IMMIGRATION, SCIENCE AND INVENTION: LESSONS FROM QUOTAS IN THE 1920S*

PETRA MOSER, NYU, NBER, AND CEPR
SHMUEL SAN, NYU

DECEMBER 28, 2019

The United States first adopted immigration quotas for “undesirable” nationalities in 1921 and 1924 to stem the inflow of low-skilled Eastern and Southern Europeans (ESE) and preserve the “Nordic” character of its population. This paper investigates whether these quotas inadvertently hurt American science and invention. Hand-collected data on the countries of birth, as well as the immigration, education, and employment histories of more than 80,000 American scientists reveal a dramatic decline in the arrival of ESE-born scientists after 1924. An estimated 1,170 ESE-born scientists were missing from US science by the 1950s. To examine the effects of this change on invention, we compare changes in patenting by US scientists in the pre-quota fields of ESE-born scientists with changes in other fields in which US scientists were active inventors. Methodologically, we apply *k*-means clustering to scientist-level data on research topics to assign each scientists to a research field, and then compare changes in patenting for the pre-quota fields of ESE-born US scientists with the pre-quota fields of other US scientists. Baseline estimates indicate that the quotas led to 68 percent decline in US invention in ESE fields. Decomposing this effect, we find that the quotas reduced not only the number of US scientists working in ESE fields, but also the number of patents per scientist. Firms employing ESE immigrants before the quotas experienced a disproportionate decline in invention. The quotas damaging effects on US invention persisted into the 1960s.

KEY WORDS: IMMIGRATION, SCIENCE, AND INVENTION.

JEL CODES: O34, L82, AND N42

---

* We wish to thank seminar participants at Berkeley, Boston University, Dartmouth, Duke, Georgia Tech, Harvard, Hokkaido, Iowa State, Lausanne, Luxembourg, NYU, Peking University, SUNY Binghamton, UCLA, Yale SOM, and Vancouver, as well as Barbara Biasi, Josh Lerner, Hiroyuki Okamuro, and Marco Tabellini for helpful comments. Ryan Stevens, Bang Nguyen, and Vasily Rusanov provided outstanding research assistance. Moser gratefully acknowledges financial support from NYU’s Center for Global Economy and Business and from the National Science Foundation through Grant 1824354 for Social Mobility and the Origins of US Science. San thanks the Economic History Association for financial support through an EHA Dissertation fellowship.

In the 1920s, the United States implemented nationality-based immigration quotas to keep out low-skilled immigrants from Eastern and Southern Europe and preserve the "Nordic" character of its population. This paper examines the effects of such policies on American science and invention. Did the quotes discourage foreign-born scientists from coming to the United States, even though they targeted low-skilled workers? Did they encourage or discourage invention by American scientists? And how did this change affect US invention overall?

Until the late 19th century, most immigrants to the United States had come from Britain, Ireland, Germany, and other German-speaking parts of Europe. By 1890, changes in pull and push forces shifted the sources of mass migration to Italy and Eastern Europe. These "new" immigrants met with a surge of nativist sentiment, reaching to the highest level of the US executive. Writing in the popular magazine Good Housekeeping, soon-to-be Vice President, Calvin Coolidge (1921, pp. 13-14) argued that the United States "must cease to be regarded as a dumping ground," and asked for an "ethnic law" to change the nature of immigration. A 1921 editorial in the *New York Times* warned that "American institutions are menaced; and the menace centres (sic) in the swarms of aliens whom we are imported as 'hands' for our industries."

Intended to stem the inflow of low-skilled immigrants from Eastern and Southern Europe, the 1921 Emergency Quota Act (Ch. 8, 42, Stat 5) restricted the number of immigrants per year to 3 percent of the number of residents from that country in the US Census of 1910. When this quota proved ineffective, the 1924 Johnson-Reed Act further reduced the quota to 2 percent and changed its reference population to the Census of 1890 (pub. L. 68-139, 43, Stat. 153). With these changes, immigration fell precipitously from nearly 360,000 in 1923-24 to less than 165,000 the following year. But, beyond merely reducing the number of immigrants, the 1924 quota act adjusted the ethnic mix of migration. Arrivals from Asia were banned, and immigration from Italy fell by more than 90 percent, while immigration from Britain and Ireland dropped by a mere 19 percent (Murray 1976, p. 7).

Strengthened during the Cold War, the national origins quotas ruled US immigration until they were abolished by the Immigration Act of 1965. In his "Remarks on Signing the Immigration Bill" President Lyndon B. Johnson (1965) called the quota system a "cruel and enduring wrong […] Only 3 countries were allowed to supply 70 percent of all the immigrants. […] Men of needed skill and talent were denied entrance because they came from southern or eastern Europe […] We can now believe that it will never again shadow the gate to the American Nation with the twin barriers of prejudice and privilege."

This paper uses detailed biographical data on nearly 100,000 American scientists in 1921 and 1956, matched with their patents, to examine the effects of the quotas on American science and invention. A major strength of our data is that they include the full name of each scientist, their precise birth dates and place of birth, their education and employment history, and their year of naturalization. Using birth dates and full names, we are able to establish a high-quality match between scientists and their patents.[1]

Naturalization data indicate a dramatic decline in the arrival of new ESE-born scientists after the quotas. Until 1924, arrivals of new ESE-born immigrant scientists were comparable to arrivals from Northern and Western Europe (NEW), who were subject to comparable pull and push factors of migration.[2] After the quotas, arrivals of ESE-born scientists decline significantly while arrivals from Northern and Western Europe continue to increase. Extrapolating from naturalization records we estimate that more than 800 ESE-born scientists were missing from the United States scientific workforce as a result of the quotas. At an annual level, this implies a loss of 33 scientists per year, equivalent to eliminating the physics department of a major university each year.

To estimate the effects of this change on US inventions, we compare changes in patenting per year after 1924 in the pre-quota fields of ESE-born US scientists with changes in patenting the pre-quota fields of other US scientists. This identification strategy allows us to control for changes in invention by US scientists across fields, for example, as a result of changes in research funding. Year fixed effects further control for changes in patenting over time that are shared across field. Field fixed effects control for variation in the intensity of patenting across fields, e.g., between basic and applied research.[3]

Methodologically, we apply *k*-means clustering to scientist-level data on research topics to assign each scientist to a unique research field, and then compare changes in US patents per year in the pre-quota fields of ESE-born US scientists with the pre-quota fields of other US scientists. Intuitively, *k*-means clustering works like a multi-dimensional least square algorithm, which groups together data points (here, scientists) that are most similar in terms of their observable characteristics (here, research topics). We first apply *k*-means clustering to the research topics of all 41,094 American scientists in 1956 to assign each scientist to a unique

---

[1] Starting from a standard Levenshtein (1966) distance measure (allowing one letter to differ between the scientist's name and the name on a patent), we use the scientist's age in the year of the patent application to filter out false positive. Specifically, we estimate a false positive, type I, error rate using the number of patents that were submitted when the inventor was between 0 and 18 years old as a proxy for false positive matches between scientists and patents. Exploiting data on the first, middle, and last name, discipline, and the frequency of her first and last name, we are able to reduce this error rate to 5 percent compared with more than 80 percent for the most naïve Levenshtein matching (ignoring middle names, disciplines, and name frequencies). We estimate robustness checks including common names and allowing for different middle names (Table A3).

[2] Immigrants from other parts of Europe were attracted by the same labor market conditions and faced similar costs of trans-Atlantic migration but, unlike ESE immigrants, they were not targeted by the quotas. We describe these factors in more detail in section 4.

[3] For example, scientists may patent more in an applied field, such as "radio waves," than in a more theoretical field, such as "calculus of variations." Moreover, inventors may choose to patent their inventions in some fields but not in others, depending on the effectiveness of alternative mechanisms (Moser 2012a). Research field fixed effects control for such differences in the intensity of patenting.

research field.[4] We then use the research topics of American scientists in 1921 to assign each of them to one of the fields defined on the 1956 data. This process allows us to identify research fields in which ESE-born scientists were active in 1921, before the quotas.

Baseline estimates reveal a large and persistent decline in invention by US scientist in the pre-quota fields of ESE-born scientists. US scientists produced 68 percent fewer additional patents in the pre-quota fields of ESE-born scientists after 1924 compared with the pre-quota fields of other US scientists. Time-varying effects show a large decline in invention by US scientists in the 1930s, which persisted through World War II and into the 1960s. Importantly, time-varying estimates indicate no pre-existing differences in patenting for ESE and other fields until 1924.

This large and persistent decline in invention by US scientists is robust to a broad range of alternative regression models, including quasi-maximum likelihood (QML) Poisson, and negative binomial regressions. It is also robust to alternative controls for pre-trends in patenting and to different choices of *k,* which determines the number of fields. All results are robust to alternative matchings between scientists and patents, even though these alternative methods introduce a substantial amount of noise, as we have shown above.

Invention in ESE fields declined both at the intensive margin (from more to fewer patents) and at the extensive margin (from some patents to no patents at all). After 1924 US scientists produced 45 percent fewer patents in ESE fields, and 16 percent additional ESE fields had no patents at all. Complementary tests at the level of individual scientists indicates that 40 percent *fewer scientists* were active in ESE fields after the quotas and that US scientists produced 33 percent *fewer patents per scientist*. Time-varying estimates, which compare the number of active in ESE fields with other fields show that timing of this decline closely matches the timing of the observed decline in patenting.

Importantly, estimates for US-born scientists are only slightly smaller than estimates for all American scientists (at 62 percent, compared with 68 percent). These results indicate that the benefits from reduced competition with immigrants were substantially smaller than the costs of reduced interactions with ESE-born scientists, The case of the famous Hungarian mathematician Paul Erdős illustrates how the quotas reduced interactions and knowledge spillovers from ESE-born to US-born scientists.[5] A professor at Notre Dame and a Hungarian citizen, Erdős was

[4] 41,094 American scientists in 1956, include 39,998 scientists who work in the United States in 1956 and another 911 scientists who are employed in Canada. Another 185 American scientists work outside the US and Canada in 1956, we exclude these scientists from the main tests.

[5] Even today most mathematicians and many economists know their Erdős number, the number of co-authors that separate them from Paul Erdős. In 2016, the median winner of the Fields Medal had an Erdős number of 3 (with a range from 2 to 6) compared with 5 across all of mathematics. In economics, the median Erdős number for a Nobel Laureate is 4 (with a range from 2 to 8, very close to math*)* *wwwp.oakland.edu/enp/trivia* accessed July 31, 2019.

denied a re-entry visa by the US immigration services in 1954, and not granted re-entry until 1963. During these years, Erdős professional network of collaborators shifted away from the United States to Europe. Between 1954 and 1963, 24 percent of Erdős' new co-authors were US scientists, compared with 62 percent until 1954 and 56 percent afterwards.

To further investigate the mechanism of the observed decline in invention, we examine the influence of selection into research fields. To perform these tests, we estimate placebo regressions for Canada, which did not implement comparable national origins quotas in 1924. Time-varying estimates indicate no decline in Canadian invention in ESE fields. In fact, invention by Canadian scientists in ESE-fields increases relative to US scientists after 1924. We also investigate whether the aging of scientists can explain the decline in invention in ESE fields, as the quotas reduced the inflow of young immigrants. This analysis indicates that the aging of ESE fields contributed to the decline in invention, without, however, explaining a substantial share.

Most importantly, the quotas appear to have prevented ESE-born refugees from the Nazis to flee to the United States. ESE countries were especially affected by Nazi brutality. Poland, for example, had the largest Jewish population in 1933, with more than 3 million people. By 1950 Poland had lost 98 percent of that population, with less than 50,000 remaining Jews.[6] While German-born refugees were allowed into the United States (where they encouraged innovation, Moser, Voena, and Waldinger 2014), the quotas capped the entry ESE-born refugees, causing an immense loss for US science and invention.

A final section explores the broader effects of the quotas on the firms that employ immigrant scientists. We find that firms who employed ESE-born immigrants in 1921 created 53 percent fewer inventions after the quotas. A text analysis of the titles of US patents indicates that invention also declined more broadly, beyond firms that were directly affected by the quotas. After the quotas, 23 percent fewer US patents describe inventions that relate to ESE fields compared with other fields.

We also investigate the quotas potential spillovers on other countries, and specifically Israel. Migration data show that some of the missing scientists moved to the future Israel, where they helped to build the foundation for scientific institutions that fuel Israeli innovation to this day. Migration data for Jewish scientist, which we collect from another source (the *World Jewish Register,* 1955) reveal a dramatic increase in the migration of Jewish scientists to Palestine, around the time of the quotas. Several of these scientists moved to the Technion, which had been

[6] United States Holocaust Memorial Museum Collection, accessed July 19, 2019.

founded in Haifa in 1912, and grew dramatically during the time of the quotas. Today, the Technion is Israel's premier university for technology and science.[7]

Our findings relate to the broader literature on immigration, and in particular on the effects of immigration on innovation.[8] Several recent papers examine the effects of the quota acts on low-skilled immigration (Tabellini forthcoming, Doran and Yoon 2019, Abramitzky et al. 2019).[9] Our research complements these papers by investigating the quotas *unintended effects on highly skilled immigrants* - which were not the target of the acts. Our approach also implements a different identification strategy, which allows us to examine the effects of the quotas on American science and invention across fields. Following Card (2001), other papers have used geographic variation in the pre-existing flows of immigrants to identify the effects of immigration on locations that had received many immigrants leading up to the quotas.[10] Since we are interested in examining knowledge spillovers in *idea space* we define the unit of analysis at the level of research fields.[11] This approach allows us to investigate the effects of immigration on the inventions by American scientists and other American inventors.

1. Historical Background

Until 1880, 90 percent of immigrants to America came from the British Isles and German-speaking parts of Continental Europe (Historical Statistics of the United States 1975, pp. 106-09). By the end of the 19th-century, labor markets these areas began to tighten, turning Britain and Germany into net importers of workers.

1.1. After 1890, Sources of Mass Migration Shift to Eastern and Southern Europe

---

[7] Technion Presidents Report 2018, available at https://presidentsreport.technion.ac.il/the-technion-in-numbers/, accessed on June 10, 2019.

[8] For example, Kerr and Lincoln (2010); Hunt and Gauthier-Loiselle (2010), Moser, Voena, and Waldinger (2014), Clemens, Lewis, and Postel (2018), and recent working papers by Bernstein, Diamond, McQuade, and Pousada (2019) and San (2019). Annelli et al (2019) take a different approach by examining the effects of outmigration on entrepreneurship and the creation of new firms.

[9] Using pre-existing settlement patterns as an instrument for the location decisions of new immigrants, Tabellini (forthcoming) finds that immigration triggered support for anti-immigrant legislation (and the election of more conservative legislators) even where it increased employment. Doran and Yoon (2019) find that restrictions on unskilled immigration reduced innovation, while Abramitzky et al. (2019) show that the loss of immigrant workers encouraged farmers to shift toward capital-intensive agriculture and encouraged US born workers to move to cities.

[10] Sequeira, Nunn, and Qian (2019) pursue a different identification strategy, which also exploits geographic variation. To examines the effects of European immigration *before the quotas*, during the Age of Mass Migration (1850-1920), they interact variation over time in total arrivals to the United States with variation across locations and over time in the expansion of the US railway network: New waves of immigrants were more likely to move to counties that had recently been connected to the rail network.

[11] This approach is consistent with the research of Azoulay et al (2010) on super star inventors, which suggests that knowledge spillovers are strongest in idea space, rather than in geographic space.

In the final years of the 19^{th} century, a combination of push and pull factors triggered a new wave of mass migration from Eastern and Southern Europe. One major pull factor was America's rapid industrialization, which increased US demand for unskilled workers (Rosenbloom 2002) while improved rail and steamship links to the United States facilitated immigration from Eastern and Southern Europe (Keeling 2012, p. 23). Among push factors, lower transport costs reduced the benefits from staying at home, as competition with American grain reduced rural incomes (O'Rourke 1997, pp. 775-76). Jews from Russia's Pale of Settlement came to the United States to escape oppression and violence. Across Russia, Poland, and Austria-Hungary, the hardship of military service further encouraged migration.

Due to a combination of these factors, the share of Eastern Europeans and Italians among all US immigrants exploded from a mere 8 percent in the 1870s and 18 in the 1880s to 49 percent in the 1890s, 76 in the 1900s and 80 in the 1910s. Three countries alone - Russia, Austria-Hungary, and Italy - accounted for nine in ten immigrants from Southern and Eastern Europe. None of these countries had made up more than 10 percent of European migration before 1890.

To better understand the changed nature of these new flows, the Federal Bureau of Immigration began to compile statistics on a new category "race" based on a person's "mother tongue." Collecting this new "race" variable in addition to "country of origin"[12] allowed the Bureau to distinguish "Poles" and "Hebrews" among immigrants from Russia and to separate "Poles" into Poles from Germany and Poles from Russia. The first tallies of the new race variable in 1899 showed that 26 percent of immigrants from Europe were Italians, 12 percent were "Hebrews," and 9 percent were Poles. These relative shares stayed roughly constant until the eve of World War I, with Italians averaging 24 percent and Jews and Poles 11 percent each.[13]

Most Italian immigrants were "propertyless peasants" from the rural South. Roughly two thirds of Polish immigrants were "landless peasants and the agrarian proletariat" (Nugent, 1992 p. 94). Jewish immigrants, three quarters them coming from Russia, were artisans, professionals, and urban workers from medium-sized towns ("shtetls").

In 1915, Arthur Salz, a German Jewish professor of Economics at Heidelberg summed up the role that these Eastern European immigrants played in the US economy.

---

[12] The US Immigration Act of 1903 required passenger lists to record "race." Between 1899 and 1903, "race" was recorded on supplemental passenger manifests, under the column "Mother Tongue (language or dialect)." While the language at the time used "race" to distinguish ethnicities and even religious affiliations (Christian vs. Jewish), we use the hyphenated term "race" and adhere to usage today when "race" denotes "each of the major divisions of humankind, having distinct physical characteristics, and "ethnicity" defines membership in a "social group that has a common national or cultural tradition." Oxford Living Dictionary, 2018.

[13] Italians were further divided into Northern Italians (4 percent) and Southern Italians (20 percent, Bureau of Immigration, Annual Report, 1915, pp. 101-102, cited in Keeling (2012, p. 25).

> "These men, employed in agriculture or as manual workers or day laborers in their home countries, fully supply the needs of American industry for unskilled labor. They not only supply that market, they oversupply it, and monopolize it: They are the sacred regiments of a reserve army drawn from the ranks of the willingly enslaved." (Salz 1915, pp. 110-11, cited in Keeling 2012, p. 21).

At the end of the 19th- century, nearly half of all workers in New York, Chicago, and Boston were foreign-born. Across the United States, one fifth of the labor force came from abroad. By 1910, half of all industrial workers, miners, and railroad employees in the United States were born outside of the United States. More than half of all garment-makers, and one quarter of all domestic servants were foreign-born. In New York City, one quarter of the police had been born outside of the United States (Rosenbloom 2002, p. 16, Taylor 1971, pp. 192-201).

1.2. Nativism Reaching up to the Highest Levels

Cultural differences between the old and new immigrants triggered a nativist response reaching up to the highest levels of the executive (Jones 1992, p. 176).[14] In 1911, Commissioner Williams (p. 215): wrote in the Bureau of Immigration's annual report that "We should…strive for quality rather than quantity." In the same year, the 41-volume Dillingham report proposed the introduction of a literacy test. Yet, when it was introduced in 1917 this test failed to stem Eastern European immigration because the new arrivals could read remarkably well.

In February 1921, soon-to-be Vice President Calvin Coolidge warned that the United States "must cease to be regarded as a dumping ground," and asked for an "ethnic law" to regulate migration. An editorial in the *New York Times* (February 9, 1921, p. 7) argued in favor of the proposed law

> The Immigration Bill will serve as an index, a finger that points accusation. The need for restriction is manifest. Literally millions of workmen are out of employment. American institutions are menaced; and the menace centres (sic) in the swarms of aliens whom we are importing as 'hands' for our industries, regardless of the fact that each hand has a mind and potentially a vote. With the diseases of ignorance and Bolshevism we are importing also the most loathsome diseases of the flesh. Typhus, the carrier of which is human vermin, has already been scattered among us…[15]

---

[14] The distinction between "new" and "old immigration" was first made in the Dillingham Report (e.g., vol. 1, pp. 12-14), named for its chairman US Senator William P. Dillingham, a Republican from Vermont.

[15] While the Republican party initially served as the principal channel of restrictionist agitation, the shift of "big business" to an anti-restrictionist view, and the attraction of southern and eastern European voting blocs ended the party's effectiveness as a nativist instrument (Higham 1955, 8th edition, p. 126). Describing media influence on the opposite end of the political spectrum Higham (1955, 8th edition 2011, p. 127) explains that William Randolph Hearst who "exerted no little influence in (sic) behalf of the foreign-born, for he gave them raucous support and received in return their devoted loyalty….Hearst learnt early that a newspaper with bold type, simple ideas, and passionate appeals for social justice could command the pennies and the votes of the immigrant working class. He became the knight-errant of the tenements…posing as the great American champion of the maltreated Jews in

1.3. Quotas Target Eastern and Southern Europeans

On May 1921 the Emergency Quota Act (Ch. 8, 42, Stat 5) introduced numerical limits on the number of immigrants per year, for the first time in US history. The Act also established a quota system that restricted immigrants per year to 3 percent of the number of residents from that country in the US Census of 1910. Yet, due to the dramatic inflow of immigrants from Southern and Eastern Europe between 1890 and 1910, the 1921 Act had little bite.

