). Other works have turned their attention to role of high-skilled immigrants as carriers of knowledge. For example, Ganguli (2015), shows that Russian scientists who migrated to the US after the collapse of USSR in 1991 were cited by US scientists more than those who did not migrated, suggesting that migration favoured the transmission of knowledge from origin to destination. More recently Bahar, Choudhury, and Rapoport (2019) conduct a wide cross-country study and show that receiving countries develop comparative advantages in the same technologies of the immigrants' country of origin. This finding suggests that migrants contribute to innovation activity in the receiving countries by 'importing' knowledge from their home country.

Regarding studies that focus on specific historical events in the US, evidence indicates that migration had positive effects on US inventive activity. For example, Moser, Voena, and Waldinger (2014) show that German-Jewish chemists escaping Nazi-Germany in the 1930s brought new ideas to the US scientific community that eventually contributed to emergence of new subfields in chemistry. Moser and San (2019) investigate the effect of the introduction of immigration quotas in US in the early 1920s. They show that this policy, which was originally aimed at preventing the entry of low-skilled workers coming from selected European countries, had the unintended consequence of reducing also the influx of scientists from these countries. Their estimates indicate that the quota system overall led to a sharp decline in US inventive activity in subsequent years. Doran and Yoon (2018) also look at the effect of quotas on invention during the Age of Mass Migration, but they highlight a different mechanism. They claim that the decline of low-skilled workers from countries affected by the quotas negatively impacted on the productivity of native inventors. Akcigit, Grigsby, and Nicholas (2017) analyses the long-term impact of immigrants on innovation (i.e. patenting) in the US over the period 1880-1940. Their findings show that technological fields where immigrants were most active during the Age of Mass Migration developed at faster pace in the long-run (1940-2000).

Our work, by building on the important insights of the above literature, investigates the city-level impact of immigrants' patenting in the period 1870-1940. A major strength of our analysis is that it relies on an original patent dataset that includes the fully disambiguated names of migrant inventors, their country of origin and their county and state of residency in the US. We text-mined this information from publicly available USPTO historical patent documents, which used to disclose the nationality of foreign applicants[2]. In addition, we create series of name-matching algorithms to search for patents of these inventors among the universe of all US patents, thus incorporating migrants' contribution to the US patenting system after they may have acquired the US citizenship.

Since this name-matching procedure might lead to the inclusion of false positives in our database, we use different versions in which further restrictions are imposed (i.e. the matched patent should be in the same metropolitan area, and/or should be granted not later than 10/20 years after the first patent of the immigrant inventor).

We exploit time, place, and technological variability in the patenting activity of migrants and natives to test three different channels through which migration may have affected the technological development of places. First, we test the direct impact of immigrant patenting on natives' inventive activity. Second, we use measures of migrants' country-of-origin expertise to evaluate the importance of knowledge diffusion channels (Bahar, Choudhury, and Rapoport, 2019; Ganguli, 2015). Third, we evaluate how these two channels affect the specialization profile of the places migrants move into and whether new specialization patterns

[2]These historical patent documents usually contain the following structure: "(...) Nikola Tesla, from Smiljan Lika, border of Austria-Hungary, residing at New York, N.Y., has invented ...". With this piece of information we are able to identify unambiguously if an inventor is an immigrant, which in our definition is a foreign-born individual who is resident in the US.

emerge as a result of this.

In order to tackle the first question, the impact of immigrant inventors on US inventive activity, we estimate a baseline model in which we regresses the total number of patents by native inventors in a given technology, metropolitan area, and decade on the number of patents authored by immigrants. The model includes region-level variables along with interaction dummies to controls for time-invariant technological and regional factors. OLS estimates indicate that doubling the number of patents by immigrant inventors results on a 20% increase in the number of patents of natives (a 0.2 elasticity). To address the endogeneity concerns present in this model, we adopt a twofold identification strategy: first, we instrument the number of immigrants' patents with a modified version of the shift-share (Bartik) instrument; second, we implement a lagged dynamic model.

We modify the conventional Bartik instrument in three important ways. First, we exclude from the shift component, which is given by the total number of patents in a given year-technology-country of origin, the patents of the immigrant inventors from the corresponding US region in the share component. By doing so we remove the endogenous part of the shift. Second, we use different dimensions to construct the shift and share components, which is usually not the case for the conventional Bartik. While the shift component includes a country-of-origin and technology dimension, the share uses a country-of-origin and region-of-destination dimension. Therefore, the share (which is computed before 1890, while the analysis is carried out from 1900 to 1940) is exogenous because it refers to all inventions of a given country in all technologies, rather than those in a given technology. Third, we replace the share component (i.e. share of patents) with the share of immigrants. The latter two modifications of the instrument should address the critique of Goldsmith-Pinkham, Sorkin, and Swift (2018) and Jaeger, Ruist, and Stuhler (2018). The instrument is based on the idea that immigrant inventors rely on social-ethnic ties when they have to make a localization choice. This hypothesis finds support in our data, since immigrant inventors do tend to cluster in space in ways that resemble the spatial distribution of immigrants during the age of mass migration (Abramitzky and Boustan, 2017). The estimates of the IV model are positive and significant, with an elasticity of 1.1, which is much larger than the OLS estimates, but5 robust across different specifications. The second element of our identification strategy attempts to control for simultaneity. The findings of the lagged dynamic model are in line with previous baseline findings (although larger in size).

We also provide a battery of tests to show that our results are robust to changes in the econometric specification. On the one hand, we test the robustness of the results to potential biases induced by the name-matching procedure. The coefficient estimates prove to be robust to modifications in the matching algorithm, they all range between 0.2 to 0.4 for OLS and 0.9 to 1.5 for IV, and are in line with the findings of the benchmark model (i.e. main dataset). On the other hand, we change the unit of analysis: we replace the Metropolitan Statistical Areas (MSAs) with states. Qualitatively, the findings do not change, although the elasticity estimates are lower.

After having measured the direct impact of immigrant patenting on natives' inventive activity, we evaluate our second hypothesis. In this part we test whether immigrant inventors carry knowledge which resembles the technological specialization of their country of origin, thus contributing to the recent literature on knowledge diffusion (Bahar, Choudhury, and Rapoport, 2019; Ganguli, 2015). To test this mechanism we adapt to the regional context a measure of 'foreign expertise', which has been first used by Akcigit, Grigsby, and Nicholas (2017) for the US case. This indicator is made of two components: the first one captures the technological specialization of the immigrant's country of origin; the second one counts the total number of patents of migrant inventors in a given US region and from a given country of origin (but it does not have a technological class dimension). This measure aims at capturing whether a specific piece of foreign knowledge is imported by an inventor from her country to the US city she moved to. Our results show that the migrants seems to bring with them foreign expertise that becomes relevant for the technological development of the places they immigrated to.

Finally, we test whether immigrant inventors contribute to shape the technological evolution of the receiving region. We observe that new technologies, which were not present yet in a region, emerged because of the inventive activity of immigrant inventors in those regions. Our results suggest that both immigrants' inventive activity and the knowledge they import from their home country helped US cites to enter new technological fields.

Our findings are in line with a growing literature that analyses the role high-skilled immigrants in the host country (Breschi, Lissoni, and Temgoua, 2016; Kerr, Kerr, Özden, and Parsons, 2016). Our work also contributes to the recent literature on historical migration in the US (Rodriguez-Pose and Von Berlepsch, 2014; Sequeira, Nunn, and Qian, 2020; Tabellini, 2020). More specifically, we add original evidence to the strands of studies that focused on the link between historical migration and innovation in the US (Akcigit, Grigsby, and Nicholas, 2017; Moser, Voena, and Waldinger, 2014; Moser and San, 2019). In line with these studies we find that immigrant inventors played a crucial role in the construction of the US technological system in the late nineteen and early twenty century.

We complement the above literature in two ways. First, our work generalise some of the important findings of these studies that focused on specific historical cases (e.g. Moser, Voena, and Waldinger (2014) on German chemists; Ganguli (2015) on Russian scientists) by looking at a broader set of immigrant groups and technological fields. Second, our work adds a geographical dimension to the studies that had mainly a country perspective (Akcigit, Grigsby, and Nicholas, 2017; Moser and San, 2019) and shows that immigration played an important role also at local level.

The paper is structured as follows. In section 2, we present some historical background information about the age of mass migration and invention in the US. We illustrate how immigrants related to invention and patenting in the US. In section 3, the data are presented with a description of how we built the dataset. Section 4 lays out our empirical strategy, while section 5 illustrates the main findings. Section 6 concludes with some discussion of the contribution of our work and its possible extensions.

