European Immigrants and the United States' Rise to the Technological Frontier

Costas Arkolakis
Yale

Sun Kyoung Lee
Yale

Michael Peters
Yale

Abstract

What is the role of immigrants on (American) Growth? To answer this perplex question, we undertake a massive effort of collecting, digitizing, and harmonizing micro and macro economic data from the 19th and early 20th centuries. The data originate from the historical manufacturing and demographic census of the United States, immigration records datasets and the universe of US patents. To analyze the counterfactual implications of alternative allocations of immigrants, we develop a dynamical trade model where heterogeneous firms make innovation and exporting decisions across space and time. The model predicts that the timing and the spatial allocation of immigrant arrivals affect the path of growth outcomes for each location and the aggregate US economy. We use the structural equations arising from the model to interpret empirical findings from difference-in-difference analysis for the importance of the influx of skilled immigrants on the differential growth of US counties. We illustrate how to conduct counterfactual scenarios of alternative allocation of skilled immigrants from different countries across space and time to measure the economic impact of barriers to migration to the United States economy.

1 Introduction

The transformation of the US economy in the last two hundred years has been remarkable. While being primarily rural at the beginning of the 19th century, the US had developed into an essentially industrial nation when the century came to an end. More strikingly, after lagging behind the technological frontier (represented by the UK) for most of the 19th century, the US entered the twentieth century as the global technology leader and the richest nation on the globe (Gordon, 2017). During this period, which is also referred to as the "The Second Industrial Revolution", the US economy also experienced a massive inflow of immigrants, mostly from the European continent.

To answer this question, we make a massive effort to collect, digitize and harmonize data that combine longitudinal information on immigrants and their occupations along with measures of economic outcomes such as, output, wages, and innovation at the disaggregated county-industry level. Using this information we investigate to what extent this influx of immigrants was an important contributor to the transformation of the American economic landscape between 1850 and 1940 and provide a comprehensive account of the importance of spatial mobility for the immigrants' life-cycle. To analyze this wealth of data we develop a dynamical model of innovation where firms in each location innovate taking into account the full stream of future incomes that can be generated there now and in the future. Firms hire production and research workers. Workers, immigrants or natives, can specialize in either of these activities depending on their idiosyncratic abilities that are determined by their comparative advantage. The arrival of immigrants leads to standard scale effects, as in the Romer (1990) model or as we formally argue may even further boost innovation if immigrants have a comparative advantage on this activity.

To measure the impact of immigration on historical growth we combine several historical datasets. We use original immigration records and historical passenger lists from the ships heading from Europe to the US.[1] These data sources are real treasure troves for empirical researchers, as they contain direct microdata on immigrants' *pre*-migration occupations.[2] In particular, both in the immigration records and in the passenger lists, all immigrants were required to give a detailed account of their last occupation in Europe along with

[1]The immigration database of 13 million immigrants and the passenger lists of around 5 million immigrants leaving for the US via the German port Hamburg, the so-called "Hamburg Passenger Lists" were provided to us for research purposes by the Battery Conservancy and the Archives of the city of Hamburg, respectively. See http://www.castlegarden.org and http://www.germanroots.com/hamburg.html for additional information. To the best of our knowledge, these data sources have yet to be used in empirical research.

[2]We constructed a crosswalk between these published occupational strings and the *Historical International Classification of Occupations (HISCO)*. For more details on the Historical International Classification of Occupations, see http://historyofwork.iisg.nl/index.php.

other important information such as the time of arrival in the US, the place of residence in Europe, and age. We link these immigration datasets with the restricted-use complete count US Federal demographic decennial censuses from 1850-1940 using modern record-linking techniques. To link these datasets, we exploit the fact that both the immigration records (and passenger lists) and the federal US demographic Census contain time-invariant individual information. In that way, we construct a large-scale micro *panel* data set for immigrants with information on their *pre*-migration occupations and their *post*-migration labor market outcomes and spatial mobility patterns over their whole life-time.

To measure the economic impact of immigrants across space we combine this microdata with novel measures of wages, productivity growth, and patent activity. In particular, by first digitize the published results for the historical Manufacturing Census from 1860 to 1939, we construct data on wages and productivity measures at the county-industry level. Secondly, we also complement our data with new county-industry measures of patent activity and construct a measure of a patent novelty using textual analysis methods.

The model is tractable and allows the analysis of a range of regional policies, such as the ones that improve labor productivity, improve innovation productivity or the one that increases population due to arrivals of natives or immigrants. As it integrates a standard gravity trade model of trade (Anderson and Van Wincoop (2003)) with a forward-looking growth model, it can be used to study counterfactuals that analyze the transition path of the economy in the absence of immigrant arrivals.

Furthermore, the model implies a difference-in-difference empirical specification that intuitively links the growth in different regions to its fundamentals (such as past productivity or average quality of research or regional specific characteristics) the market size of a market and also the number of research workers. As the number of research workers depends on the inflow of immigrants this suggests the use of an instrumental variables specification that can be used to estimate the impact of immigrants on measures of growth such as productivity or the number of patents. This furthermore provides precise structural underpinnings to difference-in-difference approaches suggested in the literature (see Ottaviano et al. (2013)).

This combination of micro-information on immigrants and macro-measures of productivity and spatial idea creation allows us to relate knowledge flows (as proxied by inflowing immigrants with pre-migration expertise) to data on productivity growth and patent activity for the study period. By doing so, we provide novel evidence on potential mechanisms by which past immigrant settlements could affect economic outcomes (see e.g. Nunn et al. (2017) and Akcigit et al. (2017)). Furthermore, our study period is not only interesting in itself, but it also provides an ideal laboratory to empirically identify the importance of idea flows, which feature prominently in recent theories of economic growth (Kortum, 1997; Lucas Jr and Moll, 2014; Perla and Tonetti, 2014). As communication flows and technology were far less developed than those of today, the importance of embodied knowledge transmission was arguably much more important at that time.

Furthermore, we investigate the role of spatial mobility for the immigrants' earnings life-cycle. We exploit the longitudinal and spatial aspects of our dataset more intensely. In particular, we construct "spatial-sector"-based earning measures from our newly digitized information on manufacturing wages. It might be at the heart of understanding earnings differences between natives and immigrants (see eg Abramitzky et al. (2012)) as our data shows that there is a systematic positive correlation between average wages, urbanization and immigrant shares in the 19th century.[3]

The rest of the paper is structured as follows. Section 2 presents the data and the main facts about immigrants, their skills, and patterns of innovation. Section 3 presents the dynamical trade model and its main insights. Section 4 presents reduced form evidence based on a difference-in-difference specification suggested by the model. Section 5 discussed how the counterfactuals are conducted in our model. Section 6 concludes.

[3]This finding is consistent with the findings of a literature in urban economics that finds large city-wage premia in recent data (see e.g. Roca and Puga (2017)).

2 Immigrants and Innovation in the 19th Century

Our analysis relies on four main sources of data. Three datasets are individual-level data: novel datasets on millions of immigrants entering the US in the 19th and early 20th century, the full count population census for the years 1850 to 1940 and the population of US patents since 1790. Because our theory stresses the role of immigrants for local productivity growth, we combine this data with information on productivity at the county- and city-level, which we measure from the Manufacturing Census. The combination of these datasets allows us to systematically explore the relationship between immigrants' prior expertise, immigrant's location choice, productivity growth and patent activity at the county-industry level. An overview of these data sources is contained in Figure 1.

Historical Immigration Records (1820 - 1914) We construct our immigration database from two primary sources: the Castle Garden Immigration Database and the Hamburg Passenger Lists. The *Castle Garden Immigration Database (1820-1892)* is an educational project of the Battery Conservancy. The database contains the list of all immigrants entering the US via the port of New York between 1820 to 1892. In total, the database comprises approximately 11 million individual micro-records. Through cooperation with the Battery Conservancy, we have access to the entire Castle Garden Immigration database. We complement this database with original passenger lists of all immigrants leaving from the port of Hamburg to the US between 1850 and 1914 called the *Hamburg Passenger List Database (1850-1914).* We have access to the complete records through a cooperation with the Hamburg State Archive. Importantly, these immigration records contain detailed information on the pre-migration occupations of the respective immigrants and a host of demographic information like the name, age or family structure, which we use to match the immigration records to the population census (see below).

The Full Count Population Census (1850-1940) To measure the impact of immigration, we require information on immigrants' characteristics after their arrival in the US. We do so by linking the individual immigration records from the Immigration Database with the Complete Count Federal Demographic Census. We take advantage of the complete transcription of Federal Census Records between 1850 to 1940. The US federal demographic census year records exist for all years from 1850 to 1940 every decade except for 1890 (which was lost due to fire) and contain detailed information on individuals including demographic characteristics and employment information at the sectoral and occupational level.

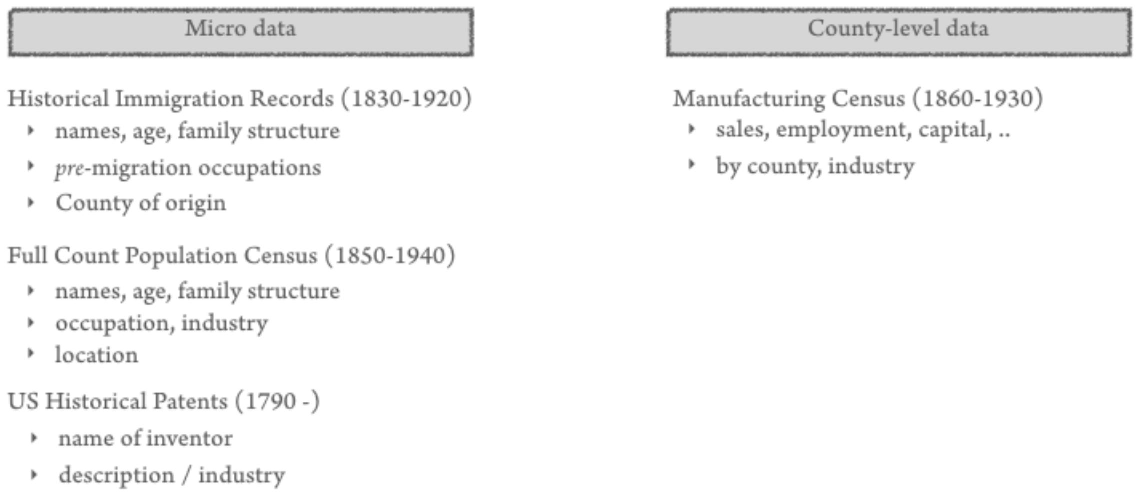


Figure 1: The Data

Historical Data on US Patents To measure innovative activity we exploit information on patenting. The United States Patent and Trademark Office (USPTO) granted millions of patents since 1790 and all patents contain information about the location of the patent and the name of the owner of the patent. We use the information on patenting in two ways. First, we study the determinants of patenting at the micro-level. This allows us to measure differences in innovative activity between natives and immigrants and how the propensity to innovate varies across occupations, sectors, and space. Second, we geo-reference the individual patents to connect them to the manufacturing census at the county-level. This allows us to study the response of patent activity at the county-level to the inflow of immigrants. We can also directly study whether the arrival of immigrants induced new patent activity in a particular location. To do so, we devise a measure of *spatial idea novelty.* This measure exploits textual analysis to measure the extent to which patents originating in a particular region are similar to the patents that have been invented in that region in the past.[4]

The Manufacturing Census (1870 - 1929) Our empirical strategy heavily relies on spatial variation in productivity growth, innovation activity and the settlement of immigrants. Measuring productivity at a fine spatial resolution in the 19th Century is difficult. First of all, there are no measures of wages at the county-level. While the available individual-level

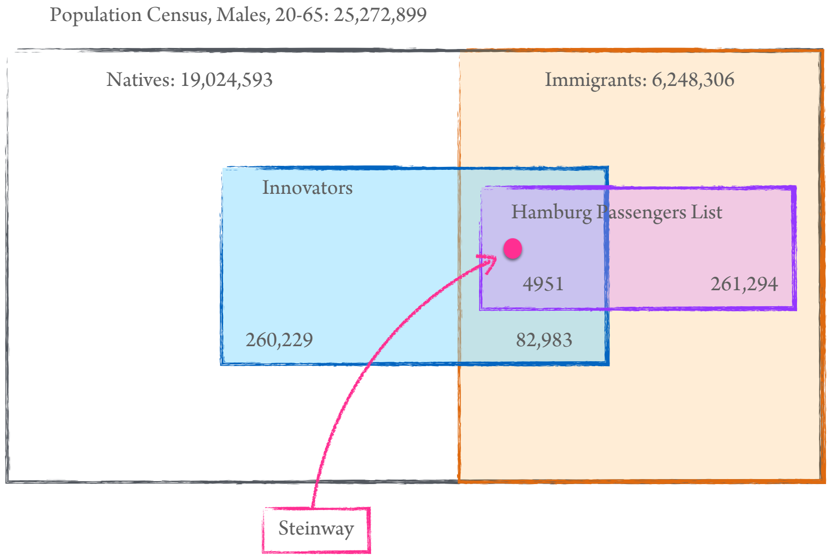
[4]Intuitively, as in **?**, we measure the similarity between two patents as the correlation in the words, respective patents use. Given this measure of patent-to-patent similarity, we then calculate the similarity of ideas invented in a region as the average patent-to-patent similarity between new patents originated in a particular region and the set of patents stemming from a particular region in the past. See Section B in the Appendix for more details.

data in the decennial Population Census contains county-identifiers, earnings have only been reported starting in 1940. Secondly, information on labor earnings stemming from the National Accounts is available in the 19th Century, but the data does not have a spatial dimension. To overcome this problem, we digitized the published results from the Census of Manufacturers. These tables are published at either the county-industry level or city-industry level and report standard information from firms' balance sheets. In particular, they report the number of manufacturing establishments, total number of workers, value of manufactured output, wage-bill and value of the capital stock. As our main measures of spatial productivity, we consider total value-added per manufacturing employee, manufacturing value-added relative to the wage bill or manufacturing revenue TFP, i.e. $Y_r / \left(K_r^{\alpha} L_r^{1-\alpha}\right)$, where Y_r is total value-added and K_r and L_r denote the capital stock and employment. We digitize this data at the *county*-industry level for the years 1860, 1870, 1880 and 1929 and at the *city*-industry level for the years 1880, 1900, 1909, 1919, 1929 and 1939. In Section B in the Appendix, we explain how we harmonize this information and combine it with other sources.

2.1 Record Matching

A key aspect of our empirical work is to match the historical immigration records to the population census and to the microdata on patents. While we relay all details of the matching procedure to Section B in the Appendix, in Figure 2 we summarize some aspects for the year 1910. For now, we focus on males between 20 and 60 years old, which we suspect to account for the majority of patent holders. In the federal census of 1910, there about 25m. We then try to match the patent grantee from the individual patent records to the population census using the information on the name and the geographic location of the inventor. In total we observe 0.5m patents, which are issued before 1910.[5] Of these, we can match 350,000 to individuals in the population. Similarly, we turn to the immigration records from the Hamburg Passenger Lists and match individual immigrants to the population census using the information on name, the date of arrival, nationality and family characteristics (all of which are available in the population census and the immigration records). We can match 266,000 records between the Hamburg Passenger Lists and the demographic US 1910 Census. Finally, the union of matches who are like ''Steinways'', i.e. individual migrants who came from the port of Hamburg, whom we can track in 1910 US Census and whom we held at least one patent in the US, accounts for 5,000 individuals.

[5]In practice we look at the population of patents issued between 1838 and 1910 where the inventors' name is available.



Notes: The figure displays the relationships of our matched data micro data in 1910 using the Population Census, and patent records and the Hamburg Passenger Lists. Because we match patents and immigration records to the Population Census, they are strict subsets of the population in the census.

Figure 2: Matching Immigrants and Patents to the Population Census in 1910

This sample forms the basis of our micro-level dataset. In Section 2.3 below we present an analysis of basic empirical regularities. Before turning to this systematic analysis we want to present a particular individual biography, which summarizes the process of this record matching procedure.

2.2 A guiding example: The Story of Heinrich Engelhard Steinweg

To illustrate how our novel data can be merged with existing data sources to shed light on the process of technology transfer from old Europe, we start by a particular example: the case of Heinrich Engelhard Steinweg, later known as Henry Engelhard Steinway, the founder of renowned piano manufacturing company *Steinway & Sons*. Heinrich Steinweg left Germany on May 28th, 1850 via the port of Hamburg. This information is declared on the *Hamburg Passenger Lists (1850-1934),* which is available to us through a cooperation with the Archives of the city of Hamburg. His shipment record also indicates his pre-immigration occupation in Germany as *Instrumentenmacher* (instrument maker).[6] As we can see from the same record, shown in Figure 3, his destination was New York and he was accompanied by four family members.

H Steinweg
in the **Hamburg Passenger Lists, 1850-1934**

Name:	H Steinweg
Occupation:	Instrumentenmacher
Birth Place:	Seesen, Hannover
Departure Date:	28 Mai 1850 (28 May 1850)
Port of Departure:	Hamburg
Destination:	New York (New York City (All Boroughs))
Port of Arrival:	New York (New York City (All Boroughs))
Ship Name:	Helena Sloman
Captain:	Paulsen, P N
Shipping Clerk:	Rob. M. Sloman
Shipping line:	Rob. M. Sloman
Ship Type:	Dampfschiff
Accommodation:	ohne Angabe
Volume:	373-7 I, VIII A 1 Band 001

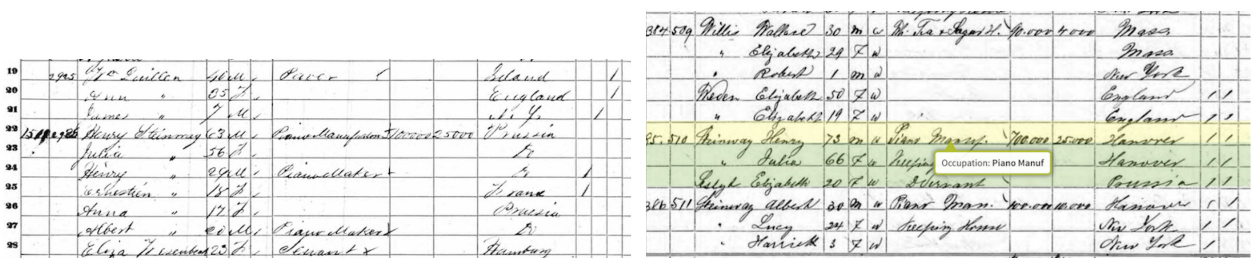
Household Members:

Name	Age
H Steinweg	
Steinweg	
Steinweg	7
Steinweg	
Steinweg	23

Notes: The figure shows the records of Heinich Steinweg in the Hamburg Passenger Lists.

Figure 3: Heinrich Steinweg in the *Hamburg Passenger Lists, 1850-1934*

We can track Mr. Steinweg, now Mr. Steinway, in the subsequent US population censuses, shown in Figure 4. Both in in 1860 and 1870, Mr. Steinway and his family are recorded to reside in New York. Furthermore, his occupation *piano manufacturer* indicates Steinway's successful transition from a piano maker in Germany to the piano manufacturer of the US.



Notes: The figure shows Mr Steinway's US census records in 1860 (left panel) and 1870 (right panel).