When Warren G Harding died of a heart attack on August 2, 1923, Coolidge became President and used his first address to Congress to argue for restrictions on immigration:

> "New arrivals should be limited to our capacity to absorb them into the ranks of good citizenship. America must be kept American. For this purpose, it is necessary to continue a policy of restricted immigration."

In May 1924, the Johnson-Reed Act (pub. L. 68-139, 43, Stat. 153) reduced the national origins quotas to 2 percent and pushed their reference population back to the Census of 1890. Senator Reed, a Republican from Pennsylvania, explained his reasoning for the Act in *New York Times* article on "Our New Nordic Immigration Policy"

> "There has come about a general realization of the fact that the races of men who have been coming to us in recent years are wholly dissimilar to the native-born Americans; that they are untrained in self-government – a faculty that it has taken the Northwestern Europeans many centuries to acquire. […] From all this has grown the conviction that it was best for America that our incoming immigrants should hereafter be of the same races as those of us who are already here, so that each year's immigration should so far as possible be a miniature America, resembling in national origins the persons who are already settled in our country […] It is true that 75 per cent of our immigration will hereafter come from Northwestern Europe; but it is fair that it should do so, because 75 per cent of us who are now here owe our origins to immigrants from those same countries." (*Literary Digest*, May 10, 1924, pp. 12-13)

To ensure enforcement, Congress appropriated funding and instructed courts to deport nationals from countries that had exceeded their quotas. With these changes, immigration fell precipitously from 357,803 in 1923-24 to 164,667 in 1924-25. Arrivals from Asia were banned, and immigration from Italy fell by more than 90 percent. At the same time, arrivals from Britain and Ireland dropped by a mere 19 percent (Murray 1976, p. 7).

The quotas were in effect for more than 40 years. During the Cold War, Congress further solidified the national origins quotas through the 1952 Immigration and Nationality Act. In late

---

Russia,….For years, therefore, the growing chain of Hearst newspapers fulminated against further restrictive legislation and also against strict enforcement of existing laws."

September 1965, however, Fidel Castro's announced that Cubans with families in the United States would be permitted to emigrate. On October 3 of the same year Lyndon B. Johnson (1965) made the following "Remarks on Signing the Immigration Bill" on New York's Liberty Island:

> This bill that we will sign today […] corrects a cruel and enduring wrong in the conduct of the American Nation […] Yet the fact is that for over four decades the immigration policy of the United States has been twisted and has been distorted by the harsh injustice of the national origins quota system. Under that system the ability of new immigrants to come to America depended upon the country of their birth. Only 3 countries were allowed to supply 70 percent of all the immigrants. […] Men of needed skill and talent were denied entrance because they came from southern or eastern Europe or from one of the developing continents. […] Today, with my signature, this system is abolished. We can now believe that it will never again shadow the gate to the American Nation with the twin barriers of prejudice and privilege.[16]

2. Data: American Scientists and their Patents

Our main data consist of hand-collected biographical information on nearly 100,000 American scientists, matched with their US patents between 1900 and 1970. A major strength of these data is that they include precise information on scientist's place of birth (allowing us to identify foreign born scientists), the scientist's date of birth (which we exploit to create a high-quality match between scientists and their patents), as well as information on education, employment, and naturalization (which we use to determine each immigrant's year of arrival in the United States).[17]

2.1. Detailed Biographies of American Scientists in 1921 and 1956

Biographical data are drawn from the 1921 and 1956 edition of the Men of Science (MoS). Originally collected by James McKeen Cattell (1860-1944), the "chief service" of the MoS was to "make men of science acquainted with one another and with one another's work" (Cattell 1921). Cattell had been the first US professor of psychology in the United States. He also served as the first editor of *Science* and remained in that role for 50 years. Cattell used this

[16] The Immigration Bill (H.R. 2580) became Public Law 89-236 (79 Stat. 911). On October 7, the first Cuban refugees came on a small boat; days later more than 700 Cubans arrived in Florida. According to a White House Statement on February 15, 1966, two months after the act, " it has already reunited hundreds of families through its preferential admissions policy for aliens with close relatives in the United States .... Another 9,268 refugees from Cuba arrived in the United States during 1965. Of these, 3,349 came in December via the airlift arranged by the United States and the Cuban governments. Some 104,430 resident aliens were naturalized as American citizens during the year" (Weekly Compilation of Presidential Documents (vol. 2, p. 220). Card (1990) examines the effects of the Mariel Boatlift on Miami's local labor market and finds that the influx of Cuban immigrants had no deflationary effect on local wages, even though it increased Miami's labor force by 7 percent, primarily among unskilled workers.

[17] Existing analyses have used names as a proxy for ethnicities (e.g., Moser 2012b). Name-based ethnicity measures, however, measure national origins with much noise and may be a biased measure of ethnic origins..

expertise to establish a compendium of American scientists that he used in his own research.[18] Cattell published the first set of scientist biographies in the American Men of Science (MoS, for short) in 1907, and continued updating the compendium until he retired, passing the baton to his son Jacques.

The MoS was "initially intended as a reference list for the Carnegie Institution of Washington....But the chief service it should render is to make men of science acquainted with one another and with one another's work" (Cattell 1921). Despite its name the MoS included both male and female scientists in Canada and the United States.

To capture the state of American science immediately before the quota act, we hand-collected all 9,544 biographies from the 1921 edition of the MoS. According to the editors, the 1921 edition is "tolerably complete for those in North America who have carried on research work in the natural and exact sciences." (Cattell and Brimhall 1921, p.v). Beyond this strict definition of science, the 1921 includes exceptional people in fields outside of the hard sciences: "There are also some whose work has been chiefly in engineering, medicine or other applied sciences, and a few whose work is in education, economics or other subjects not commonly included under the exact and natural sciences. But the book does not profess to cover these fields." (Cattell and Brimhall 1921, p.v)

Detailed biographical data for 82,094 American scientists in 1956 make it possible to observe American scientists 20 years after the quotas.[19] Beyond the Physical Sciences (volume 1), and the Biological Sciences (volume 2), the 1956 edition also includes the Social & Behavioral Sciences (volume III, 15,493 scientists).[20] We use this disciplinary division to improve the patent matching (as described below).

Both the 1921 and the 1956 edition of the MoS were subject to comprehensive input and review from "scientific societies, universities, colleges, and industrial laboratories." Cattell's son Jaques thanks them for having "assisted in supplying the names of those whom they regard as having the attainments required for inclusion in the Directory." He also thanks "the thousands of scientific men who have contributed names and information about those working in science," and "acknowledges the willing counsel of a special joint committee of the American Association

---

[18] Like many of his contemporaries, Cattell was intrigued by eugenics. Implementing his own special brand of those theories, he offered his children $1,000 each for marrying the offspring of another professor.

[19] This count excludes 6,352 duplicate mentions of scientists who appear in more than one of the three volumes of the MoS (1956). We also exclude 2,015 scientists whose entry consists only of a reference to another edition of the MoS, and we omit 534 scientists whose entry is a reference to the 3rd edition of the *Directory of American Scholars* (1957), an analogue to the MoS for the humanities, by the same editor, Jaques Cattell.

[20] Each scientist could choose the volume in which they wanted to be included, and "depending on the field emphasized in his specialty, his wish was followed in so far as possible." A "'see reference' has been inserted in the other volume, so that the scientist's name appears in both volumes" (Cattell 1956, Editor's Preface). We only count each scientist once and use information on their research topics (below the level of the volume and below the level of the discipline) to define the research field of each scientist.

for the Advancement of Science and the National Academy of Science National Research Council “which acted in an "advisory capacity“ (Cattell 1956, Editor‘s Preface).

2.1.1. Date and Place of Birth

A major advantage of the MoS is that they list the scientist’s date and place of birth. For example, the entry for Professor George Michael Volkoff in the 1956 edition tells us that Volkoff was born in Moscow, Russia, on February 23, 1914:

> Volkoff, Prof. G(eorge) M(ichael), Dept. of Physics, University of British Columbia, Vancouver 8, B.C. Can. PHYSICS. Moscow, Russia, Feb. 23, 1914, Can. Citizen; m.40, c.3. B.A., British Columbia, 34. M.A. 36, hon D.Sc, 45: Royal Soc. Can. Fellow, California, 39-40, Ph.D. (theoretical physics), 40. Asst. prof. physics, British Columbia, 40-43; assoc. research physicist, Montreal lab, Nat. Research Council Can, 43-45, research physicist and head theoret. Physics branch, Atomic Energy Proj. Montreal and Chalk River, 45-46, PROF. PHYSICS, BRITISH COLUMBIA, 46- Ed.’Can. Jour. Physics.’ 50- Mem. Order of the British Empire, 46. A.A; Asn. Physics Teachers; Physical Soc; fel. Royal Soc. Can; Can. Asn. Physicists. Theoretical nuclear physics; neutron diffusion; nuclear magnetic and quadrupole resonance.

Information on birth dates enables us to identify false positive matches, which we use to improve the matching process (as described below). Birth years are available for 99.23 percent of the 82,094 MoS in 1956; exact birth dates (including the day and month) are available for 98.93 percent. Volkoff’s place of birth is “Moscow, Russia,” which makes him “ESE-born.” . Birth places are known for 81,682 of 82,094 American scientists in 1956 (99.5 percent), and 79,114 of 79,507 US scientists (99.5 percent). Among US scientists in 1956, 2,066 (2.5 percent) were born in Eastern or Southern Europe. Another 4,029 scientists (4.9 percent) were born in Northern or Western Europe, 70,927 (86.4 percent) were born in the United States, and another 3,117 (3.8 percent) were born in Canada (Table 1). The most common birthplaces for ESE-born US scientists are Russia, Poland, and Hungary with 613, 319, and 256 scientists, respectively, followed by Czechoslovakia (201) and Italy (173 scientists, Figure A2). In 1921, birth places are known for 9,449 American scientists (99.0 percent). Like in 1956, Russia, Poland, and Hungary were also the most common birthplaces for ESE-born American scientists in 1921 respectively (Figure A1).

2.1.2. Naturalization Records

As a first proxy for the year when immigrant scientists arrived in the United States, we exploit data on a scientist’s year of naturalization in the MoS. Elias Klein, for example, was born in Wilno, Poland (today’s Vilnius, Lithuania) in 1890, and became a US citizen in 1912. By 1956, 6,118 foreign-born scientists had become naturalized citizens of the United States. Data on the year of the naturalization is available for 2,775 of these scientists, including 745 ESE-born scientists and 1,296 WNE-born.

2.1.3. Measuring Entry into Science

To determine the year in which a scientist entered US science, we exploit detailed information each scientist's education and career history from the MoS. Using these data, we create two alternative measures to determine a person's entry into US science based on 1) the year in which they received their first university degree and 2) the start year of their first employment in the United States.

The country where a scientist received his university education is known for nearly all (99.4 percent) of 82,094 American scientists in 1956. The Polish-born Elija Klein, for example, received his undergraduate degree from "Valparaiso," (the Valparaiso University in Valparaiso, Indiana) in 1911. Using education to pinpoint the year of Klein's scientific activity in the United States, we determine Klein's year of entry into US science to be 1911. On average, American scientists attended 2.9 educational institutions; this yields a total of 238,895 entries on "education" and 7,175 unique institutions. Using publicly available data, we are able to assign 85.5 percent of these institutions to a country, allowing us to determine the country where a scientist received his education for 99.4 percent of scientists.

Nearly half all ESE-born American scientists, 980 of 2,066 ESE-born American scientists, or 47.4 percent earned their undergraduate degree in the United States. Another 65 earned their undergraduate degree in Canada (3.1 percent). Almost two thirds of all ESE-born US scientists (1,310, or 63.4 percent) earned a graduate degree (PhD or Master's degree) at a US institution and another 85 in Canada (4.1 percent). By comparison, scientists born in Western and Northern Europe were less likely come to the United States for their education: 1,376 of 4,029 WNE-born scientists received their undergraduate degree in the United States (34.1 percent), and another 262 earned in Canada (6.5 percent). 2,111 received their PhD or Masters in the United States (52.4 percent) and 259 in Canada (6.4 percent).

Using additional data on the start year of a scientist's first US employment, we are able to determine the year of entry into US science for 99.7 percent of our scientists. Biographies in the MoS (1956) include a total of 465,918 entries to describe scientists employment. On average a scientists in the MoS held 5.7 unique jobs; these jobs sum to a total of 117,606 institution of employment, including universities, firms, and public sector institutions, such as the US Geological Survey. To determine the country in which these 117,606 institutions were located, we develop a three-step algorithm. First, we create a cross-walk that matches universities, as well as cities and states to countries; this cross-walk implements the manual matching that we developed to identify scientists' countries of birth and university education.[21] Second, after cleaning the strings and punctuation, we match the string of words in the career institution to strings of words already matched to countries in the cross-file. 319,477 institutions are matched in this step (68.6 percent). The third and final step, revisits career institutions that remain unmatched after the first two steps, and matches individual words within the string of the career institution with birth places and educational institutions. For example, the string "Harvard Physics" is matched to Harvard and therefore to the United States. Another 84,349 institutions (18.1 percent of all career institutions) can be assigned to a unique country in this final step. Through this three-step algorithm, we are able to assign 403,826 of the 465,918 institutions in our data (86.7 percent) to a unique country.

2.1.4. Research Topics

The 1921 and 1956 editions of the MoS include detailed information about each scientist's discipline and about their specific topics. We use information scientist's discipline and topics to assign each scientist to a unique research field.

Volkoff, for example, lists "physics" as his discipline. Definitions of disciplines range from the extremely broad (such as "chemistry" or "physics") to very specific (such as "crystallographic chemistry" and "mathematical electrophysics"). All but 29 of 91,635 scientists in our data list their discipline; 82.3 percent of scientists list only one discipline.

Entries on research topics are much more detailed and informative. Volkoff describes his topics as "theoretical nuclear physics; neutron diffusion; nuclear magnetic and quadrupole resonance." These data are available for 96.8 percent of the 91,635 scientists. The median scientist lists 3 topics in addition to her discipline, with a range from 1 to 30 topics.

Our analysis of patent data focuses on the physical sciences, a field in which a large share of innovations were patented during this time (Moser 2012a), making patents a good proxy for innovations. Information on the scientist's discipline is available for 41,086 American sciences in the physical sciences in 1956 (99.98 percent), information on topics is available for 39,865 American scientists in the physical sciences (97.0 percent).[22]

2.2. US Patents, 1900-1970

[21] If a career institution was manually matched to more than one country, the cross-file assigns that city to the country that is the most frequent match. For example, a research institute that includes the word "Moscow" is assigned to Russia and not to Moscow, Indiana in the United States.

[22] Only 2 scientists have no information on both disciplines and topics, and therefore were dropped from the analysis. Among 1,230 scientists with missing information on topics, 1,143 born were in US or Canada (3.1 percent of native-born scientists) and 87 born elsewhere (2.2 percent of the foreign-born scientists). Among the foreign-born scientists, 14 born in ES Europe (1.4 percent) and 44 born in WN Europe (2.0 percent).

Changes in inventive output are measured by changes in the number of successful US patent applications per year and field. Patent data include 3,082,720 patents issued by the United States Patent Office (USPTO) between 1900 and 1970. To construct these data, we collect patent identification numbers, the full name of inventors, as well as the application and the publication (issue) date for each patent from Google Patents. To assign patents to USPTO classes and subclasses we merge Google patents with the USPTO Historical Masterfile.[23]

To measure invention as close to their creation possible, we use application (rather than issue) dates to define the timing of invention. The application date marks the date when the inventor signs his name on the patent application. This is much closer to the actual date of the invention than the "issue" or publication date of a patent, which is typically delayed by several years. For example, Thomas Edison's (1847-1931) last patent, for a "holder for article to be electroplated" (US patent 1,908,830) was granted on May 16, 1933, two years after Edison's death, but filed on July 6, 1923. Application dates are available for 2,806,038 in 2,909,518 patents issued between 1900 and 1970 (96 percent). For patents with missing application date, we proxy the application date by subtracting the median lag between application and publication dates (2.4 years) from the publication date.[24]

2.3. Matching Scientists with Patents

To match scientists with patents, we start from a standard Levenshtein (1966) measure (allowing for one different letter between the name of a scientist and the name of the inventor on a patent),[25] and then use data on the scientist's age to filter out false positives. First, we exclude any patents whose execution date falls before the birth of the inventor or after their 80th birthday.[26] In our data, 70.3 percent of all potential matches occur between the ages of 0 and 80, leaving 2,443,476 successful patent applications by 82,094 scientists.

---

[23] Available at https://www.uspto.gov/learning-and-resources/electronic-data-products/historical-patent-data-files, accessed October 7, 2019.

[24] Citations from later patents contain useful information about the quality of patents. For example, detailed field trial data on hybrid corn show that citations are a good predictor for tangible improvements in yields and other characteristics of new patented varieties (Moser, Ohmstedt, and Rhode 2018). Yet, we choose not to use citations as a quality measure in this paper, because citations are not systematically recorded on patent documents until 1947, which means that citations-based measures are extremely noisy for the period that we study.

[25] As a measure of approximate string similarity, the Levenshtein measure matches the string of the scientist's name with the string of the inventors' name in a patent document. The algorithm's key component is that it allows for a certain number of "errors" in the matching. These errors define the "distance" between the matches. In our application, we allow the distance to be one letter. See Moser, San, and Stevens (2019) for a detailed description of the matching process, as well as links to python codes for data matching and cleaning.

[26] This is not to say that scientists cannot patent after the age of 80, but even the most successful inventors, like Edison, slow down after 70. Edison's last patent (issued two years after his death in 1933), lists an application date in 1923, when the inventor was 76 years old. Edison was productive for an exceptionally long time. In total he held 1,093 patents, most of them with application dates between 1880 and 1890 (Thomas A. Edison papers, Rutgers, available at https://edison.rutgers.edu/patents.htm, accessed May 27, 2019).

Next, we use patents that the inventor would have filed between the ages of 0 and 18 as a proxy for false positives. While there is no age restriction on patents, applications by children are exceptional and inventors apply for a very small number of patents before they turn 20 (see Figure A3).[27] We use these patterns to eliminate matches that are likely to be false positives. For example, James Leroy Anderson, a theoretical physicist from the University of Maryland is matched with a patent for a "Torch Cutting Machine" (patent number 2031583) by James L. Anderson of the Air Reduction Company in 1931. Born in 1926, James Leroy Anderson would have had to apply for this patent at age 5, and we assume that it is a false positive match. Under the assumption that false positive matches are distributed uniformly across the age of inventors, we can use patent applications by children as a measure to estimate the rate of false positive (type I) errors in our matching. (Appendix Figure A6 illustrates the calculation of the error rate and our assumption of a uniform error).

$$Error\ Rate = \frac{Sum(Error_{18-80})}{Sum(Correct_{18-80} + Error_{18-80})} = \frac{Mean(patents_{0-18})}{Mean(patents_{18-80})} \quad (1)$$

A naïve Levenshtein matching yields a type I error rate of 79.6 percent across all disciplines, suggesting that nearly four in five "matches" are false positive. Notably, the error rate is much lower in the physical sciences (73 percent) than in the biological and social sciences (with 88 and 89 percent, respectively, Table A1). This is consistent with historical research which suggests that the share of innovations that are patented varies strongly across industries, with high patenting rates for mechanical inventions and chemicals in this period (Moser 2012a). By comparison, inventions that scientists made in the biological or social sciences would not have been patentable at the time.

To reduce the rate of false positives, we first match scientists with patents using information on the middle name and middle initial. Specifically, we count a scientist-patent pair as a "middle name match" if two conditions are met: First, the MoS and the patent must list the same number of names (e.g., three names including a middle name vs two names including no middle names). By this rule, "Robert Burnett King" and "Robert King" are no middle name match. The second condition for a middle name match is that the scientist and the patentee have either the same full middle name or the same first initial. For example, "Earl Manning" - "Earl Manning" and "Aarons W. Melvin" - "Aarons Wolf Melvin" are middle name matches. However, "Robert A. Lester"- "Robert Lee Lester", and "Arthur Dwight Smith"- "Arthur Dean

[27] The middle-school inventor Marissa Streng, for example, was invited to speak on the Tonight Show with Jimmy Fallon after she patented a dog dryer (USPTO 8371246, https://www.uspto.gov/kids/inventors-kids.html, accessed May 27, 2019).

Smith" are no middle name match. Adding these rules for matching middle names, the rate of false positive errors declines from 73.4 to 16.4 percent in the physical sciences. Notably, the error rate for the biological and social sciences stays high at 66.1 and 79.6 percent, respectively (Table A1 and Figure A4). To further reduce the rate of false positives, we exclude the top quintile of the most common names, like John or James Smith.[28] Excluding common names (in addition to matching on middle names) further reduces error rate for the full sample from 79.6 to 63.2 percent.

Controlling for middle names and dropping the top quintile of frequent names reduces the error rate to just above 5 percent for the physical sciences (Table A1 and Figure A5). Error rates for the biological and social sciences remain high at 32 and 63 percent, respectively (Table A1 and Figure A3), consistent with substantial differences in the propensity to patents (Moser 2012). Most advances in the biological sciences were not patentable until the 1980s (when the USPTO granted the first patent for oil-slick eating bacteria). In the social and psychological sciences, scientific advances have not been patentable until recently.[29]

Focusing on the physical sciences, we are able to match 107,376 successful patent applications between 1910 and 1956 with 12,590 unique scientists, including 387 ESE scientists and 821 WNE scientists.