# 2 The Age of Mass Migration in the US: immigration, invention and patenting

More than 30 million people migrated to the US from all around the world between the 1830s and 1920s (Hatton, Williamson, et al., 1998). A large majority consisted of Europeans from different geographical origins who entered US in large consecutive waves. The first wave gathered up strength through the 1830s and 1840s, bringing mainly northern Europeans from Ireland, Germany and England and picked up in 1850. In this year ten percent of the US population was foreign born. A second wave reached its peak in 1880, and was made up mainly of Germans and Scandinavians. At this time about 90% of foreign born immigrants came from Northern and Western Europe, while Southern and Eastern Europeans represented less than 5%. After the 1880s the trend was reversed, a large wave of Italians and Eastern Europeans moved to the US, representing 40% of all foreign borns by the turn of the century (Abramitzky and Boustan, 2017). Overall, the share of the foreign born population rose up to 14% by 1870 and remained stable around this level until 1920. The arrival of immigrants came to an abrupt halt in 1914 because of the outbreak of the world war. However, as soon as the war was over, immigration flows increased again. In the next decade (1920s), the Age of Mass Migration came to an end when in 1924 the US Congress passed a law that introduced country-specific quotas (Goldin, 1994).

Along with the millions of low-skilled immigrants entering the US during these six decades, in the order of thousands were or became inventors and patentees. According to historians in this period of time, immigrants were (as they are today) disproportionally represented among inventors (Khan, 2005). Recent evidence confirm these estimates (Akcigit, Grigsby, and Nicholas, 2017). Although it may appear at first surprising, especially if contrasted with the typical immigrant profile of that time, which was mainly a low skilled individual, this is less so if one considers how the inventive activity was organised in the late nineteen century in the US. Inventive activity before the early twentieth century was primarily an individual endeavour, which required relatively little capital (Hughes, 2004). Inventions were often the outcome of a trial and error process and fortunate accidents which allowed to come up with smart solutions that fixed specific technical problems (Sokoloff, 1988). Formal training was also not a necessary condition and even the most prolific inventors had little formal education and did not rely on scientific methods to run experiments (e.g. Edison and the Wright Brothers) (Hughes, 2004). Historian suggest that about forty percent of foreign-born inventors and about 25% of natives did not have formal education (Khan, 2005).

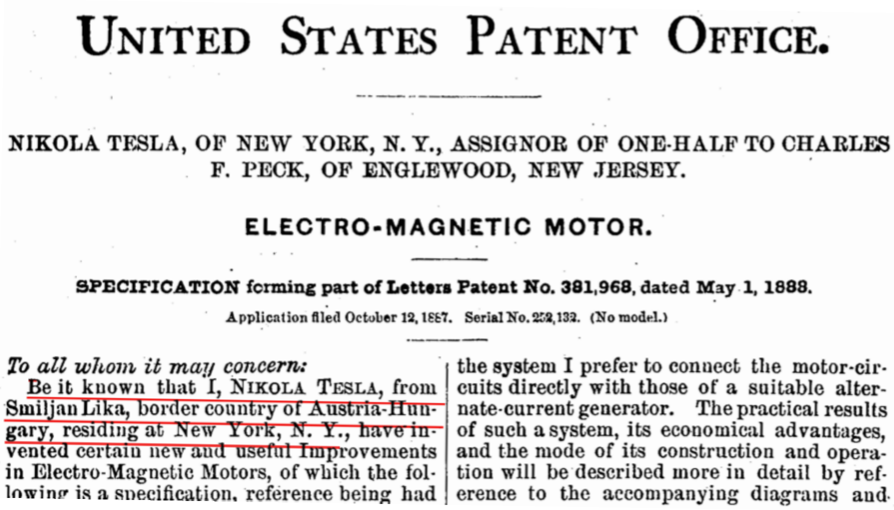
Another important aspect to take into account is that the US patent system, in contrast to the British or French, had very low barriers to entry. Registering a patent in the US was affordable and relatively cheap compared to UK. Moreover, technological invention was given a central stage in the US social and economic life, to the extent that patenting was mentioned in the US Constitution and accordingly promoted and enforced. An additional feature of the US patent system favoured particularly the participation of disadvantageous groups, as it required that a patent to be granted to an applicant, she should be the true and first inventor worldwide. This contrasted with England and other European countries, where a patent was granted also to imported foreign inventions. This latter practice clearly favoured wealthy traders and companies who could afford purchasing technology abroad and patent them domestically (Sokoloff, 1988). As result, the barriers to patenting were particularly low in the US, a condition which clearly favoured immigrants. As Khan (2005) states "the notion of patenting and inventive activity as means of achieving eminence, especially for disadvantage groups, is borne out by the experience of foreign-born inventors" (p. 2014).

The biographies and background of immigrant inventors are however very heterogeneous. We could classify them in two broad categories. A first group includes those who arrived to the US during their childhood: they were raised and trained in the US. For example Elihu Thomson, the prolific inventor and founding father of several successful companies (e.g. General Electric, Thomson SA), migrated from the UK to the US in his childhood. A second category refers to foreign-born inventors who were already trained or active in a specific scientific field before moving to the US. Nikola Tesla is perhaps the most paradigmatic example of this group. With his inventions he gave key contributions to the nascent US electrical industry, besides many other related fields. He migrated to the US with already relevant experience in telephony and electrical engineering and had formal tertiary education. When he arrived in the US he soon built a reputation of prolific inventor, which allowed him to work with and sell patents to the high tech companies of his time (e.g. Edison, Westinghouse Electric and General Electric) (Hughes, 2004; Tesla, 2011).

# 3 Identification of Immigrant Inventors in Patents

Since we focus on the impact of particular type of immigrants, i.e. those who arrived in the US with a baggage of relevant working or intellectual experience, most of the available databases and empirical approaches that are common in the literature are not a suitable

Figure I: Nationality Information contained in Historical Patent Documents


UNITED STATES PATENT OFFICE.

NIKOLA TESLA, OF NEW YORK, N. Y., ASSIGNOR OF ONE-HALF TO CHARLES F. PECK, OF ENGLEWOOD, NEW JERSEY.

ELECTRO-MAGNETIC MOTOR.

SPECIFICATION forming part of Letters Patent No. 381,968, dated May 1, 1888.

Application filed October 12, 1887. Serial No. 252,132. (No model.)

*To all whom it may concern:*

Be it known that I, NIKOLA TESLA, from Smiljan Lika, border country of Austria-Hungary, residing at New York, N. Y., have invented certain new and useful Improvements in Electro-Magnetic Motors, of which the following is a specification, reference being had the system I prefer to connect the motor-circuits directly with those of a suitable alternate-current generator. The practical results of such a system, its economical advantages, and the mode of its construction and operation will be described more in detail by reference to the accompanying diagrams and

option. This is because they usually identify migrants without distinguishing where they acquired their knowledge. For instance, when migrants are identified using the ethnic origin of their surnames it is not possible to know whether they arrived to the US during their childhood, and were therefore trained and raised in the US.

In this section we describe the construction of a new dataset that identifies migrants in historical patent documents at the USPTO. We exploit the fact that old historical patent documents, prior the 1940s, include information about the nationality of the inventors by disclosing the place they come from if they are foreign. Consider for instance Figure I below, which shows patent document number 381,968 granted to Nikola Tesla,[3] who arrived to the US in 1884 from Europe and started working at Edison's company almost immediately after. Note that patent documents were describing not only the place of residence of the inventor (New York) but also its nationality (Austro-Hungarian).

The creation of this database can be divided into three distinct stages. The first challenge consisted on identifying historical patent documents of migrants inventors from the pool of all patented inventions granted at the USPTO prior the 1940s. Since manually scanning all documents for foreign inventors would render the task unfeasible, we relied instead on an automated algorithm to identify potential candidates. We trained an algorithm to identify

[3]See entire patent document here : `https://patents.google.com/patent/US381968`

patents who could be attributed to an immigrant inventor based on the vocabulary used in its description. Words such “a subject of”, “a citizen of”, or “kingdom” are usually associated with the description of the location of foreign inventors in patents. These should appear in combination with words such as “residing in” and the name of an US location. This algorithm is analogous to the one described and documented in Petralia, Balland, and Rigby (2016) but tailored to this particular problem[4]

It is likely, however, that the subset of patents identified as coming from migrants (as well as the information extracted from them) contains mistakes. This could happen if a certain combination of keywords results in our algorithm identifying the presence of a migrant when it is actually not the case. For instance, the word “England” may refer to the location of the inventor (“New England”) instead of his nationality, thus increasing the probability of falsely identifying the presence of a migrant in the patent. The second step of the procedure consisted on correcting possible mistakes made by the algorithm. To do so we manually checked all patents that were flagged as produced by an immigrant inventor (approximately 36,000) and whenever necessary we corrected misspells or added the missing information. From this procedure we obtained 15,055 manually checked patent-inventor observations.