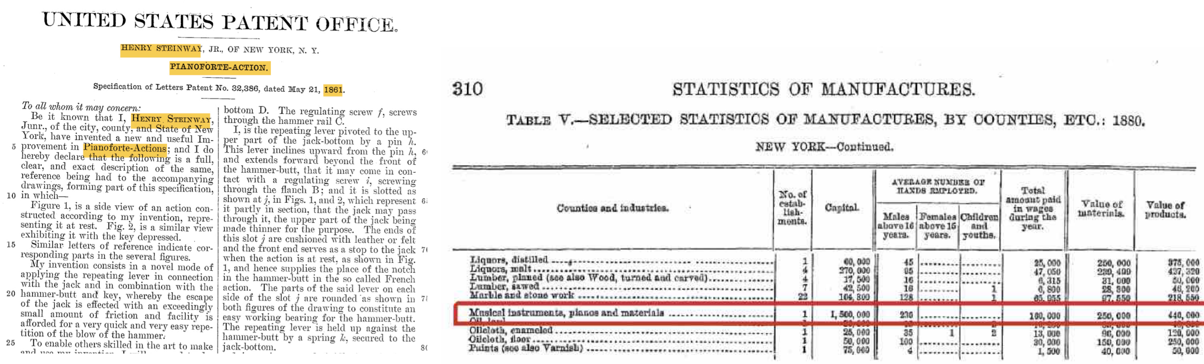
Figure 4: Henry Steinway in the *US Census Schedules 1860 and 1870*

That this successful career trajectory might have been in part due to Mr Steinweg's prior knowledge is consistent with micro data on patenting. Using digitized historical patent

[6]Henry Steinway started working on producing pianos early on with immediate success. But the unstable political climate following the revolutions of 1848 and the limited economic opportunities for a man working outside a guild let him to immigrate to the US. See Claudius Torp, ‘‘Heinrich Engelhard Steinway.’’ In Immigrant Entrepreneurship: German-American Business Biographies, 1720 to the Present, vol. 2, edited by William J. Hausman. German Historical Institute.

data from the United States Patent and Trademark Office, we could extract several patents granted to him and his sons' names. For example, Steinway's famous piano-forte patent, dated 1862, is shown in the left panel of Figure 5.

Finally, we can use our newly digitized data from the US Census of Manufactures to learn about the economic magnitude of Mr. Steinway's success. While the US Census of Manufacturers data is not at the plant level (but reported at industry-by-county cells), the information is detailed enough to identify the main manufacturing plant of Steinway & Sons in Queens, NY. As the right panel of Figure 5 shows, the Steinway family had an enormous impact on manufacturing production in the New York area. The digitized Census of Manufacturers for the year 1880, for example, reveals that this single piano manufacturing plant was one of the most capital intensive sectors in New York City with more than $1.5 millions of capital and sales close to a half a million dollars.



UNITED STATES PATENT OFFICE.

HENRY STEINWAY, JR., OF NEW YORK, N. Y.

PIANOFORTE-ACTION.

Specification of Letters Patent No. 32,386, dated May 21, 1861.

To all whom it may concern:

Be it known that I, HENRY STEINWAY, Junr., of the city, county, and State of New York, have invented a new and useful Improvement in Pianoforte-Actions; and I do hereby declare that the following is a full, clear, and exact description of the same, reference being had to the accompanying drawings, forming part of this specification, in which—

Figure 1, is a side view of an action constructed according to my invention, representing it at rest. Fig. 2, is a similar view exhibiting it with the key depressed.

Similar letters of reference indicate corresponding parts in the several figures.

My invention consists in a novel mode of applying the repeating lever in connection with the jack and in combination with the hammer-butt and key, whereby the escape of the jack is effected with an exceedingly small amount of friction and facility is afforded for a very quick and very easy repetition of the blow of the hammer.

To enable others skilled in the art to make

bottom D. The regulating screw *f*, screws through the hammer rail C.

I, is the repeating lever pivoted to the upper part of the jack-bottom by a pin *h*. This lever inclines upward from the pin *h*, and extends forward beyond the front of the hammer-butt, that it may come in contact with a regulating screw *i*, screwing through the flanch B; and it is slotted as shown at *j*, in Figs. 1, and 2, which represent it partly in section, that the jack may pass through it, the upper part of the jack being made thinner for the purpose. The ends of this slot *j* are cushioned with leather or felt and the front end serves as a stop to the jack when the action is at rest, as shown in Fig. 1, and hence supplies the place of the notch in the hammer-butt in the so called French action. The parts of the said lever on each side of the slot *j* are rounded as shown in both figures of the drawing to constitute an easy working bearing for the hammer-butt. The repeating lever is held up against the hammer-butt by a spring *k*, secured to the jack-bottom.

310 STATISTICS OF MANUFACTURES.

TABLE V.—SELECTED STATISTICS OF MANUFACTURES, BY COUNTIES, ETC.: 1880.

NEW YORK—Continued.

Counties and industries.	No. of establishments.	Capital.	Average number of hands employed. Males above 16 years.	Females above 15 years.	Children and youths.	Total amount paid in wages during the year.	Value of materials.	Value of products.
Liquors, distilled	1	60,000	45			25,000	250,000	375,000
Liquors, malt	4	270,000	95			47,050	230,400	437,320
Lumber, planed (see also Wood, turned and carved)	4	17,500	16			6,315	31,000	50,000
Lumber, sawed	7	42,500	16		1	6,800	28,300	46,200
Marble and stone work	22	104,300	128		1	65,955	97,550	218,550
Musical instruments, pianos and materials	1	1,500,000	230			160,000	250,000	440,000
Oil, lead	[illegible]	[illegible]	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]
Oilcloth, enameled	1	25,000	35	1	2	13,000	96,000	120,000
Oilcloth, floor	1	50,000	100			30,000	150,000	250,000
Paints (see also Varnish)	1	75,000	4			1,500	40,000	50,000

Notes: The figure shows the original text of Mr Steinweg's piano forte patent (left panel) and the entry for the Steinway & Sons piano factory in the manufacturing census.

Figure 5: Mr Steinway's Pianoforte patent and Steinway & Sons in the *US Census of Manufacturers 1880*

2.3 Patenting & Immigration in the early 20th Century: Systematic Evidence

In this section we provide systematic evidence on the patterns of patenting from our matched micro data summarized in Table 2. Using this data we highlight five empirical regularities. These regularities are not only interesting per se but they inform the structure of our theory presented below. We find that:

1. Innovators were rare and showed a clear life-cycle profile

2. Immigrants were equally important for the process of innovation as natives. But there is substantial heterogeneity in patent activity by immigrants' nationality

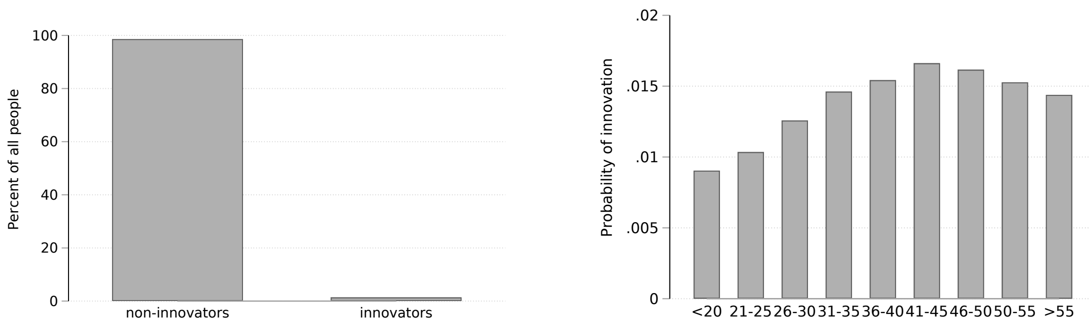
3. Immigrant and native innovators specialized in different occupations and industries

4. Innovation is highly concentrated in cities

5. Immigrants settled in urban areas

6. Innovation activity is tightly correlated with occupation-based measures of skills

Figure 6 shows two facts about patent activity in the 19th century, which are reminiscent of patent activity today. First of all, patenting was relatively rare - on average about 1.5% of all individuals held any patents and the vast majority of patentees held a single patent.[7] In the right panel we show that patenting had a clear life-cycle profile. In particular, the probability of patenting peaked at around 45 years.



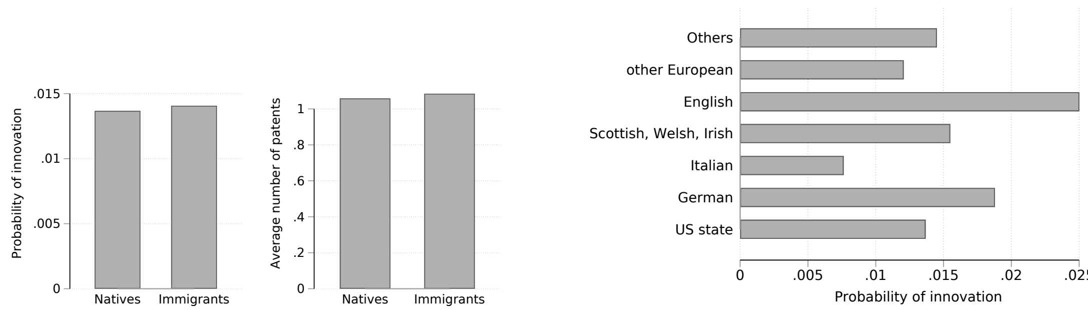
Notes: In the left panel we show the share of individuals in the Population Census who own at least one patent (‘‘Innovators’’). In the right panel we show the life-cycle of patenting, i.e. the share of individuals with at least one patent by age.

Figure 6: Regularity 1: Patent Propensity & The Patent Lifecycle

Figure 7 documents our second fact, i.e. that immigrants played a similar role for the process of innovation as natives. In the left panel, we show that the share of innovators among immigrants and the number of patents conditional on patenting are very similar between natives and immigrants. In the right panel we show that there is substantial heterogeneity in patent proclivity across nationalities. While the English and the Germans were more likely to patent as natives, the Italians had a lower propensity to patent.

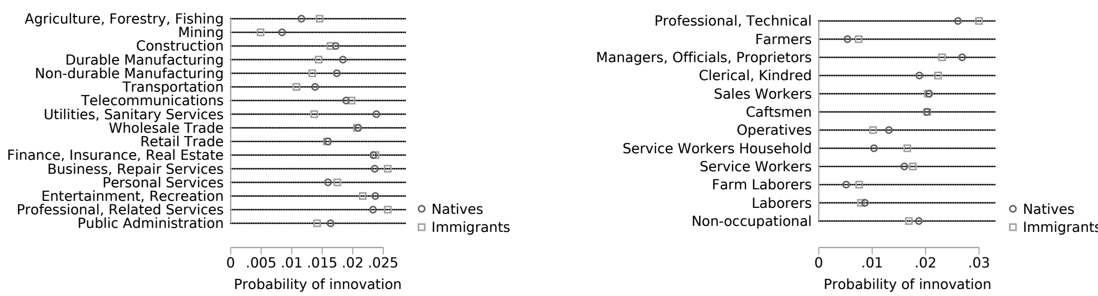
Not only was there substantial heterogeneity in patent activity across nationalities, but immigrant and native innovators were also active in different sectors and occupations. This is seen in Figure 8, where we report the distribution of innovators across sectors (left panel)

[7] In the Appendix we document that the distribution of the number of patents conditional on patenting was highly skewed - most patentees had a single patent but some innovators were very productivity owning more than 30 patents.



Notes: The figure on the left shows the share of individuals in the Population Census who own at least one patent (“Innovators”) by their immigration status and the average number of patents by the immigration status of inventor in 1910. In the right panel we display the share of innovators by the nationality of the innovator.

Figure 7: Regularity 2: Patent Activity of Immigrants

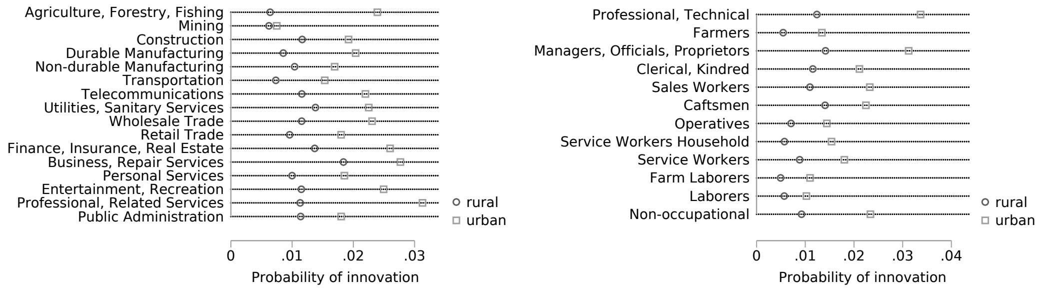


Notes: The figure shows the probability of patenting by sector (left panel) and occupation (right panel) for both immigrants and natives.

Figure 8: Regularity 3: Heterogeneity in Patent Activity Across Sectors and Occupations

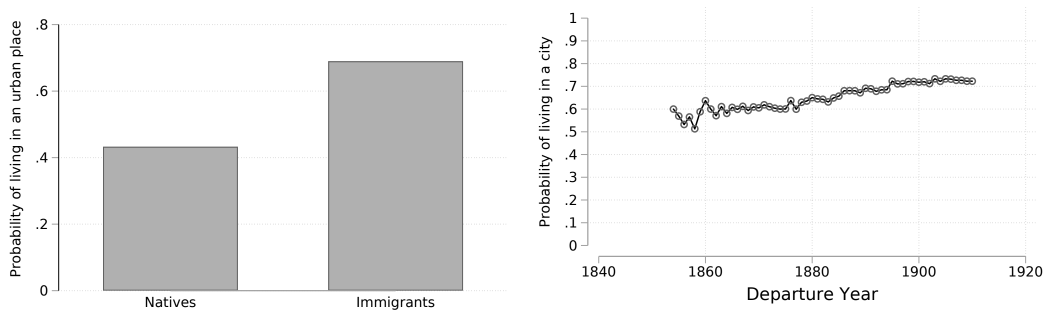
and occupations (right panel) for both immigrants and natives. The figure shows that immigrants’ input into the innovation process varies across both industries and occupations. While 25% of all native innovators are in the agricultural sector, this is only true for 15% of all immigrant patentees. Conversely, immigrant inventors are more likely to be in durable manufacturing, non-durable manufacturing and retail trade. A similar pattern emerges at the occupational level. While one third of all immigrant innovators are craftsmen, this it only true for 20% of all natives. Natives are in turn more likely to be farmers and farm laborers and managers, officials and proprietors. This heterogeneity reflects both the different occupational distribution between immigrants and natives and differences in the probability of patenting within sectors and occupations.

Our fourth regularity concerns the spatial distribution of patenting. As expected, most patent holders reside in urban areas. This is seen in Figure 9, which shows the probability of



Notes: The figure shows the probability of patenting by sector (left panel) and occupation (right panel) for in urban and rural areas.

Figure 9: Regularity 4: The Spatial Concentration of Patenting



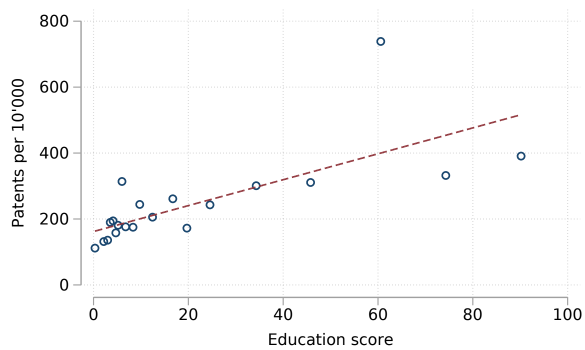
Notes: The left panel shows the probability of living in an urban area by immigration status in 1910. The right panel probability of living in an urban area status in 1910 for immigrants by their date of arrival as reported in the Hamburg Passenger Lists.

Figure 10: Regularity 5: Spatial Sorting of Immigrants

patenting within sectors (left panel) and occupation (right panel) by the urban status of the patent holder. It is seen that living in an urban area increases the probability of patenting by a factor 2 or 3. Reassuringly, this is not the case in the agricultural sector of in the mining industry.

Fifth, we show in Figure 10 that immigrants pre-dominantly live in cities. In the left panel we show that roughly 70% of the foreign born population lives in cities, while this is only the case for 40% of the native population. In the right panel we show that this ‘‘urban bias’’ of immigrants is extremely persistent. From our matched data we can calculate the urban share among immigrants by their date of arrival. While there is slight negative slope, which is expected given that most immigrants arrived in the cities, the urban share is still much higher among the foreign born population even if their lived in the US for decades.

Sixth, we show that - like today - patenting is strongly correlated with skill. Information



Note: The figure shows the number of patents (per 1000 workers) in each occupation, where occupations are ranked by their education ranked as reported in the Population Census.

Figure 11: Regularity 6: Patent Activity By Occupation Skill

on individuals' skill in the early 20th century is scarce. However, the US census contains information on educational attainment and earnings potential at the occupational level. This allows us to rank occupations by these measures and connect them to patent activity (aggregated to the occupational level). As we shows in the Appendix, the earnings-based and education-based measures of skills are highly correlated. In Figure 11 we connect the educational ranking of occupations to a measure of innovation output - the number of patents per 10.000 people working in a particular occupation. The figure shows a clear positive relationship, i.e. occupations with a higher educational score do indeed generate more patent activity.

In Tables 1 and 2 we show these patters in a regression format. Table 1 reports the results of estimate the specification

$$P_{it} = \beta \times Immigrant_{it} + X_{it}'\gamma + u_{it}, \tag{1}$$

where P_{it} is a dummy variable for whether or not individual i has a patent, $Immigration_{it}$ is a dummy variable for the immigrant status and X_{it} contains various controls. Hence, 1 studies the extensive margin of patenting. In Table 2 we focus on the intensive margin, where we replace P_{it} with the log number of patenting, conditional on patenting.

Table 1 documents the importance of spatial sorting for the extensive margin of patenting. Column 1 shows that - in accordance with Figure 7 - immigrants are as likely to patent as natives. In column 2 we specifically ask whether ''native movers'', i.e. natives, that live in state different from their birthplace, are relative more likely to patent. Column 3, which includes a set of occupation and sector fixed effects, shows that there are not important compositional differences between immigrants and natives. Columns 4 - 6 then document the

	(1)	(2)	(3)	(4)	(5)	(6)
			Probability of patenting			
Dummy Immigrant	0.000395	0.00236	0.000772	-0.000548	-0.00299***	-0.00624***
	(0.00119)	(0.00128)	(0.000742)	(0.000658)	(0.000433)	(0.000911)
Dummy Native Mover		0.00604***	0.00409***	0.00379**	0.00420***	0.00281***
		(0.00150)	(0.00116)	(0.00111)	(0.000955)	(0.000493)
Age			0.000155***	0.000154***	0.000139***	0.000162***
			(0.0000132)	(0.0000137)	(0.0000131)	(0.0000117)
Dummy Urban				0.00901***	0.00724***	0.00120***
				(0.000963)	(0.000831)	(0.000160)
Age			Yes	Yes	Yes	Yes
Occupation FE			Yes	Yes	Yes	Yes
Sector FE			Yes	Yes	Yes	Yes
State FE					Yes	
County FE						Yes
R^2	0.000	0.000	0.006	0.007	0.009	0.013
N	25272899	25272899	25272898	25272898	25272898	25272897

Notes: Robust standard errors in parentheses. *, **, *** denotes significance at 10%, 5% and 1% respectively. The sample comprises all males between 20 and 65 in 1910. "Dummy Immigrant" is a dummy variable for whether the individual is foreign born. "Dummy Native Mover" is a dummy variable for whether or not the individual is a native but lives in different state than his birthplace.