3. Effects on Entry into US Science

Proponents of the quotas, like President Coolidge, aimed to clear the United States from "diseases of ignorance" by restricting the inflow of Eastern and Southern Europeans. In this section, we examine whether the quotas had the opposite effect by discouraging entry into US science. While it is impossible to say with certainty how many ESE-born scientists would have entered US Science without the quota acts, comparisons with scientists from Western and Northern Europe (WNE) allow us to estimate counterfactual immigration flows. WNE-born immigrants were attracted by the same labor markets as the ESE-born, and they faced comparable costs of trans-Atlantic migration. Unlike ESE-born immigrants, however, WNE immigrants were not targeted by the quotas.

---

[28] The three most common surnames in the United States are Smith, Johnson, and Williams (with a share of 0.98, 0.76, and 0.63 percent of the surnames, respectively, in the US Census of 2000). The three most common first names (including names for both men and women) are James, John, and Robert, with 3.15, 3.14, and 2.96 percent of first names, respectively, in the 1880-2013 Social Security Administration data. To calculate the frequency of a scientist's name we multiply the probability for her first name by the probability of her last name. Based on these calculations, the three most common names for male scientists are James Smith, John Smith, and Robert Smith.

[29] Surveys of research laboratories, such as Cohen, Nelson and Walsh (2000) document enormous differences in firms' reliance on patents across industries. In these surveys, chemistry is typically the most "patent-friendly" industry. Moser (2012) uses exhibition data on innovations with and without patents between 1851 and 1915 to estimate variation in the share of innovations that are patented across industries and over time, and shows that this period saw a major shift towards patenting for innovations in chemicals.

3.1. Nearly 1,200 Missing Scientists

Naturalization records reveal a sharp decline in the arrival of ESE-born scientists in the United States after the quotas. Before the quotas, 18 ESE-born US scientists arrive per year between 1920 and 1924. After the quotas, arrival decline by half, to 9 per year between 1925 and 1930 (Appendix Figure A7).[30] While arrivals of ESE scientists declined after 1924, arrivals of WNE scientists increased by 22 percent, from 17 per year between 1920 and 1924 to 21 per year between 1925 and 1930.

Using naturalization data to estimate the number of missing scientists indicates a loss of 1,170 ESE-born US scientists (Table 2). At an annual basis, this number implies a loss of 38 per year, equivalent to eliminating one major physics department each year. The key assumption of this estimate is that, in the absence of the quotas, the ratio of ESE-born and WNE-born scientists arriving in the United States would have been unchanged. For years between 1910 to 1924 this ratio was, in fact, relatively stable with an average of 488/554 for 1910-1924. After the quotas, the total number of WNE scientists is 1,330 for years 1925-1955. If the ratio of ESE/WNE scientists had been constant, the number of ESE scientists in 1925-1955 would have been 488/554 * 2,838= 2,500. Yet, the actual number of 1,330 ESE-born scientists arriving in the United States was only 1,330 between 1925 and 1955, which implies a loss of 1,170 missing scientists.

4. Effects on Invention: Empirical Strategy

To estimate the causal effects of the quotas on US invention, we compare changes in patenting after the quotas in the fields of ESE-born scientists with other fields. Fields of ESE-born scientists are defined by the research topics of scientists *before* the quotas, which we collect from the MoS in 1921. Under the assumption that changes in patenting after 1924 would have been comparable in ESE and other fields of US science, this simple difference-in-difference test estimates the causal effects of the quotas on patenting.

4.1. Defining Research Fields Using K-Means Clustering

Detailed data on the precise research topics of each scientist create a unique opportunity to assign scientists to fields. This approach offers important advantages compare with using disciplines, which are available observable from the MoS. Volkoff, for example, lists his discipline as "physics," but another 4,882 scientists who study extremely dissimilar topics also list physics. "Chemistry" is an even larger and more varied discipline, with 7,091 scientists

[30] Data on the year of a scientist's naturalization are available for 2,775 ESE-born American scientists. Under US laws, immigrants are eligible for naturalization five years after their arrival in the United States. Using this rule, we estimate a scientist's year of arrival by subtracting five years from their year of naturalization.

(Appendix Figure A8). At the opposite extreme of the size distribution, 384 of 781 disciplines within the physical sciences include just one single scientist, and another 119 include only two.[31]

To assign each scientist to a meaningful and unique field, we apply k-means clustering to a "bag of words" that includes both the discipline, as well as unique data on the *topics* of each scientists. For example, Professor Volkoff's entry lists the following topics:

> "Theoretical nuclear physics; neutron diffusion; nuclear magnetic and quadrupole resonance."

K-means clustering allows us to use this information to match each scientist with other scientists who work on related research. *K-means* is one of the most basic and intuitive unsupervised machine learning classification algorithms.[32] A "cluster" (in our setting a research field) refers to a collection of data points (here scientists) that are grouped together because they include similar observable characteristics (here research topics). To group scientists into clusters, the k-means algorithm (implemented through python's *scikit-learn* library) assigns researchers to one of k centroids by minimizing the distance between the observations and the centroid. The number of clusters k is a choice variable; we set k=100 for simplicity, and report robustness checks with alternative choices of k.

To measure distance between the research topics of scientists, we represent each scientist's research topics in the Euclidian space. To do so, we first concatenate all fields and topics of a scientist into a list of words ("document"), removing punctuation and stop words. Then, our "corpus" of documents represented by a matrix with one row per document and one column per word occurring in the corpus, where entries counting occurrences of words in each document. Because frequent words like "theory" or "research" carry less information than rarer words like "neutron" or "polymer", we made a transformation to this matrix that assigns less weight on frequent words. Specifically, an entry in the transformed matrix is $tf\_idf(w,d) = tf(w,d) \times idf(w)$, where $tf(w,d)$ is the frequency of word w in document d, $idf(w) = \log[1 + n/1 + df(w)] + 1$, n is the number of documents, and $df(w)$ is the number of documents that contain word w (Baeza-Yates and Ribeiro, 2011).

In a process that is similar to OLS, k-means algorithm starts with a group of randomly selected centroids, and then performs iterative calculations to minimize the mean of the sum of the squared distances between the centroids and the data. The process stops when further changes

[31] Another issue with using disciplines is that 9.2 percent of scientists report two or more disciplines (2,322 of 41,086 in the physical sciences, 7,558 of 82,067 overall). To create Appendix Figure A8 of the original MOS discipline variable, we use the first discipline for each scientist. Refining fields with k-mean clustering allows us to use data for all disciplines, as part of the bag of words that describe a scientist's research topics.
[32] Unsupervised classification algorithms make inferences from datasets about the best classification of the data points without referring to known, or labelled, classes.

to the location of the centroids yields no further decline in the minimized sum of squared distances. The *'means'* in *k*-means refers to averaging the data; that is, finding the centroid that minimizes the average distance between the data points and the centroid.

Compared with other methods of text analysis, a key benefit of k-means is its stability to the (random) choice of the original centroids. K-means also delivers training results relatively quickly, even for large data sets. A potential disadvantage is that clusters are assumed to be spherical and evenly sized. In our data clusters are nicely distributed, which suggests that this assumption is not a problem (Figure 2). The median cluster (number 58, "Vitamin") includes 303 scientists, the average cluster includes 410.9 scientists, with a standard deviation of 514.7.[33]

To check whether the content of the cluster assignments is sensible, we use Google to "name" our clusters and check whether the cluster assignments make intuitive sense. To perform this check, we search Google for the 10 most frequent words in each cluster and name the cluster with the first result of that search. Next, we pick clusters that are relatively easy to understand to check our assignment. Cluster 59, for example, is named "aircraft;" it includes words that seem to be sensible research topics in that field: aeronautical, aircraft, engineering, structures, design, control, flight, research, stability, guided." Volkoff's research, from our example above falls into cluster 39, which includes the words "nuclear, physics, energy, spectroscopy, cosmic, rays, scattering, reactor, reactions, neutron," and receives the name "neutron radiation," which the Oxford Living Dictionary defines as "Neutrons released from the nucleus during interactions such as nuclear fission or fusion."[34]

Compared with the disciplines that are directly listed in the MoS, k-means clustering is better able to capture the essence of a scientists' research topics. To illustrate this point, consider the examples of Caesar Fragola and Elder de Turk. In 1956, Fragola worked at Sperry Gyroscope Corporation in Long Island, NY. His field is engineering, and he lists the following topics: "aircraft instrumentation engineering; development of aircraft flight and navigation instruments; individual components and complete system components for stabilized remotely located aircraft compasses and flight directors." The second scientist, de Turk worked in Naval Air Test Center. De Turk's discipline is physics, and he lists his topics as "design and development at aircraft instruments; test of gravity meters; test, development and evaluation of aircraft armament systems." The original classification by discipline would have missed the connection between these two fields, while the k-means algorithm connects the two scientists to the substantive field of "aircraft."

[33] A residual cluster (number 25) includes 4,881 scientists. The top ten words in this cluster are "chemistry", "organic", "geology", "engineering", "analysis", "development", "research", "methods", "oil", and "chemical." We include the residual cluster in the main specifications, and exclude it in robustness checks. Excluding the residual cluster, the average cluster includes 366.5 scientists, with a standard deviation of 260.8.

[34] See Appendix Table A2 for these two clusters, as well as eight other examples of typical clusters.

Compared with the USPTO classification, the main advantage of using the scientists research topics to define fields is that we can identify scientists who work in the fields of ESE-born scientists, even if they do not patent. This is of particular importance when we examine flows of scientists in fields in which innovations are rarely patented (like in the biological or the social sciences until fairly recently). K-means clustering allows us to examine changes in the number of scientists per field across all fields, irrespective of their propensity to patent.

4.2. ESE versus Other Fields

ESE fields are research fields that include at least 1 ESE-born American scientist in 1921. For instance, ESE-born scientists account for 16.7 percent of American scientists in 1921 in field 50, "Fluid dynamics." By 1956, the share of ESE scientists in "fluid dynamics is 8.0. Volkoff's field 39 ("Neutron radiation") has zero ESE scientists in 1921 and is therefore assigned to the control. By 1956 18 ESE scientists are active in "neutron radiation," 2.4 percent of all scientists in 1956. The average field includes 1.64 ESE-born scientists in 1921, with a standard deviation of 8.44. The median field in 1921 includes no ESE scientists (Figure A9).

Comparisons of research fields in 1921 and 1956 indicate a strong persistence in the relative size of fields, and some persistence between fields that were ESE-fields in 1921 and fields that are ESE fields in 1956. The correlation between the counts of ESE-born scientists per field in 1921 and 1956 is 0.89 (significant at 1 percent) and 0.50 for logs (p-value < 0.01, Appendix Figure A10). The correlation between the share of ESE scientists in 1921 and 1956 is 0.30 (p-value < 0.01, Appendix Figure A11).

In the main specifications we exclude five "new" fields have no scientists in 1921: "Solid-state chemistry" (field 27), "Electronic engineering" (field 53), "Aircraft" (field 59), "Polymer" (field 74) and "Nylon" (field 97). We include these new fields in robustness checks and show that results are robust to including or excluding them.

To identify the causal effects of the quotas on American invention, we compare changes in patenting after 1924 by American scientists in the pre-quota research fields of ESE-born American scientists with changes in patenting in other fields in which no ESE scientists were active in 1921. Under the assumption that changes in patenting would have been comparable in these fields without the quotas, this comparison identifies the causal effects of the quotas on American invention.

To investigate this identifying assumption, we first compare the observable characteristics of research fields with and without ESE scientists in 1921. Fields with and without ESE scientists are comparable in terms of the number of scientists per field (Figure 2). They are also comparable in terms of the demographic characteristics of pre-quota scientists: the average age of the scientists (44.7 and 44.4 in ESE and other fields, respectively) and the share of female scientists (1.1 percent and 1.2 percent, respectively). Likewise, the share of pre-quota WNE born scientists is similar (5.4 percent in ESE fields compared to 5.1 percent in other fields). Finally, the share of "star" scientists (scientists marked as leading scientists by other scientists in their field[35]) is also comparable (11.5 percent and 10.4 in ESE and other fields, respectively). The difference is not statistically significant in each of these variables. The most significant difference between ESE and other fields lies in the share of scientists who were born in Eastern and Southern Europe (Table 3). Below we present additional tests of the identification strategy, including time-varying effects, alternative specifications of pre-trends, and Placebo tests for Canada, which did not adopt the quotas.

5. Effects on Invention

Data on annual patents by US scientists reveal a clear decline in ESE fields relative to other fields after the quotas (Figure 3). Before the quotas, American scientists patented more in ESE fields compared with other fields. Between 1910 and 1924, American scientists filed 256 successful patent applications per year in the fields of ESE-born scientist compared with 142 in other fields. After the quotas invention in other fields first overtakes invention in ESE fields in 1929. Patenting in ESE fields remains below other fields through the 1960s..

5.1. *Effects on Invention by American Scientists*

To investigate the causal effects of the quotas on US invention, we estimate OLS regressions:

$$\ln(y_{it}) = \beta \cdot ESE_i \cdot post_t + \gamma_i + \delta_t + \epsilon_{it} \quad (2)$$

where the dependent variable $\ln(y_{it})$ represents the natural logarithm of the number of US patents by American scientists in field i and year t.[36] The variable ESE_i indicates fields in which ESE-born scientists pursued research before the quotas. The indicator $post_t$ denotes years after 1924. Field fixed effects γ_i control for differences in patenting across fields that stay constant over time. For example, scientists in a theoretical field, such as the "calculus of variations" (cluster 89), patent fairly little both before and after the quotas, with a total of 0.07 patents per

---

[35] The first editor of the MoS, J. McKeen Cattell, constructed this measure to capture the perception of his peers: "In each of the twelve principal sciences the names were arranged in the order of merit by ten leading students of the science. The average positions and the probable errors were calculated, so that in each science the order of merit was determined together with its validity. The names were then combined in one list by interpolation, the numbers in each science being taken approximately proportional to the total number of workers in that science."

[36] About one fifth of all field-year pairs (21.7 percent) have zero patents. To include them in the log regressions, we add 0.01 to all observations. Regressions with smaller numbers (0.001 and 0.0001) increase the size of the estimated effects. Below we report robustness checks with Poisson, probit, logit, and other alternatives to the log regressions.

scientist, the lowest number of patents across all fields, while scientists in an applied field, such as "radio waves" (cluster 1) have a high number of patents (with 10.3 patents per scientist, the largest number of patents for any field (Appendix Figure A12). Year fixed effects δ_t control for variation in patenting over time that is shared across fields, e.g., as a result of a reduction in research output or increased secrecy during World War II.[37]

The identifying assumption of equation (2) is that, in the absence of the quotas, changes in patenting would have been comparable across ESE and other fields, controlling for year and field fixed effects. If it is satisfied, the coefficient β estimates the causal effects of the quotas on invention by American scientists. (To investigate this assumption, we have compared observable characteristics for ESE and other fields in Table 1 above. We also estimate alternative specifications with controls for pre-trends as well as time-varying effects below.)

OLS estimates of equation (2) imply a substantial decline in invention by American scientists in ESE fields. After 1924, American scientists produced 63 percent fewer additional inventions in the pre-quota fields of ESE scientists compared with other fields (with an estimate of -0.905 for β on $ESE \times Post,$ significant at the 1 percent level, Table 4, column 1).[38]

This decline in invention is robust to controlling for field-specific pre-trends, as well as to excluding the largest fields, excluding the fields with the largest share of ESE scientists, or including newly developing fields. Controlling for field-specific pre-trends American scientists produced 62 percent fewer additional patents after 1924 in the pre-quota fields of ESE scientists (Table 4, columns 2, significant at 5 percent). The decline in invention is also is robust to excluding the five largest fields (Table 4 column 3, with a percentage change), to dropping fields with the highest share of ESE-born scientists (column 5, with a percentage change of 68). Finally, the estimated decline is robust to including newly developing fields that did not have any scientists in 1921 (column 7, with a percentage change of 69).

5.2. Time-varying Estimates, 1910-1970

To examine whether the decline in US invention in ESE fields may have preceded the quotas, and to investigate the timing of the decline in patenting after the quotas, we estimate

$$\ln(y_{it}) = \beta_t ESE_i + \gamma_i + \delta_t + \epsilon_{it} \quad (3)$$

[37] Gross (2019), for example, shows that government orders to keep secret over 11,000 patent application during World War II were effective in keeping sensitive technologies out of public view.

[38] All regression tables report percentage changes along with coefficients. For example, the percentage change for Table 4 column (1) is calculated as 1-exp(-1.142)=1-0.32=0.68.

where β_t is a vector of time-varying estimates for the quotas' effect on American science. 1918-1920 is the excluded period; all other variables are as defined above.

Time-varying estimates are close to zero before the quotas and yield no evidence for a pre-existing differential trend (Figure 4). After the quotas, time-varying estimates decline to imply 66 percent fewer additional patents in the pre-quota fields of ESE-born scientists for 1933-1935. Estimates remain consistently large and negative between 69 and 83 percent throughout the 1960s, with an estimated decline of 79 percent in 1969-70. These results suggest that the quotas may have led to a permanent reduction in US invention in the fields of ESE scientists.

5.3. Robustness to Alternative Matching Rules and Definitions of Fields

All results are robust to alternative matching rules, even though these alternatives introduce some noise. Re-estimating the baseline specification with the full data set, including the most common names, yields an estimated 69 percent decline in patenting (Table A3, column 2, significant at 1 percent). Allowing for scientists and patentees to have different middle name increases the estimate to 73 percent (column 3, significant at 1 percent). Including common names and allowing for different middle names reduces the estimate to 57 percent (column 4, significant at 1 percent, compared with 63 percent in the baseline, column 1).

Importantly, our results also do not depend on the choice of 100 clusters (k=100). Re-estimating our analysis with 50 fields implies a 61 percent decline in invention (Table A4, column 1), 75 fields yield a 60 percent decline (column 2), and 125 fields yield a 63 percent decline (column 4). All very close to the estimated decline of 63 percent decline in our preferred specification (Table A4, column 3).

5.4. Robustness to Poisson, Negative Binomial, and other Econometric Models

21.7 percent of the field-year pairs in our data include zero patents. In the log specifications, we preserve these observations by adding a tiny number (0.01). We have also re-estimated the regressions with other small numbers, such as 0.1, 0.001, or 0.0001; all of these specifications confirm a decline in invention (Table A5, columns 3-6). The decline in invention is robust to alternative count data models. QML Poisson estimates confirm the large decline in invention, with a 53 percent decline in invention (Table A5, column 1, significant at 1 percent). Negative binomial regressions imply a 60 percent (Table A5, column 2, significant at 1 percent).

5.5. Intensity: Invention Declines more in Fields with Higher Shares of ESE Scientists

Intensity regressions examine whether fields with a larger share of ESE-born scientists before the quotas experienced a larger decline in invention after 1924. Specifically, we estimate

$$\ln(y_{it}) = \beta \cdot \%ESE_i post_t + \gamma_i + \delta_t + \epsilon_{it} \quad (4)$$

where the explanatory variable $\%ESE_i$ represents the share of ESE-born scientists in field *i* in the 1921 edition of the *Men of Science*, the last year before the quotas.

OLS estimates confirm that research fields that were more exposed to ethnicity-based restrictions on immigration experienced a larger decline in patenting after 1924. Fields that had a 10 percent higher share of ESE-born scientists in 1921 experienced a 70 percent decline in patenting after the quotas (Table 5, column 5, significant at 5 percent).

6. MECHANISMS

How did the national origins quotas reduce patenting in the United States? To investigate this question, we begin by decomposing the decline in invention into changes at the intensive and extensive margin. We then examine changes in invention by native-born scientists (who may have benefitted from reduced competition with immigrants). We also examine selection into research fields and the aging of scientists in ESE fields as potential mechanisms.

6.1. Changes at the Extensive and Intensive Margin of Invention

As a first test of the process by which the quotas reduced US invention, we decompose the overall effects into changes at the intensive and extensive margin. First, we estimate the baseline log-level OLS model excluding field-year pairs with zero patents. This specification ignores changes at the extensive margin and instead estimates only the effect of the quotas on the intensive margin (more or less innovative activity per field). OLS estimates indicate a 45-percent decline in invention in ESE fields at the intensive margin (Table 6, column 4, significant at 1 percent). Next, we examine whether the quotas reduced the number of ESE fields in which American scientists were active inventors. Specifically, we estimate extensive margin regressions in which the outcome variable equals one if field *i* has at least one patent in year *t*. OLS, probit, and logit models all yield negative and statistically significant estimates indicating a 10 to 16 percent decline in the number of research-active ESE fields (Table 6, columns 1-3, significant at 1 and 5 percent).[39]

Next, we decompose the change in invention *by scientists* into two part: the change in the number of scientists per year and field and the change in the number of patents *per scientist.* To

[39] Back-of-the-envelope calculations that combine estimates of the extensive and intensive margins, imply a 55 to 61 percent decline in American invention in the pre-quota fields of ESE scientists after the quotas. Adding the 45 percent decline in the intensive margin (column 3) and the 16 percent decline from the extensive margin (column 4), yields a total decline of 61 percent, just slightly less than the baseline estimate.

perform these tests we use detailed biographical data on scientists' education and career histories to determine when a scientist was professionally active. These data allow us to count active scientists per field and year. OLS estimates of these data reveal a 40-percent decline in the number of active scientists for ESE fields relative to other fields (Table 6, column 6, significant at 1 percent). Analogous regressions for patents per scientists as the outcome variable indicate a 33-percent in patents per scientist in ESE fields compared with other fields (Table 6, column 7, significant at 5 percent). Combining the effects at the extensive and intensive margin indicates that, the total number of patents by US scientists declined in 60 percent in ESE fields relative to other fields (Table 6, column 8, significant at 1 percent).[40]

6.2. A Decline in Invention by US-born Scientists

Ex ante, the effects of immigration on native-born scientists may be ambiguous, if native-born scientists compete with immigrants for jobs and opportunities to patent. Borjas and Doran (2012), for example, document that US mathematicians published less and in worse journals once they had to compete with Russian immigrant scientists after 1990. Alternatively, native-born scientists have benefitted from exposure to new types of knowledge and methods that immigrants brought to the United States. Consistent with such positive spillover effects, Moser, Voena, and Waldinger (2014) show that US inventors became more productive in the fields of German-Jewish émigrés after the Nazis expelled Jews from German Universities in 1933.[41] If the costs of competition out-weighted the benefits of knowledge spillovers, native-born US scientists should have patented *more* after the quotas in fields of ESE-born scientists.