Finally, we had to correct for the fact that our automated detection algorithm would not detect the patents of immigrants that have obtained the US citizenship after residing in the US for some time. This is because foreign citizenship was not disclosed in patent documents if the immigrant had obtained the US citizenship. We tackled this issue by text-mining all patents documents in the period 1840-1940 to search for the names of the 15,055 manually identified migrants. We allowed for minor discrepancies in the name matching algorithm to take into account the possibility of minor misspellings, which were later manually checked. This resulted in a final database containing 49,841 manually-checked inventor-patent combinations with information about the place of residence of the inventor, the country of origin, the year the patent was granted, and the technological profile the patent. Even though we manually checked that all matches were not due to misspells, it could be the case that some of the additional patents found at this stage are not of the migrant inventor in question but

[4] A detailed example can be fount at: `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/3ZLC8E`

from somebody else with the exact same name. We applied several criteria to restrict this possibility. If we include all inventors that match the originally manually collected name we obtain the 49,841 inventor-patent combinations we mentioned before. If we restrict to name matches that occur within a 20 year window from the original (manually identified) name this number goes down to 47,186; and to 40,582 if we use a 10 year window instead. In addition, we restrict to name matches for which the state of residence also matches within a 20 or 10 year window, which results on a sample of 36,414 and 33,209 inventor-patent combinations respectively. Our results are robust to these different matching approaches.

Figure II shows the total number of patents of immigrant inventors during the period. We observe a growing trend in patenting which peaks in 1916, possibly capturing the effect of WWI on both patenting activity and inflow of migrants. After that, a new peak is reached in 1926, right after the introduction of immigration quotas, which ended the open door immigration policy in the US. This time dynamics follows closely the inflow of migrants during that period of time (Gibson and Lennon, 1999).

Figure II: Migrants' Patents over the Period

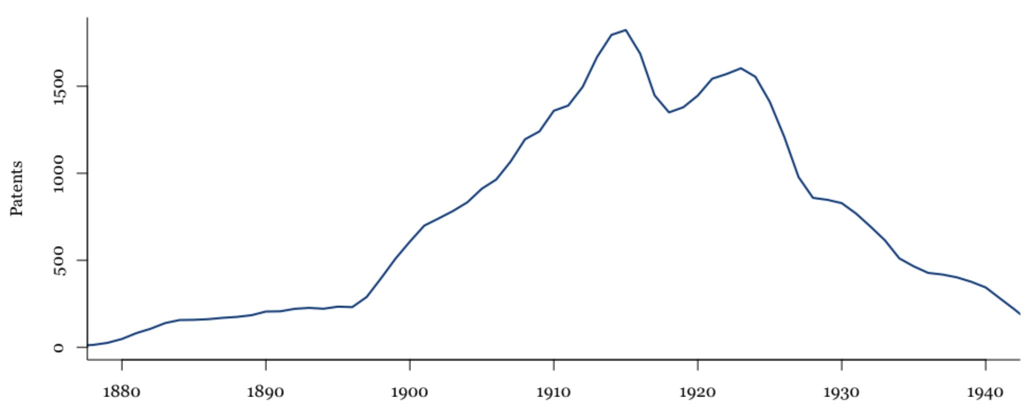



Table I shows the most prolific nationalities. Not surprisingly, this ranking resembles to a large extent the distribution of the immigrant population in the US, with Great Britain & Ireland at the top of the list, followed by Germany (Gibson and Lennon, 1999). All major European countries which had large flow of emigrants to US are listed, i.e. Sweden, Italy, Russia and central European countries. We have grouped countries following the USPTO aggregation criteria. More specifically, Great Britain & Ireland includes Ireland, Wales, Scotland and England, Austria-Hungary includes Austria, Hungary, Croatia, Czechia, Slovakia and Slovenia, while Russia includes also Lithuania and Latvia. This is because the USPTO referred to these territories exchangeably, sometimes referring to cities like Vienna as part of Austria and others as part of the Austro-Hungarian empire.

Table I: Patents by Nationality

| | Origin | Patents | Share |
|---|---|---|---|
| 1 | GREAT BRITAIN & IRELAND | 18,093 | 0.368 |
| 2 | GERMANY | 6,430 | 0.131 |
| 3 | SWEDEN-NORWAY | 6,092 | 0.124 |
| 4 | AUSTRIA-HUNGARY | 3,569 | 0.073 |
| 5 | RUSSIA | 3,290 | 0.067 |
| 6 | ITALY | 2,461 | 0.050 |
| 7 | CANADA | 2,081 | 0.042 |
| 8 | SWITZERLAND | 1,489 | 0.030 |
| 9 | DENMARK | 1,147 | 0.023 |
| 10 | FRANCE | 1,136 | 0.023 |

Turning to the geography of these migrant inventors, Figure III shows the most popular migrant group per county. Migrant inventors tended to cluster in space resembling closely the geographical footprint of other migrants from the same nationality (Abramitzky and Boustan, 2017). Not surprisingly, large urban areas are highly represented, with cities like New York and Chicago ranking at the top. Even though the east coast is the epicenter of migrant inventive activities (and patenting in general), large communities of German and Scandinavian immigrants were active throughout the Mid-West.

Finally, Table II shows the technological composition of migrants' (and US natives) patenting activity in the period. Note that Germans were relatively more oriented to the production of Mechanical and Electrical & Electronic technologies than US natives. In addition, North-Europeans and Russians were relatively more predominant than US natives in Electrical & Electronic, one of the fastest growing technological domain of the time (Hughes, 2004).

Figure III: Largest Group of Migrant Inventors

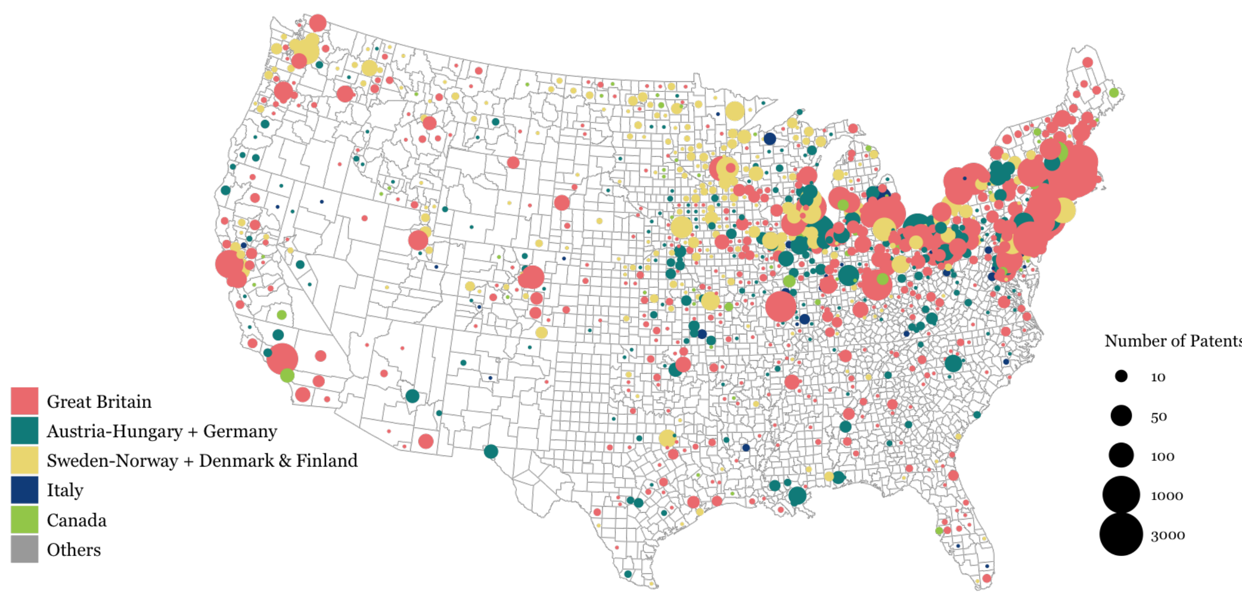


Table II: Type of Technology by Nationality

| | GB | DE | SW-NO | AT-HU | RU | IT | CA | US |
|---|---|---|---|---|---|---|---|---|
| Others | 0.357 | 0.340 | 0.351 | 0.416 | 0.499 | 0.466 | 0.419 | 0.441 |
| Mechanical | 0.369 | 0.402 | 0.430 | 0.366 | 0.283 | 0.346 | 0.360 | 0.366 |
| Electrical and Electronic | 0.127 | 0.144 | 0.115 | 0.100 | 0.123 | 0.076 | 0.080 | 0.074 |
| Drugs and Medical | 0.013 | 0.008 | 0.004 | 0.007 | 0.010 | 0.018 | 0.014 | 0.013 |
| Computers and Communications | 0.024 | 0.010 | 0.017 | 0.027 | 0.012 | 0.019 | 0.048 | 0.017 |
| Chemical | 0.109 | 0.096 | 0.083 | 0.085 | 0.073 | 0.074 | 0.079 | 0.089 |

Notes: US shares are calculated using HistPat data (Petralia, Balland, and Rigby, 2016) .