Table 1: Immigration & Patenting: Extensive Margin

extent of spatial sorting. Once we control for a dummy of living in an urban area (column 4), state fixed effects (column 5) and county fixed effects (column 6), immigrants appear less and less likely to patent. This pattern is consistent with the fact that immigrants are more likely to live in urban areas (Figure 10) and that innovation activity is much higher in urban areas (Figure 9).

In Table 2 we show that this *not* the case for the intensive margin of innovation. Using the same specification as in Table 1 we show that immigrants are slightly more productive innovators conditional on patenting and that this productivity advantage is not only driven by their spatial sorting: even when controlling for the location of the innovator, the average number if patents is about 0.5%-1% higher than for natives. This is about the same order of magnitude as for moving natives.

	(1)	(2)	(3)	(4)	(5)	(6)
			log Number of patents			
Dummy Immigrant	0.0139**	0.0166**	0.0144***	0.0128***	0.0105***	0.00448***
	(0.00494)	(0.00484)	(0.00407)	(0.00355)	(0.00286)	(0.000924)
Dummy Native Mover		0.00626**	0.00466*	0.00417*	0.00889***	0.00554***
		(0.00203)	(0.00174)	(0.00163)	(0.00228)	(0.000999)
Age			0.0000594	0.0000680	0.0000722	0.000161***
			(0.0000631)	(0.0000591)	(0.0000535)	(0.0000264)
Dummy Urban				0.0177**	0.0128***	0.00328***
				(0.00514)	(0.00310)	(0.000831)
Age			Yes	Yes	Yes	Yes
Occupation FE			Yes	Yes	Yes	Yes
Sector FE			Yes	Yes	Yes	Yes
State FE					Yes	
County FE						Yes
R^2	0.001	0.001	0.006	0.008	0.014	0.024
N	348163	348163	348160	348160	348160	348062

Notes: Robust standard errors in parentheses. *, **, *** denotes significance at 10%, 5% and 1% respectively. The sample comprises all males between 20 and 65 in 1910. ‘‘Dummy Immigrant’’ is a dummy variable for whether the individual is foreign born. ‘‘Dummy Native Mover’’ is a dummy variable for whether or not the individual is a native but lives in different state than his birthplace.

Table 2: Immigration & Patenting: Intensive Margin

3 Theory

The country is divided in R regions denoted by r. Time is discrete. We assume the existence of iceberg trade costs between locations. Thus, the final good price of goods from location r in location j is given by

$$p_{rjt} = \tau_{rj} p_{rt},$$

where p_{rt} denotes the price in location r. All workers have identical Constant Elasticity of Substitution (CES) preferences over goods of all locations r arriving in region j, c_{rjt}, at period t given by

$$C_{jt} = \left(\sum_r c_{rjt}^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{2}$$

where ε is the elasticity of substitution across varieties.

Total spending (and income) of the representative agent in region r is denoted by E_{rt}. Workers can work as production and research workers with associated wages in location r, w_r^P and w_r^R. The native population in each region is denoted by L_{rt} and the number of immigrants is denoted by I_{rt}.

Each region produces a final tradable good, which we denote by Y_{jt}. The production of this final good in each location requires a unit continuum of differentiated, non-tradable varieties i, $x_t(i)$, so that

$$Y_{rt} = Z_{rt}^A \left(\int_{i=0}^{1} x_t(i)^{\frac{\sigma-1}{\sigma}} \, di \right)^{\frac{\sigma}{\sigma-1}}, \tag{3}$$

where Z_{rt} is a regional exogenous productivity term. The price of the good produced in in the region is given by

$$p_{rt} = \frac{U_{rt}}{Z_{rt}^A},$$

where U_{rt} denotes the cost of production of the intermediate input bundle and is given by

$$U_{rt} = \left(\int_{i=0}^{1} u_{rt}(i)^{1-\sigma} \, di \right)^{\frac{1}{1-\sigma}},$$

where $u_{rt}(i)$ is the price per unit of variety i in region r at time t.

3.1 Firms and Innovation

Firms are monopolists for their differentiated varieties. Firms differ by location and efficiency. The production function for varieties in region r with efficiency z is given by

$$x_{rt}(z) = z\left(q_{rt}(z)\right)^{\frac{1}{\sigma-1}} h_{rt}(z) \tag{4}$$

where h_{rt} denotes the total amount of efficiency units hired for production by a firm in region r time t, q_{rt} denotes the quality of firm z in region r and z is an exogenous, firm-specific efficiency.[8]

While z is an exogenous firm-characteristic, which is constant, quality q evolves endogenously. In particular, firms can increase their current quality $q_{rt}(z)$ by a factor $i_{rt}(z)$ according to

$$q_{rt+1}(z) = q_{rt}(z)\, i_{rt}(z)\,. \tag{5}$$

Note that we, in principle, allow for depreciation of the firm quality i.e. if firms do not innovate enough, their productivity might decline, $i_t < 1$. As we show below, this will not be the case along a Spatial Balanced Growth Path. To innovate, firms need to hire researchers. More specifically, we assume that to increase their quality by $i_{rt}(z)$, firms pay a cost of

$$c_{rt}^{I}(q; z) = \frac{1}{\zeta_r Z_t^I} \frac{z^{\sigma-1} q}{Q_{rt}^{\lambda}} \frac{i_t^{\iota}}{\iota} w_{rt}^{R}, \tag{6}$$

where Q_{rt} is the average productivity in region r at time t

$$Q_{rt} = \mathbb{E}_{rt}\left[z^{\sigma-1} q(z)\right] = \int z^{\sigma-1} q_{rt}(z)\, dF_{rt}(z) \tag{7}$$

and $F_{rt}(z)$ is the cross-sectional distribution of firm efficiencies in region r at time t. Equation (6) stresses that the costs of innovation depend on both firm-level and regional characteristics. On the regional level, they are determined by the prevailing research wage in location r, w_{rt}^R, a fixed region-specific "innovation efficiency" ζ_r and a spill-over term Q_{rt}^{λ}. The parameter λ governs the extent to which research cost fall in the existing level of productivity Q_{rt}. As we show below, if $\lambda = 1$ the model becomes an endogenous growth model and if $\lambda < 1$ the model is a semi-endogenous growth model. We also show that this distinction makes precise predictions on the long-run effect of immigrant inflows on regional economic activity in the long-run. As in Atkeson and Burstein (2010) we also assume that innovation costs

[8] The scaling $q^{\frac{1}{\sigma-1}}$ is a normalization that allows to write profits in a linear form, ultimately.

are linear in firm efficiency, $z^{\sigma-1}q$. This scaling implies that the model is consistent with Gibrat's Law where growth is independent of size. Finally, Z_t^I is a time-varying efficiency shifter determining the cost of innovation, which is common across all locations.

Firm optimization The constant elasticity aggregator across intermediate input producers implies that firms' prices are given by a constant markup over the production cost

$$u_{rt}\left(z\right)=\frac{\sigma}{\sigma-1}\frac{w_{rt}^P}{zq_{rt}\left(z\right)^{\frac{1}{\sigma-1}}}. \tag{8}$$

Here w_{rt}^P is the wage for production workers in region r at time t. Standard arguments imply that firm z in region r at time t has a variable profit of

$$\pi_{rt}\left(z\right)=\frac{1}{\sigma}\left(\frac{u_{rt}\left(z\right)}{U_{rt}}\right)^{1-\sigma}E_{rt}=\frac{1}{\sigma}\frac{z^{\sigma-1}q\left(z\right)}{Q_{rt}}E_{rt}. \tag{9}$$

Firms' innovation decisions are of course dynamic in nature. Letting the real interest rate be r_t, the value function of a firm in region r at time t with a productivity $z^{\sigma-1}q\left(z\right)$ is given

$$V_{rt}\left(z^{\sigma-1}q\right)=\pi_{rt}\left(z^{\sigma-1}q\right)+\max_{i_t}\left[\frac{1}{1+r_t}V_{rt+1}\left(z^{\sigma-1}qi_t\right)-\frac{1}{\zeta_r Z_t^I}\frac{z^{\sigma-1}q}{Q_{rt}^{\lambda}}\frac{i_t^{\iota}}{\iota}w_{rt}^R\right]. \tag{10}$$

Hence, $\frac{1}{1+r_t}V_{rt+1}\left(z^{\sigma-1}qi_t\right)$ is the expected value of being a firm with productivity qi_t in period $t+1$.

Proposition 1. *Consider the value function $V_{rt}\left(q\right)$ in (10). The value function is linear homogeneous in q, i.e. $V_{rt}\left(z^{\sigma-1}q\right)=z^{\sigma-1}qv_{rt}$, where*

$$v_{rt}=\frac{1}{\sigma}\frac{E_{rt}}{Q_{rt}}+\frac{\iota-1}{\iota}\frac{w_{rt}^R}{\zeta_r Z_t^I}\frac{i_t^{\iota}}{Q_{rt}^{\lambda}}, \tag{11}$$

and the optimal rate of innovation i_{rt} is given by

$$i_t=\left(\frac{v_{rt+1}}{1+r_t}\frac{Q_{rt}^{\lambda}\zeta_r Z_t^I}{w_{rt}^R}\right)^{\frac{1}{\iota-1}}. \tag{12}$$

Proof. See Section A.1 in the Appendix. □

Proposition 1 shows that innovation incentives i_{rt} are equalized across all firms in location r. This is an implication of the homogeneity of the value function. The policy function for firms' innovation incentives in (12) shows that the optimal innovation rate depends on the discounted future value $\frac{v_{rt+1}}{1+r_t}$ relative to the cost of innovation $\frac{Q_{rt}^{\lambda}\zeta_r Z_t^I}{w_{rt}^R}$. Note also that by combining (11) and (12) we can express the value function as a forward looking difference equation

$$v_{rt} = \frac{1}{\sigma}\frac{E_{rt}}{Q_{rt}} + \left(\frac{\iota - 1}{\iota}\right)\left(\frac{Q_{rt}^{\lambda}\zeta_r Z_t^I}{w_{rt}^R}\right)^{\frac{1}{\iota-1}}\left(\frac{v_{rt+1}}{1+r_t}\right)^{\frac{\iota}{\iota-1}}. \tag{13}$$

The Endogenous Law of Motion for Aggregate Quality

Using equation (5) and aggregating among firms given the innovation equation (11) we obtain

$$Q_{rt+1} = Q_{rt} i_{rt}$$

Combining this equation, with (20) we can directly relate the the demand for research workers to the resulting growth rate

$$\frac{Q_{rt+1}}{Q_{rt}} = \left(H_{rt}^R \zeta_r Z_t^I Q_{rt}^{\lambda-1}\iota\right)^{1/\iota}, \tag{14}$$

i.e. the evolution of local productivity Q_{rt} is fully determined from the equilibrium amount of researchers H_{rt}^R.

3.2 Aggregate Labor Demand

Above we have characterized the optimal decisions of firms conditional on wages. We now aggregate these decisions, conditional on prevailing wages, to construct aggregate labor demand for innovation and production.

An Aggregate Production Function Our economy aggregates - at the production side - to a standard macro-spatial model. In particular, we can define an aggregate production function of each location that is linear on the total supply of efficiency units of labor in region r and time t and its aggregate productivity is determined by aggregate object of the economy. This result is formalized in the following Lemma

Lemma 1. *Consider the model above. Let H_{rt}^P be the total supply of efficiency units of production workers in region r at time t. Aggregate output of the tradable good in region r is given by*

$$Y_{rt} = A_{rt} \times H_{rt}^P,$$

where the endogenous TFP term A_{rt} is given by

$$A_{rt} = Z_{rt}^A \left(Q_{rt}\right)^{\frac{1}{\sigma-1}}. \tag{15}$$

Proof. See Section A.3 in the Appendix. □

Lemma 1 and (14) show that the evolution of spatial labor supply $\left\{H_{rt}^{P}, H_{rt}^{R}\right\}$ fully summarize the evolution of aggregate output $\{Y_{rt}\}$: $\left\{H_{rt}^{R}\right\}$ determines the evolution of productivity $\{Q_{rt}\}$ from (14) and $\left\{H_{rt}^{P}\right\}$ determines aggregate output as in a standard model of trade.

Aggregate Demand for Production and Innovation Workers To aggregate the model we characterize the aggregate demand for production and innovation labor.

Notice that given the CES demand one can show that production workers receive a constant share of aggregate income. Hence, total spending, E_{rt}, can be written as a function of production worker income

$$E_{rt} \equiv w_{rt}^{P} H_{rt}^{P} \frac{\sigma}{\sigma - 1}. \tag{16}$$

Regional trade flows are statically determined by considering firm prices. Each consumer consumes products from different locations. Given the assumption of CES demand, the market share of location r in the basket of location j at time t is given by

$$\lambda_{rjt} = \frac{\int \left(u_{rt}\left(z\right) dF_{rt}\left(z\right)\right)^{1-\sigma}}{\sum_{r'} \int \left(u_{r't}\left(z\right) dF_{r't}\left(z\right)\right)^{1-\sigma}} = \frac{\left(\frac{w_{rt}^{P}}{A_{rt}}\right)^{1-\sigma} \int z^{\sigma-1} q\left(z\right) dF_{rt}\left(z\right)}{\sum_{r'} \left(\frac{w_{r't}^{P}}{A_{r't}}\right)^{1-\sigma} \int z^{\sigma-1} q_{r'}\left(z\right) dF_{r't}\left(z\right)}. \tag{17}$$

We assume that trade is balanced period-by-period, i.e. that product markets clear every period. In other words we have that in equilibrium $E_{rt} = \sum_{j} \lambda_{rjt} E_{jt}$ and thus aggregate labor demand for production in region r time t is given by

$$H_{rt}^{P} = \frac{1}{w_{rt}^{P}} \sum_{j} \lambda_{rjt} w_{jt}^{P} H_{jt}^{P}. \tag{18}$$

Furthermore, notice that the market for labor for innovation implies that the total number of efficiency units of workers in innovation must be equal to

$$H_{rt}^{R} = \frac{1}{\iota} \frac{1}{\zeta_{r} Z_{t}^{I}} \frac{\int z^{\sigma-1} q\left(z\right) dF_{rt}\left(z\right)}{Q_{rt}^{\lambda}} i_{t}^{\iota} = \frac{1}{\iota} \frac{i_{t}^{\iota}}{\zeta_{r} Z_{t}^{I}} \frac{1}{Q_{rt}^{\lambda-1}} \tag{19}$$

Combining this equation with the first order condition for innovation equation (11) we obtain the labor demand for innovation given by

$$H_{rt}^{R} = \frac{1}{w_{rt}^{R}} \frac{1}{\iota} \left(\frac{v_{rt+1}}{1 + r_{t}}\right)^{\frac{\iota}{\iota-1}} \left(\frac{Q_{rt}^{\lambda} \zeta_{r} Z_{t}^{I}}{w_{rt}^{R}}\right)^{\frac{1}{\iota-1}} Q_{rt}. \tag{20}$$

3.3 Aggregate Labor Supply

In our model, individuals have two margins for their labor supply decisions: they decide which sector to work in and which location to migrate to. In terms of timing, we assume that individuals first decide on their geographical location r and then on their preferred sector of employment.

Labor supply across sectors We model sectoral labor supply with a simple Roy structure. Individuals are characterized by a single attribute - their immigration status, which we denote by $n \in \{N, I\}$, where $n = N$ denotes "Natives"and $n = I$ denotes "Immigrants". We assume that individual i draws a vector of efficiency

$$\left\{x^P, x^R\right\},$$

where x^P and x^R denotes the efficiency units as a production worker and a research worker. We assume that x^P and x^R are drawn independently from the following Frechet distribution

$$F_{jn}\left(x\right) = e^{-h_n^j x^{-\theta}}. \tag{21}$$

Here h_n^j parametrizes the average human capital of an individual with nationality $n$in sector $j \in \{R, P\}$ and θ parametrizes the labor supply elasticity.

Standard arguments imply that the share of people of type n working in sector $j = R, P$ in region r is given by

$$s_{rnt}^j = \frac{h_n^j \left(w_{rt}^j\right)^\theta}{h_n^P \left(w_{rt}^P\right)^\theta + h_n^R \left(w_{rt}^R\right)^\theta}. \tag{22}$$

Similarly, the aggregate level of human capital provided by workers of type n in region r towards sector $j \in \{R, P\}$ is given by

$$H_{rnt}^j = L_{rnt}\Gamma\left(1 - \frac{1}{\theta}\right)\left(h_n^j\right)^{\frac{1}{\theta}}\left(s_{rnt}^j\right)^{\frac{\theta-1}{\theta}}, \tag{23}$$

where L_{rnt} is the number of people of type n in region r. Hence, the aggregate supply of efficiency units provided to sector j in region r is given by

$$H_{rt}^j = L_{rt}\Gamma\left(1 - \frac{1}{\theta}\right)\left[\left(1 - \varpi_{rI}\right)\left(h_N^j\right)^{\frac{1}{\theta}}\left(s_{rNt}^j\right)^{\frac{\theta-1}{\theta}} + \varpi_{rI}\left(h_I^j\right)^{\frac{1}{\theta}}\left(s_{rIt}^j\right)^{\frac{\theta-1}{\theta}}\right] \; j \in \{I, N\}, \tag{24}$$

where $\varpi_{rI} = L_{rI}/L_r$ is the population share of immigrants in region r. Note that the respective employment shares s^j_{rN} only depend on relative wages (see (22)). Hence, the aggregate supply of human capital towards the research and the production sector also only depends on the relative wage within region r.[9]

For future reference we distinguish two special cases as delineated in Arkolakis et al. (2018). If skills are inelastically provided (which corresponds to the case of $\theta \to 1$), the share of people of group n working in sector j and the total amount of human capital is given by

$$s^j_{rnt} = \frac{h^j_n}{h^R_n + h^P_n} \text{ and } H^j_{rtn} = L_{rtn} h^j_n.$$

The second polar case is the case of homogeneity, i.e. $\theta \to \infty$. In that case, workers' occupation choice problem takes a simple cutoff-rule

$$\text{work in sector } R \text{ if and only } h^R_{rtn} w^R_{rt} \geq h^P_{rtn} w^P_{rt}.$$

If, for example $h^R_{rtn} = h^P_{rtn} = 1$, and with an interior solution where all locations produce and innovate we get that wages are equalized across sectors, i.e.

$$w_{rt} = w^R_{rt} = w^R_{rt}$$

and that the share of people working in the two sectors are fully demand determined.[10]

3.4 Dynamical Equilibrium

We exploit these aggregation results to define the equilibrium as a macro system of discrete blocks of equations, statics and dynamic. To characterize the equilibrium of the model we need to consider the market clearing of both product and labor markets. 20

Definition 1. *A dynamical equilibrium are sequences of production and research wages $\{w^P_{rt}, w^R_{rt}\}_{rt}$, per-quality-unit value functions $\{v_{rt}\}_{rt}$, innovation choices $\{i_{rt}\}_{rt}$, labor allocations $\{H^P_{rt}, H^R_{rt}\}_{rt}$, regional qualities $\{Q_{rt}\}_{rt}$, and consumption demands $\{c_{jrt}\}_{jrt}$ such that given an initial level of regional quality $\{Q_{r0}\}_r$, labor and good markets clear at each point time,*

1. *firms' innovation choices $\{i_{rt}\}_{rt}$ are consistent with $\{v_{rt}\}_{rt}$, i.e. solve (12)*

[9] Note that if natives and immigrants are identical, i.e. $h^j_N = h^j_I$, (24) reduces to the usual expression $H^j_r = L_r \Gamma\left(1 - \frac{1}{\theta}\right)\left(h^j\right)^{\frac{1}{\theta}}\left(s^j_r\right)^{\frac{\theta-1}{\theta}}$, as sectoral employment shares will be equalized given that they face the same prices.