Estimates for native-born American scientists reveal a substantial decline in American invention in response to the quotas, at levels that are only slightly below the baseline estimates. After the quotas restrict the inflow of ESE-born scientists to the United States, native-born US scientists produce 62 percent fewer inventions in the fields of ESE-born scientists compared with other fields (Table 7, column 1, significant at 5 percent). Effects on native-born American scientists are robust to excluding the largest fields, as well as to excluding fields with the highest share of ESE scientists, and including new clusters (Table 7, columns 2-4).

To further examine the mechanisms by which the quotas reduced US invention, we decompose the change in invention by US-born scientists into the change in the *number of scientists* who are active in ESE fields and the change in the number of patents *per scientist.* OLS estimates show that invention declined at both margins. After the quotas, 40 percent fewer

[40] Note that this scientist-level analysis includes data only until 1955, the last year in which we observe scientists in the MoS (1956). The analysis is also limited to 99.5 percent of scientists in the physical sciences for whom the date of entry into science is known (through their education and employment histories.)
[41] See Figure A18 for similar results using the empirical strategy of this study.

scientists worked in ESE fields compared with other fields (Table 7, columns 5, significant at 1 percent level). Moreover, the number of patents *per scientist* declined by 31 percent (Table 7, columns 6, significant at 1 percent level).

Taken together, these results show that the quotas hurt rather than helped the productivity of native-born American scientists. Thus, our research qualifies earlier findings by Borjas and Doran (2012) who had shown that the inflow of Soviet mathematicians after the collapse of the Soviet Union lowered the productivity of American scientists, measured by journal publications. Compared with publications, patents are not subject to capacity constraints, allowing the benefits from knowledge spillovers to outweigh the costs of competition.

6.3. Changes in Professional Networks of Co-Authorships

A key mechanism for spillovers are knowledge flows through collaborations and mentorships. The experience of the Hungarian mathematician Paul Erdős illustrates how such collaborations were impacted by the quotas. After the Anschluss of Austria in 1938, Erdős came to Princeton for a six-month fellowship, where he was soon dismissed as "uncouth and unconventional." Erdős then moved to other US universities, writing most of his 1,500 papers with co-authors.[42] In 1954, the US Citizenship and Immigration Services denied Erdős a re-entry visa, citing his Hungarian citizenship. Erdős left his position at Notre Dame and returned to Hungary, repeatedly, but unsuccessfully requesting reconsideration. When his request was finally granted in 1963, Erdős resumed to visit American universities to teach, but never again made the United States his permanent home.

Data that we collected on the home countries of Erdős co-authors indicate that Erdős' influential network of co-authors shifted away from the United States after he was denied entry (Appendix Figure A15). Between 1935 and 1954, 62 percent of Erdős co-authors were based in the United States. After 1954, this share declined to 32 percent. It only recovered after 1963, when Erdős was allowed to enter the United States again. When Erdős died in 1966, a *New York Times* obituary explained that he "founded the field of discrete mathematics, which is the foundation of computer science." Our analysis of Erdős' co-authors indicates that much more of this knowledge would have stayed in the United States, without the quota system.

[42] Erdős' collaborations are particularly well documented through the Erdös number, which measures the distance of an author to Erdös, in terms of co-authors (Goffman, 1969, p. 791). Erdős' direct coauthors have Erdös number 1. Their co-authors have Erdős number 2, and so on. (If there is no chain of co-authorships connecting someone with Erdös, then that person's Erdös number is infinite.) In mathematics, most winners of the Fields Medal, the Nevanlinna Prize, the Abel Prize, the Wolf Prize in Mathematics, and the Steele Prize for Lifetime Achievement, have low Erdös numbers. In computer science, influential people with low Erdös numbers include Bill Gates (4). In biology, Eugen V Konnin, of the National Center for Biotechnology Information, has an Erdős' number of 2 (https://oakland.edu/enp/erdpaths/ accessed July 31, 2019).

An analysis of co-inventor networks in the MoS (1921 and 1956) shows that the quotas reduced patenting by native-born co-inventors of ESE-born scientists, as well as the co-inventors of co-inventors (Figure 6). Before the quotas, between 1910 and 1924, scientists in the professional network of ESE-born scientists produced a comparable number of patents as did scientists in the network of WNE-born scientists (with 948 patents for ESE and 1,167 for WNE, Figure 6). After the quotas, however, native-born collaborators of ESE-born scientists produced many fewer patents than collaborators of WNE scientists (14,763 and 24,416 between 1925 and 1970, respectively). Ballpark estimates based on the comparison with WNE scientists imply that restrictions on the number of ESE-born scientists costs their US collaborators to forego 5,071 patents (compared with a counterfactual level of 19,834).

6.4. Changes in Entry into ESE Fields

Next, we use data on scientist's employment histories to examine whether the quotas may have discouraged scientists from entering into ESE fields after the quotas. In these tests, we exploit data on a scientist's university degrees and their employment to create two complementary measures for the number of scientists who are active in ESE fields compared with other fields. First, we use the start year of the scientist's first US degree to determine the start year of that scientists work life in the United States. The second measure uses the start year of a scientist's first US degree or job. Using this data we re-estimate equation (2) with the outcome variable $\ln(y_{it})$ as the natural logarithm of the number of scientists in field i active in the US at year t. Estimates for the physical sciences indicate that the quotas led to a 46-47 percent reduction in entry into ESE fields (Table 8, columns 3-4). Since these tests do not require patent data, we can perform them for *all* fields of American science, including the biological and social sciences. These estimates confirm a broad-based decline in entry in ESE fields after the quotas. Using data on a scientist's education and employment, we find that the quotas led to a 23-24 percent reduction in entry into the pre-quota fields of ESE-born scientists (Table 8, columns 1-2).

6.5. Selection into ESE Fields: Placebo Estimates for Canada

A potential alternative explanation for the decline in invention after the quotas is that ESE-born scientists may have selected into fields in which patenting was declining after 1924 even without the quotas. To investigate selection, we estimate placebo regression for Canada scientists. Since Canada did not adopt comparable quotas in 1924, there should be no decline in invention in ESE fields if the decline in invention was in fact due to the quotas. If, however, the decline was due to selection, we should see the same decline in invention in Canada and the United States.

Placebo estimates show that – unlike scientists in the United States – scientists in Canada did not produce fewer patents in ESE fields after the quotas (Table A6). Estimates for time-varying coefficients are close to zero, between positive 32 percent in 1963-65 and negative 32 percent in 1948-50, and they are not statistically significant for any year (Appendix Figure A17). These estimates imply that the observed decline in invention in ESE fields *cannot* be explained by the intrinsic characteristics of the ESE fields.

Triple-differences estimates confirm that American scientists became less productive relative to Canadian scientists in ESE fields after 1924. These estimates compare changes in patenting by Canadian with American scientists after 1924 in ESE fields with other fields:

$$\ln(y_{ict}) = \beta\, ESE_i\, US_c\, post_t + \gamma_{ic} + \delta_{it} + \theta_{ct} + \epsilon_{ict} \quad (5)$$

where y_{ict} measures successful patent applications in field *i* by scientists in country *c* and application year t. The variable US_c equals 1 for scientists who are employed in the United States in 1956 and 0 for scientists who work in Canada. The variables $\gamma_{ic}, \delta_{it}, and\ \theta_{ct}$ denote field-country, field-year and country-year fixed effects, respectively.

Triple-difference estimates indicate that American scientists produced 70 percent fewer patents in the pre-quota fields of ESE scientists after 1924 compared with US scientists (Table 9, column 1). This estimated decline is also robust to excluding the 5 largest fields (Table 9, column 3, with a percentage change of 66), to dropping fields with highest ESE share (column 5, with a percentage change of 74), and to including new fields in the control (column 7, with a percentage change of 75). Controlling for country-field-specific pre-trends yields similar estimates (Table 9, columns 2,4,6 and 8).

Time-varying estimates are close to zero before the quotas, and not statistically significant (Figure 7). The estimated difference between US and Canadian invention become negative after the quotas. In 1933-35, US scientists produce 72 percent fewer additional patents in ESE fields compared with other fields and compared with Canadian scientists (Figure 7, significant at the five percent level). Estimates remain consistently large and negative between 62 and 86 percent throughout 1960s, with an estimate of 83 percent in 1969-70. The timing and intensity of these changes indicate that the quotas moved US invention to an equilibrium of lower productivity in the pre-quota fields of ESE scientists compared to other fields compared to the parallel difference in Canada.

6.6. Effects of an Ageing Work Force

In addition to knowledge spillovers, another potential mechanism for the decline in invention is that the scientific workforce in ESE fields may have aged as the quotas reduced the inflow of younger ESE scientists. Immigrants tend to be younger (e.g., Annelli et al 2019),[43] and our analysis shows that older scientists (above 40) are on average less productive, in terms of patenting (Figure A3, in the section on the patent matching above). Our biographical data also show that, by 1956, ESE-born American scientists were 3 years older than other American scientists (Appendix Table A1). Taken together, these issues suggest that ageing may have been a factor in reducing the creation of new patents in ESE fields.

To investigate this issue, we re-estimate the baseline specification with an additional interaction term for the age profile of ESE scientists.

$$\ln(y_{it}) = \beta_1 \cdot ESE_i \cdot post_t + \beta_2 \cdot ESEAge_i \cdot post_t + \gamma_i + \delta_t + \epsilon_{it} \quad (6)$$

where the variable *ESEAge* represents three alternative measures for the aging of ESE scientists: first, the share of ESE scientists in field *i* who are older than 40 years in 1956 (Table 10, column 1), second the share of ESE scientists who are older than 65 in 1956 (column 2), and third, the average age of ESE scientists by in field *i*. All other variables are as defined in equation (2).

This analysis shows that aging cannot explain the observed decline in patenting in ESE fields after 1924. Estimates with alternative controls for the age of ESE scientists leave the estimated decline in invention between 63 and 65 percent (Table 10 Columns 2-4), only slightly less than the baseline estimate of 68 percent. Controlling for all of the three variables together, leaves the estimate at 64 percent (column 5), very close to the baseline estimate of 68 percent. Taken together these estimates indicate that aging cannot explain the decline in invention in ESE fields, suggesting that the decline in invention is due to reduced knowledge spillovers and other costs of restricting high-skilled immigration.

6.7. Visa Restrictions or Fear of Discrimination?

Crude ethnicity-based visa restrictions may have affected high-skilled scientists along with foreign “hands”, even though US immigration officials had intended to encourage “quality” immigration (Williams 1911, p. 215). But ESE scientists may also have avoided the United States by choice, if they feared discrimination. To investigate whether ESE-born immigrants voluntarily avoided the United States, we examine the migration decisions of immigrants who initially entered the Americas in Canada. Specifically, we examine whether ESE-born scientists who immigrated to Canada after 1924 were more likely to move to the United States compared

[43] In an analysis of the effects of emigration from Italy, Annelli et al (2019) find that, for each 1,000 emigrants, Italy creates 10 fewer young-owned firms. 60 percent of the observed decline in firm creation is generated by the emigration of young Italians.

with ESE-born immigrants who arrived in Canada after 1924 and compared with WNE immigrants.

These data indicate that ESE-born scientists to Canada moved to the United States at higher rates after 1924 (Table A7): 20 of 30 ESE scientists who had immigrated to the Americas via Canada after 1924 had moved to a job in the United States by 1956, up from 3 in 7 ESE immigrants who had arrived in Canada before 1924.[44] By comparison, the share of WNE movers remained stable after 1924: 22 of 41 WNE scientists who had immigrated into Canada before 1924 had moved to the United States by 1956, and 59 of 120 afterwards. These patterns suggest that ESE-born scientists were in fact kept out of the United States by the quotas, rather than a fear of discrimination.[45]

7. Aggregate Effects on Invention in the United States and Abroad

We have shown that the quotas greatly reduced the number of ESE-born scientists in the United States and that they discouraged innovation by American scientists, both immigrants and native-born. To complement these results, we now take a step back to investigate the broader effects of the quotas on American firms and aggregate invention.

7.1. Effects on Firms Employing Immigrants

Today, the impact of immigration quotas on innovative firms is a major point of contention, yet it is impossible to evaluate the long-run effects of such policies on US firms today. Our data on the employment histories of immigrants allows us to shed light on this question by examining the effects of the national origins quotas on the firms that had employed immigrants before the quotas.

The empirical strategy of these tests is analogous to the main regressions: To identify the causal effects of the quota on firms that employ immigrants we compare changes in patenting after 1924 by firms that employed ESE-born scientists before the quotas with changes in patenting for firms that employed other scientists who were not ESE-born. This approach allows

[44] In this test, we currently define a scientist's year of immigration and the destination country by start year of a scientist's first degree in the United States or Canada. Ongoing research refines this variable using information on each scientist's complete employment history, which is available from the MoS.

[45] Although Canada did not implement comparable migration quotas, it was affected by severe anti-Semitism during this period. "Canadian Jews, immigrant and Canadian born alike, confronted widespread discrimination in employment and housing. On the eve of the war in 1939, a 'Report on Anti-Semitic Activities' compiled by the Canadian Jewish Congress noted that employment opportunities for Jews in English-speaking Canada were severely attenuated. Few of the country's teachers and none of its school principals were Jews. Both federal and provincial public services frowned on hiring Jews. Banks, insurance companies, and large industrial commercial interests openly discriminated against Jews […] Jewish doctors, even Canadian trained, rarely received hospital appointments and university and professional schools limited the access of Jewish students and did not hire Jewish faculty (Abela and Troper, 2012, preface).

us to control for unobservable changes in patenting that may have affected patenting by any firm that employed scientists in 1921. Here, the identifying assumption is that, in the absence of the quotas, changes in patenting after 1924 who have been similar for firms that employed ESE-born scientists in 1921 and for firms in which other scientists were active inventors in 1921.[46]

Figure 5 shows a clear decline in invention after the quotas for US firms that had employed ESE-born US scientists before the quotas. Until the quotas were passed, firms that employed ESE-born scientists and firms that employed other scientists produced a comparable number of inventions. Between 1910 and 1924, inventors in ESE firms filed 1,119 successful patent per year compared with 1,205 in other firms. After the quotas, patenting declined for firms that employed immigrants. Between 1925 and 1970, inventors in ESE firms filed 2,449 successful patents per year, less than half the 5,559 patents by other firms. Moreover, the time pattern of these changes suggests that the quotas damaging effects were long-lasting.

7.2. Text of Patent Titles – Aggregate Effects on Invention

As a final test, we use information on the text of patent titles to assign patents to research fields, and examine whether ESE fields experienced an overall decline in patenting. Specifically, we extend the predictions of the k-means model in the main analysis, fitted on the research topics of scientists in 1956 to assign each patent title to a field of science.[47]

These data further corroborate the decline in patenting. Before the quotas, US inventors patented at the same rate in ESE and other fields. Between 1910 and 1924, US inventors filed 1,130 successful patent applications per year in the fields of ESE-born scientist compared with 1,137 in other fields. After the quotas, US inventors patented less in ESE fields with 2,353 patents per year in ESE fields compared with 3,056 in other fields (Figure A14).

7.3. *Gains for Palestine/ Israel*

In section 3.1 above we estimated a loss of 1,170 ESE-born scientists for the United States. Some of these missing scientists moved to Palestine. Migration patterns for Jewish scientists (from the *World Jewish Register* (1955) reveal a dramatic increase in the migration of Jewish scientists to Palestine, around the time of the quotas. Over the 10 years between 1910-1919, only

---

[46] To perform this test, we construct data on the assignee (owner of each patent). For patents after 1926, assignment data are available from Kogan et al.'s (2017) cross-file between firms and patents issues after 1926. We extend these data to include patents issued before 1926 through a matching algorithm. If an assignee string is matched to more than one firm, the cross-file assigns that string to the firm that is the most frequent match. Next, we create a match between MoS scientists and firms.

[47] Between 1910-1970, US inventors filled 2,748,078 successful patents. The corpus of all titles of these patents creates a very large set of words, much larger than the corpus of research topics of our MoS scientists. As a result, 89 percent of the patent titles are allocated to a residual cluster. Our analysis in this section examines 301,206 patents that can be assigned to the other clusters, excluding the residual..

1.4 Jewish scientists who had been born in Eastern or Southern Europe immigrated to Palestine per year. This number increased to 8.8 scientists per year between 1920-1925 (Figure 9). Data on immigration to the US show a moderate increase, from 0.7 scientists per year in 1910-1919 to 2.3 in 1920-1925. The number of Jewish ESE scientists immigrated to Palestine arrived at its peak in 1925, right after the implementation of the 2-percent quota, with 15 scientists immigrated at that year. The number of Jewish ESE scientists immigrated to Palestine (or to Israel after its establishment in 1948) between 1926-1950 remained high at 2.3 scientists per year, compared to 0.7 scientists immigrated to the US (Figure 8). These scientists helped to create the backbone of major universities that built the foundation for Israel's scientific workforce,

8. Conclusions

This paper has examined detailed biographical data on more than 80,000 American scientists to examine the effects of ethnicity-based immigration rules on American science and invention. Migration data indicate that the quotas caused a dramatic decline in the arrival of ESE scientists in the United States. Using comparisons with arrivals from Western and Northern Europe (which were on a comparable trend before the quotas) we estimate that roughly 850 ESE-born scientists were "missing" from the United States scientific workforce as a result of the quota. At an annual level this is equivalent to roughly 30 missing scientists per year, the graduating cohort of PhDs of a major university.

With the support from relief organizations, like the Emergency Committee in Aid of Displaced Foreign Scholars, many ESE-born scientists found refuge in other countries. Yet,

> "measured against the millions who were murdered […] the number saved was pitifully small. During the twelve years of Nazi terror, from 1933 to 1945, the United Kingdom opened its doors to 70,000, and allowed another 125,000 into British-administered Palestine. Other states, with long histories of immigration, did even less. Argentina took 50,000, Brazil 27,000 and Australia 15,000. Some Latin American states, where life-granting visas were bought and sold like any other commodity, admitted but the trickle of Jews who could pay for their salvation." (Abela and Troper 2012)

Beyond this immense human loss, we find that the quotas created major costs for American innovation that persisted through World War II and the Cold War into the 1960s. After the quotas restricted the inflow of ESE-born scientists, American scientists produced more than 60 percent fewer additional patents in the pre-quota fields of ESE-born scientists throughout the 1960s. Equivalent analysis of aggregate levels of invention indicate a 30 percent decline in US invention as a result of the quotas.

Our ongoing research that links scientists with their and their parents' census records, indicates find that many of the US-born scientists in our data were the children of immigrants from ESE countries. For example, the MoS (1956) indicates that Dr. Richard Phillips Feynman of the California Institute of Technology, born in New York, NY on May 11, 1918, was a US-born scientist. Feinman became a member of the National Academy and received the prestigious Einstein Award in 1954. Feynman's father was born in Belarus and moved to the United States when he was 5 years old, Feynman's mother was born in Poland. Had the quotas been established earlier, Feinman's parents would have been kept out of the United States.

References

Abella, Irving and Harold Troper. 2012. *None Is Too Many: Canada and the Jews of Europe, 1933-1948*. Toronto: University of Toronto Press, first edition, 1983.

Abramitzky, Ran, Philipp Ager, Leah Boustan, Elior David Cohen, and Casper Worm Hansen. 2019. "The Effects of Immigration on the Economy: Lessons from the 1920s Border Closure."

Ager, Philipp and Hansen, Casper Worm, 2017. "Closing Heaven's Door: Evidence from the 1920s U.S. Immigration Quota Acts" https://ssrn.com/abstract=3059439

Anelli, Massimo, Gaetano Basso, Giuseppe Ippedico, and Giovanni Peri "Youth Drain, Entrepreneurship and Innovation" NBER Working Paper No. 26055, July 2019

Azoulay, Pierre, Joshua S. Graff Zivin, Jialan Wang. "Superstar Extinction" *The Quarterly Journal of Economics*, Volume 125, Issue 2, May 2010, Pages 549–589.

Baeza-Yates, R. and Ribeiro-Neto, B., 1999. Modern information retrieval (Vol. 463). New York: ACM press.

Borjas, George J. and Kirk B. Doran. "The Collapse of the Soviet Union and the Productivity of American Mathematicians." *The Quarterly Journal of Economics*, Volume 127, Issue 3, August 2012, Pages 1143–1203.

Bernstein, Shai, Rebecca Diamond, Tim McQuade and Beatriz Pousada. 2019. "The Contribution of High-Skilled Immigrants to Innovation in the United States." Stanford University working paper.