# 4 Empirical Strategy

## 4.1 Impact on the inventive activity of US regions

In order to investigate the contribution of immigrant inventors to the inventive activity of US regions we estimate the following model.

$$nat_t^{rk} = \beta_1 mig_t^{rk} + \gamma_t^r + \psi_t^k + \varphi^{rk} + \eta_t^{rk}, \tag{1}$$

where $nat_t^{rk}$ is the total number of patents (in logs[5]) by native inventors in technology $k$, region $r$, and period $t$. Note that in the benchmark regressions, we use Metropolitan Statistical Areas (MSAs) and 10-year windows, for region and period respectively. Our variable of interest, $mig_t^{rk}$, is the log of the number of patents authored by immigrants. Lastly, $\gamma_t^r$, $\psi_t^k$, $\varphi^{rk}$, and $\eta_t^{rk}$ are the three interaction dummies and the error term. Note that $\gamma_t^r$ captures all the region-level variables such as value added, population, population density (etc.), $\psi_t^k$ controls for the state of the technology, and $\varphi^{rk}$ for the (time-invariant) technological specialization of the region.

This basic empirical setup, as described in equation (1) is highly endogenous – even though the model is saturated with all possible dummies. In fact, idiosyncratic changes in the conditions of a region-technology combination (for instance the opening of a research lab by a corporation or a university) would affect both $nat_t^{rk}$ and $mig_t^{rk}$ and bias the estimate of $\beta_1$. For this reason, in the next section we describe how we identify the impact of migrants on regional innovation in the US.

[5]In the benchmark regressions we keep all observations, including region-technology combinations with zero patenting. We, thus, measure the log of patent count as $log(patents + 1)$. For robustness, we repeat the analysis dropping all observations with zeros and find that results are, albeit weaker, consistent (see Appendix C).

## 4.2 Identification

We deal with the inherent endogeneity of the empirical model in (1) in two ways: first, we instrument $mig_t^{rk}$ using a modified version of a shift-share (Bartik) instrument and, second, we exploit the panel nature of our data to re-write (1) into a dynamic empirical model.

### 4.2.1 Shift-share instrument

Shift-share instruments are well grounded in the migration literature (see Card, 2001) and widely applied in the recent literature on immigration and innovation (see Hunt and Gauthier-Loiselle, 2010; Ganguli, 2015). The instrument is usually composed of two parts: the inflow of immigrants from a given country to a destination country (e.g. the shift), and the share of immigrants of that country residing in a specific city in the previous period (e.g. the share). In our case the instrumental variable is constructed as follows:

$$IV_t^{rk} = \sum_c \frac{MIG_{t_0}^{cr}}{MIG_{t_0}^{c}}(MIG_t^{ck} - MIG_t^{crk}), \tag{2}$$

where $MIG$ is the non-log version of the endogenous variable ($log(MIG_t^{crk}) = mig_t^{crk}$). The shift component of the instrument ($MIG_t^{ck} - MIG_t^{crk}$) is the the total flow of patents in period $t$, from an immigrant born in country $c$, in technology $k$. Note that, however, this total flow excludes those patents in region $r$ ($MIG_t^{crk}$) to remove the endogenous portion of the shift. We further highlight, in fact, that in our setting we have an additional dimension (that is technological class $k$), which is typically not available to most studies on migration using shift-share instruments. We can therefore exploit this feature in the construction of the instrument: while for the shift we use the flow with country-of-origin×technology dimension, for the share we use country-of-origin×region-of-destination. This share (which is computed with $t_0 < 1890$, when the analysis is carried out from 1900 to 1950) is exogenous because it does not contain migrants in technological class $k$ specifically, but inventions from country $c$ in all technological classes.

This should address the critique of Goldsmith-Pinkham, Sorkin, and Swift (2018) or Jaeger, Ruist, and Stuhler (2018) who point out that the share component of the instrument is generally problematic, as adjustments from previous migration may still be ongoing. Here we suggest that the next wave of migrants with specialization in technology $k$ would migrate where there are existing communities of fellow countrymen, because of social ties, hence irrespective of technological specialization of the previous wave. The ongoing adjustments should be exogenous to the competence brought by the migrant in technology $k$.

To go a step further, for our benchmark results we substitute the share of patents by migrants in (2) with the share of all migrants from country $c$ (inventors and non-inventors) from the population census of 1890[6].

$$\widetilde{IV}_t^{rk} = \sum_c \frac{CENSUS_{t_0}^{cr}}{CENSUS_{t_0}^{c}} (MIG_t^{ck} - MIG_t^{crk}), \tag{3}$$

Hereafter, we denote the log of the IV variables as $iv_t^{rk}$ and $\tilde{iv}_t^{rk}$, respectively.

### 4.2.2 Dynamic model

As a complementary identification strategy, we also attempt to account for potential endogeneity of $nat_t^{crk}$ and $mig_t^{crk}$ with a dynamic empirical model. We re-write (1), as

$$\Delta nat_{t-1\to t}^{rk} = \theta nat_{t-1}^{rk} + \beta_1 mig_{t-1}^{rk} + \gamma_t^r + \psi_t^k + \eta_t^{rk}. \tag{4}$$

That is, we now relate the growth (log difference) in patenting activities of natives to patents of migrants in the previous period. Crucially, we also include a lagged dependent variable so that changes in the environment (shocks in $\eta_t^{rk}$) affecting both native and migrant patenting are absorbed by $nat_{t-1}^{rk}$. In addition, we instrument $mig_{t-1}^{rk}$ with $\tilde{iv}_{t-1}^{rk}$.

We finally note that in the dynamic setting we cannot include region×technology dummies ($\varphi^{rk}$) without biasing the results (Nickell bias).[7] This may raise the additional concern that (although we cluster standard errors by region and technology) modest temporal variation may inflate significance without fixed effects. As a further check, we re-design the dynamic model to exploit the whole time-span of our data, but in cross-sectional form:

[6] When we use patents data for the share, we sum all the patents published by migrants from 1870 to 1890. This is because patent production is a flow variable. When we compute the share using census data on migrants, we use the stock of foreign born in 1890 instead.

[7] For completeness, we will report these results nonetheless.

$$\Delta nat^{rk}_{t_1 \to t_2} = \theta nat^{rk}_{t_1} + \beta_1 mig^{rk}_{t_1} + \delta^r + \iota^k + \eta^{rk}, \quad (5)$$

where $t_1 = [1890, 1930)$ and $t_2 = [1930, 1950)$. The corresponding instruments also use $t_1 = [1890, 1930)$ for the shift and, as before, $t_0 = [1870, 1890)$[8] for the share.

# 5 Empirical Results

## 5.1 Impact on the inventive activity of US regions

In Table III we report the results of estimating Equation (1). This is the most basic setup we estimate, where we simply relate contemporaneous patents of natives to patents of migrants. In columns 1–4 we estimate the model with OLS and various combination of variables. The most complete estimation (with all interacted dummies) in column (4) suggests an elasticity of about 0.2; that is doubling migrants' patents increases patents of natives by 20%.

As discussed in Section 4.2, the OLS estimate of the contemporaneous model are likely to be biased by endogeneity. Columns (5) and (6) report the instrumental variable estimates of the model in Equation (1). The shift-share instrument ($\tilde{iv}^{rk}_t$) uses past population by country of origin for its share component, as described in Equation (3). The Kleibergen-Paap F statistics are well-above the usual cut-off point of 10.