[10] Note that if natives and immigrants differ in h^j_{rtn}, generically one group will be fully specialized.

2. *the evolution of qualities* $\{Q_{rt}\}_{rt}$ *is consistent with firms' innovation choices* $\{i_{rt}\}_{rt}$*, i.e. solve (14),*

3. *the per-quality-unit value functions* $\{v_{rt}\}_{rt}$ *solve (11).*

4. Labor markets clear, i.e. labor demand and supply for production and research labor equalize given equations (20), (18), and (24).

Definition 1 makes clear that the general equilibrium consists of a set of dynamic and static equations for an arbitrary number of regions. To solve this daunting dynamic fixed point problem we follow a strategy of modularization, as in Adao et al. (2019), in order to determine sets of equations that can be independently solved taking a subset of the variables of the system at a time. The difference with our approach is that one of the equations of the paper, the dynamic-innovation module described below, contains dynamic difference equations and not just static. The three modules are as follows:

1. *The Research Module:* Given production worker wages $\left\{w_{rt}^{P}\right\}_{r}$, average product quality $\{Q_{rt}\}_{r}$ and the value of innovation $\{v_{rt+1}\}_{r}$, we can use the labor supply equations (24) and the labor demand for research workers (20) to determine $\left\{H_{rt}^{P}, H_{rt}^{R}, w_{rt}^{R}\right\}_{r}$,

2. *The Production Module:* Given production worker labor supply $\left\{H_{rt}^{P}\right\}_{r}$ and average product quality $\{Q_{rt}\}_{r}$, production wages $\left\{w_{rt}^{P}\right\}_{r}$ are determined from the goods markets. Using (18) we obtain that

$$w_{rt}^{P}H_{rt}^{P} = \sum_{j} \frac{\left(\frac{w_{r}^{P}}{A_{rt}}\tau_{rj}\right)^{1-\varepsilon}}{\sum_{r}\left(\frac{w_{r}^{P}}{A_{rt}}\tau_{rj}\right)^{1-\varepsilon}} w_{jt}^{P}H_{jt}^{P}. \tag{25}$$

3. *The Dynamic-Innovation Module:* Given wages and labor allocations $\left\{H_{rt}^{P}, H_{rt}^{R}, w_{rt}^{P}, w_{rt}^{R}\right\}_{r}$ and the level of quality $\{Q_{rt}\}_{r}$, we can solve for i_{rt} from (14). Given i_{rt} we can solve for $\{v_{rt}, v_{rt+1}\}$ from (11) and (12).

Notice that this system can be further simplified. Starting from the value function, equation 13,

$$v_{rt} = \frac{1}{\sigma}\frac{E_{rt}}{Q_{rt}} + \left(\frac{\iota-1}{\iota}\right)\left(\frac{Q_{rt}^{\lambda}\zeta_{r}Z_{t}^{I}}{w_{rt}^{R}}\right)^{\frac{1}{\iota-1}}\left(\frac{v_{rt+1}}{1+r_{t}}\right)^{\frac{\iota}{\iota-1}}.$$

Now use (14), and $E_{rt} = \frac{\sigma}{\sigma-1}w_{rt}^{P}H_{rt}^{P}$, to obtain and also equation (20) to obtain

$$v_{rt}Q_{rt} = (\sigma-1)^{-1}w_{rt}^{P}H_{rt}^{P} + (\iota-1)\,w_{rt}^{R}H_{rt}^{R} \tag{26}$$

This equation connects equilibrium wage payments with the total payments to the two types of labor. Furthermore, notice that it allows to express the value of the firm as a function of wages by using equation 20 and the law of motion of aggregate quality, equation (14), one can show that

$$w_{rt}^{R}H_{rt}^{R}\iota = \frac{v_{rt+1}Q_{rt+1}}{1+r_t}$$

We now turn to special cases of our model to develop intuition for its main properties. In particular, we consider a specific factor case, where workers provide their skills inelastically to the two production sectors. This case of

no mobility across sectors corresponds to $\kappa \to 1$ where $H_{rt}^{P} = \bar{H}_{rt}^{P}$, $H_{rt}^{R} = \bar{H}_{rt}^{R}$ (see Arkolakis et al. (2018)). Notice that in the second case equation (14) implies that the law of motion of Q_{rt} does not depend on anything else other than $\zeta_r Z_t^I$. In other words, the law of motion of Q_{rt} can be completely determined by this equation and it is country-by-country specific.

Assuming that indeed $\zeta_r Z_t^I$ are fixed and we start from a steady state $Q_{rt} = \bar{Q}_r$. Then using (26) the value function of the firm is given by

$$v_{rt} = \frac{1}{\sigma-1}\frac{w_{rt}^{P}\bar{H}_{rt}^{P}}{\bar{Q}_r} + \left(\frac{\iota-1}{\iota}\right)\iota^{\frac{1}{\iota-1}}\frac{v_{rt+1}}{1+r_t}.$$

The solution of the system is then vectors of wages and value function $\{v_{rt}\}$, $\{w_{rt}^{P}\}$ that solve the above equation and

$$w_{rt}^{P}\bar{H}_{rt}^{P} = \sum_j \frac{\left(\frac{w_{rt}^{P}}{A_{rt}}\tau_{rj}\right)^{1-\varepsilon}}{\sum_r \left(\frac{w_{rt}^{P}}{A_{rt}}\tau_{rj}\right)^{1-\varepsilon}} w_{jt}^{P}\bar{H}_{jt}^{P},$$

with $A_{rt} = Z_{rt}^{A}\left(\bar{Q}_{rt}\right)^{\frac{1}{\sigma-1}}$.

3.5 The Spatial Balanced Growth Path

We now characterize the balanced growth path (BGP) of this economy. Along the BGP, the distribution of wages and spending across regions is stationary. This requires that productivity grows at the same rate in all regions. The characterization of the BGP is contained in the following proposition.

Proposition 2. *Consider the economy above and consider a BGP. Along the BGP, wages $\left\{w_{rt}^{P}, w_{rt}^{R}\right\}_r$ and spending $\{E_{rt}\}_r$ grow at the rate of TFP A_{rt} and the population distribution is stationary.*

1. *The growth rate of TFP A_{rt} is constant across regions and is given by*

$$1+g_A = (1+\overline{g}_Z)\,(1+\overline{g}_M)^{\frac{1}{1-\lambda}\frac{1}{\sigma-1}}\,,$$

   where $\overline{g}_Z$ is the exogenous growth rate of Z_{rt} and $\overline{g}_M$ is the exogenous growth rate of research productivity M_t.

2. *The growth rate of productivity Q_{rt} is given by*

$$1+g_Q = (1+\overline{g}_M)^{\frac{1}{1-\lambda}}$$

3. *The value function v_{rt} is given by*

$$v_{rt} = \frac{1}{1-\frac{\iota-1}{\iota}\frac{1+g_A}{1+r}}\frac{1}{\sigma}\frac{E_{rt}}{Q_{rt}},$$

   where E_{rt} is total spending on goods in region r

4. *The spending share on production workers and researchers is equalized across space, i.e.*

$$\frac{H^R_{rt}w^R_{rt}}{H^P_{rt}w^P_{rt}} = \frac{1}{\sigma-1}\frac{\frac{1}{\iota}\frac{1+g_A}{1+r}}{1-\frac{\iota-1}{\iota}\frac{1+g_A}{1+r}}. \tag{27}$$

5. *The distribution of productivity across space satisfies*

$$\frac{Q_{rt}}{Q_{jt}} = \left(\frac{\zeta_r}{\zeta_j}\times\frac{\frac{E_{rt}}{w^R_{rt}}}{\frac{E_{rt}}{w^R_{rt}}}\right)^{\frac{1}{1-\lambda}} \quad \textit{for all } r,j \tag{28}$$

Proof. See Section A.4 in the Appendix. □

The main implication of Proposition 2 is contained in (28): the long run distribution of productivity across space is endogenously determined. Using the fact that $E_{rt}\propto H^P_{rt}w^P_{rt}$ (see (16)) and $H^P_{rt}w^P_{rt}\propto H^R_{rt}w^R_{rt}$ (see (27)), we can express (28) as

$$\ln\left(\frac{Q_r}{Q_j}\right) = \underbrace{\frac{1}{\lambda}\ln\left(\frac{\zeta_r}{\zeta_j}\right)}_{\text{Exogenous differences in research efficiency}} + \underbrace{\frac{1}{\lambda}\ln\left(\frac{H^R_{rt}}{H^R_{jt}}\right)}_{\text{Supply of researchers}}. \tag{29}$$

Hence, region r has high productivity relative to region j if it is relatively efficient to produce new ideas, i.e. $\zeta_r>\zeta_j$, and it is able to attract relatively more researchers, i.e. $H^R_{rt}>H^R_{jt}$. The relative abundance of researchers is of course endogenous and determined from both the trade equilibrium and the system of equations governing labor mobility. Moving costs or the degree of openness across regions as part of the trade module will therefore affect the long-run distribution of productivity across space.

4 Counterfactuals

To conduct the counterfactuals we build on the procedure of Dekle et al. (2007) on conditioning the unobservables on actual data. We define $\hat{x} = x'/x$ and apply this definition to all the equilibrium equations of the model. For the trade module, equation 25,

$$\hat{w}_{rt}^P \hat{H}_{rt}^P w_{rt}^P H_{rt}^P = \sum_j \hat{x}_{ijt} x_{ijt} \hat{w}_{jt}^P \hat{H}_{jt}^P w_{jt}^P H_{jt}^P, \tag{30}$$

where

$$\hat{x}_{ijt} = \frac{\left(\frac{\hat{w}_r^P}{\hat{A}_{rt}} \hat{\tau}_{rj}\right)^{1-\varepsilon}}{\sum_r x_{ijt} \left(\frac{\hat{w}_r^P}{\hat{A}_{rt}} \hat{\tau}_{rj}\right)^{1-\varepsilon}}$$

and

$$\hat{A}_{rt} = \hat{Z}_{rt}^A \left(\hat{Q}_{rt}\right)^{\frac{1}{\sigma-1}}.$$

The innovation module uses equations *(11) and (12)*

$$\hat{v}_{rt} v_{rt} = \frac{1}{\sigma} \frac{E_{rt}}{Q_{rt}} \frac{\hat{E}_{rt}}{\hat{Q}_{rt}} + \frac{\iota - 1}{\iota} \left(\frac{\hat{v}_{rt+1}}{1+r_t} v_{rt+1}\right)^{\frac{\iota}{\iota-1}} \left(\frac{\hat{Q}_{rt}^\lambda \zeta_r \hat{Z}_t^I}{\hat{w}_{rt}^R} \frac{Q_{rt}^\lambda Z_t^I}{w_{rt}^R}\right)^{\frac{1}{\iota-1}} \iff$$

$$\hat{v}_{rt} v_{rt} = \frac{1}{\sigma-1} \frac{w_{rt}^P H_{rt}^P}{Q_{rt}} \frac{\hat{w}_{rt}^P \hat{H}_{rt}^P}{\hat{Q}_{rt}} + \frac{\iota-1}{\iota} \left(\frac{\hat{v}_{rt+1}}{1+r_t}\right)^{\frac{\iota}{\iota-1}} \left(\frac{\hat{Q}_{rt}^\lambda \hat{Z}_t^I}{\hat{w}_{rt}^R}\right)^{\frac{1}{\iota-1}} (v_{rt+1})^{\frac{\iota}{\iota-1}} \left(\frac{Q_{rt}^\lambda \zeta_r Z_t^I}{w_{rt}^R}\right)^{\frac{1}{\iota-1}}$$

Notice that, using equation 20 we have

$$H_{rt}^R = \frac{1}{\iota} \left(\frac{v_{rt+1}}{1+r_t} \frac{1}{w_{rt}^R}\right)^{\frac{\iota}{\iota-1}} \left(Q_{rt}^\lambda \zeta_r Z_t^I\right)^{\frac{1}{\iota-1}} Q_{rt}$$

so that

$$\hat{v}_{rt} v_{rt} = \frac{1}{\sigma-1} \frac{w_{rt}^P H_{rt}^P}{Q_{rt}} \frac{\hat{w}_{rt}^P \hat{H}_{rt}^P}{\hat{Q}_{rt}} + (\iota-1)\left(\hat{v}_{rt+1}\right)^{\frac{\iota}{\iota-1}} \left(\frac{\hat{Q}_{rt}^\lambda \hat{Z}_t^I}{\hat{w}_{rt}^R}\right)^{\frac{1}{\iota-1}} \frac{H_{rt}^R w_{rt}^R}{Q_{rt}} \iff$$

$$\hat{v}_{rt} v_{rt} Q_{rt} = \frac{1}{\sigma-1} w_{rt}^P H_{rt}^P \frac{\hat{w}_{rt}^P \hat{H}_{rt}^P}{\hat{Q}_{rt}} + (\iota-1)\left(\hat{v}_{rt+1}\right)^{\frac{\iota}{\iota-1}} \left(\frac{\hat{Q}_{rt}^\lambda \hat{Z}_t^I}{\hat{w}_{rt}^R}\right)^{\frac{1}{\iota-1}} H_{rt}^R w_{rt}^R \tag{31}$$

Same equations imply that

$$v_{rt}Q_{rt} = \frac{1}{\sigma - 1} w^P_{rt} H^P_{rt} + \frac{\iota - 1}{\iota} H^R_{rt} w^R_{rt}$$

With this equation we can calibrate the initial

$$v_{r0}Q_{r0},$$

with knowledge of $w^P_{r0}H^P_{r0}$, $H^R_{r0}w^R_{r0}$.

Also notice that equation 20 in changes is

$$\hat{H}^R_{rt} = \left(\frac{\hat{v}_{rt+1}}{1 \hat{+} r_t} \frac{1}{\hat{w}^R_{rt}} \right)^{\frac{\iota}{\iota - 1}} \left(\hat{Q}^\lambda_{rt} \hat{Z}^I_t \right)^{\frac{1}{\iota - 1}} \hat{Q}_{rt} \tag{32}$$

The second equation for the innovation module (14) can be directly written in changes

$$\frac{\hat{Q}_{rt+1}}{\hat{Q}_{rt}} = \hat{i}_t = \left(\hat{H}^R_{rt} \hat{Z}^I_t \hat{Q}^{-(1-\lambda)}_{rt} \right)^{1/\iota} \tag{33}$$

Thus, given knowledge of $\left\{ w^j_{rt} H^j_{rt} \right\}$ for $j = P, R$and changes in $\left\{ \hat{Z}^A_{rt}, \hat{Z}^I_t \right\}$ we can solve for $\hat{Q}_{rt+1}, \hat{v}_{rt+1}, \hat{w}^R_{rt}, \hat{w}^P_{rt}, \hat{H}^R_{rt}, \hat{H}^P_{rt}$ as we summarize in the next proposition

Proposition 3. *Conditional on the initial levels of* $\left\{ w^j_{rt} \right\}, \left\{ H^j_{rt} \right\}$ *for* $j = P, R$ *and* $t = 0$ *and changes in* $\left\{ \hat{Z}^A_{rt}, \hat{Z}^I_t \right\}$ *the changes in* $\left\{ \hat{Q}_{rt}, \hat{v}_{rt}, \hat{w}^R_{rt}, \hat{w}^P_{rt}, \hat{H}^R_{rt}, \hat{H}^P_{rt} \right\}$ *can be determined with the solution of equations 30, 31, 33, and 32, and the labor supply equations for* $\left\{ \hat{H}^R_{rt}, \hat{H}^P_{rt} \right\}$.

We plan to implement simulation to measure the impact of immigration on local and aggregate growth.

Algorithm for Computing the Model

To simulate the model we assume that there is a terminal steady state where all parameters are constant as $t \to \infty$. As we have data for the periods between 1850-1940 we may not assume that the model was in the steady state during that time. To simulate the counterfactuals for the model past the last year we have data we require trade shares, λ_{ijT}, and change in aggregate quality per region, $\hat{Q}_{rT}$, for some year T and we assume the model transits to a steady state after that. The actual algorithm for computing the equilibrium is based on a shooting algorithm similar to the ones used for computing transition paths of semi-endogenous growth models.

5 Reduced form evidence

In this section, we provide direct evidence for the mechanism in our model. In particular, we show that immigration inflows are positively related to regional productivity growth and patent activity.

5.1 Constructing an instrument for the allocation of immigrants

The "Card" Instrument

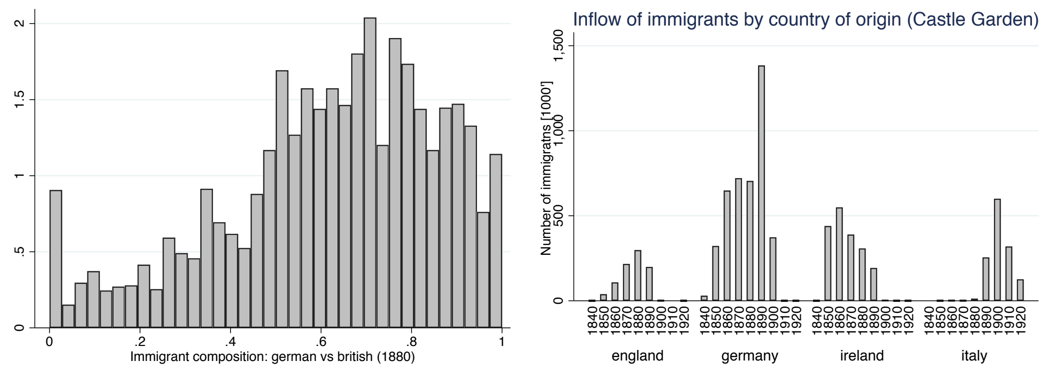
In Figure 12 we show the conceptual idea for our instrument, which is based on the work by David Card. We construct predicted immigrant flows from the time-series of the aggregate inflow of immigrants from different countries origin interacted with the existing cross-sectional distribution of immigrants prior to our sample. The time-series variation is shown in the right panel of Figure 12. Two aspects are interesting. First of all, there is a substantial time-series variation *within* immigrant groups. While the flow of Irish immigrants is declining after 1860, immigration from Germany is increasing and has a peak in the decade between 1880 and 1890. At the same time, there is a large cross-sectional variation in the composition of the immigrant population across counties in the US. As an example, we depict the share of german immigrants relative to immigrants from the UK across US counties in 1880 in the left panel of Figure 12. It is clearly seen that the cross-sectional variation is large. While many counties have a large "surplus" in German immigrants relative to British immigrants other counties are much more populated by British immigrants relative to Germans.