Card, David, 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *ILR Review*, *43*(2), 245–257.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration." *Journal of Labor Economics*, Vol. 19, No. 1. January, pp. 22-64.

Carmin, Itzhak J. 1955. *World Jewish Register, A Biographical Compendium of Notable Jews in the Arts, Sciences, and Professions.* Monde.

Cattell, James McKeen and Brimhall, Dean R. 1921. *American Men of Science; a Biographical Directory*, 3rd edition, Garrison, NY, The Science Press, 1921.

Cattell, Jacques, *1956. American Men of Science: A Biographical Directory. Volumes I- III.* New York: R. R. Bowker Company. 1956.

Cohen, Wesley M., Richard R. Nelson and John P. Walsh "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)" NBER Working Paper No. 7552, Issued in February 2000.

Coolidge, Calvin, 1921. "Whose Country is This?" *Good Housekeeping*, volume 72 number 2

February 1921, pages 13-14, 109
Doran, Kirk, and Chungeun Yoon. 2019. "Immigration and Invention: Evidence from the Quota Acts."
Goffman, Caspar. 1969. "And What is Your Erdös Number? *The American Mathematical Monthly,* Volume 76, Issue 7. p. 791.
Gross, D.P., 2019. The Consequences of Invention Secrecy: Evidence from the USPTO Patent Secrecy Program in World War II.
Higham, John. 1963. *Strangers in the Land*. Rutgers, the State University Press.
Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics*, 2 (2): 31-56.
Johnson, Lyndon B., 1965. "Remarks at the Signing of the Immigration Bill Liberty Island, New York" October 3, 1965 in *Public Papers of the Presidents of the United States: Lyndon B. Johnson, 1965.* Volume II, entry 546, pp. 1037-1040. Washington, D. C.: Government Printing Office, 1966.
Jones, Maldwyn Allen, "American Immigration," *The Chicago History of American Civilization*, first edition 1960, second edition 1992.
Kelley, Ninette and Michael Trebilcock, 1998. *The Making of the Mosaic: A History of Canadian Immigration Policy*. Toronto: University of Toronto Press.
Knowles, Valerie. 1997. *Strangers at Our Gates: Canadian Immigration and Immigration Policy, 1540-1997*. Toronto: Dundurn Press.
Kogan, L., Papanikolaou, D., Seru, A. and Stoffman, N., 2017. Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, *132*(2), pp.665-712.
Koven, Steven G, and Frank Götzke, American Immigration Policy: Confronting the Nation's Challenges (Springer, 2010), p. 133Nugent, Walter. *Crossings: The Great Transatlantic Migrations, 1870-1914*. Bloomington: Indiana University Press, 1992.
Levenshtein, Vladimir. 1966. "Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*. 10 (8), February. pp. 707-10.
Moser, Petra, 2012a. "Innovation without Patents – Evidence from World's Fairs" *The Journal of Law and Economics*, Volume 55, No. 1, February 2012, pp. 43-74.
Moser, Petra, 2012b. "Taste-Based Discrimination. Evidence from a Shift in Ethnic Preferences after WWI." *Explorations in Economic History*, Volume 49, Issue 2, April 2012, pp. 167–188.
Moser, Petra, Alessandra M. Voena, and Fabian Waldinger. 2014 *American Economic Review*, 2014, Volume 104(10), pp. 3222-55.
Moser, Petra, Joerg Ohmstedt, and Paul W. Rhode. 2018 "Patent Citations—An Analysis of Quality Differences and Citing Practices in Hybrid Corn." *Management Science* 64:4, 1926-1940
Moser, Petra, Shmuel San, and Ryan Stevens, 2019. "Linking People and Patents. Improved Matching Methods and Code." NYU working paper.
Murray, Robert K. 1976. *The 103rd Ballot: Democrats and the Disaster in Madison Square Garden*. New York: Harper & Row.
O'Rourke, Kevin H. 1997. "The European Grain Invasion, 1870-1913." *Journal of Economic History*, 57 (4), pp. 775-801
Rosenbloom, Joshua L. 2002, *Looking for Work, Searching for Workers: Labor Markets during American Industrialization.* Cambridge University Press.

Salz, Arthur. “Auswanderung und Schiffahrt mit besonderer Beruecksichtigung der oesterreichischen Verhaeltnisse.” *Archiv fuer Sozialwissenschaften und Sozialpolitik* 39, part 1 (1995).

San, Shmuel. 2019. “Labor Supply and Directed Technical Change: Evidence from the Abrogation of the Bracero Program in 1964.” NYU working paper.

Sequeira Sandra, Nathan Nunn, and Nancy Quian. “Immigrants and the Making of America.” *Review of Economic Studies*. Forthcoming.

Tabellini, Marco. “Gifts of Immigrants, Woes of Natives: Lessons from the Age of Mass Migration” *Review of Economic Studies*, Forthcoming.

Taylor, Philip. The Distant Magnet: European Migration to the U.S.A. New York: Harper and Row, 1971.

U.S. Bureau of the Census, “Immigrants by Country:1820-1870” in *Historical Statistics of the United States*, series C 89-119, 1975, Washington, DC.

TABLE 1 – PERSONAL CHARACTERISTICS OF SCIENTISTS BY PLACE OF BIRTH

	All Scientists	ESE	WNE	Other
N Scientists	82,094	2,066	4,029	75,999
Age in 1956	47.02	50.22	48.76	46.84
Married	85.23%	82.96%	83.97%	85.36%
Children	1.61	1.25	1.38	1.63
Female	3.26%	3.58%	2.61%	3.28%

Notes: *All Scientists* includes all scientist who work in the United States in 1956. Within this group, ESE refers to scientists who are born in Eastern or Southern Europe; WNE are scientists who were born in Western or Northern Europe; Other refers to all other scientists. Data constructed from individual entries in the MoS (1956). Eastern-Southern Europe (ESE) includes Armenia, Austria-Hungary, Bulgaria, Caucasus, Cyprus, Czechoslovakia, Estonia, Georgia, Greece, Hungary, Italy, Latvia, Lithuania, Macedonia, Malta, Moldova, Poland, Portugal, Romania, Russia, Slovakia, Spain, Ukraine and Yugoslavia. Western-Northern Europe (WNE) includes Austria, Belgium, Denmark, England, Finland, France, Germany, Iceland, Ireland, Luxembourg, Netherlands, Norway, Scotland, Sweden, Switzerland, and Wales.

TABLE 2 – MISSING ESE-BORN SCIENTISTS IN THE UNITED STATES

	US Scientists				Counterfactual ESE -born scientists post 1924	Missing # ESE-born scientists post 1924
	ESE-born		WNE-born			
	pre 1924	post 1924	pre 1924	post 1924		
All disciplines						
US Naturalization	250	403	244	962	986	583
US education	353	927	336	1684	1,769	842
US education or employment	428	1435	515	2892	2,403	968
US naturalization, education, or employment	488	1330	554	2838	2,500	1,170
Physical sciences						
US naturalization	148	250	144	624	641	391
US education	153	438	151	881	893	455
US education or employment	189	692	273	1,569	1,086	394
US naturalization, education, or employment	235	637	304	1,539	1,190	553

Notes: Estimates of the number of missing ESE-born American scientists after the quota act of 1924, which successfully reduced the inflow of immigrants from Eastern and Southern Europe. Estimates based on "naturalization" use the year when a scientist became a naturalized US citizen as way to estimate the arrival year for foreign-born US scientists, by subtracting the time it takes to become a naturalized US citizen (five years) from the scientist's year of arrival. Estimates based on US education use the start year of the scientist's first US degree to estimate the year of arrival. Estimates based on US education or employment use the start year of the scientist's first US degree or job (the earliest). Finally, measures based on all three sets of information (naturalization , education, and employment) are the earliest year among the three.

TABLE 3 – BALANCING TABLE. COMPARING ESE WITH OTHER FIELDS

	Fields			
	ESE	Other	Difference	p-value
ESE-born	0.035	0.000	0.035	0.000
WNE-born	0.054	0.051	0.003	0.825
Age	44.72	44.41	0.314	0.854
Female	0.011	0.012	-0.002	0.830
Star scientists	0.115	0.104	0.010	0.662

Notes: Pre-quotas 1921 scientists.

TABLE 4 – EFFECTS OF THE QUOTAS ON INVENTION BY AMERICAN SCIENTISTS, BASELINE ESTIMATES

	ln(patents)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ESE x post	-1.142***	-1.103**	-1.192***	-1.247**	-1.285***	-1.362**	-1.288***	-1.292**
	(0.359)	(0.535)	(0.380)	(0.557)	(0.378)	(0.559)	(0.358)	(0.531)
	Baseline		Excl. 5% largest fields		Excl. fields w top 5% ESE share		Incl. new fields	
Percentage change	-0.68	-0.67	-0.70	-0.71	-0.72	-0.74	-0.72	-0.73
Mean patents before 1924	4.15	4.15	3.47	3.47	4.22	4.22	3.97	3.97
N (fields x years)	5,795	5,795	5,490	5,490	5,551	5,551	6,100	6,100
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field-specific pre-trends	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Difference-in-differences regressions compare changes in patenting per year in the pre-quota fields of ESE scientists with changes in other research fields of American scientists: $\ln(y_{it}) = \beta \cdot ESE_i \cdot Post_t + \gamma_i + \delta_t + \epsilon_{it}$ where $\ln(y_{it})$ is the natural logarithm of the number of US patents by scientists worked in the US in 1956 in field i and year t, ESE_i indicates fields with ESE scientists in 1921, $Post_t$ indicates years after 1925, and γ_i and δ_t are field and year fixed effects, respectively. Even columns also control for field-specific linear pre-trends. Standard errors are clustered at the field level. Scientists matched with US patents as described in the text. To be a match, we require the scientist to be between the ages of 18 and 80 at the time of the patent application and have the same name first, last, and middle name. This data reports summary statistics for the main specifications, which exclude the top quintiles of the most frequent names. Robustness checks (Table A4) use the full data, including scientists with common name

TABLE 5 – EFFECTS OF THE QUOTAS ON INVENTION BY AMERICAN SCIENTISTS, INTENSITY REGRESSIONS

	ln(patents)			
	(1)	(2)	(3)	(4)
%ESE x post	-1.205**	-1.409**	-2.660**	-1.446**
	(0.567)	(0.613)	(1.148)	(0.598)
	Baseline	Excl. 5% largest fields	Excl. fields w top 5% ESE share	Incl. new fields
Percentage change	-0.70	-0.76	-0.93	-0.76
Mean patents before 1924	4.15	3.47	4.22	3.97
N (fields x years)	5,795	5,490	5,551	6,100
Year FE	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of difference-in-differences regressions compare changes in patenting per year in the pre-quota files of ESE scientists with changes in other research fields of native American scientists: $\ln(y_{it}) = \beta \cdot \%ESE_i Post_t + \gamma_i + \delta_t + \epsilon_{it}$ where $\ln(y_{it})$ is the natural logarithm of the number of US patents by US-born scientists worked in the US in 1956 in field i and year t, $\%ESE_i$ represents the share of American scientists who were born in ESE countries in 1921, $Post_t$ indicates years after 1925, and γ_i and δ_t are field and year fixed effects, respectively. Standard errors are clustered at the field level.

TABLE 6 – EFFECTS OF THE QUOTAS ON PATENTS BY AMERICAN SCIENTISTS, EXTENSIVE VS INTENSIVE MARGIN

	Fields					Scientists		
	Extensive			Intensive	Combined	Extensive	Intensive	Combined
	𝕀(patents > 0)			ln(patents)	ln(patents)	ln(scientists)	ln(patents/ scientist)	ln(patents)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ESE x post	-0.160***	-0.777***	-1.339**	-0.589***	-1.142***	-0.515***	-0.394**	-0.923***
	(0.049)	(0.275)	(0.552)	(0.196)	(0.359)	(0.102)	(0.156)	(0.326)
	OLS	Probit	Logit	OLS (dropping zeros)	OLS (baseline)	OLS	OLS	OLS
Percentage change	-0.16	-0.11	-0.10	-0.45	-0.68	-0.40	-0.33	-0.60
Mean outcome before 1924	0.53	0.49	0.49	7.77	4.15	66.32	0.05	3.83
N (fields x years)	5,795	5,246	5,246	4,742	5,795	4,275	4,275	4,275
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table decompose the effect of the quotas on the number of patents into extensive and intensive margins in two ways. Columns 1-5 decompose the effect in terms of active fields per year (extensive margin) and number of patents per active field (intensive margin). Columns 6-8 decompose the effect in terms of number of scientists per field and year (extensive margin) and number of patents per scientists (intensive margin). Specifically, columns 1-3 show the estimates of three binary outcome models, Linear probability model (OLS), probit and logit, where the outcome equals one for field-year pairs with at least one patent. Column 4 estimates the baseline OLS specification, dropping field-year pairs with no patents. Column 5 reports the baseline OLS specification including all field-year pairs (with the outcome : $\ln(y_{it} + 0.01)$). Column 6 reports the estimate of a specification with the (log) number of scientists in a field-year pair as outcome. We use the information on the start year of first US education or job to determine the first year a scientists starts to be active in her field. Column 7 estimates the effect on the number of patents per scientist in a field-year per (only patents by scientists that were active at the application year of the patent). Column 8 estimate the total effect on patents (for the same set of patents). Standard errors are clustered at the field level.

TABLE 7 – EFFECTS OF THE QUOTAS ON INVENTION BY AMERICAN SCIENTISTS, US-BORN SCIENTISTS

	ln(patents)				ln(scientists)	ln(patents/ scientist)	ln(patents)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ESE x post	-0.979**	-1.029**	-1.103***	-1.119***	-0.506***	-0.366**	-0.819**
	(0.374)	(0.396)	(0.396)	(0.371)	(0.101)	(0.164)	(0.345)
	Baseline	Excl. 5% largest fields	Excl. fields w top 5% ESE share	Incl. new fields	OLS	OLS	OLS
Percentage change	-0.62	-0.64	-0.67	-0.67	-0.40	-0.31	-0.56
Mean patents before 1924	3.61	3.04	3.68	3.45	61.61	0.05	3.52
N (fields x years)	5,795	5,490	5,551	6,100	4,275	4,275	4,275
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of difference-in-differences regressions compare changes in patenting per year in the pre-quota files of ESE scientists with changes in other research fields of native American scientists: $\ln(y_{it}) = \beta \cdot ESE_i \cdot Post_t + \gamma_i + \delta_t + \epsilon_{it}$ where $\ln(y_{it})$ is the natural logarithm of the number of US patents by US-born scientists worked in the US in 1956 in field i and year t, ESE_i indicates fields with ESE scientists in 1921, $Post_t$ indicates years after 1925, and γ_i and δ_t are field and year fixed effects, respectively. Column 5 reports the estimate of a specification with the (log) number of Us-born scientists in a field-year pair as outcome. We use the information from the start year of first US education or job to determine the first year a scientists starts to be active in her field. Column 6 estimates the effect on the number of patents per US-born scientist in a field-year per (only patents by scientists that were active at the application year of the patent). Column 7 estimate the total effect on US-born patents (for the same set of patents). Standard errors are clustered at the field level.

TABLE 8 – EFFECTS OF THE QUOTAS ON THE NUMBER OF NEW AMERICAN SCIENTISTS IN ESE VS OTHER FIELDS

	ln(scientists)			
	All disciplines		Physical sciences	
	(1)	(2)	(3)	(4)
ESE x post	-0.274**	-0.259**	-0.638***	-0.623***
	(0.133)	(0.130)	(0.206)	(0.204)
	Education	Education + Employment	Education	Education + Employment
Percentage change	-0.24	-0.23	-0.47	-0.46
Mean scientists before 1924	12.07	12.47	5.42	5.65
N (fields x years)	3,600	3,600	3,800	3,800
Year FE	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes

Notes: Difference-in-differences regressions compare changes in the number of new scientists per year in the pre-quota fields of ESE scientists with changes in other research fields of American scientists: $\ln(y_{it}) = \beta \cdot ESE_i \cdot Post_t + \gamma_i + \delta_t + \epsilon_{it}$ where $\ln(y_{it})$ is the natural logarithm of the flow of new scientists in field i and year t, ESE_i indicates fields with ESE scientists in 1921, $Post_t$ indicates years after 1925, and γ_i and δ_t are field and year fixed effects, respectively. We use the complete education and employment history of the scientists to build the measure of new scientists by year. In columns (1) and (4), y_{it} is the number of scientists belong to field i started to study at her first US institution at year t. In columns (2) and (5), we use the analogues measure using the start year of the first US job. The outcome measure in columns (3) and (6) combines all information available and uses the start year of either a US degree or a job. Columns (1)-(3) estimated using all scientists in the MoS, while columns (4)-(6) are only for the physical sciences. Standard errors are clustered at the field level.

TABLE 9 – EFFECTS OF THE QUOTAS ON AMERICAN INVENTION, TRIPLE DIFFERENCES

	ln(patents)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ESE x US x post	-1.191***	-1.084**	-1.089***	-1.071**	-1.346***	-1.337**	-1.370***	-1.313**
	(0.356)	(0.507)	(0.363)	(0.517)	(0.371)	(0.533)	(0.357)	(0.507)
	Baseline		Excl. 5% largest fields		Excl. fields w top 5% ESE share		Incl. new fields	
Percentage change	-0.70	-0.66	-0.66	-0.66	-0.74	-0.74	-0.75	-0.73
Mean patents before 1924	2.08	2.08	1.74	1.74	2.12	2.12	1.99	1.99
N (clusters x countries x years)	11,590	11,590	10,980	10,980	11,102	11,102	12,200	12,200
Year-field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-field-specific pre-trends	No	Yes	No	Yes	No	Yes	No	Yes

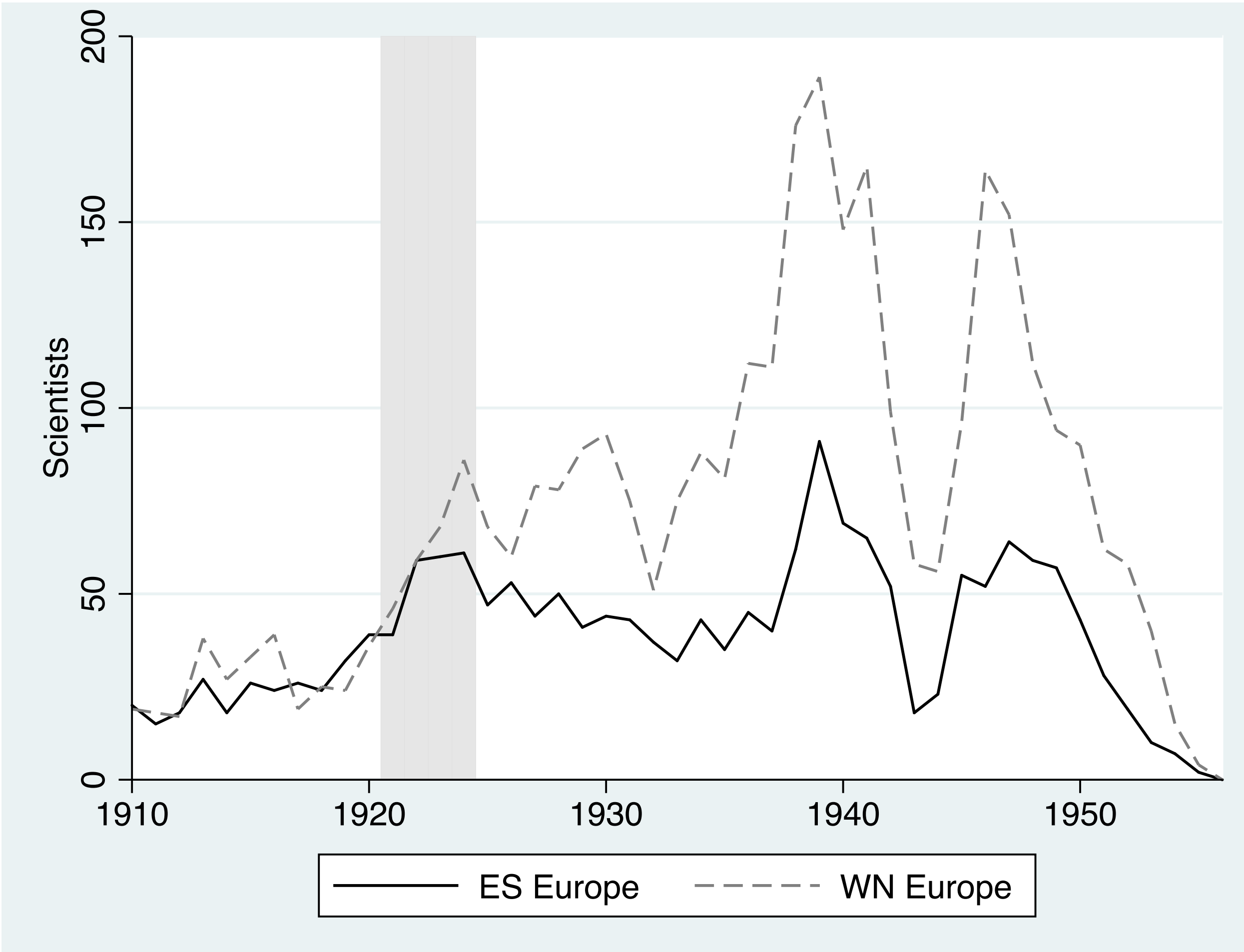
Notes: Triple-differences regressions compare changes in patenting by Canadian with American scientists after 1924 in ESE fields with other fields: $\ln(y_{ict}) = \beta\, ESE_i\, US_c\, Post_t + \gamma_{ic} + \delta_{it} + \theta_{ct} + \epsilon_{ict}$ where $\ln(y_{ict})$is the natural logarithm of the number of US patents by scientists worked in the country c (Canada/US) in 1956 in field i and year t, ESE_i indicates fields with ESE scientists in 1921, $Post_t$ indicates years after 1925, US_c equals one for US and zero for Canada, and $\gamma_{ic}, \delta_{it}, and\ \theta_{ct}$ are field-country, field-year and country-year fixed effects, respectively. Even columns also control for field-country-specific linear pre-trends. Standard errors are clustered at the field level.