Note that the point estimate (an elasticity of about 1.1 in Column 6 ) is significantly larger than the corresponding OLS estimate[9], we find that (as we show in the reminder of this section) this magnitude is remarkably robust across specifications. Overall, results suggests that a large role was played by migrants in the innovation environment of the United States of the early 20th Century.

In section 4.2, we propose an alternative econometric specification that could better control for the simultaneity of $nat^{rk}_t$ and $mig^{rk}_t$ (see Equation 4). Table IV reports the coefficients estimated using this dynamic setting. The benchmark results for Table IV are

[8] $t_0 = 1890$ for $\tilde{iv}^{rk}_{t_1}$.

[9] The most likely explanation for an IV estimate larger than the OLS is that the IV corrects for measurement error.

Table III: The relationship between US and immigrant patenting

| | Dependent variable: Patents of natives | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of migrants | 2.651*** | 0.273*** | 1.069*** | 0.219*** | 5.623*** | 1.173*** |
| | (0.090) | (0.028) | (0.109) | (0.043) | (1.395) | (0.297) |
| $Adj.R^2$ | 0.135 | 0.779 | 0.548 | 0.750 | | |
| Obs. | 758940 | 758940 | 758940 | 758940 | 758940 | 758940 |
| F (first stage) | | | | | 30.684 | 46.832 |
| Dummies | t | k,r,t | kr,t | kr,kt,rt | t | kr,kt,rt |

*Notes: all variables are in logs. Dependent variable: number of patents by natives ($nat_t^{rk}$). Explanatory variable: patenting activity by migrants ($mig_t^{rk}$). Instrumental variable: $\tilde{iv}_t^{rk}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

columns (4) and (6) for OLS and IV, respectively. The estimated coefficients in these cases are comparable in magnitude (and in fact larger) to the corresponding coefficients of Table III. Note that, unlike in Table IV, column (4) and (6) in this Table do not include city×technology dummies. This is to avoid introducing Nickell bias, which is indeed present in column (3) and, to a lesser extent in column (2).

The final specification we discuss for this section is the one reported in Equation 5. Similarly to the growth-level setup, whose estimates are reported in Table IV, this empirical model differs in that we aggregate the whole dataset in three time periods ($t_0 = [1870, 1890)$, $t_1 = [1890, 1930)$, $t_2 = [1930, 1950)$). As we use $t_0$ for the instrument, $t_1$ for the level variables, and the difference between $t_2$ and $t_1$ for the growth variables, we functionally have a cross-sectional dynamic model. The results of this exercise are reported in table V. Coefficients are in line with previous estimates, even though the most complete IV estimate (column 4) suggests a smaller elasticity of about 0.8.

Table IV: The relationship between US and immigrant patenting: alternative specification

| Dependent variable: Patents of natives (growth between $t$-1 and $t$) | | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of natives (t-1) | -0.234*** | -0.426*** | -0.960*** | -0.428*** | -0.289*** | -0.469*** |
| | (0.023) | (0.021) | (0.018) | (0.021) | (0.032) | (0.025) |
| Patents of migrants (t-1) | 0.580*** | 0.403*** | 0.174*** | 0.414*** | 1.450*** | 1.310*** |
| | (0.040) | (0.033) | (0.015) | (0.029) | (0.463) | (0.439) |
| $Adj.R^2$ | 0.123 | 0.234 | 0.385 | 0.264 | | |
| Obs. | 607152 | 607152 | 607152 | 607152 | 607152 | 607152 |
| F (first stage) | | | | | 25.694 | 38.175 |
| Dummies | t | k,r,t | kr,t | kt,rt | t | kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t-1\to t}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t-1}$) and migrants ($mig^{rk}_{t-1}$). Instrumental variable: $\tilde{iv}^{rk}_{t-1}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

## 5.2 Robustness

The results of Section 5 indicate a strong influence of migrant patenting activity on that of natives, with an elasticity estimated around 0.2–0.4 with OLS, and around 0.8–1.3 with IV. We discuss here that these estimated values are robust to a number of empirical design choices. A first set of issues relates to the extraction of migrants in the dataset. Given the degrees of freedom we have in this process, we have to decide whether to have a broad dataset (with many patents matched to foreign born inventors, but potentially a significant number of false positive) or a narrow data (where matched patents are more accurate, but with potentially more false negatives). We then decide to have a flexible dataset, which we use in this paper in three version: (1) the one we use in the benchmark regressions of Section 5. (2) An extended version, which we create by matching name and location of

Table V: The relationship between US and immigrant patenting: cross-section

| Dependent variable: Patents of natives (growth between $t_1$ and $t_2$) | | | | |
|---|---|---|---|---|
| | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) |
| Patents of natives ($t_1$) | -0.475*** | -0.606*** | -0.535*** | -0.640*** |
| | (0.022) | (0.018) | (0.030) | (0.022) |
| Patents of migrants ($t_1$) | 0.803*** | 0.417*** | 1.369*** | 0.835*** |
| | (0.045) | (0.045) | (0.255) | (0.226) |
| | | | | |
| $Adj.R^2$ | 0.389 | 0.554 | | |
| Obs. | 149604 | 149604 | 149604 | 149604 |
| F (first stage) | | | 39.645 | 83.825 |
| Dummies | | k,r | | k,r |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t_1 \to t_2}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t_1}$) and migrants ($mig^{rk}_{t_1}$). Instrumental variable: $\tilde{iv}^{rk}_{t_1}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time: $t_0 = [1870, 1890)$, $t_1 = [1890, 1930)$, $t_2 = [1930, 1950)$. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

migrant inventors in the 20 years following their first match. This is done since migrants might acquire citizenship and therefore become undetectable to us in subsequent patents. (3) A restricted version, where we drop all patents that could not be matched with HistPat (Petralia, Balland, and Rigby, 2016) (as an additional sanity check on the location of the patent). The three versions of the dataset produce very similar results. In Appendix A, we report two versions of Table III, which use instead the extended and restricted versions of the dataset. The versions of Table IV and V are also very consistent, with the elasticity estimates of 0.2–0.4 for OLS and 0.9–1.5 for IV. Given their similarity to benchmark regressions, we do not include them in the appendix, as we opt to mostly report results that are sufficiently different in order save space. These tables can be found in the Supplementary Material of the paper (Tables S.1–S.4).

Similarly, we also leave to the Supplementary Material the analysis carried out using 5-year, instead of 10-year, windows (S.5 and S.6), as well as the analysis using $iv^{rk}$ instead of $\tilde{iv}^{rk}$. Also in these cases, we find results which are perfectly in line with the benchmark (S.5–S.9).[10]

Next, a typical issue in regional studies is how to best define the geographical unit of analysis. In the US, Metropolitan Statistical Areas (MSAs) are often the unit of choice, since they capture the countries' main urban agglomerations, with borders defined by their economic interactions, such as commuting links. For this reason, cities so-defined constitute the geographical units of our benchmark analysis. However MSA do not capture the entirety of the economy and may miss geographical links that go beyond the cities. In Appendix B, we report the main analysis of Section 5, but conducted using states as territorial unit $r$. It can be seen that, while we can observe that the main findings are robust, the elasticity estimates are lower in this case, still around 0.2–0.4 for OLS, but around 0.4–0.6 for IV.

Finally, we change drastically the size of the dataset by dropping all observations in which the patent count is zero. The analysis using the benchmark parameters shows that the OLS results are perfectly robust, and that the IV results are robust in the first specification, while they become insignificant in the second and the third. We note however that these specifications have weak first stages. If instead, we pick regressions where the first stage is strong (such as using the extended data at the state level), results are again positive and significant, but with smaller magnitudes (elasticities: 0.1–0.2). All these results are reported in Appendix C.

## 5.3 Impact of foreign expertise on the inventive activity of US regions

The empirical findings of Sections 5 and 5.2 indicate that the knowledge of the migrant plays a strong role stimulating innovation in the United States. This knowledge it is likely to have originated in the migrants' country of origin, since foreign born who lived for long

[10] The only noteworthy deviation is an elasticity of 0.6 in the cross-sectional model instrumented $iv^{rk}$.

(and possibly studied) in the United States become citizen, and thus are not picked up as migrants by our algorithm.

In line with a growing literature on contemporary immigration and knowledge diffusion (Miguelez and Temgoua, 2019; Bahar, Choudhury, and Rapoport, 2019), our analysis suggests that migrants acted as carriers of knowledge across distant places. However, our measurement of knowledge flow from migrants ($mig_t^{rk}$) leaves an important question open: are the migrants bringing just their own knowledge or are they a bridge to the set of competence of their country of origin?