Given these two sources of variations, we construct a *predicted* immigrant stock in county r as follows. Let I_{r1860} be the total immigrants in county r in 1860. Also, let I_{r1860}^{n} denote the total number of immigrants with nationality n in county r in 1860. We then construct the predicted stock of immigrants in county r for all $t > 1860$, I_{rt}^{P}, as

$$I_{rt}^{P} = I_{rt-1}^{P} + \sum_{n \in \{GER,UK,ITA,IRL\}} IF_{t}^{n} \times \frac{I_{r1860}^{n}}{\sum_{r} I_{r1860}^{n}}, \tag{34}$$

where IF_{t}^{n} is the aggregate inflow of immigrants in year t as depicted in Figure 12. Intuitively, we assign the inflowing immigrants from country n according to their cross-sectional distribution in 1860 and then accumulate the regional immigrant stock with these inflows. As we only use (34) as an instrument, we abstract from mortality, which in principle might be specific to particular regions and nationalities.

Because many of our regression utilize the share of immigrants in a particular locality as



Note: The left panel shows the cross-sectional distribution $\frac{I_r^{GER}}{I_r^{GER}+I_r^{UK}}$, where I_r^{GER} (I_r^{UK}) is the number of german (english) immigrants living in region r in the year 1880. In the right panel we display the time-series of immigration inflows by decade for the four main countries of origin.

Figure 12: Constructing the Instrument

a regressor, we also construct a predicted immigrant share as

$$s_{Irt}^{P} = \frac{I_{rt}^{P}}{I_{rt}^{P} + L_{Nrt}},$$

where L_{Nrt} is the native population in region r at time t.

The Quota-Instrument

As an alternative instrument for the one illustrated above, we also exploit the Quota-Act, which was imposed in the 1921 and then in a more severe form in 1924 and limited the inflow of immigrants, in particular from countries of Easter and Southern Europe (see also Ager and Hansen (2017) and Doran and Yoon (2018)). The law limited the number of immigrants from any country to 3% of the number of residents from that same country living in the United States as reported in the Census in 1910. This meant that the quota had very different effects for different nationalities. In particular, ‘‘traditional’’ immigrant groups like the Germans, Irish, British or Scandinavians had a relatively high quota because the stock of people from these countries was already quite large. In contrast, ‘‘new’’ immigrant groups like the Italians or people from Eastern Europe experienced a binding quota because the flow of new immigrants was large relative to the stock. Of course, this discrepancy was precisely the political motivation behind the imposition of the Quota-Act. The Quota-Act was very effective in reducing the flow of immigration, in particular from countries in Southern and Eastern Europe.

We exploit this policy change in the following way. Let there be n nationalities. For each of these groups, we know the quota Q_n. Then we predict the inflow of immigrants that should have migrated to the US between 1920 and 1930 in the absence of the Quota by extrapolating aggregate immigration trends from the past with a flexible time trend . Let us call this prediction M_n. For each county r in the US we then calculate the statistic

$$Z_r = \frac{1}{L_{r1900}} \sum_n (M_n - Q_n) \frac{L_{rn1900}}{\sum_r L_{rn1900}},$$

where L_{rn1900} denotes the number of people with nationality n in region r as of 1900. Hence, $(M_n - Q_n) \frac{L_{rn1900}}{\sum_r L_{rn1900}}$ captures the gap reduction in predicted migration flows from county n, which would have affected region r, if the inflowing immigrants made their location choice in the the US in proportion to the stock of immigrants in 1900. By summing over all nationalities n and dividing by L_{r1900} we express this migration shortfall relative to county population in 1900.

Because the Quota came only in effect in 1921, we exploit this source of variation only in the period after 1921. Hence, we instrument the immigrant population in county r at time t with the variable $Z_r \times Post^{1920}$, where $Post^{1920}$ is a dummy, which takes the number 1 if $t > 1920$, i.e. which switches on once the Quota gets implemented. Because Z_r varies across regions and $Post^{1920}$ across time, we can use this instrument for the immigrant population even in a specification with county fixed effects in case we have at least two periods of data.

5.2 Results: Immigrants and Patent Activity

Our theory makes strong predictions for the relationship between regional immigration inflows and the creation of patents. We adopt the following measurement approach. Let $\mathcal{P}_{rt}$ be the *stock* of patents filed in region r up to year t. We assume that the level of knowledge Q_{rt} is proportional to the number of parents, i.e.[11]

$$Q_{rt} = \kappa \mathcal{P}_{rt}.$$

Hence, the model implies that (see (14))

$$\frac{\mathcal{P}_{rt}}{\mathcal{P}_{rt-1}} = \frac{Q_{rt}}{Q_{rt-1}} = \left(H^R_{rt-1} \zeta_r Z^I_t Q^{\lambda-1}_{rt-1} \iota\right)^{1/\iota} = \left(\kappa^{\lambda-1} \iota H^R_{rt-1} \zeta_r Z^I_{t-1} \mathcal{P}^{\lambda-1}_{rt-1}\right)^{1/\iota}.$$

[11] The *flow* of new patents in year t, i.e. the number of patents filed at year t, N^{Pat}_{rt}, is therefore given by $N^{Pat}_{rt} = \mathcal{P}_{rt} - \mathcal{P}_{rt-1}$.

This implies that

$$\ln \mathcal{P}_{rt} = \underbrace{\delta_r}_{\zeta_r} + \underbrace{\delta_t}_{Z^I_{t-1}} + \frac{\lambda - 1 + \iota}{\iota} \ln \mathcal{P}_{rt-1} + \frac{1}{\iota} \ln H^R_{rt-1}. \quad (35)$$

Using again the approximation for $\ln H^R_{rt-1}$, equation (35) suggests the regression

$$\ln \mathcal{P}_{rs+1} = \delta_t + \mu \varpi_{rIt} + \rho \ln \mathcal{P}_{rst} + \beta \ln L_r + X'_{rt}\gamma + u_{rst}, \quad (36)$$

where the theory implies that $\rho = \frac{\lambda-1+\iota}{\iota}$, $\beta = \frac{1}{\iota}$, and $\mu > 0$ if and only if $\left(\frac{h^R_I}{h^R_N}\right)^{\frac{1}{\theta}} \left(\frac{s^R_{rIt-1}}{s^R_{rNt-1}}\right)^{\frac{\theta-1}{\theta}} > 1$. The time fixed effects control for the state of the research technology, $\ln Z^I_{t-1}$, and $X'_{rt}\gamma$ contains a set of observable regional characteristics which control for the systematic variation in research efficiency across space, ζ_r.

Our main parameter of interest is μ because it is informative about the relative comparative advantage between immigrants and natives. The theory gives us guidance about the possible bias. If immigrants are more likely to settle in locations with high innovation potential ζ_r, the coefficient on μ will be upward biased. In fact, our analysis of patenting on the micro-level suggests this to be the case. We address this concern in three ways. First, if we think locations to be defined at the state or county level, (35) implies that a simple fixed effect estimator addresses the endogeneity concerns. Second, we also estimate (36) using two different instruments. On one hand, we use the "Card-IV" as highlighted above. On the other hand, we exploit the imposition of immigration quotas during the Quota-Act in the early 1920s.

The results of estimating (36) are reported in Table (3). In columns 1 - 3 we report the OLS specification for various choices of spatial control variables. Columns 1 and 2 show that there is a strong positive relationship between the share of foreign borns and the growth of patents. The coefficient declines slightly once we control for the share of the urban population. In column 3, which is our preferred specification as it follows most directly from our theory, we control for a whole set of regional fixed effects. The fact that the coefficient of the share of foreign-born is close to zero suggests that - through the lens of our theory - that the innovation potential between natives and immigrants is roughly identical. Recall that this is consistent with our empirical results using microdata, in particular, Tables (1) and (1), where we showed that immigrants are less likely to patent (after regional FE are controlled for) but have more patents conditional on innovation.

In columns 4 and 5 we report the same specification as in column 3 with our instrument variable strategy. In column 4 we use the imposition of the Quota Act in the 1920s as an

	ln Patentstock				
		OLS		Quota Act	Card IV
$FB\ share_{ct}$	1.256***	1.183***	-0.172	-0.682	0.0175
	(0.138)	(0.120)	(0.110)	(0.630)	(0.851)
$(ln)\ Population_{ct}$	0.241***	0.172***	0.302***	0.302***	0.297***
	(0.0255)	(0.0260)	(0.0246)	(0.0397)	(0.0403)
$(ln)\ Patentstock_{ct-1}$	0.773***	0.807***	0.240***	0.225***	0.259***
	(0.0211)	(0.0198)	(0.0153)	(0.0260)	(0.0371)
$(ln)\ Urban\ Population_{ct}$		0.0255*	-0.00436	0.000307	-0.00655
		(0.0145)	(0.0128)	(0.0132)	(0.0134)
R^2	0.962	0.967	0.996	0.457	0.469
N	2959	2038	2038	1848	1710
$Time\ FE$	No	No	Yes	Yes	Yes
$County\ FE$	No	No	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *, **, *** denotes significance at 10%, 5% and 1% respectively. We look at the counties with a non-zero stock of patents for the three decades between 1900 and 1930. ‘‘FB share’’ denotes the share of foreign born. The patentstock is the accumulated number of parents filed in county c since 1790.

Table 3: Immigrant Inflows and Regional Patent Activity State

	ln Patentstock				
		OLS		Quota Act	Card IV
$FB\ share_{ct}$	1.256***	1.183***	0.462***	1.141	-1.268*
	(0.138)	(0.120)	(0.0763)	(1.838)	(0.647)
$(ln)\ Population_{ct}$	0.241***	0.172***	0.158***	0.153***	0.174***
	(0.0255)	(0.0260)	(0.0139)	(0.0267)	(0.0341)
$(ln)\ Patentstock_{ct-1}$	0.773***	0.807***	0.855***	0.860***	0.838***
	(0.0211)	(0.0198)	(0.00675)	(0.0222)	(0.0227)
$(ln)\ Urban\ Population_{ct}$		0.0255*	0.0335***	0.0203	0.0694***
		(0.0145)	(0.00760)	(0.0387)	(0.0181)
R^2	0.962	0.967	0.979	0.968	0.960
N	2959	2038	2038	2038	1936
$Time\ FE$	No	No	Yes	Yes	Yes
$State\ FE$	No	No	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *, **, *** denotes significance at 10%, 5% and 1% respectively. We look at the counties with a non-zero stock of patents for the three decades between 1900 and 1930. ‘‘FB share’’ denotes the share of foreign born. The patentstock is the accumulated number of parents filed in county c since 1790.

Table 4: Immigrant Inflows and Regional Patent Activity State

instrument. In column 5 we exploit the size of the pre-determined immigration population and the time-variation in immigrant inflow as explained above. The estimates are less precise but we cannot reject that the effect is also equal to zero.

6 Conclusions

We have developed an empirical and theoretical framework to analyze the role of human capital and spatial policies on economic growth. Our big data historical approach allows us to analyze decades of data on the American economy and the associated effects of the influx of immigrants. We couple these data with a model of forward-looking innovating firms that allows us to evaluate the empirical data using structural relationships that arise from the theory. In future work, we plan to exploit the micro aspect of the data to fully understand the process of knowledge creation by immigrants and to provide more definitive conclusions on the impact of immigration on American growth.

References

Abramitzky, R., Boustan, L. P., Eriksson, K., August 2012. Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. American Economic Review 102 (5), 1832--56.
URL https://ideas.repec.org/a/aea/aecrev/v102y2012i5p1832-56.html

Adao, R., Arkolakis, C., Esposito, F., 2019. Spatial linkages, global shocks, and local labor markets: Theory and evidence.

Ager, P., Hansen, C. W., 2017. Closing heaven's door: Evidence from the 1920s us immigration quota acts. Available at SSRN 3059439.

Akcigit, U., Grigsby, J., Nicholas, T., 2017. Immigration and the rise of american ingenuity. Tech. rep., National Bureau of Economic Research.

Anderson, J. E., Van Wincoop, E., 2003. Gravity with gravitas: A solution to the border puzzle. American Economic Review 93 (1), 170--192.

Arkolakis, C., Ramondo, N., Rodríguez-Clare, A., Yeaple, S., 2018. Innovation and production in the global economy. American Economic Review 108 (8), 2128--73.

Atkeson, A., Burstein, A., 2010. Innovation, firm dynamics, and international trade. Journal of Political Economy 118 (3), 433--489.

Dekle, R., Eaton, J., Kortum, S., 2007. Unbalanced trade. Tech. rep., National Bureau of Economic Research.

Doran, K., Yoon, C., 2018. Immigration and invention: Evidence from the quota acts.

Feigenbaum, J. J., 2016. Automated census record linking: A machine learning approach.

Gordon, R. J., 2017. The rise and fall of American growth: The US standard of living since the civil war. Princeton University Press.

Kortum, S., 1997. Research, patenting, and technological change. Econometrica 65 (6), 1389--1419.

Lucas Jr, R. E., Moll, B., 2014. Knowledge growth and the allocation of time. Journal of Political Economy 122 (1), 1--51.

Nunn, N., Qian, N., Sequeira, S., 2017. Migrants and the making of america: The short and long run effects of immigration during the age of mass migration.
URL https://ideas.repec.org/a/ces/ifodic/v15y2017i3p50000000000049.html

Ottaviano, G. I., Peri, G., Wright, G. C., 2013. Immigration, offshoring, and american jobs. American Economic Review 103 (5), 1925--59.

Perla, J., Tonetti, C., 2014. Equilibrium imitation and growth. Journal of Political Economy 122 (1), 52--76.

Roca, J. D. L., Puga, D., 2017. Learning by working in big cities. The Review of Economic Studies 84 (1), 106--142.

Romer, P. M., 1990. Endogenous technological change. The Journal of Political Economy 98 (5), S71--S102.

Ye, J., Han, S., Hu, Y., Coskun, B., Liu, M., Qin, H., Skiena, S., 2017. Nationality classification using name embeddings. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. CIKM '17. ACM, New York, NY, USA, pp. 1897--1906.
URL http://doi.acm.org/10.1145/3132847.3133008

A Theoretical Appendix

A.1 Proof of Proposition 1

Consider the value function in (10), given by

$$V_{rt}\left(z^{\sigma-1}q\right)=\pi_{rt}\left(z^{\sigma-1}q\right)+\max_{i_t}\left[\frac{1}{1+r_t}V_{rt+1}\left(z^{\sigma-1}qi_t\right)-w_{rt}^{R}\frac{z^{\sigma-1}}{\zeta_r^I Z_t^I}\frac{q}{Q_{rt}^{\lambda}}\frac{i_t^{\iota}}{\iota}\right].$$

We conjecture that $V_{rt}(q)$ is linear in q, i.e. takes the form

$$V_{rt}(q)=\upsilon_{rt}q, \tag{37}$$

and we will determine υ_{rt}. Using, (37) and (9) we get that

$$\upsilon_{rt}z^{\sigma-1}q=\frac{1}{\sigma}\frac{z^{\sigma-1}q}{Q_{rt}}\frac{E_{rt}}{N_{rt}}+\max_{i_t}\left[\frac{1}{1+r_t}\upsilon_{rt+1}iz^{\sigma-1}q-w_{rt}^{R}\frac{1}{\zeta_r^I Z_t^I}\frac{z^{\sigma-1}q}{Q_{rt}^{\lambda}}\frac{i_t^{\iota}}{\iota}\right],$$

so that indeed

$$\upsilon_{rt}=\frac{1}{\sigma}\frac{1}{Q_{rt}}\frac{E_{rt}}{N_{rt}}+\max_{i_t}\left[\frac{1}{1+r_t}\upsilon_{rt+1}i-w_{rt}^{R}\frac{1}{\zeta_r^I Z_t^I}\frac{1}{Q_{rt}^{\lambda}}\frac{i_t^{\zeta}}{\zeta}\right].$$

The optimality condition for i_t reads

$$\frac{1}{1+r_t}\upsilon_{rt+1}=w_{rt}^{R}\frac{1}{\zeta_r^I Z_t^I}\frac{1}{Q_{rt}^{\lambda}}i_t^{\iota-1}.$$

This implies that the optimal innovation rate is given by

$$i_t=\left(\frac{1}{1+r_t}\frac{\upsilon_{rt+1}}{w_{rt}^{R}}Q_{rt}^{\lambda}\zeta_r^I Z_t^I\right)^{\frac{1}{\iota-1}}, \tag{38}$$

and that the value function is given by

$$\upsilon_{rt}=\frac{1}{\sigma}\frac{1}{Q_{rt}}\frac{E_{rt}}{N_{rt}}+\left(\frac{\iota-1}{\iota}\right)w_{rt}^{R}\frac{1}{\zeta_r^I Z_t^I}\frac{1}{Q_{rt}^{\lambda}}i_t^{\iota},$$

where i_t is given in (38). Substituting for i_t in (38) we can also express the value function as a forward looking difference equation

$$\begin{aligned}\upsilon_{rt} &= \frac{1}{\sigma}\frac{1}{Q_{rt}}\frac{E_{rt}}{N_{rt}}+\left(\frac{\iota-1}{\iota}\right)w_{rt}^{R}\frac{1}{\zeta_r^I Z_t^I}\frac{1}{Q_{rt}^{\lambda}}\left(\frac{1}{1+r_t}\frac{\upsilon_{rt+1}}{w_{rt}^{R}}Q_{rt}^{\lambda}\zeta_r^I Z_t^I\right)^{\frac{\iota}{\iota-1}}\\ &= \frac{1}{\sigma}\frac{1}{Q_{rt}}\frac{E_{rt}}{N_{rt}}+\left(\frac{\iota-1}{\iota}\right)\left(\frac{Q_{rt}^{\lambda}\zeta_r^I Z_t^I}{w_{rt}^{R}}\right)^{\frac{1}{\zeta-1}}\left(\frac{\upsilon_{rt+1}}{1+r_t}\right)^{\frac{\iota}{\iota-1}}.\end{aligned}$$

A.2 Deriving the labor supply relationships

To derive the results in Section 3.3, we rely heavily on the max stability of the Frechet distribution. In particular, if the S dimensional vector $[x_s]_s$ is iid Frechet distributed across

s, with

$$P\left(x_s \leq z\right) = F_s\left(z\right) = e^{-h_s z^{-\theta}},$$

the variable

$$a = \max_s \left[\lambda_s x_s\right] \tag{39}$$

is distributed according to

$$P\left(a \leq \alpha\right) = e^{-\Lambda^{\theta}\alpha^{-\theta}}$$

where

$$\Lambda = \left(\sum_s \lambda_s^{\theta} h_s\right)^{1/\theta}$$

Using this result we can derive the expressions for average income, average human capital and the relative employment shares. In particular,

$$E\left[a\right] = \Gamma\left(1 - \frac{1}{\theta}\right) \times \Lambda$$

Deriving average income Total income of individual i is given by

$$y^i = \left\{w^P x^P, w^R x^R\right\}.$$

Hence,

$$P\left(y^i \leq \overline{y}\right) = e^{-W^{\theta}\overline{y}^{-\theta}}$$

where

$$W^i = \left(\left(w^P\right)^{\theta} h_i^P + \left(w^R\right)^{\theta} h_i^R\right)^{1/\theta}.$$

Hence, if for example individual i if skill type k and nationality n in region r, W^i is given by

$$W_{rn}^k = \left(\left(w_r^P\right)^{\theta} h_n^{Pk} + \left(w_r^R\right)^{\theta} h_n^{Rk}\right)^{1/\theta}.$$

Deriving the total supply of human capital by occupation (equation (23)) To derive the aggregate supply of human capital of individuals of type (n, k) in occupation j, note that the *average* number of efficiency units provided to occupation o is

$$E\left[z_n^{ok} | y_n^{jk} = \max_j \left\{y_n^{jk}\right\}\right] = E\left[z_n^{ok} | z_n^{ok} = \max_j \left\{\frac{w^j}{w^o} z_n^{jk}\right\}\right].$$

Using (39) with $\lambda_j = \frac{w^j}{w^o}$ this object is given by

$$
\begin{aligned}
E\left[z_n^{ok} | y_n^{jk} = \max_j \left\{y_n^{jk}\right\}\right] &= \Gamma\left(1-\frac{1}{\theta}\right)\left(\sum_j \left(\frac{w^j}{w^o}\right)^{\theta} h_n^{jk}\right)^{1/\theta} \\
&= \Gamma\left(1-\frac{1}{\theta}\right)\frac{\left(\sum_j \left(w^j\right)^{\theta} h_n^{jk}\right)^{1/\theta}}{w^o}
\end{aligned}
$$

Also note that the share of people of type (n,k) working in occupation j in region r is given by

$$
s_{rn}^{jk} = P\left(w_r^j z_n^{jk} = \max_o \left\{w_r^o z_n^{ok}\right\}\right) = \frac{h_n^{jk}\left(w_r^j\right)^{\theta}}{\sum_j h_n^{jk}\left(w_r^j\right)^{\theta}} = h_n^{jk}\frac{\left(w_r^j\right)^{\theta}}{\left(W_{rn}^k\right)^{\theta}}.
$$

Hence, letting the mass of workers of type (n,k) in region r working in occupation j be

$$
L_{rn}^{jk} = L_{rn}^k s_{rn}^{jk},
$$

we get that

$$
\begin{aligned}
H_{rn}^{jk} &= L_{rn}^{jk} E\left[z_n^{ok} | y_n^{jk} = \max_j \left\{y_n^{jk}\right\}\right] \\
&= L_{rn}^k s_{rn}^{jk} \Gamma\left(1-\frac{1}{\theta}\right)\frac{W_{rn}^k}{w_r^j} \\
&= L_{rn}^k \Gamma\left(1-\frac{1}{\theta}\right) s_{rn}^{jk}\left(\frac{h_n^{jk}}{s_{rn}^{jk}}\right)^{1/\theta} = L_{rn}^k \Gamma\left(1-\frac{1}{\theta}\right)\left(h_n^{jk}\right)^{1/\theta}\left(s_{rn}^{jk}\right)^{\frac{\theta-1}{\theta}}.
\end{aligned}
$$

Alternatively we can also express this object as

$$
H_{rn}^{jk} = L_{rn}^k \Gamma\left(1-\frac{1}{\theta}\right) s_{rn}^{jk}\frac{W_{rn}^k}{w_r^j} = L_{rn}^k \Gamma\left(1-\frac{1}{\theta}\right) h_n^{jk}\left(\frac{w_r^j}{W_{rn}^k}\right)^{\theta-1}.
$$

These expressions are exactly (23).