TABLE 10 – EFFECTS OF THE QUOTAS ON AMERICAN INVENTION, CONTROLLING FOR ESE-SCIENTISTS AGE

	ln(patents)				
	(1)	(2)	(3)	(4)	(5)
ESE x post	-1.153***	-1.053***	-1.075***	-0.990***	-1.014***
	(0.370)	(0.362)	(0.384)	(0.367)	(0.377)
Share above 40 x post		-0.011			-0.016
		(0.007)			(0.013)
Share above 65 x post			-0.007		-0.009
			(0.015)		(0.017)
Average age x post				-0.035	0.026
				(0.027)	(0.055)
Percentage change	-0.68	-0.65	-0.66	-0.63	-0.64
Mean patents pre-1924	4.22	4.22	4.22	4.22	4.22
N (fields x years)	5,551	5,551	5,551	5,551	5,551
Year FE	Yes	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes	Yes

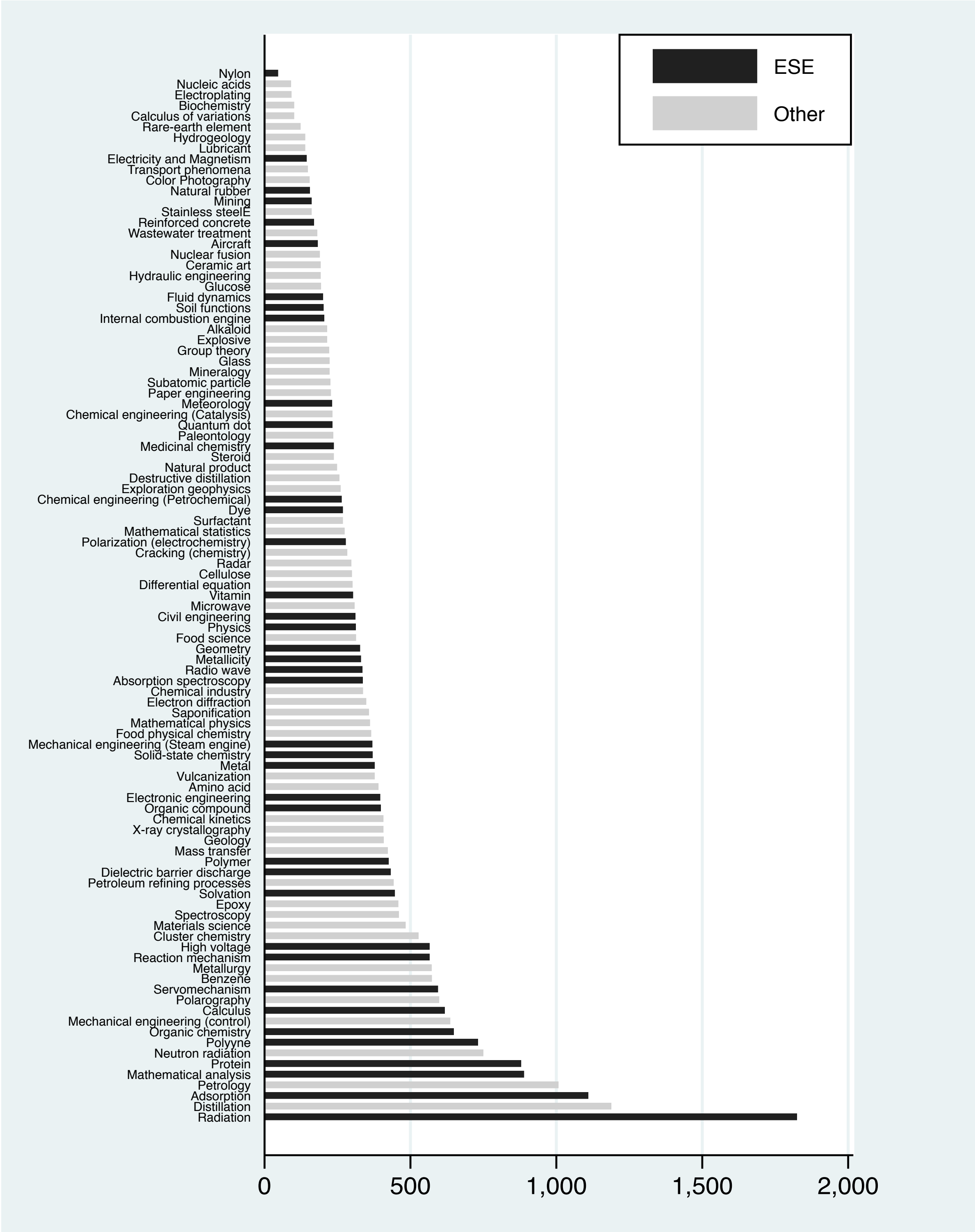
Notes: This table show the results of the baseline difference-in-differences specification, with the addition of variables capturing various dimensions of the age profile of ESE scientists in 1956 within a field: $\ln(y_{it}) = \beta_1 \cdot ESE_i \cdot Post_t + \beta_2 \cdot ESE - age_i \cdot Post_t + \gamma_i + \delta_t + \epsilon_{it}$. $ESE - age_i$ is the share of ESE scientists who are older than 40 years in 1956 (column 2), the share of ESE scientists who are older than 65 in 1956 (column 3), the average age of ESE scientists by field (column 4), and all three age variables together (column 5). All other variables are as defined in previous tables. This table includes fields with at least one ESE scientists in 1956. Standard errors are clustered at the field level.

FIGURE 1 – ARRIVALS OF AMERICAN SCIENTISTS FROM ES VS. WN EUROPE



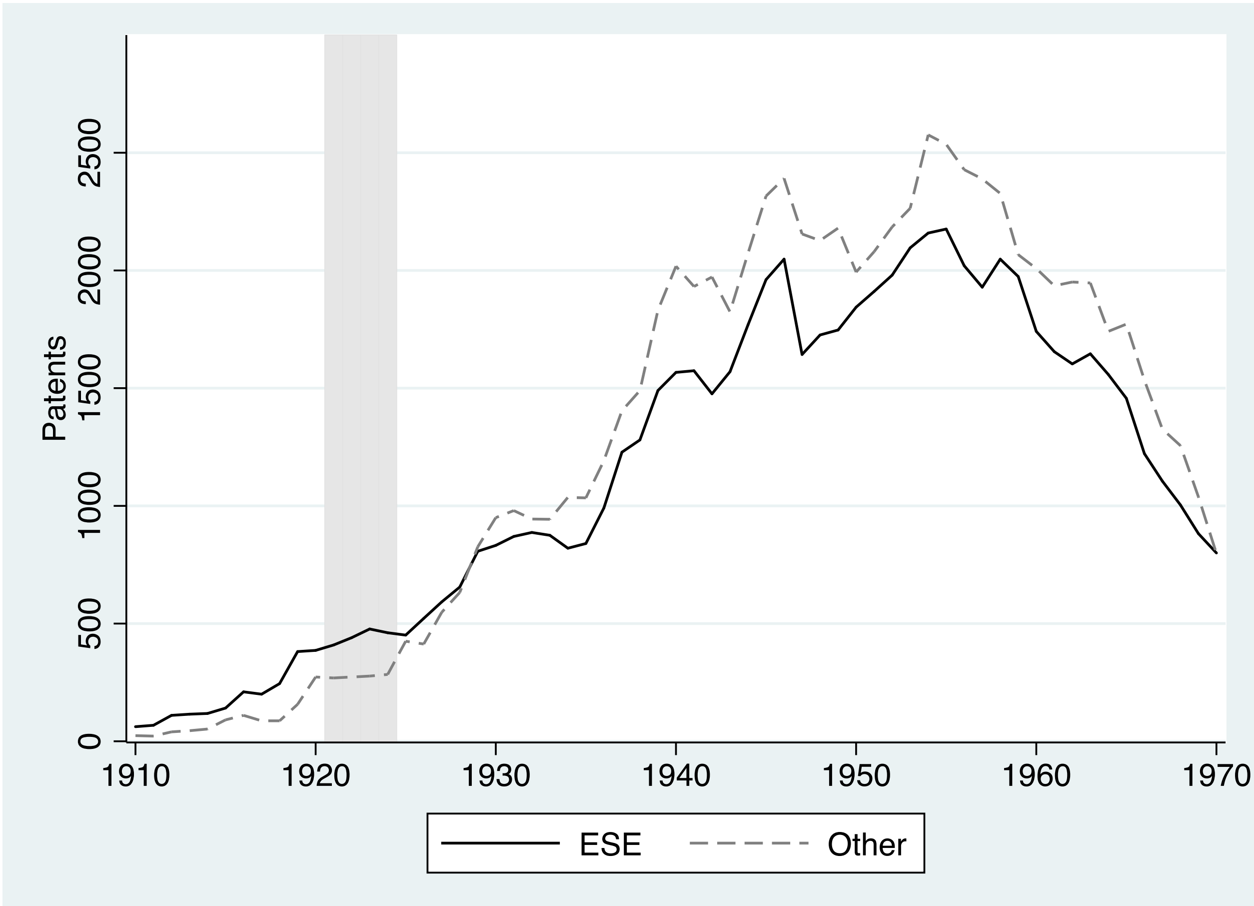
Notes: Arrivals per year of ESE-born American scientists compared with American scientists born with WN Europe. Years of arrivals are proxied by information on US naturalization, education and employment.

FIGURE 2– SCIENTISTS IN ESE AND OTHER FIELDS



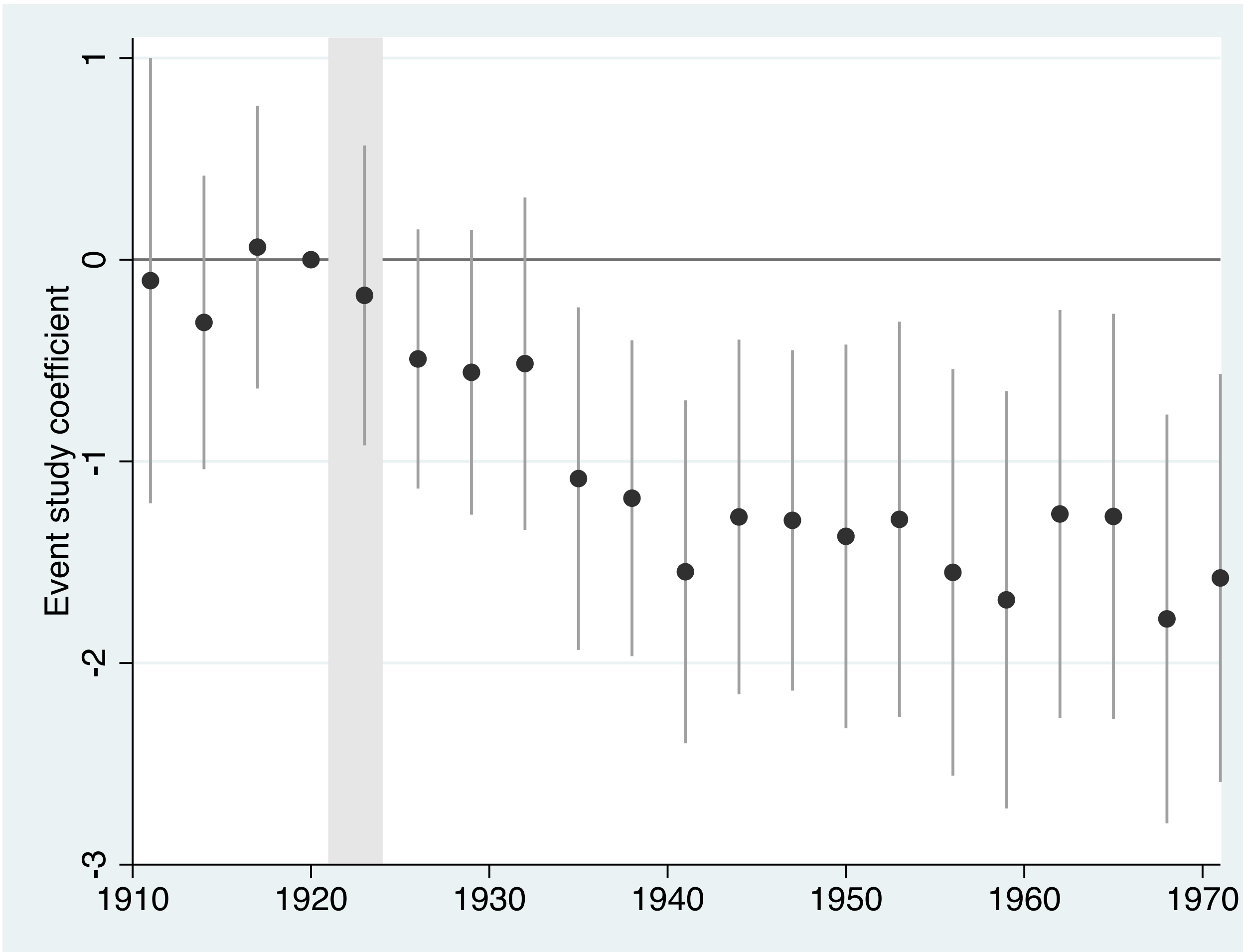
Notes: Number of scientists in 1956 by field for fields. ESE fields (in black) are fields in which Eastern European-born scientists were research-active in 1921.The figure excludes the residual cluster (25, "Chemistry") which includes 4,811scientists.

FIGURE 3 – PATENTS BY SCIENTISTS PER YEAR IN ESE AND OTHER FIELDS



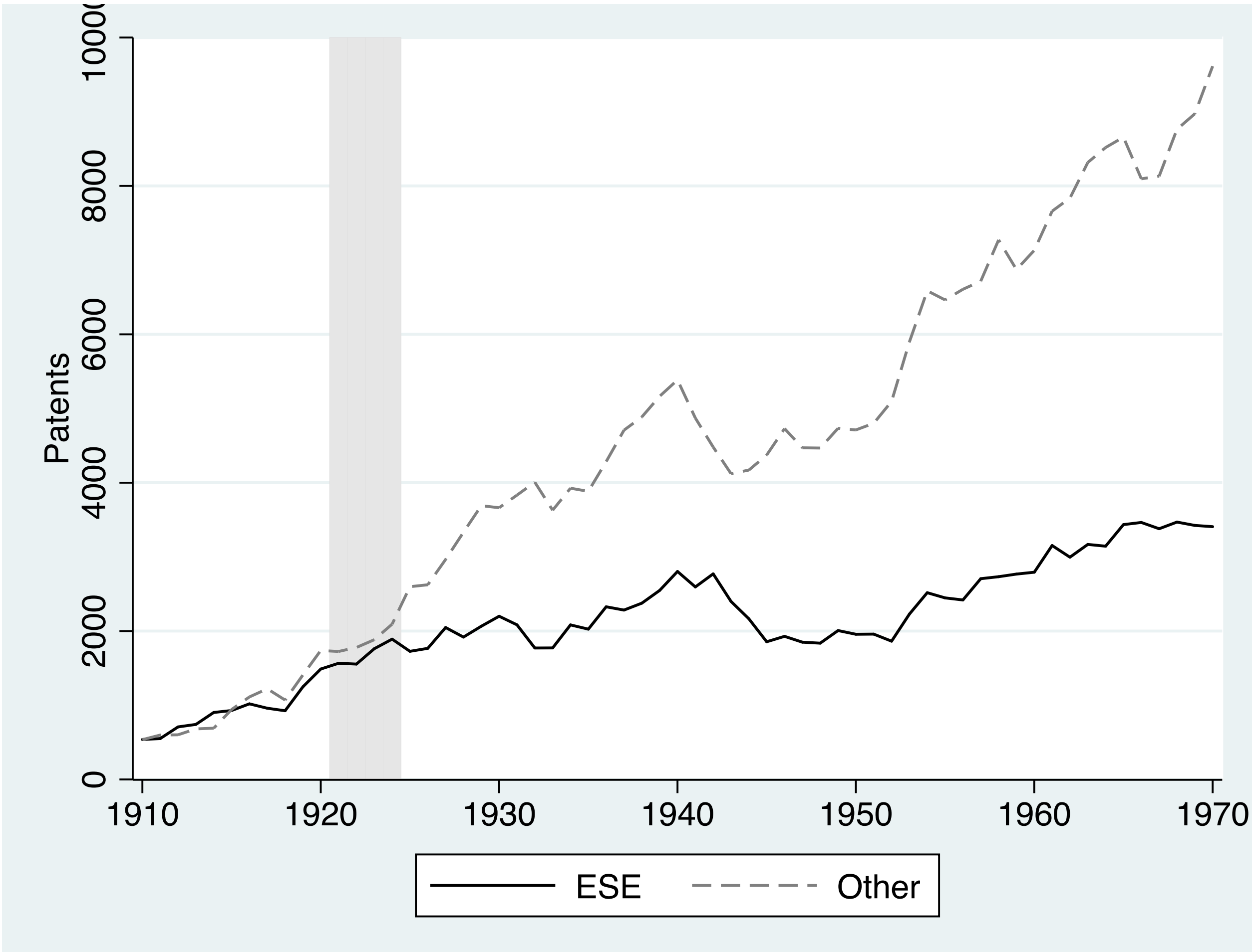
Notes: Patents by scientists per year (measured in the year of the patent application or filing) in the pre-quota fields of *ESE* scientists (solid line) and other fields (interrupted line).

FIGURE 4 –TIME-VARYING EFFECTS ON INVENTION BY AMERICAN SCIENTISTS



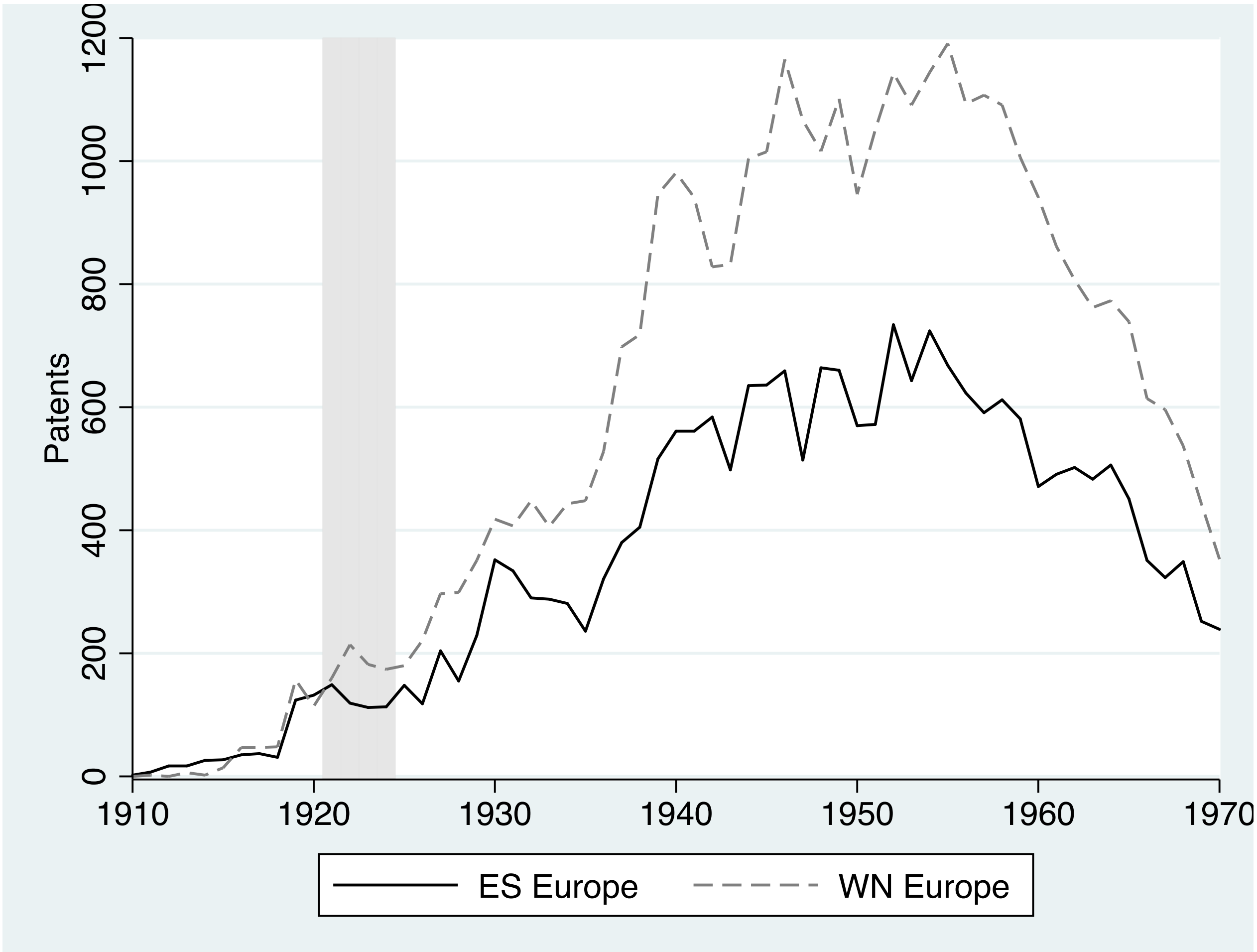
Notes: Time-varying estimates of β_t in the OLS regression $\ln(y_{it}) = \beta_t ESE_i + \gamma_i + \delta_t + \epsilon_{it}$ where $\ln(y_{it})$ is the natural logarithm of the number of US patents by American field i and year t. The variable ESE_i indicates the research fields of ESE scientists in 1921, and γ_i and δ_t are field and year fixed effects, respectively, and 1918-1920 is the excluded period. Standard errors are clustered at the field level.