We introduce here a variable that can help make the distinction. We take this variable from Akcigit, Grigsby, and Nicholas (2017), but adapt it to our regional context: $E_t^{rk}$ is the foreign expertise on technology $k$ that migrants bring to region $r$.

$$E_{t_1}^{rk} = \sum_c \frac{PAT_{t_0}^{ck}}{PAT_{t_0}^{c}}(MIG_{t_1}^{cr} - MIG_{t_1}^{crk}), \tag{6}$$

where $PAT_{t_0}^{ck}$ is the production of patents of country $c$ in technology $k$ *at home.* $(MIG_{t_1}^{cr} - MIG_{t_1}^{crk})$ is the flow of patents by migrant inventors from country $c$, in region $r$ (excluding those in the target technology $k$).

This indicator of expertise, which differs from the one proposed by Akcigit, Grigsby, and Nicholas (2017) because it varies also by region $r$, is similar in spirit to our instrumental variable (Equations 2 and 3). This measure of expertise inverts the indices $r$ and $k$ for the share and the shift component (apart from using inventions by non-migrants in the share component). While this may appear minor at first sight, it is substantial: controlling for $mig^{rk}$ or $\tilde{iv}^{rk}$, expertise captures the connections US cities have with technology $k$ to countries that are specialized in that technology, *beyond having experts that migrated from those countries.* In this way, we can distinguish between the knowledge that was brought *directly* by migrants through their own competence and the knowledge brought *indirectly* through links with the home country.

The specification estimated by Akcigit, Grigsby, and Nicholas (2017) is comparable to our model in (5). We then write:

$$\Delta nat_{t_1 \to t_2}^{rk} = \theta nat_{t_1}^{rk} + \beta_1 mig_{t_1}^{rk} + \beta_2 e_{t_1}^{rk} + \delta^r + \iota^k + \eta^{rk}, \tag{7}$$

where $e_{t_1}^{rk} = log(E_{t_1}^{rk})$. In Table VI, the reader can find the estimates of this specification. We observe that expertise and patents of migrants are significant at the same time in all specifications, but one. While this may be hinting that *direct* knowledge of migrants is more important in stimulating innovation, in robustness analysis, we find that expertise is significant in all specifications. For instance, if we repeat the analysis with the restricted data, we find that the estimated coefficient of expertise in column (4) is larger, and its standard error smaller (likely due to a greater accuracy in the location of migrant inventors), thus resulting in a statistically significant estimate (see Appendix D). We conclude – from this analysis – that the role of migration appears to be both *direct* (through the knowledge embedded in the migrants themselves) and *indirect* (through the links that the migrants provide with their home country).

## 5.4 The impact on technological evolution of US regions

The direct and indirect impact of migration on US innovation has the additional (but equally important) consequence to change the technological evolution of cities. While Akcigit, Grigsby, and Nicholas (2017) note that migration has driven the technological trajectory of the US, and Moser, Voena, and Waldinger (2014) observe this in a specific technological field (i.e. chemistry), in this paper we show that this process happens at the regional level, with migration shaping the technological evolution of cities.

To highlight this point with more emphasis, we run here the analysis at the extensive margin. That is, instead of focusing on regions that have a specific technology, and study how the presence of migrant inventors influences its growth, in this section we look uniquely at regions where a technology is missing.

In this empirical design we then drop all observations where in $t_1$ there is a patent by a native (that is if $NAT_{t_1}^{rk} > 0$). We then look at period $t_2$ to see if innovative activities in that technology have appeared. Formally, we write:

$$appear_{t_1 \to t_2}^{rk} = \mathbb{1}\left[NAT_{t_2}^{rk} > 0 \mid NAT_{t_1}^{rk} = 0\right]. \tag{8}$$

Table VI: The role of expertise in innovation

| Dependent variable: Patents of natives (growth between $t_1$ and $t_2$) | | | | |
|---|---|---|---|---|
| | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) |
| Patents of natives ($t_1$) | -0.517*** | -0.617*** | -0.537*** | -0.639*** |
| | (0.023) | (0.018) | (0.029) | (0.022) |
| Patents of migrants ($t_1$) | 0.390*** | 0.318*** | 0.883*** | 0.730*** |
| | (0.046) | (0.044) | (0.288) | (0.237) |
| Expertise ($t_1$) | 0.858*** | 0.340*** | 0.532*** | 0.130 |
| | (0.147) | (0.101) | (0.174) | (0.084) |
| $Adj.R^2$ | 0.413 | 0.557 | | |
| Obs. | 149604 | 149604 | 149604 | 149604 |
| F (first stage) | | | 42.127 | 66.117 |
| Dummies | | k,r | | k,r |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t_1 \to t_2}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t_1}$) and migrants ($mig^{rk}_{t_1}$), and expertise ($e^{rk}_{t_1}$). Instrumental variable (for $mig^{rk}_{t_1}$): $\tilde{iv}^{rk}_{t_1}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time: $t_0 = [1870, 1890)$, $t_1 = [1890, 1930)$, $t_2 = [1930, 1950)$. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

The corresponding econometric model is comparable to (7):

$$appear^{rk}_{t_1 \to t_2} = \beta_1 mig^{rk}_{t_1} + \beta_2 e^{rk}_{t_1} + \delta^r + \iota^k + \eta^{rk}. \tag{9}$$

We again find that migrants play a direct and indirect role, with both independent variables $mig^{rk}$ and $e^{rk}$ estimated to be positive and significant (see Table VII).

Table VII: The extensive margin: Appearance of new city-technology combinations

| Appearance of technological class k, in region r, time $t_2$ | | | | |
|---|---|---|---|---|
| | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) |
| Patents of migrants ($t_1$) | 0.322*** | 0.194*** | 18.885*** | 11.296*** |
| | (0.048) | (0.039) | (4.354) | (3.326) |
| Expertise ($t_1$) | 1.211*** | 0.472*** | 0.479** | 0.233* |
| | (0.230) | (0.118) | (0.229) | (0.129) |
| | | | | |
| $Adj.R^2$ | 0.017 | 0.159 | | |
| Obs. | 95452 | 95452 | 95452 | 95452 |
| F (first stage) | | | 21.334 | 15.938 |
| Dummies | | k,r | | k,r |

*Notes: all variables are in logs. Dependent variable: appearance of patenting activity by natives ($appear^{rk}_{t_1 \to t_2}$). Explanatory variables: patenting activity by migrants ($mig^{rk}_{t_1}$) and expertise ($e^{rk}_{t_1}$). Instrumental variable (for $mig^{rk}_{t_1}$): $\tilde{iv}^{rk}_{t_1}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time: $t_0 = [1870, 1890)$, $t_1 = [1890, 1930)$, $t_2 = [1930, 1950)$. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

# 6 Conclusion

In this paper we examined the impact of immigrants' patenting on the inventive activity of US native inventors from an historical perspective. We find that US regions greatly benefited from the presence of immigrant inventors: they gave rise to spatially localised knowledge spillovers who had positive effects on the patenting activity of native inventors. We show that the contribution of immigrant inventors was also indirect: they acted as brokers of knowledge between their country of origin and the regions in the US they happened to migrate to. Therefore the positive effect of the immigrants' foreign expertise to the growth of US regional patenting is additional to the direct effect (i.e. patenting of immigrants). This diffusion mechanism is illustrated by the historian Thomas Hughes, when he presents the involvement of Charles Steinmetz and other German physicists and mathematicians at the General Electric research laboratories. He argues that besides their inventive activity their greatest contribution was, in his words, to have "introduced American engineers to advance mathematical modes of analyzing alternative current light and power systems. These modes greatly enhanced the problem solving abilities of engineering colleagues at GE" (Hughes, 2004) (page.161). These mathematical modes and the scientific method underpinning them were learned by the German researchers while working, experimenting and studying at their companies' or universities' labs in their home country. These immigrants embodied such tacit knowledge and carried it with them while migrating to the US, where they shared it with their fellow colleagues and researchers. The knowledge spillovers generated by immigrants also shaped the technological evolution of US regions. Our results indeed show that US regions entered in new technological fields thanks to the knowledge imported by immigrants. This evidence aligns well with recent findings on contemporary migration (Bahar, Choudhury, and Rapoport, 2019) and it overall suggests that policies restricting migration may prevent regional economies to tap into international knowledge flows, which proved to be relevant for the technological renewal of these regions.