A.3 Proof of Lemma 1

The production function for the final good in (3) is given by

$$
Y_{rt} = Z_{rt} X_{rt} = Z_{rt}\left(\int_{i=0}^{1} x_{rt}(z)^{\frac{\sigma-1}{\sigma}} dF(z)\right)^{\frac{\sigma}{\sigma-1}}.
$$

Letting U_{rt} be the price index for the bundle X_{rt}, we get that

$$
\frac{u_{rt}(z)\, x_{rt}(z)}{U_{rt} X_{rt}} = \left(\frac{u_{rt}(z)}{U_{rt}}\right)^{1-\sigma} \Longrightarrow x_{rt}(z) = \left(\frac{u_{rt}(z)}{U_{rt}}\right)^{-\sigma} X_{rt}.
$$

We have that $u_{rt}(z) = \frac{\sigma}{\sigma-1} w^p_{rt}/q(z)^{1/(\sigma-1)} z$ and $U_{rt} = \left(\int_0^1 u_{rt}(z)^{1-\sigma} F(z)\, dz\right)^{1/(1-\sigma)}$. Substituting the production function (4) and $\frac{u_{rt}(z)}{U_{rt}} = \left(\frac{z^{\sigma-1}q(z)}{Q}\right)^{\frac{1}{1-\sigma}}$, we get that total employment of production workers at firm i is given by

$$h_{rt}(z) = z^{-1} q_{rt}(z)^{-\frac{1}{\sigma-1}} x_{rt}(z) = z^{-1} q_{rt}(z)^{-\frac{1}{\sigma-1}} \left(\frac{z^{\sigma-1} q_{rt}(z)}{Q_r}\right)^{\frac{\sigma}{\sigma-1}} X_{rt} = z^{\sigma-1} q_{rt}(z) \left(\frac{1}{Q_r}\right)^{\frac{\sigma}{\sigma-1}} X_{rt}$$

Hence, total labor demand is

$$\int h_{rt}(z)\, dF(z) = \left(\frac{1}{Q_{rt}}\right)^{\frac{\sigma}{\sigma-1}} X_{rt} \int z^{\sigma-1} q(z)\, dF(z) = \left(\frac{1}{Q_{rt}}\right)^{\frac{\sigma}{\sigma-1}} X_{rt} Q_{rt} = (Q_{rt})^{\frac{-1}{\sigma-1}} X_{rt}.$$

Labor market clearing implies that $\int h_{rt}(z)\, dF(z) = H^P_{rt}$. Hence,

$$X_{rt} = (Q_{rt})^{\frac{1}{\sigma-1}} H^P_{rt}.$$

Substituting into the production function yields

$$Y_{rt} = Z_{rt} (Q_{rt})^{\frac{1}{\sigma-1}} H^P_{rt}.$$

A.4 Characterization of the Balanced Growth Path (BGP)

In this section, we characterize the details of the balanced growth path (BGP). Along the BGP the allocation of people is constant across space and all aggregate variables grow at some constant rate, g^i where i is the relevant variable and i could be potentially different for different i. Along the balanced growth path interest rates are also constant, $r_t = r$. We assume that $\frac{Z^A_{rt+1}}{Z^A_{rt}} = 1 + \bar{g}_Z$, i.e. the exogenous component of productivity grows at rate g (which is the same for all regions). We also assume the aggregate research productivity M_t grows at a constant rate, i.e. $\frac{M_{t+1}}{M_t} = 1 + \bar{g}_M$

To have a balanced growth path we need that aggregate productivity A_{rt} grows at the same rate in all regions. Hence, see Lemma 1, we need that

$$1 + g_A = \frac{A_{rt+1}}{A_{rt}} = \frac{Z^A_{rt+1}}{Z^A_{rt}} \left(\frac{Q_{rt+1}}{Q_{rt}}\right)^{\frac{1}{\sigma-1}} = (1 + \bar{g}_Z)(1 + g_Q)^{\frac{1}{\sigma-1}}.$$

The BGP growth rate of productivity Q is therefore given by

$$1 + g_Q = \left(\frac{1 + g_A}{1 + g_Z}\right)^{\sigma-1}. \tag{40}$$

Using (14), this implies that the innovation rate in region r, i_{rt}, has to be constant across locations and time. Using (12) this implies that

$$1+g_Q = i = \left(\frac{1}{1+r}\frac{\zeta_r^I Z_t^I}{Q_{rt}^{1-\lambda}}\frac{v_{rt+1}Q_{rt}}{w_{rt}^R}\right)^{\frac{1}{\iota-1}}. \tag{41}$$

Moreover, the value function v_{rt} given by equation (11) can be written as

$$\frac{v_{rt}Q_{rt}}{w_{rt}^R} = \frac{1}{\sigma}\frac{E_{rt}}{w_{rt}^R} + \frac{\iota-1}{\iota}\frac{Q_{rt}^{1-\lambda}}{\zeta_r^I Z_t^I} i^{\iota}. \tag{42}$$

First note that w_{rt}^R and E_{rt} grow at the same rate in a stationary equilibrium. Hence, conjecture that along a BGP both $\frac{v_{rt}Q_{rt}}{w_{rt}^R}$ and $\frac{Q_{rt}^{1-\lambda}}{Z_t^I}$ are constant. This implies that

$$1+g_Q = (1+\overline{g}_M)^{\frac{1}{1-\lambda}}. \tag{43}$$

(40) therefore implies that the growth rate of TFP is given by

$$1+g_A = (1+\overline{g}_Z)(1+\overline{g}_M)^{\frac{1}{1-\lambda}\frac{1}{\sigma-1}}. \tag{44}$$

Under our price normalization, wages w_{rt}^R and total spending E_{rt} are growing at the rate of aggregate TFP g_A. Hence, for $\frac{v_{rt}Q_{rt}}{w_{rt}^R}$ to be constant, the growth rate of the value function given by

$$\frac{v_{rt+1}}{v_{rt}} = (1+g_v) = \frac{w_{rt+1}^R/w_{rt}^R}{Q_{rt+1}/Q_{rt}} = \frac{1+g_A}{1+g_Q} = \frac{(1+\overline{g}_Z)(1+\overline{g}_M)^{\frac{1}{1-\lambda}\frac{1}{\sigma-1}}}{(1+\overline{g}_M)^{\frac{1}{1-\lambda}}} = (1+\overline{g}_Z)(1+\overline{g}_M)^{\frac{1}{1-\lambda}\left(\frac{2-\sigma}{\sigma-1}\right)}. \tag{45}$$

In fact, the value function can be solved explicitly along the BGP.

Proposition 4. *Let $r > g_A$. Along the BGP the value function is given by*

$$\frac{v_{rt}Q_{rt}}{w_{rt}^R} = \frac{1}{1-\frac{\zeta-1}{\zeta}\frac{1+g_A}{1+r}}\frac{1}{\sigma}\frac{E_{rt}}{w_{rt}^R}.$$

Proof. From (41) and the fact that v_{rt} grows at at rate $1+g_v$ given in (45), we get that

$$1+g_Q = i = \left(\frac{1}{1+r}\frac{\zeta_r^I Z_t^I}{Q_{rt}^{1-\lambda}}\frac{v_{rt+1}Q_{rt}}{w_{rt}^R}\right)^{\frac{1}{\iota-1}} = \left(\frac{1}{1+r}\frac{\zeta_r^I Z_t^I}{Q_{rt}^{1-\lambda}}\frac{(1+g_v)\,v_{rt}Q_{rt}}{w_{rt}^R}\right)^{\frac{1}{\zeta-1}}.$$

Hence, we can solve for $\frac{Q_{rt}^{1-\lambda}}{\varphi_r^I M_t}$ as

$$\frac{Q_{rt}^{1-\lambda}}{\zeta_r^I Z_t^I} = \frac{1}{1+r}\frac{1}{(1+g_Q)^{\zeta-1}}\frac{(1+g_v)\,v_{rt}Q_{rt}}{w_{rt}^R}.$$

Substituting into (42) yields

$$\begin{aligned}\frac{v_{rt}Q_{rt}}{w_{rt}^R} &= \frac{1}{\sigma}\frac{E_{rt}}{N_{rt}w_{rt}^R} + \frac{\iota-1}{\zeta}\frac{1}{1+r}\frac{1}{(1+g_Q)^{\zeta-1}}\frac{(1+g_v)\,v_{rt}Q_{rt}}{w_{rt}^R}i^{\zeta} \\ &= \frac{1}{\sigma}\frac{E_{rt}}{N_{rt}w_{rt}^R} + \frac{\iota-1}{\iota}\frac{1+g_Q}{1+r}\frac{(1+g_v)\,v_{rt}Q_{rt}}{w_{rt}^R}\end{aligned}$$

so that

$$\frac{v_{rt}Q_{rt}}{w_{rt}^R} = \frac{1}{1-\frac{\iota-1}{\iota}\frac{1+g_Q}{1+r}(1+g_v)}\frac{1}{\sigma}\frac{E_{rt}}{N_{rt}w_{rt}^R} = \frac{1}{1-\frac{\iota-1}{\iota}\frac{1+g_A}{1+r}}\frac{1}{\sigma}\frac{E_{rt}}{N_{rt}w_{rt}^R}.$$

□

The spatial productivity distribution Along the BGP, the spatial distribution of productivity Q_{rt} is stationary as all regions grow at the same rate. The *level* of productivity is, however, determined endogenously. In particular, (41) implies that

$$\frac{\varphi_r^I M_t}{Q_{rt}^{1-\lambda}}\frac{v_{rt+1}Q_{rt}}{w_{rt}^R} = \frac{\varphi_j^I M_t}{Q_{jt}^{1-\lambda}}\frac{v_{jt+1}Q_{jt}}{w_{jt}^R} \text{ for all } r, j. \tag{46}$$

This implies that relative productivities Q_{rt} and Q_{jt} are given by

$$\frac{Q_{rt}}{Q_{jt}} = \left(\frac{\varphi_r^I}{\varphi_j^I} \times \frac{\frac{v_{rt}Q_{rt}}{w_{rt}^R}}{\frac{v_{jt}Q_{jt}}{w_{jt}^R}}\right)^{\frac{1}{1-\lambda}}.$$

Hence, long-run differences in productivity across space Q_{rt}/Q_{jt} are governed by

$$\ln\left(\frac{Q_{rt}}{Q_{jt}}\right) = \underbrace{\frac{1}{1-\lambda}\ln\left(\frac{\varphi_r^I}{\varphi_j^I}\right)}_{\text{Exogenous differences in research efficiency}} + \underbrace{\frac{1}{1-\lambda}\ln\left(\frac{\frac{v_{rt}Q_{rt}}{w_{rt}^R}}{\frac{v_{jt}Q_{jt}}{w_{jt}^R}}\right)}_{\text{Endogenous value of innovation}}. \tag{47}$$

Using Proposition 4, (47) can be written as

$$\ln\left(\frac{Q_{rt}}{Q_{jt}}\right) = \underbrace{\frac{1}{1-\lambda}\ln\left(\frac{\varphi_r^I}{\varphi_j^I}\right)}_{\text{Exogenous differences in research efficiency}} + \underbrace{\frac{1}{1-\lambda}\ln\left(\frac{\frac{E_{rt}}{N_{rt}w_{rt}^R}}{\frac{E_{jt}}{N_{jt}w_{jt}^R}}\right)}_{\text{Endogenous Market Size}}. \tag{48}$$

Finally, note that we can express (48) also in terms the labor supply. To do so note that along the BGP payments to researchers and production workers are equalized.

Proposition 5. *Consider a BGP. Then*

$$\frac{w_{rt}^R H_{rt}^R}{w_{rt}^P H_{rt}^P} = \frac{1}{\sigma - 1} \frac{\frac{1}{\iota}\frac{1+g_A}{1+r}}{1 - \frac{\iota - 1}{\iota}\frac{1+g_A}{1+r}}.$$

Proof. The total demand for efficiency units in the research sector is given by

$$\frac{1}{\zeta_r^I Z_t^I} \frac{q}{Q_{rt}^\lambda} \frac{i_t^\iota}{\iota} w_{rt}^R$$

$$H_{rt}^R = \int_q \frac{1}{\zeta_r^I Z_t^I} \frac{q}{Q_{rt}^\lambda} \frac{i^\iota}{\iota} dF_{rt}(q) = \frac{Q_{rt}^{1-\lambda}}{\zeta_r^I Z_t^I} \frac{i^\iota}{\iota}.$$

Using that $i = 1 + g_Q$ and $\frac{Q_{rt}^{1-\lambda}}{\zeta_r^I Z_t^I} = \frac{1}{1+r} \frac{1}{\left(1+g_Q\right)^{\zeta - 1}} \frac{(1+g_v) v_{rt} Q_{rt}}{w_{rt}^R}$ (see proof of Proposition 4), we get that

$$H_{rt}^R = N_{rt} \frac{1+g_Q}{1+r} \frac{(1+g_v) v_{rt} Q_{rt}}{w_{rt}^R} \frac{1}{\iota} = \frac{1+g_A}{1+r} \frac{v_{rt} Q_{rt}}{w_{rt}^R} \frac{1}{\iota}.$$

Using the result in Proposition 4, this implies that

$$H_{rt}^R w_{rt}^R = \frac{1+g_A}{1+r} \frac{1}{1 - \frac{\iota - 1}{\iota}\frac{1+g_A}{1+r}} \frac{1}{\sigma} \frac{1}{\iota} E_{rt}.$$

Hence, the payments to researchers are a constant fraction of revenue along the BGP. And because production workers also receive a constant fraction of revenue (see (16)), we get

$$\frac{H_{rt}^R w_{rt}^R}{w_{rt}^P H_{rt}^P} = \frac{1}{\sigma - 1} \frac{\frac{1}{\iota}\frac{1+g_A}{1+r}}{1 - \frac{\iota - 1}{\iota}\frac{1+g_A}{1+r}}.$$

□

Using Proposition 5, the long run distribution of productivity in (48) can also be expressed in terms of (endogenous) amount of resources in research

$$\ln\left(\frac{Q_{rt}}{Q_{jt}}\right) = \underbrace{\frac{1}{1-\lambda} \ln\left(\frac{\varphi_r^I}{\varphi_j^I}\right)}_{\text{Exogenous differences in research efficiency}} + \underbrace{\frac{1}{1-\lambda} \ln\left(\frac{H_{rt}^R}{H_{jt}^R}\right)}_{\text{Resources employed in research}}, \tag{49}$$

or the number of production workers and the relative cost of research

$$\ln\left(\frac{Q_{rt}}{Q_{jt}}\right) = \underbrace{\frac{1}{1-\lambda} \ln\left(\frac{\varphi_r^I}{\varphi_j^I}\right)}_{\text{Exogenous differences in research efficiency}} + \underbrace{\frac{1}{1-\lambda} \ln\left(\frac{H_{rt}^P}{H_{jt}^P}\right)}_{\text{Market size by production workers}} + \underbrace{\frac{1}{1-\lambda} \ln\left(\frac{w_{rt}^P / w_{rt}^R}{w_{jt}^P / w_{jt}^R}\right)}_{\text{Relative cost of research}} \tag{50}$$

B Empirical Appendix

In this section, we provide details on the procedure and quality of the linking microdata. We link four types of data: (1) we match individuals in the US census over time, (2) & (3) we match immigrants to the US Census (both from the Hamburg Passenger Lists and Castle Garden), and (4) we match the information on patenting to the US Census. We discuss the quality of these matches in turn in the sections below.

B.1 Census-to-Census Matching of individuals over time

We have access to the complete individual-level complete-count US demographic federal census records from 1850-1940.[12] By linking our novel immigration records to the US census records, we can measure the entire life-cycle of immigrants since they entered the US.

Our record linking procedure has the following characteristics: (1) we rely on all *complete-count* US Federal Census records with *occupation* and *industry information* (a newly transcribed variable from the original census records), (2) we link people at *more than two points in time,* and (3) we use both *individual and household level information* to improve the matching of individuals.

These three elements are important for this study. The availability of the occupation information both before and after entering the US enables us to measure individuals' skills and to investigate novel economic problems, such as occupational transitions along immigrants' life-cycle. Also, while many record matching methods match individuals only at two points in time, this horizon may not be long enough to systematically analyze spatial mobility along the life-cycle. We match individuals multiple points in time (up to three or four times) so that we follow individuals for multiple decades.

Finally, most existing historical record matching practices drop non-unique potential matches. This may introduce a systematic bias of matched records, as for example records with relatively common names would be systematically excluded. We, therefore, develop new methods to find link records across sources and time.