FIGURE 5 – PATENTS PER YEAR IN ESE AND OTHER FIELDS, FIRMS



Notes: Patents by firms

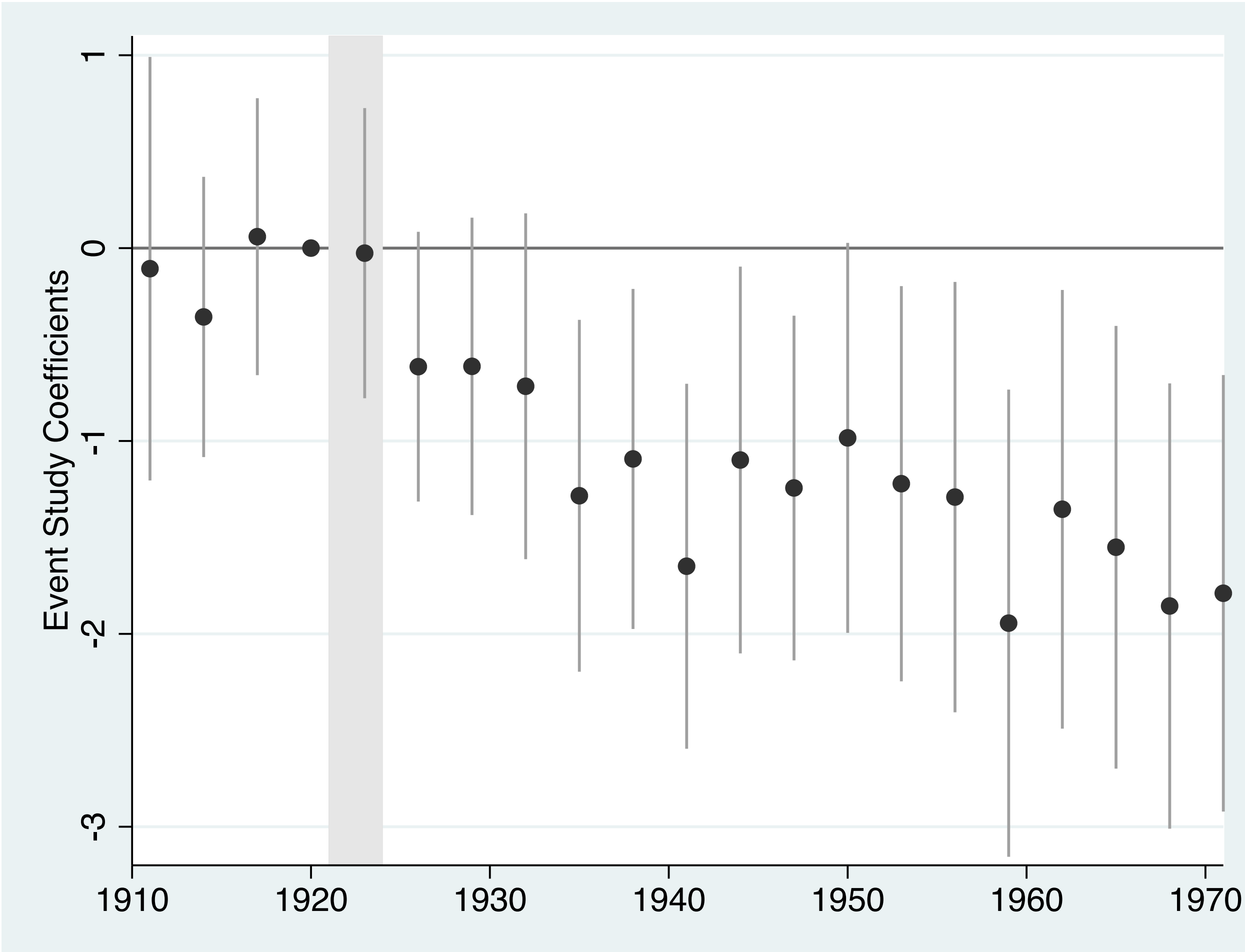
FIGURE 6– PATENTS BY CO-INVENTORS AND



CO-INVENTORS OF CO-INVENTORS OF ESE AND WNE SCIENTISTS

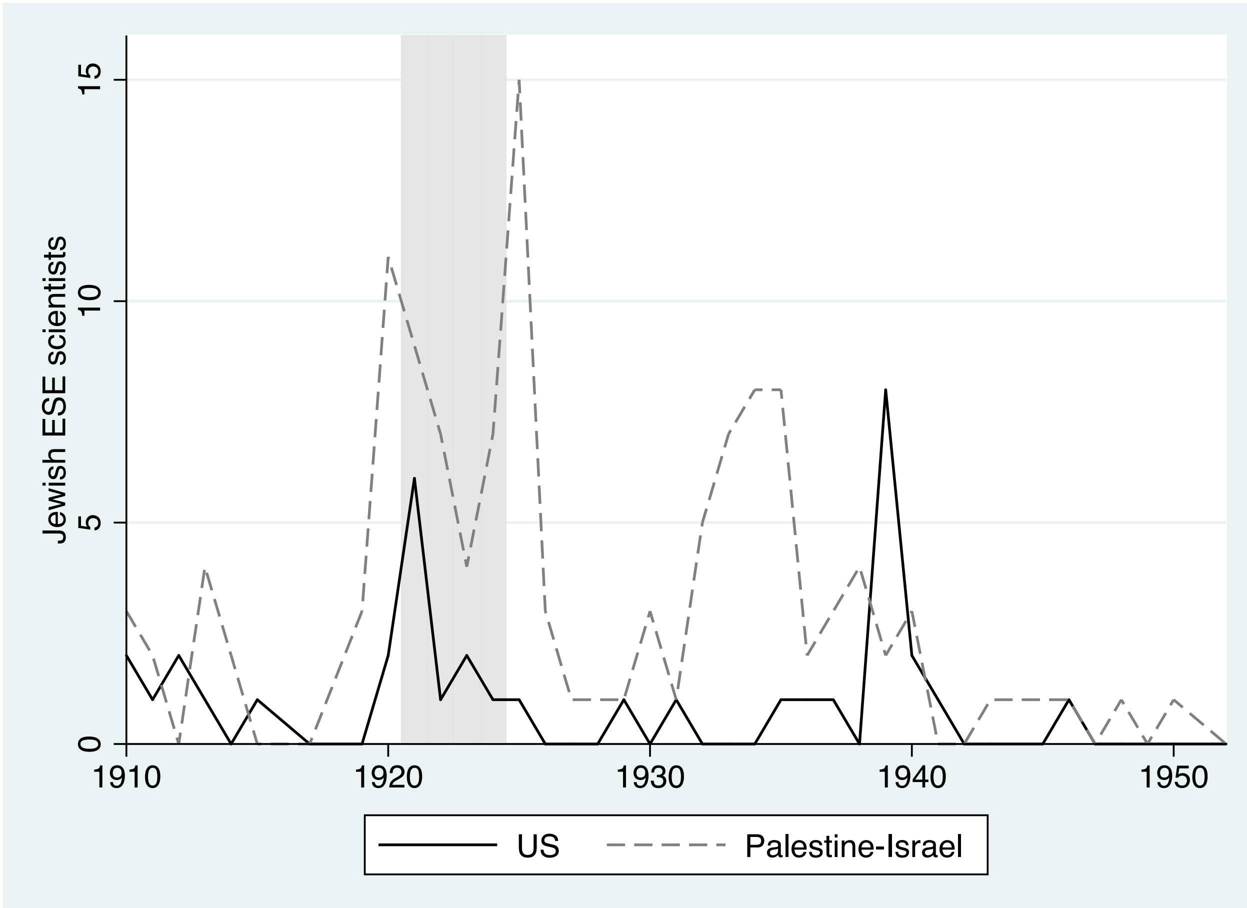
Notes: Patents by US-born scientists that have at least one common patent with ESE and WNE scientists ("co-inventors") and patents by US-born scientists that have at least one common patents with these co-inventors ("co-inventors on co-inventors").

FIGURE 7– ESTIMATES OF THE TIME-VARYING DIFFERENCES BETWEEN AMERICAN AND CANADIAN SCIENTISTS IN THE EFFECTS OF THE QUOTAS ON INVENTION



Notes: Triple-differences event study regression: $\ln(y_{ict}) = \beta_t\, ESE_i\, US_c + \gamma_{ic} + \delta_{it} + \theta_{ct} + \epsilon_{ict}$. $\ln(y_{ict})$ is the natural logarithm of the number of US patents by scientists worked in in country c in 1956 in field i and year t, β_t is a tri-annual indicator variable, ESE_i indicates fields with ESE scientists in 1921. The variable US_c takes the value of 1 for scientists who are employed in the United States in 1956 and 0 for scientists who work in Canada. $\gamma_{ic}, \delta_{it}, and\ \theta_{ct}$ are field-country, field-year and country-year fixed effects, respectively. 1918-1920 is the excluded period. The graph shows the point estimate and the 95 percent confidence interval of the coefficients β_t. Standard errors are clustered at the field level.

FIGURE 8– JEWISH ESE SCIENTISTS TO US AND PALESTINE-ISRAEL BY YEAR OF IMMIGRATION



Notes: Number of Jewish ESE-born scientists immigrated to the US and Palestine (or Israel after 1948) by year and destination of immigration. Data from the "science" part of the *World Jewish Register* (1955).

Appendix Tables

Table A1 – Patents Matching

	Total	Physical Sciences	Biological Sciences	Social Sciences
Scientists in MoS	82,094	41,096	25,505	15,493
Matches 18-80 years				
Scientists w at least 1 patent	46,158	28,453	11,629	6,076
Patents	1,960,438	1,148,855	506,358	305,225
Patents per scientist	23.88	27.96	19.85	19.70
Error Rate	79.6%	73.4%	88.1%	88.6%
Matches 18-80 years, same middle name				
Scientists with at least 1 patent	28,994	21,705	5,191	2,098
Patents	292,675	246,800	30,602	15,273
Patents per scientist	3.57	6.01	1.20	0.99
Error Rate	24.9%	16.4%	66.1%	79.6%
Matches 18-80 years, same middle name, drop top 20 percent of frequent names				
Scientists with at least 1 patent	19,079	15,721	2,681	677
Patents	184,484	171,612	10,345	2,527
Patents per scientist	2.25	4.18	0.41	0.16
Error Rate	7.5%	5.1%	32.0%	63.2%

Notes: Patents matching for 1956 MoS scientists. Type I error rate is the share of false-positive patents in the total number of patents. We use the number of patents that were submitted when the inventor was between 0 and 18 years old as a proxy for false-positive matches. A scientist and a patent are “same middle name” match if they have the same number of names, and all names are compatible. For example, “John Smith” - “John Smith” and “John G. Smith” - “John George Smith” are middle name matches. However, “John G. Smith” - “John Smith”, “John G. Smith”- “John Robert Smith”, “John Richard Smith”- “John Robert Smith”, and “John G. Smith” - “John G. R. Smith” are not. The probability of a name is calculated by multiplying the probability of the first name by the probability of the last name. Data on surname frequencies are from Census 2000 data contain surnames occurring 100 or more times. Data on first names frequencies are from the Social Security Administration for U.S. people born from 1880 to 2013 contain names occurring 5 or more times in each year.

TABLE A2 – EXAMPLES OF FIELDS

Field	9	19	29	39	49
title	Servomechanism	Chemical engineering (Catalysis)	Organic chemistry	Neutron radiation	Internal combustion engine
scientists	594	232	648	749	204
field 1	electrical engineering	chemical engineering	organic chemistry	physics	mechanical engineering
field 2	physics	engineering	Chemistry	nuclear physics	engineering
field 3	engineering	chemistry	physical organic chemistry	nuclear chemistry	chemical engineering
field 4	chemistry	industrial and chemical engineering	organic and polymer chemistry	chemistry	chemistry
field 5	electrical and chemical engineering		Biochemistry	experimental physics	physics
word 1	electrical	chemical	Organic	nuclear	combustion
word 2	engineering	engineering	Chemistry	physics	engines
word 3	power	process	Synthetic	energy	internal
word 4	electric	development	Polymer	spectroscopy	mechanical
word 5	machinery	industrial	Medicinal	cosmic	engineering
word 6	circuits	chemistry	Steroids	rays	fuels
word 7	transmission	catalysis	Research	scattering	fuel
word 8	servomechanisms	plastics	pharmaceuticals	reactor	engine
word 9	electronics	kinetics	Syntheses	reactions	jet
word 10	measurements	organic	Medicinals	neutron	gas
Field	59	69	79	89	99
title	Aircraft	Mathematical analysis	Vulcanization	Calculus of variations	Adsorption
scientists	182	889	377	101	1109
field 1	aeronautical engineering	mathematics	Chemistry	mathematics	physical chemistry
field 2	engineering	applied mathematics	organic chemistry	pure mathematics	chemistry
field 3	aeronautics	physics	chemical engineering	applied mathematics	physics
field 4	physics	actuarial mathematics	physical chemistry	mathematical analysis	physical organic chemistry
field 5	mechanical engineering	engineering	Physics	physics	oceanography
word 1	aeronautical	mathematics	Rubber	calculus	physical
word 2	aircraft	analysis	Chemistry	variations	chemistry
word 3	engineering	topology	Synthetic	mathematics	properties
word 4	structures	functions	Plastics	equations	kinetics
word 5	design	mathematical	Latex	differential	thermodynamics
word 6	control	applied	Organic	theory	adsorption
word 7	flight	series	compounding	analysis	chemical
word 8	research	functional	polymerization	functions	catalysis
word 9	stability	numerical	Technology	mathematical	surface
word 10	guided	spaces	Accelerators	problems	structure

Notes: This table present the title, the number of scientists in 1956, the 5 most common original MoS fields, and the 10 most frequent words of 10 out of 100 fields obtained from the k-mean clustering. To get the titles, we perform a Google search for the most frequent words in each field and name each cluster with the first result of that search.

TABLE A3 – EFFECTS OF THE QUOTAS ON AMERICAN INVENTION, SENSITIVITY TO THE PATENT MATCHING PROCESS

	ln(patents)			
	(1)	(2)	(3)	(4)
ESE x post	-1.142***	-1.284***	-1.403***	-0.927***
	(0.359)	(0.345)	(0.275)	(0.220)
	Baseline	Incl. common names	Incl. different middle names	Incl. common names and different middle names
Percentage change	-0.68	-0.72	-0.75	-0.60
Mean patents before 1924	4.15	6.38	7.24	39.51
N (fields x years)	5,795	5,795	5,795	5,795
Year FE	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes

Notes: This table check the sensitivity of our results to the patents-matching process. As we show in the text, dropping top quantile of common names and keeping only patents that matched also in the middle name entry, significantly increase the accuracy of the data (see appendix Table A2). Column 2 reports the estimates of the baseline specification (equation 2) if we do not drop patents by scientists with common names. In column 3 we keep all patents matched by first and last name, even if the middle name is not matched. Column 4 keeps both types of patents.

TABLE A4 – EFFECTS OF THE QUOTAS ON AMERICAN INVENTION, SENSITIVITY TO THE CLUSTERING PROCESS

	ln(patents)			
	(1)	(2)	(3)	(4)
ESE x post	-0.932**	-1.022***	-1.142***	-1.141***
	(0.429)	(0.382)	(0.359)	(0.341)
K clusters	50	75	100	125
Percentage change	-0.61	-0.64	-0.68	-0.68
Mean patents before 1924	8.37	5.50	4.15	3.51
N (clusters x years)	2,867	4,392	5,795	6,832
Year FE	Yes	Yes	Yes	Yes
Cluster FE	Yes	Yes	Yes	Yes

Notes: This table check the sensitivity of our results for the choice of the number of clusters in the K-mean clustering. In each column, we choose different number of clusters (K) and re-estimate the baseline specification (equation 2).

TABLE A5 – EFFECTS OF THE QUOTAS ON AMERICAN INVENTION,
ROBUSTNESS TO THE ECONOMETRIC MODEL

	patents		ln(patents + ϵ)			
	(1)	(2)	(3)	(4)	(5)	(6)
ESE x post	-0.756***	-0.910***	-0.775***	-1.142***	-1.510***	-1.877***
	(0.272)	(0.237)	(0.277)	(0.359)	(0.453)	(0.553)
	Poisson	Negative Binomial	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.001$	$\epsilon = 0.0001$
Percentage change	-0.53	-0.60	-0.54	-0.68	-0.78	-0.85
Mean patents before 1924	4.15	4.15	4.15	4.15	4.15	4.15
N (fields x years)	5,795	5,795	5,795	5,795	5,795	5,795
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Columns 1-2 report Poisson and negative binomial models for count data of the form: $\mathbb{E}\,[\ln(y_{it})] = \beta \cdot ESE_i \cdot Post_t + \gamma_i + \delta_t$, where the operator $\mathbb{E}[\cdot]$ represents the mean conditioned on all the variables in the right hand side of the equation. Columns 3-6 report the results of the OLS baseline specification : $\ln(y_{it} + \epsilon) = \beta \cdot ESE_i \cdot Post_t + \gamma_i + \delta_t + u_{it}$ with various values for ϵ.

TABLE A6 – EFFECTS OF THE QUOTAS ON CANADIAN INVENTION

	ln(patents)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ESE x post	0.049	-0.019	-0.103	-0.176	0.061	-0.025	0.081	0.021
	(0.151)	(0.171)	(0.131)	(0.148)	(0.158)	(0.180)	(0.148)	(0.167)
	Baseline		Excl. 5% largest fields		Excl. fields w top 5% ESE share		Incl. new fields	
Percentage change	0.05	-0.02	-0.10	-0.16	0.06	-0.02	0.08	0.02
Mean patents before 1924	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
N (fields x years)	5,795	5,795	5,490	5,490	5,551	5,551	6,100	6,100
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field-specific pre-trends	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Placebo difference-in-differences regressions estimates the baseline regressions for patents by scientists worked in Canada in 1956 (instead of the US as in the baseline regressions).

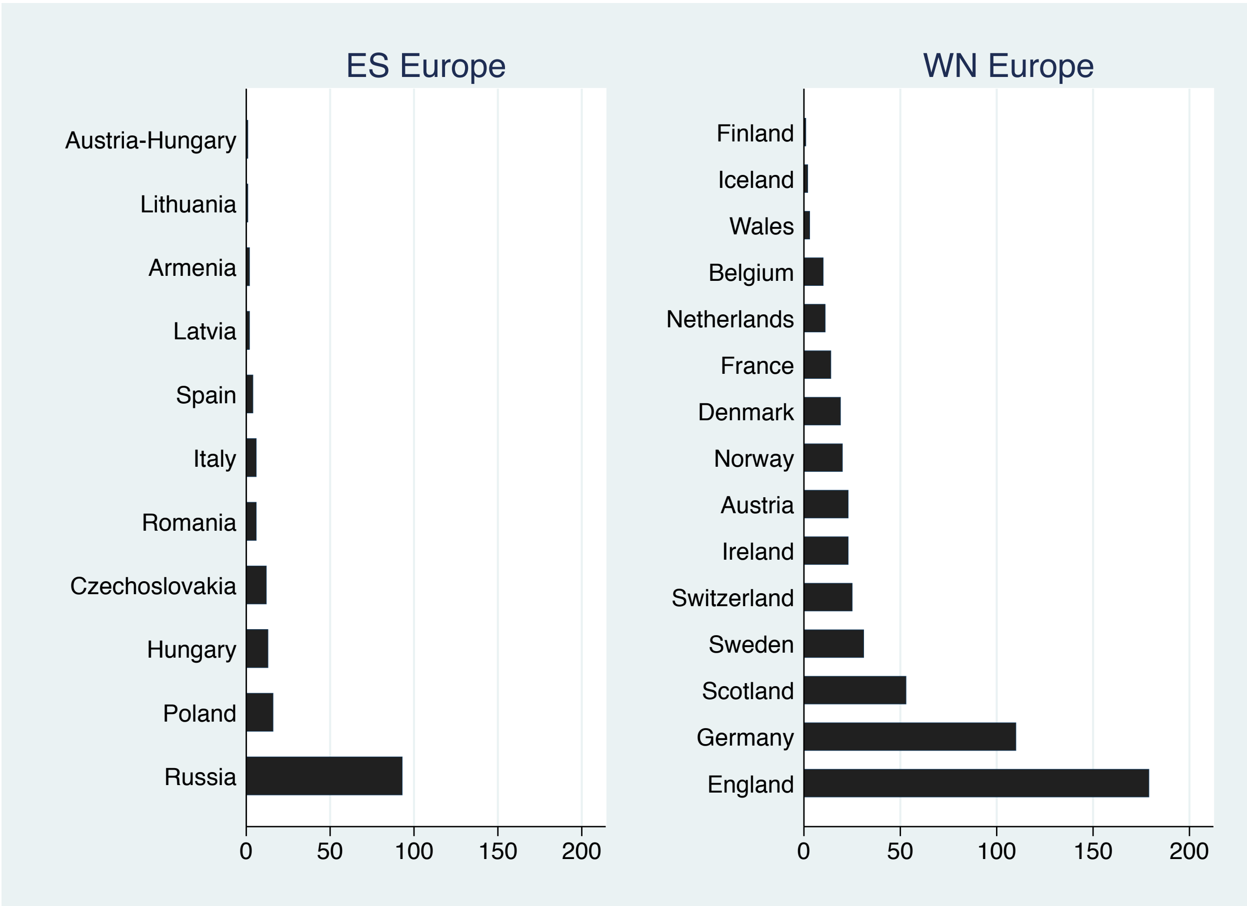
TABLE A7 – COUNTRY AT 1956 BY BIRTH COUNTRY, AND YEAR AND COUNTRY OF FIRST NORTH AMERICAN DEGREE

	ESE				WNE			
Country of first American degree	US		Canada		US		Canada	
Year of first American degree	1910-24	1925-56	1910-24	1925-56	1910-24	1925-56	1910-24	1925-56
Scientists	175	434	7	30	205	849	41	120
Country at 1956								
US	175	427	3	20	201	836	22	59
Canada	0	1	4	10	2	7	18	61
Other	0	6	0	0	2	6	1	0

Notes: ESE refers to scientists who are born in Eastern or Southern Europe; *WNE* are scientists who were born in Western or Northern Europe. Country of first American degree is determined using detailed information on the educational institutions of the scientists in MoS 1956. Year of first American degree is the start year of the first degree the scientist got from North American institution. Country at 1956 is based on the main current employee of a scientist.