## References


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

Akcigit, U., J. Grigsby, and T. Nicholas (2017): "Immigration and the rise of american ingenuity," *American Economic Review*, 107(5), 327–31.

Bahar, D., P. Choudhury, and H. Rapoport (2019): "Migrant inventors and the technological advantage of nations," .

Bandiera, O., I. Rasul, and M. Viarengo (2013): "The making of modern America: Migratory flows in the age of mass migration," *Journal of Development Economics*, 102, 23–47.

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.

Breschi, S., F. Lissoni, and C. N. Temgoua (2016): "Migration and innovation: a survey of recent studies," in *Handbook on the Geographies of Innovation*. Edward Elgar Publishing.

Card, D. (2001): "Estimating the return to schooling: Progress on some persistent econometric problems," *Econometrica*, 69(5), 1127–1160.

Doran, K., and C. Yoon (2018): "Immigration and invention: Evidence from the quota acts," *Unpublished manuscript, https://www3. nd. edu/˜ kdoran/Doran_Quotas. pdf*.

Ganguli, I. (2015): "Immigration and Ideas: What Did Russian Scientists "Bring" to the United States?," *Journal of Labor Economics*, 33(S1), S257–S288.

Gibson, C., and E. Lennon (1999): "HISTORICAL CENSUS STATISTICS ON THE FOREIGN-BORN POPULATION OF THE UNITED STATES: 1850TO1990," .

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.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018): "Bartik instruments: What, when, why, and how," Discussion paper, National Bureau of Economic Research.

Hatton, T. J., and Z. Ward (2018): "International Migration in the Atlantic Economy 1850–1940," *Handbook of Cliometrics*, pp. 1–29.

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

Hughes, T. P. (2004): *American genesis: a century of invention and technological enthusiasm, 1870-1970*. University of Chicago Press.

Hunt, J. (2011): "Which immigrants are most innovative and entrepreneurial? Distinctions by entry visa," *Journal of Labor Economics*, 29(3), 417–457.

Hunt, J., and M. Gauthier-Loiselle (2010): "How much does immigration boost innovation?," *American Economic Journal: Macroeconomics*, 2(2), 31–56.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018): "Shift-share instruments and the impact of immigration," Discussion paper, National Bureau of Economic Research.

Kerr, S. P., W. Kerr, Ç. Özden, and C. Parsons (2016): "Global talent flows," *Journal of Economic Perspectives*, 30(4), 83–106.

Kerr, W. R. (2007): "The ethnic composition of us inventors: Evidence building from ethnic names in us patents," .

Kerr, W. R., and W. F. Lincoln (2010): "The supply side of innovation: H-1B visa reforms and US ethnic invention," *Journal of Labor Economics*, 28(3), 473–508.

Khan, B. Z. (2005): *The Democratization of Invention: patents and copyrights in American economic development, 1790-1920*. Cambridge University Press.

Miguelez, E., and C. N. Temgoua (2019): "Inventor migration and knowledge flows: A two-way communication channel?," *Research Policy*, p. 103914.

Moser, P., and S. San (2019): "Immigration, science, and invention: Evidence from the 1920s quota acts," *Unpublished manuscript*.

Moser, P., A. Voena, and F. Waldinger (2014): "German Jewish émigrés and US invention," *American Economic Review*, 104(10), 3222–55.

Petralia, S., P.-A. Balland, and D. L. Rigby (2016): "Unveiling the geography of historical patents in the United States from 1836 to 1975," *Scientific Data*, 3.

Rodriguez-Pose, A., and V. Von Berlepsch (2014): "When migrants rule: the legacy of mass migration on economic development in the United States," *Annals of the Association of American Geographers*, 104(3), 628–651.

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

Sokoloff, K. L. (1988): "Inventive activity in early industrial America: evidence from patent records, 1790-1846," .

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

Tesla, N. (2011): *My inventions and other writings*. Penguin.

# Appendix

# A Analysis with extended and restricted dataset

Table A.1: Robustness of Table III: Extended dataset

| | Dependent variable: Patents of natives | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of migrants | 2.131*** | 0.274*** | 1.027*** | 0.270*** | 4.571*** | 1.526*** |
| | (0.055) | (0.026) | (0.051) | (0.023) | (0.928) | (0.254) |
| $Adj.R^2$ | 0.225 | 0.780 | 0.568 | 0.751 | | |
| Obs. | 758940 | 758940 | 758940 | 758940 | 758940 | 758940 |
| F (first stage) | | | | | 30.662 | 46.713 |
| Dummies | t | k,r,t | kr,t | kr,kt,rt | t | kr,kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patents by natives ($nat_t^{rk}$). Explanatory variable: patenting activity by migrants ($mig_t^{rk}$). Instrumental variable: $\tilde{iv}_t^{rk}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

Table A.2: Robustness of Table III: Restricted dataset

| | Dependent variable: Patents of natives | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of migrants | 2.557*** | 0.257*** | 1.087*** | 0.236*** | 5.508*** | 1.123*** |
| | (0.100) | (0.024) | (0.107) | (0.046) | (1.348) | (0.304) |
| $Adj.R^2$ | 0.132 | 0.769 | 0.530 | 0.742 | | |
| Obs. | 757120 | 757120 | 757120 | 757120 | 757120 | 757120 |
| F (first stage) | | | | | 40.957 | 53.540 |
| Dummies | t | k,r,t | kr,t | kr,kt,rt | t | kr,kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patents by natives ($nat_t^{rk}$). Explanatory variable: patenting activity by migrants ($mig_t^{rk}$). Instrumental variable: $\tilde{iv}_t^{rk}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

# B Analysis at the state level

Table B.1: Robustness of Table III: State level

| | Dependent variable: Patents of natives | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of migrants | 2.383*** | 0.124*** | 0.681*** | 0.193*** | 4.409*** | 0.380*** |
| | (0.084) | (0.028) | (0.071) | (0.036) | (0.718) | (0.108) |
| $Adj.R^2$ | 0.198 | 0.882 | 0.701 | 0.820 | | |
| Obs. | 106335 | 106335 | 106335 | 106335 | 106335 | 106335 |
| F (first stage) | | | | | 46.182 | 54.298 |
| Dummies | t | k,r,t | kr,t | kr,kt,rt | t | kr,kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patents by natives ($nat_t^{rk}$). Explanatory variable: patenting activity by migrants ($mig_t^{rk}$). Instrumental variable: $\tilde{iv}_t^{rk}$. States are used for the regional dimension. State and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

Table B.2: Robustness of Table IV: State level

| Dependent variable: Patents of natives (growth between $t$-1 and $t$) | | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of natives (t-1) | -0.150*** | -0.345*** | -0.745*** | -0.357*** | -0.186*** | -0.389*** |
| | (0.018) | (0.025) | (0.034) | (0.027) | (0.026) | (0.030) |
| Patents of migrants (t-1) | 0.316*** | 0.214*** | 0.136*** | 0.221*** | 0.669*** | 0.626*** |
| | (0.043) | (0.025) | (0.018) | (0.031) | (0.193) | (0.195) |
| $Adj.R^2$ | 0.087 | 0.235 | 0.290 | 0.333 | | |
| Obs. | 85068 | 85068 | 85068 | 85068 | 85068 | 85068 |
| F (first stage) | | | | | 39.136 | 47.397 |
| Dummies | t | k,r,t | kr,t | kt,rt | t | kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t-1\rightarrow t}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t-1}$) and migrants ($mig^{rk}_{t-1}$). Instrumental variable: $\tilde{iv}^{rk}_{t-1}$. States are used for the regional dimension. State and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

Table B.3: Robustness of Table V: State level

| Dependent variable: Patents of natives (growth between $t_1$ and $t_2$) | | | | |
|---|---|---|---|---|
| | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) |
| Patents of natives ($t_1$) | -0.409*** | -0.463*** | -0.474*** | -0.493*** |
| | (0.029) | (0.032) | (0.049) | (0.037) |
| Patents of migrants ($t_1$) | 0.589*** | 0.219*** | 0.911*** | 0.407*** |
| | (0.066) | (0.040) | (0.217) | (0.136) |
| $Adj.R^2$ | 0.278 | 0.647 | | |
| Obs. | 20961 | 20961 | 20961 | 20961 |
| F (first stage) | | | 50.362 | 79.779 |
| Dummies | | k,r | | k,r |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t_1 \to t_2}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t_1}$) and migrants ($mig^{rk}_{t_1}$). Instrumental variable: $\tilde{iv}^{rk}_{t_1}$. Time: $t_0 = [1870, 1890)$, $t_1 = [1890, 1930)$, $t_2 = [1930, 1950)$. States are used for the regional dimension. State and technology cluster robust standard errors in parentheses. First stage relevance reported with Kleibergen-Paap F statistic. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