B.2 Matching Immigrants from the Hamburg Passenger Lists to the US Census

Our goal is to extract as many unique matches as possible by imposing rules such that we can eliminate false positives. From both datasets (i.e. Hamburg Passenger Lists and US Demographic Census), we take advantage of the overlap in information sets to extract unique matches. Table 5 shows the final number of matches after completing the pruning procedure. The variable ‘‘Post First Stage Matches’’ shows the number of possible matches as outputted from the first stage linking procedure where age, first and last name similarities were primary linking criteria.

[12]We use restricted complete-count US demographic census from 1850 to 1940 to link individual-level records by implementing machine learning approach (random forest classification). This record linking methodology is similar in spirit to Minnesota Population Center (MPC) Record record linkage project of the 1850-1930 sample census records to the 1880 complete count census records. However, MPC implemented a support vector machines (SVM) to automate the record linking. Feigenbaum (2016) discusses a machine learning approach to census record linking and compares different possible matching algorithms.

Table 5: Hamburg to US Demographic Census Statistics

Record Type	Germany	Italy	UK	Austria	Poland	Hungary	Total
Immigration Record	49,390	2,804	684	129,332	96,578	156,861	437,649
1910 Census Record	1,343,333	887,044	676,429	706,524	20,029	287,229	3,920,588
Post First Stage Matches	227,604	4,983	4,194	259,734	3,741	142,558	642,814
Final Matches	26,136	1,237	251	39,615	2,091	28,711	98,041

The pruning procedure itself uses three rules applied iteratively to the data. In particular, since applying any one rule may delete potential matches that could have been found via another rule we apply rules in every possible order and extract unique matches after each application. We also extract any unique matches that were discovered after the application of the first stage. In the case that the application of a rule introduces a duplicate observation in either the census or Hamburg data, we keep the earlier iteration of that observation. This permutation-based approach extracts about 25% of the possible unique matches as a pose to a linear approach which would yield only about 3-5% of the possible matches.

The three rules we apply are as follows. First, we restrict the Jaro-Winkler distance for the last name to be greater than or equal to 0.85 (1 being the maximum and 0 being the minimum of Jaro-Winkler distance). For individuals over the age of 25 when they immigrated, we impose that their marital status must be stable across datasets. Finally, we impose the year of immigration in the Census data and the departure year in the Hamburg data may not be more than three years apart.

B.3 Matching Immigrants from the Castle Garden Immigration Records to the US Census

Similar to Section B.2, we link immigrants from the Castle Garden Database to US Census. Primarily, we use individual-level first and last name, age and year of immigration to link immigrants from the Castle Garden database to the US demographic census. We match four major immigrants groups by sending countries (Germany, Italy, Ireland, and the UK) between 1850 and 1930. This procedure yields millions of matched immigrant records which we can track their lives for multiple decades.

B.4 Matching Patents to Individuals

We link the universe of US Patent published pre-1910. Each of these patents has (via HistPat) an associated county and state, and (most, but not all, have) an associated inventor name. The inventor name is first parsed into a first and last name, where the last name is the last word in the inventor name (once words like “Jr”, “Sr” are removed), and the first name is simply all proceeding words (and therefore it may include middle name). Then, given this first and last name and county and state, consider the universe of individuals who, as of the 1910 census, were residing in that county and state. This assumes that individuals do not move after patenting. We consider individuals of age over 20 to be a potential inventor at the time of inventing one’s first patent.

Measure the Jaro-Winkler distance between the first name of patents and the listed first name of individuals (males) in the census, and similarly for the last name. For each patent, consider the matched census individual as simply that who has the highest match value. This produced a dataset of 349,198 inventors with information both from the US Patents and US Demographic Census (Inventor Name, County and State of Patents, Census Individual) linked triples.

C Analysis of Hamburg Passenger Lists (HPL)

In this section, we provide a comprehensive analysis of the information contained in the HPL. The HPL database contains passenger lists of ships that departed from the port of Hamburg, Germany from 1850-1914 (the data basically stops before World War I). The database was obtained from the Hamburg State Archive, and we have approximately 4.6 million records of individuals, with approximately 77% of whom were headed to the United States. The database usually contains information such as name, gender, age/birth year, occupation, nationality, departure date, and ship name and we discuss some key information of our interests in the following.

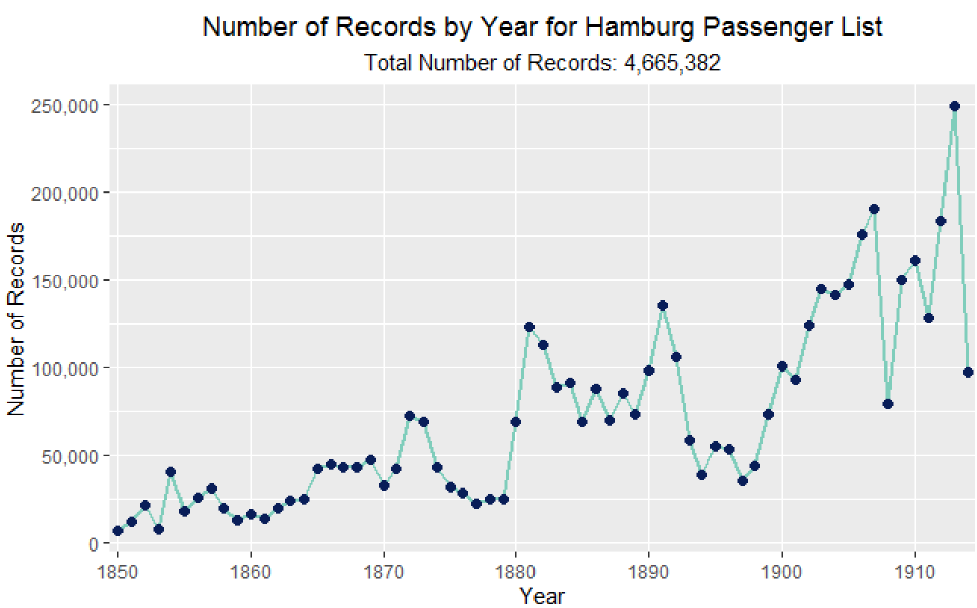


Figure 13: Number of Passengers over Time

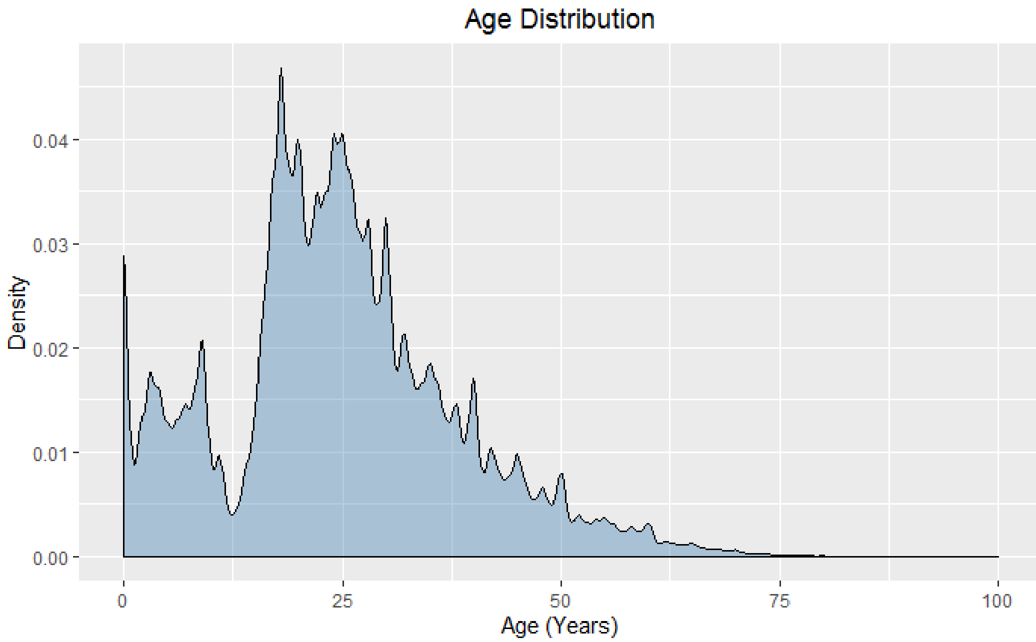


Figure 14: Age Distribution of Passengers

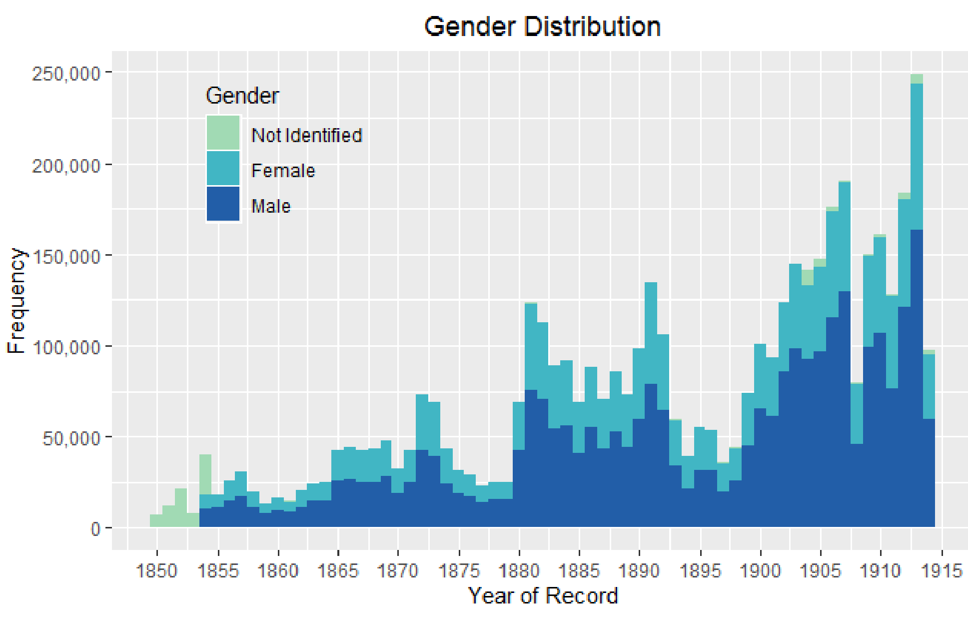


Figure 15: By Gender

Occupation (Beruf) Transcribed information on people's occupations from the Hamburg Passenger Lists (HPL) were cleaned, translated, standardized and coded following the 1950

Census Bureau occupation information classification system and the Historical International Standard classification of occupations (HISCO) to enhance comparability across years. 53% of HPL records have occupational responses and as in Figure 17, occupation was almost always available for working-age males. The most frequently appearing occupations are farm laborers, general laborers, managers, followed by manufacturing workers such as tailors, shoemakers, and cabinetmakers.

Non-Empty Records	**As Percentage of Total Records**	**Top 10 Occupations Recorded**	**Top 10 Occupations Percentage**
2,544,925	52.55%	1,923,315	75.58%

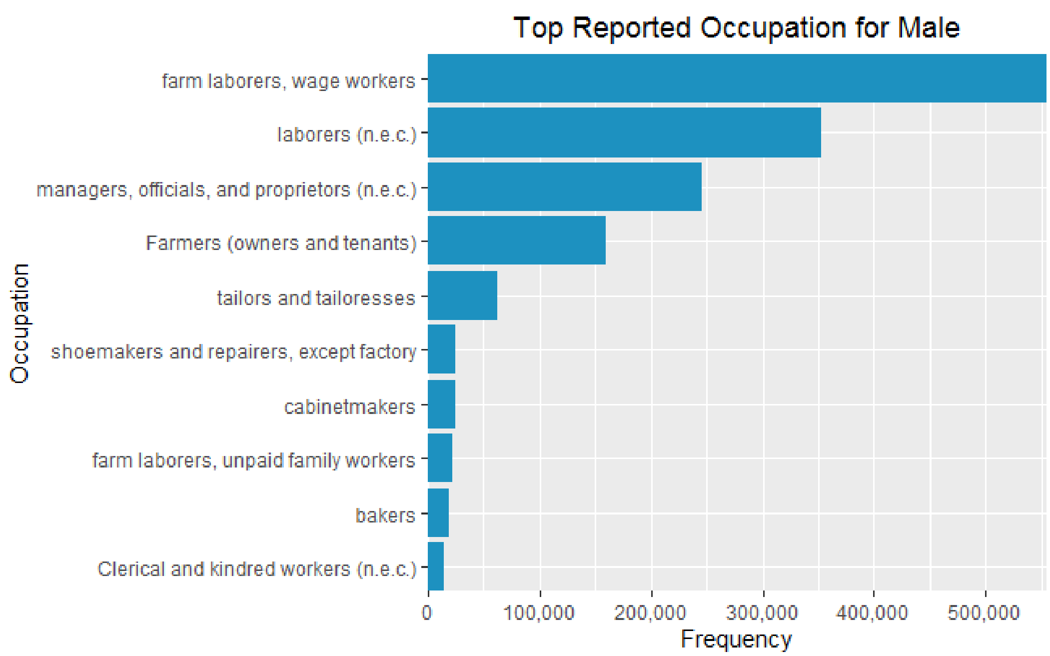


Figure 16: Top Reported Occupations

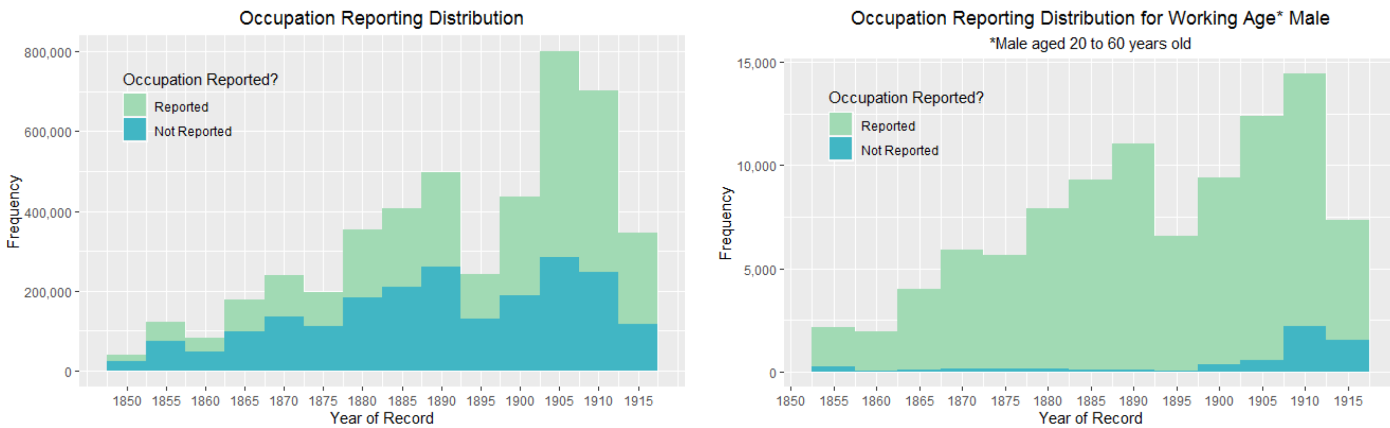


Figure 17: Occupation Reporting Pattern

Destinations (Zielort) HPL records passengers' destinations, and this information is available for approximately 25% of records and we infer passengers' destinations based on ship route and departure information which are universally available. As in Figure 18, destinations from the port of Hamburg vary significantly and majority of passengers headed to North America (such as New York, Chicago, and Quebec in Canada) whereas some headed to South America (such as Buenos Aires and Rio de Janeiro) and Europe (such as London in the UK). As in Figure 19, passengers' destination was almost never reported between 1860 and 1870 and the destination became more available since then.

# Destination Available Records	% of Destination-available Records	Top 1 Place (New York)	% US Destination
1,168,347	24.36%	533,329	77.31%

Table 6: Destination Summary

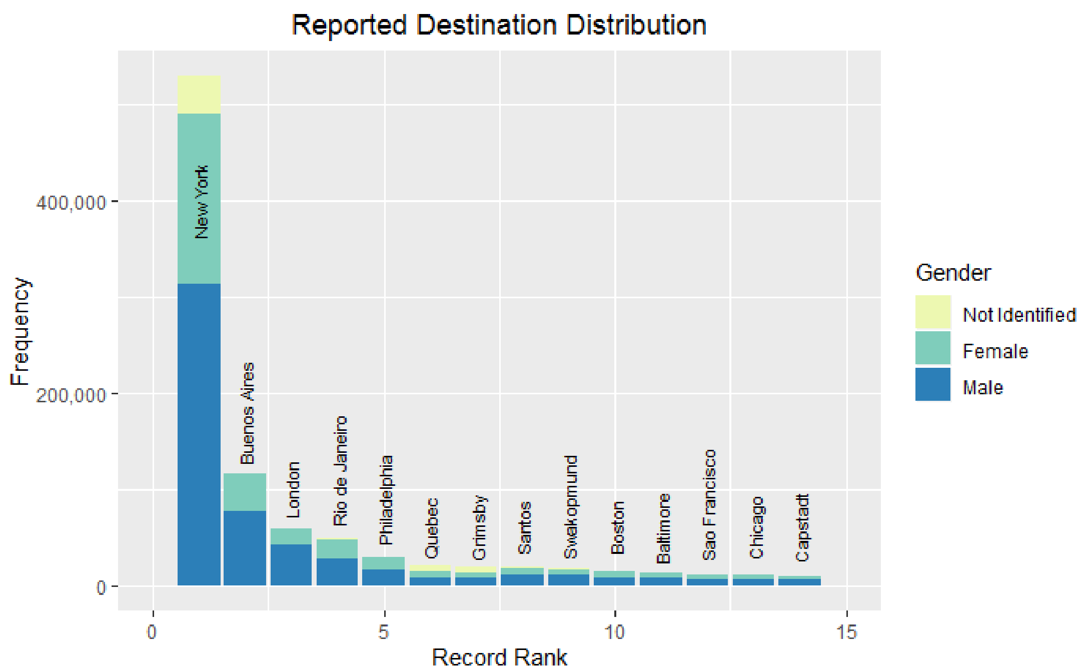


Figure 18: Top Destinations

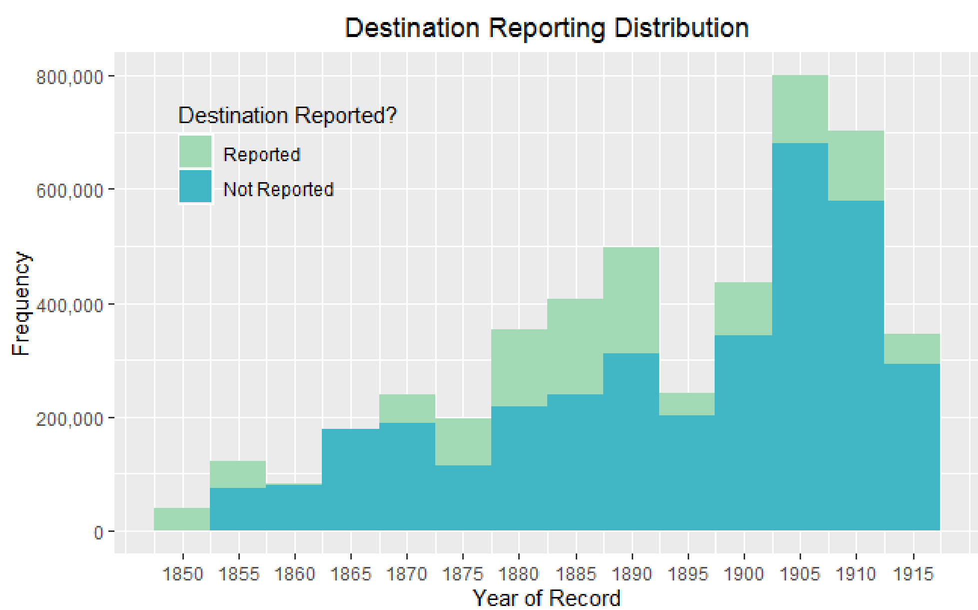


Figure 19: Destination Reporting Over Time

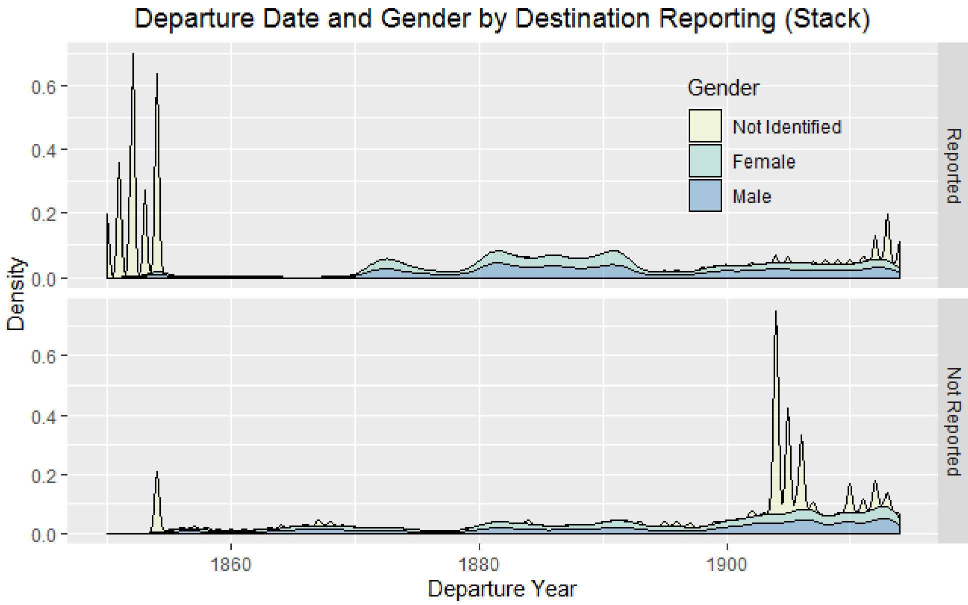


Figure 20:

Origins

- Nationality

Nationality is reported for 45.85% of the records, and this information became mostly available since 1898. The major nationality of HPL records was Russia, Austria, Hungary, the United States, followed by Poland (Table 7, Figure 22 and Figure 21). We do not directly observe nationality in our pre-1898 data, so we use an imputation procedure based on passengers' names and last residences.[13]

	% Available	**Russia**	**Austria**	**Hungary**	**United States**
Records	2,138,970	816,412	386,902	334,945	168,893
Percentage (%)	45.85*	38.17	18.09	15.66	7.90

Table 7: Nationality Summary

[13]We impute nationality by combining first and last name, age and marital status of individuals (Ye et al. (2017)).

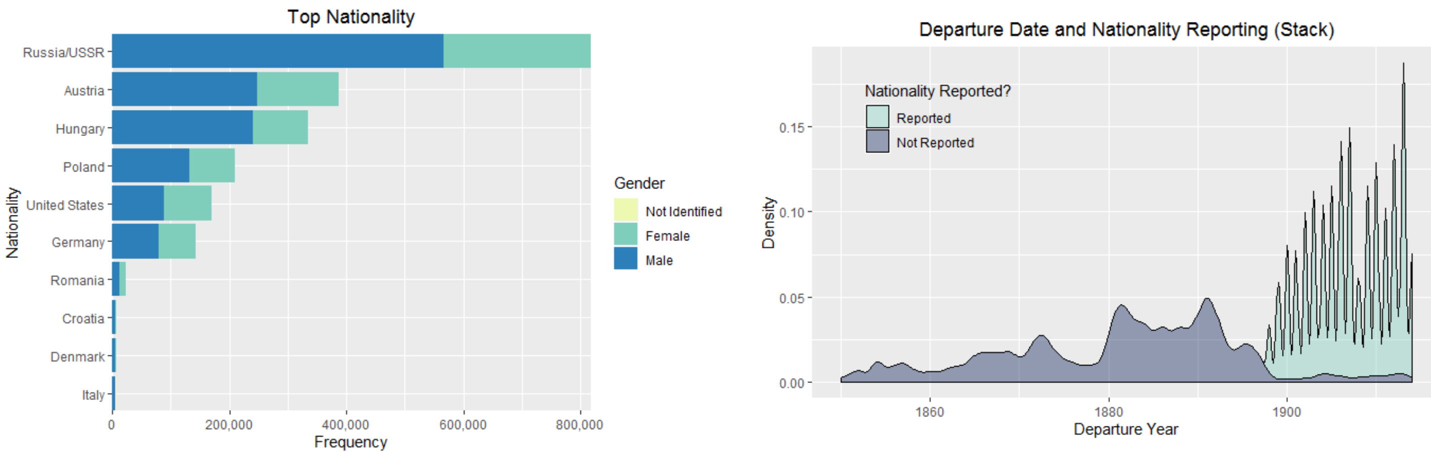


Figure 21: Reported nationalities under current administrative division

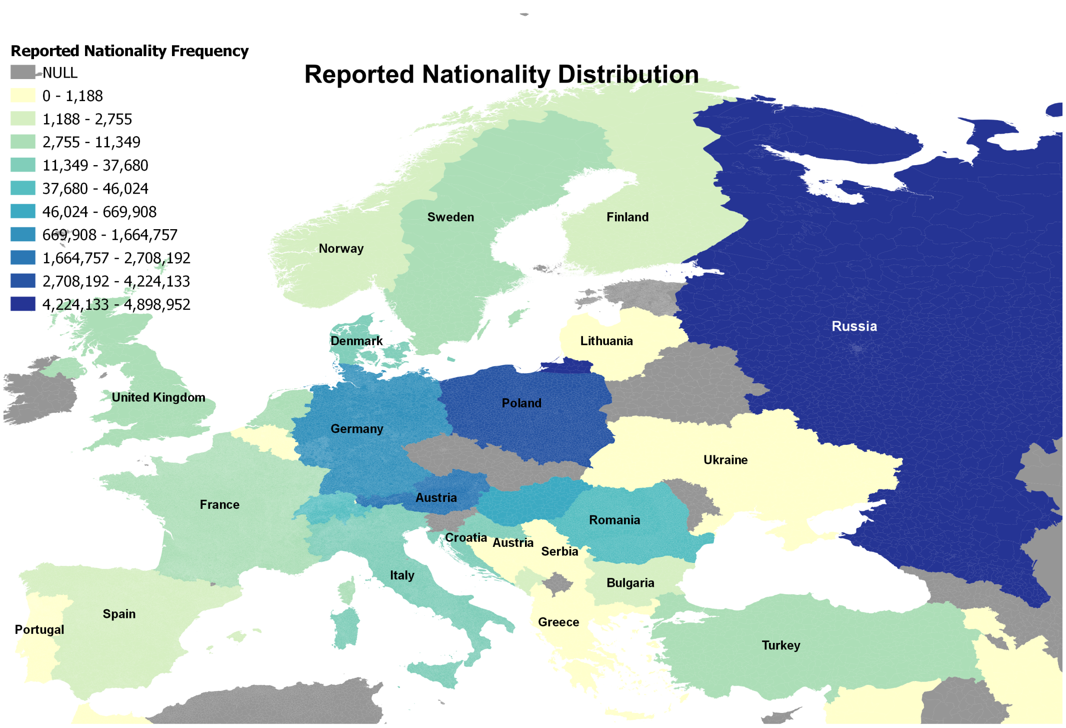


Figure 22: Reported Nationality Distribution (European Area)

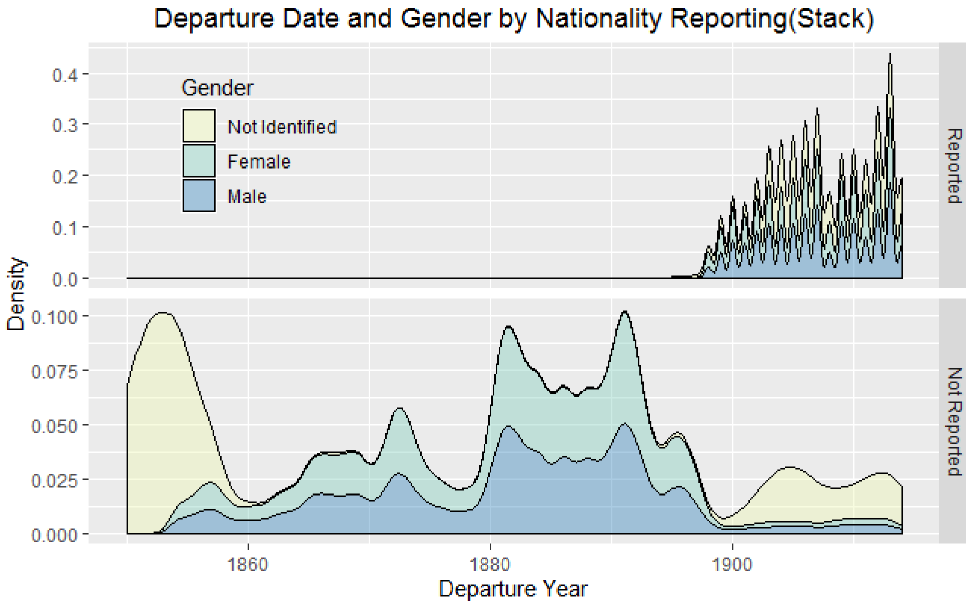


Figure 23: Nationality is reported evenly among gender groups

- Birthplace (Geburtsort)

Birthplace information was available for only 1.89% of the HPL records and it is only available for pre-1860 passengers. Due to its limited coverage of information, we do not use birthplace information as one of the linking criteria.

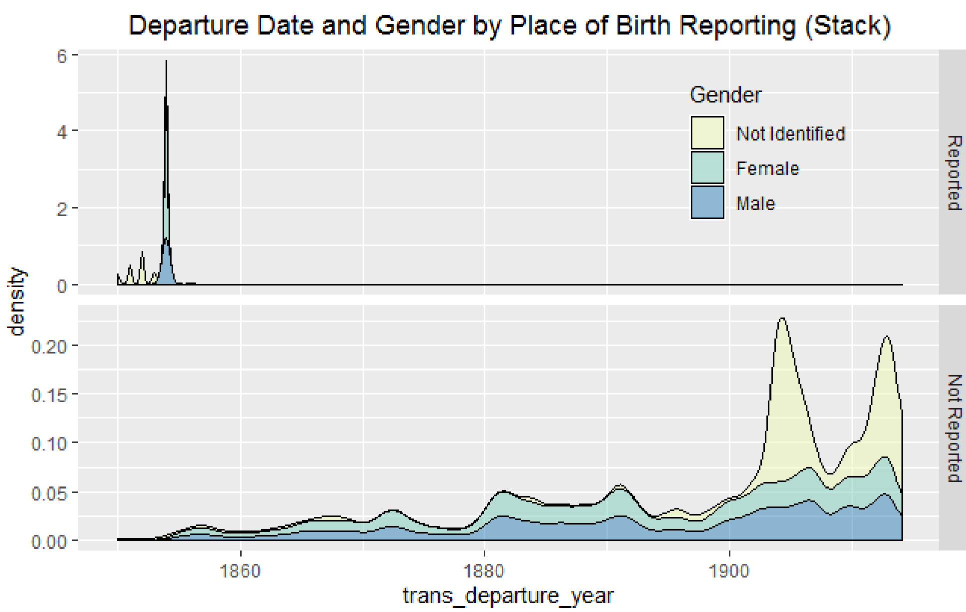


Figure 24: Almost all places of birth are reported before 1860

- Place of Residence(Wohnort)

HPL also records passengers' last residential locations (Table 8). We map transcribed texts of last residence information, harmonize and classify them at the country level using IPUMS geographic classification system. The most frequent last residences are Germany, Poland, Austria, and Hungary (Figure 25).

#Obs Last Residence Aavailable	% of Available Records	Reported Missing	Top 10 Places Reported	Top 10 Places Percentage
2,242,447	48.07%	127,787	1,474,037	65.73%

Table 8: Top places of Residence and frequency

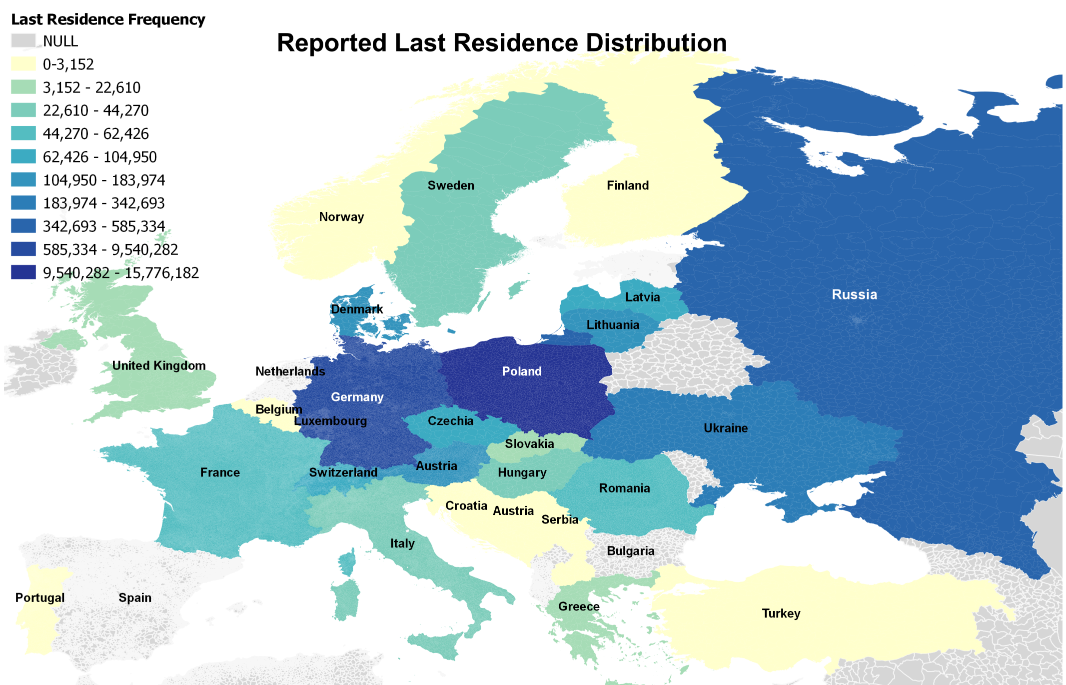


Figure 25: Reported Residence Distribution

D Occupational Classifications

D.1 Hamburg Passenger Lists (HPL) Data Harmonization

As the occupation information in the *Hamburg Passenger Lists* is only available in German and they are neither integrated nor coded with any standardized classification system (recall Mr Steinweg's occupation "Instrumentenmacher" in Hamburg Passenger Lists record), we translate the reported German occupational string and codify them in the occupational system of the US Federal Census. Similarly, the occupational classification available in the *Castle*

Garden Database, while available in English, is also not classified. Therefore, we construct occupational crosswalks between Castle Garden occupation strings and occupational measures which are consistently available in the US demographic census (i.e. variables "OCCHISCO" and "OCC1950").

D.2 Industry Classification and Harmonization

We also match the reported sectoral groups in the Census of Manufacturers to the standard classification of industries that follows the 1950 Census Bureau industrial classification system.[14] The digitized *Census of Manufactures* is more detailed than 1950 Census Bureau industrial classification system. To enhance comparability, we create industry crosswalk from reported industry strings in the Census of Manufacturers to the 1950 Census Bureau industrial classification. The same applies to the historical patent data, where we use a combination of the detailed patent descriptions and the patent classifications to assign individuals patents to particular sectors of production.

Finally, we merge all the aforementioned datasets at a unified spatial level. While the Population Census and the Patent Data is already geo-referenced, we perform geo-referencing using the original geographical information in the Manufacturing Census. We integrate all county aggregates from 1860 to 1940 at consistent levels of aggregation. As counties and state boundaries have changed over time, we take the county shapefiles from National Historical Geographic Information System and create the geographic location-based crosswalks of counties across time.

References

Abramitzky, R., Boustan, L. P., Eriksson, K., August 2012. Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. American Economic Review 102 (5), 1832--56.
URL https://ideas.repec.org/a/aea/aecrev/v102y2012i5p1832-56.html

Adao, R., Arkolakis, C., Esposito, F., 2019. Spatial linkages, global shocks, and local labor markets: Theory and evidence.

Ager, P., Hansen, C. W., 2017. Closing heaven's door: Evidence from the 1920s us immigration quota acts. Available at SSRN 3059439.

Akcigit, U., Grigsby, J., Nicholas, T., 2017. Immigration and the rise of american ingenuity. Tech. rep., National Bureau of Economic Research.

[14]Details are available here: https://usa.ipums.org/usa-action/variables/IND1950#description_section

Anderson, J. E., Van Wincoop, E., 2003. Gravity with gravitas: A solution to the border puzzle. American Economic Review 93 (1), 170--192.

Arkolakis, C., Ramondo, N., Rodríguez-Clare, A., Yeaple, S., 2018. Innovation and production in the global economy. American Economic Review 108 (8), 2128--73.

Atkeson, A., Burstein, A., 2010. Innovation, firm dynamics, and international trade. Journal of Political Economy 118 (3), 433--489.

Dekle, R., Eaton, J., Kortum, S., 2007. Unbalanced trade. Tech. rep., National Bureau of Economic Research.

Doran, K., Yoon, C., 2018. Immigration and invention: Evidence from the quota acts.

Feigenbaum, J. J., 2016. Automated census record linking: A machine learning approach.

Gordon, R. J., 2017. The rise and fall of American growth: The US standard of living since the civil war. Princeton University Press.

Kortum, S., 1997. Research, patenting, and technological change. Econometrica 65 (6), 1389--1419.

Lucas Jr, R. E., Moll, B., 2014. Knowledge growth and the allocation of time. Journal of Political Economy 122 (1), 1--51.

Nunn, N., Qian, N., Sequeira, S., 2017. Migrants and the making of america: The short and long run effects of immigration during the age of mass migration.
URL https://ideas.repec.org/a/ces/ifodic/v15y2017i3p50000000000049.html

Ottaviano, G. I., Peri, G., Wright, G. C., 2013. Immigration, offshoring, and american jobs. American Economic Review 103 (5), 1925--59.

Perla, J., Tonetti, C., 2014. Equilibrium imitation and growth. Journal of Political Economy 122 (1), 52--76.

Roca, J. D. L., Puga, D., 2017. Learning by working in big cities. The Review of Economic Studies 84 (1), 106--142.

Romer, P. M., 1990. Endogenous technological change. The Journal of Political Economy 98 (5), S71--S102.

Ye, J., Han, S., Hu, Y., Coskun, B., Liu, M., Qin, H., Skiena, S., 2017. Nationality classification using name embeddings. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. CIKM '17. ACM, New York, NY, USA, pp. 1897--1906. URL http://doi.acm.org/10.1145/3132847.3133008