# C Analysis dropping observations equal to zero

Table C.1: Robustness of Table III: Drop zeros (benchmark)

| | Dependent variable: Patents of natives | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of migrants | 1.432*** | 0.058** | 0.400*** | 0.229*** | 3.176*** | 0.246** |
| | (0.058) | (0.026) | (0.042) | (0.026) | (0.783) | (0.103) |
| $Adj.R^2$ | 0.386 | 0.918 | 0.769 | 0.890 | | |
| Obs. | 6118 | 2820 | 6043 | 3154 | 5946 | 2811 |
| F (first stage) | | | | | 8.413 | 30.352 |
| Dummies | t | k,r,t | kr,t | kr,kt,rt | t | kr,kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patents by natives ($nat_t^{rk}$). Explanatory variable: patenting activity by migrants ($mig_t^{rk}$). Instrumental variable: $\tilde{iv}_t^{rk}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

Table C.2: Robustness of Table IV: Drop zeros (benchmark)

| Dependent variable: Patents of natives (growth between $t$-1 and $t$) | | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of natives (t-1) | -0.130*** | -0.305*** | -0.987*** | -0.279*** | -0.188* | -0.282*** |
| | (0.035) | (0.049) | (0.058) | (0.044) | (0.108) | (0.069) |
| Patents of migrants (t-1) | 0.084*** | 0.010 | 0.000 | 0.035* | 0.228 | 0.046 |
| | (0.025) | (0.022) | (0.019) | (0.021) | (0.280) | (0.222) |
| $Adj.R^2$ | 0.078 | 0.390 | 0.501 | 0.488 | | |
| Obs. | 1827 | 1759 | 988 | 1605 | 1786 | 1599 |
| F (first stage) | | | | | 8.938 | 7.227 |
| Dummies | t | k,r,t | kr,t | kt,rt | t | kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t-1\rightarrow t}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t-1}$) and migrants ($mig^{rk}_{t-1}$). Instrumental variable: $\tilde{iv}^{rk}_{t-1}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

Table C.3: Robustness of Table V: Drop zeros (benchmark)

| Dependent variable: Patents of natives (growth between $t_1$ and $t_2$) | | | | |
|---|---|---|---|---|
| | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) |
| Patents of natives ($t_1$) | -0.184*** | -0.273*** | 0.036 | -0.273*** |
| | (0.036) | (0.030) | (0.165) | (0.052) |
| Patents of migrants ($t_1$) | 0.223*** | 0.125*** | -0.352 | 0.132 |
| | (0.050) | (0.026) | (0.392) | (0.174) |
| $Adj.R^2$ | 0.039 | 0.579 | | |
| Obs. | 4278 | 4206 | 4242 | 4189 |
| F (first stage) | | | 21.112 | 8.918 |
| Dummies | | k,r | | k,r |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t_1 \rightarrow t_2}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t_1}$) and migrants ($mig^{rk}_{t_1}$). Instrumental variable: $\tilde{iv}^{rk}_{t_1}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time: $t_0 = [1870, 1890)$, $t_1 = [1890, 1930)$, $t_2 = [1930, 1950)$. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

Table C.4: Robustness of Table III: Drop zeros (extended)

| | Dependent variable: Patents of natives | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of migrants | 0.923*** | 0.050*** | 0.327*** | 0.189*** | 2.126*** | 0.096*** |
| | (0.053) | (0.009) | (0.028) | (0.029) | (0.255) | (0.028) |
| $Adj.R^2$ | 0.398 | 0.917 | 0.750 | 0.824 | | |
| Obs. | 11168 | 8380 | 11158 | 8743 | 10900 | 8321 |
| F (first stage) | | | | | 30.795 | 85.924 |
| Dummies | t | k,r,t | kr,t | kr,kt,rt | t | kr,kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patents by natives ($nat_t^{rk}$). Explanatory variable: patenting activity by migrants ($mig_t^{rk}$). Instrumental variable: $\tilde{iv}_t^{rk}$. States are used for the regional dimension. State and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

Table C.5: Robustness of Table IV: Drop zeros (extended)

| Dependent variable: Patents of natives (growth between $t$-1 and $t$) | | | | | | |
|---|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Patents of natives (t-1) | -0.191*** | -0.349*** | -0.765*** | -0.301*** | -0.381** | -0.337*** |
| | (0.020) | (0.023) | (0.055) | (0.024) | (0.188) | (0.038) |
| Patents of migrants (t-1) | 0.104*** | 0.061*** | 0.019 | 0.075*** | 0.453 | 0.157*** |
| | (0.016) | (0.010) | (0.014) | (0.008) | (0.376) | (0.050) |
| $Adj.R^2$ | 0.107 | 0.370 | 0.429 | 0.547 | | |
| Obs. | 5848 | 5830 | 4702 | 5578 | 5795 | 5552 |
| F (first stage) | | | | | 9.369 | 42.526 |
| Dummies | t | k,r,t | kr,t | kt,rt | t | kt,rt |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t-1\rightarrow t}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t-1}$) and migrants ($mig^{rk}_{t-1}$). Instrumental variable: $\tilde{iv}^{rk}_{t-1}$. States are used for the regional dimension. State and technology cluster robust standard errors in parentheses. Time t is in decades. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

Table C.6: Robustness of Table V: Drop zeros (extended)

| Dependent variable: Patents of natives (growth between $t_1$ and $t_2$) | | | | |
|---|---|---|---|---|
| | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) |
| Patents of natives ($t_1$) | -0.375*** | -0.290*** | -0.962*** | -0.321*** |
| | (0.045) | (0.032) | (0.191) | (0.051) |
| Patents of migrants ($t_1$) | 0.417*** | 0.122*** | 1.356*** | 0.208** |
| | (0.049) | (0.015) | (0.329) | (0.087) |
| $Adj.R^2$ | 0.109 | 0.721 | | |
| Obs. | 5047 | 5030 | 5018 | 5008 |
| F (first stage) | | | 21.931 | 27.362 |
| Dummies | | k,r | | k,r |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t_1 \to t_2}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t_1}$) and migrants ($mig^{rk}_{t_1}$). Instrumental variable: $\tilde{iv}^{rk}_{t_1}$. States are used for the regional dimension. State and technology cluster robust standard errors in parentheses. Time: $t_0 = [1870, 1890)$, $t_1 = [1890, 1930)$, $t_2 = [1930, 1950)$. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*

# D Additional analysis on expertise

Table D.1: Re-computation of Table VI using the restricted dataset

| Dependent variable: Patents of natives (growth between $t_1$ and $t_2$) | | | | |
|---|---|---|---|---|
| | OLS | OLS | IV | IV |
| | (1) | (2) | (3) | (4) |
| Patents of natives ($t_1$) | -0.540*** | -0.630*** | -0.560*** | -0.652*** |
| | (0.024) | (0.019) | (0.030) | (0.023) |
| Patents of migrants ($t_1$) | 0.428*** | 0.338*** | 0.883*** | 0.758*** |
| | (0.047) | (0.046) | (0.284) | (0.246) |
| Expertise ($t_1$) | 0.862*** | 0.370*** | 0.580*** | 0.170** |
| | (0.136) | (0.101) | (0.152) | (0.077) |
| $Adj.R^2$ | 0.455 | 0.575 | | |
| Obs. | 148876 | 148876 | 148876 | 148876 |
| F (first stage) | | | 43.323 | 74.703 |
| Dummies | | k,r | | k,r |

*Notes: all variables are in logs. Dependent variable: growth of patenting activity by natives ($\Delta nat^{rk}_{t_1 \to t_2}$). Explanatory variables: patenting activity by natives ($nat^{rk}_{t_1}$) and migrants ($mig^{rk}_{t_1}$), and expertise ($e^{rk}_{t_1}$). Instrumental variable (for $mig^{rk}_{t_1}$): $\tilde{iv}^{rk}_{t_1}$. Cities (Metropolitan Statistical Areas) are used for the regional dimension. City and technology cluster robust standard errors in parentheses. Time: $t_0 = [1870, 1890)$, $t_1 = [1890, 1930)$, $t_2 = [1930, 1950)$. First stage relevance reported with Kleibergen-Paap F statistic. Significance is denoted with *** p<0.01, ** p<0.05, * p<0.1.*