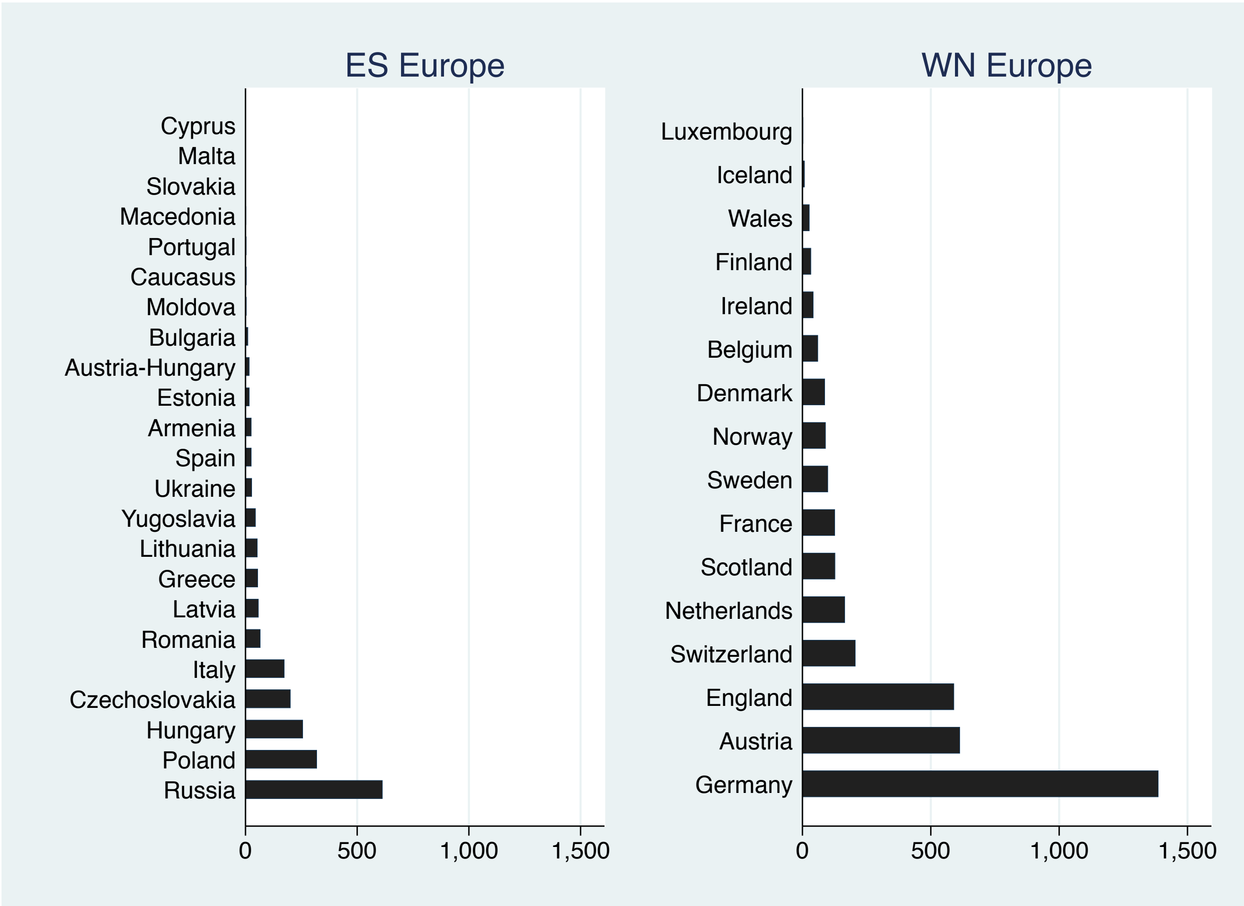
Appendix Figures

Figure A1- Birth Places of American Scientists in 1921



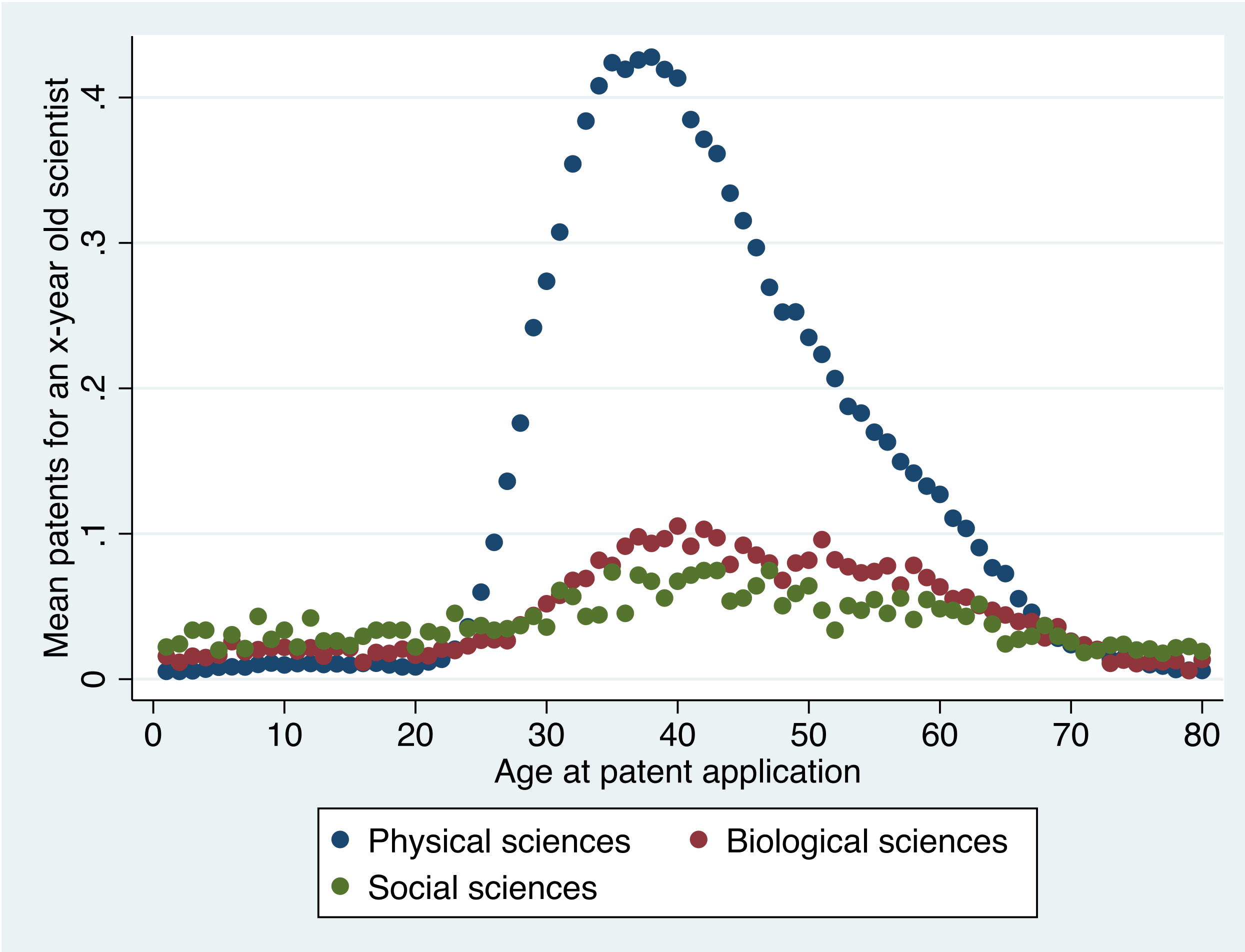
Note: European-born scientists in the 1921 edition of MoS.

FIGURE A2 - BIRTH PLACES OF AMERICAN SCIENTISTS IN 1956



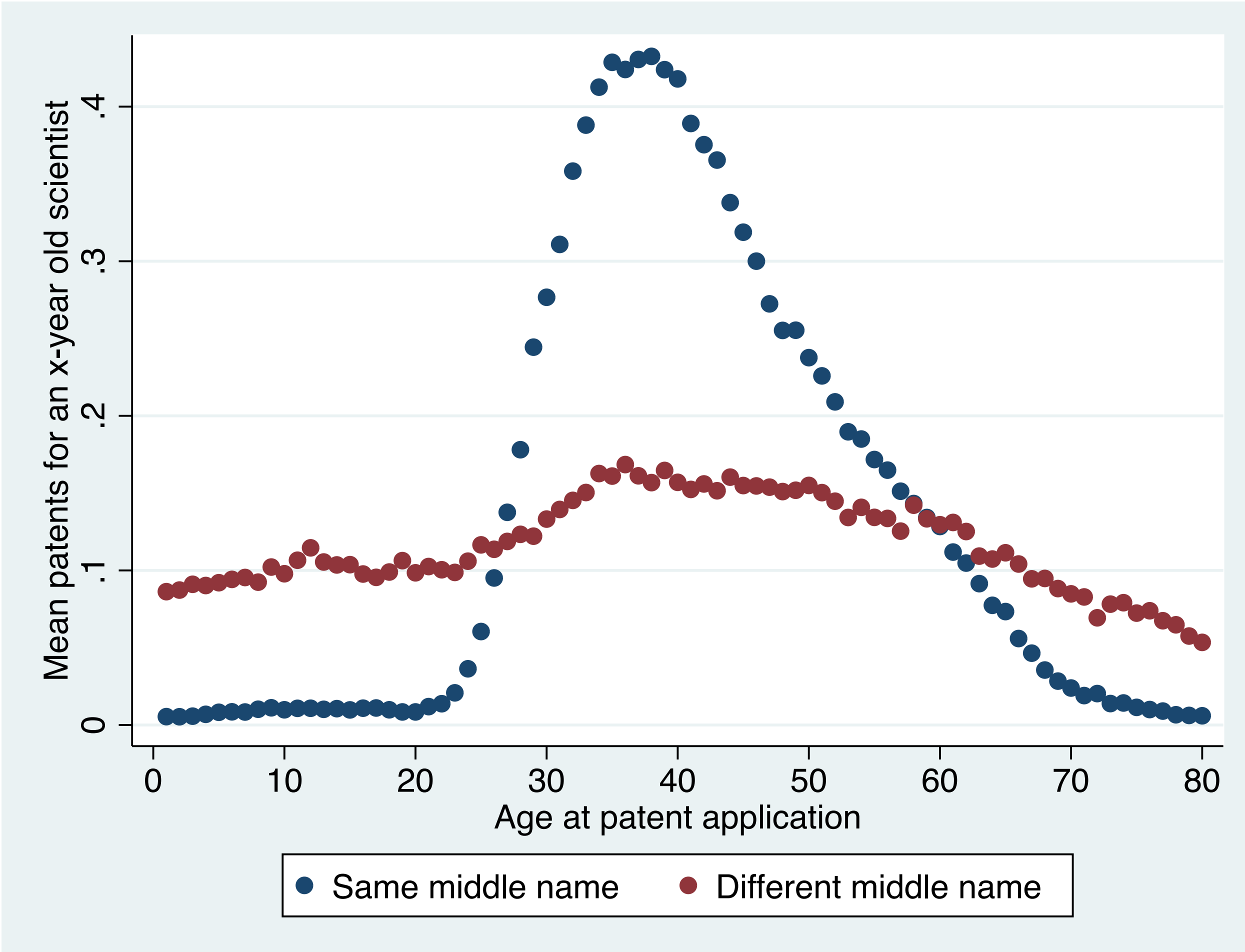
Note: European-born scientists in the 1956 edition of MoS.

FIGURE A3 – THE AGE PROFILE OF INVENTION: PATENTS PER SCIENTIST AND AGE FOR THE PHYSICAL, BIOLOGICAL AND SOCIAL SCIENCES



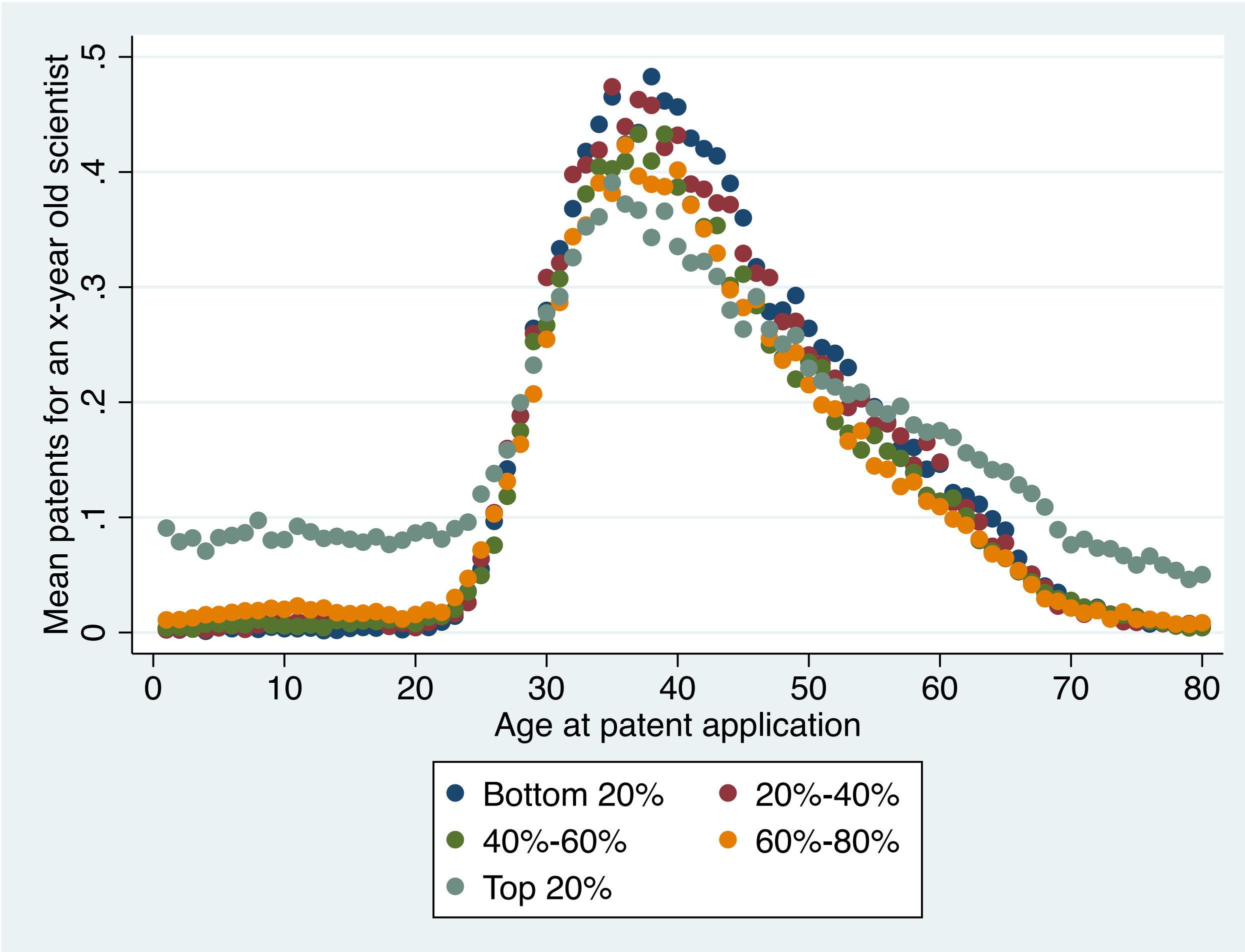
Notes: Average patents per scientist for scientists who are x-year old at the year of the patent application. This average is calculated separately for the three disciplines. Patents matched on first, middle and last names, excluding the top quintile of common names.

FIGURE A4 – MEAN PATENTS PER SCIENTIST AND AGE BY MATCHING TYPE



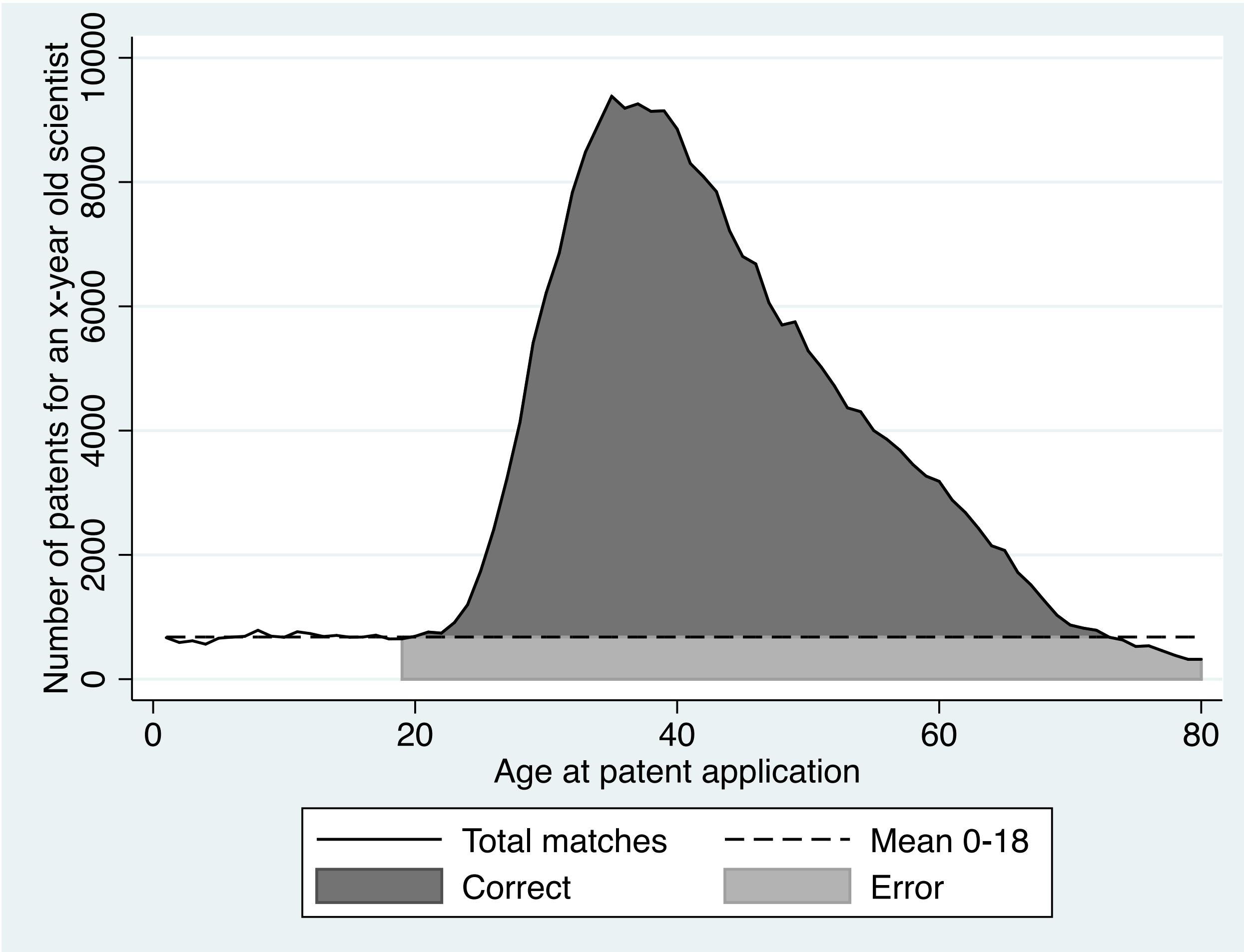
Notes: Average patents per scientist for scientists in the physical sciences who are x-year old at the year of the patent application. Patents matched on first and last names, excluding the top quintile of common names. The average is calculated separately for patents which are matched also on middle name and patents which do not match on middle name. A scientist and a patent are "same middle name" match if they have the same number of names, and all names are compatible.

FIGURE A5 – MEAN PATENTS PER SCIENTIST AND AGE BY GROUPS OF NAME FREQUENCY



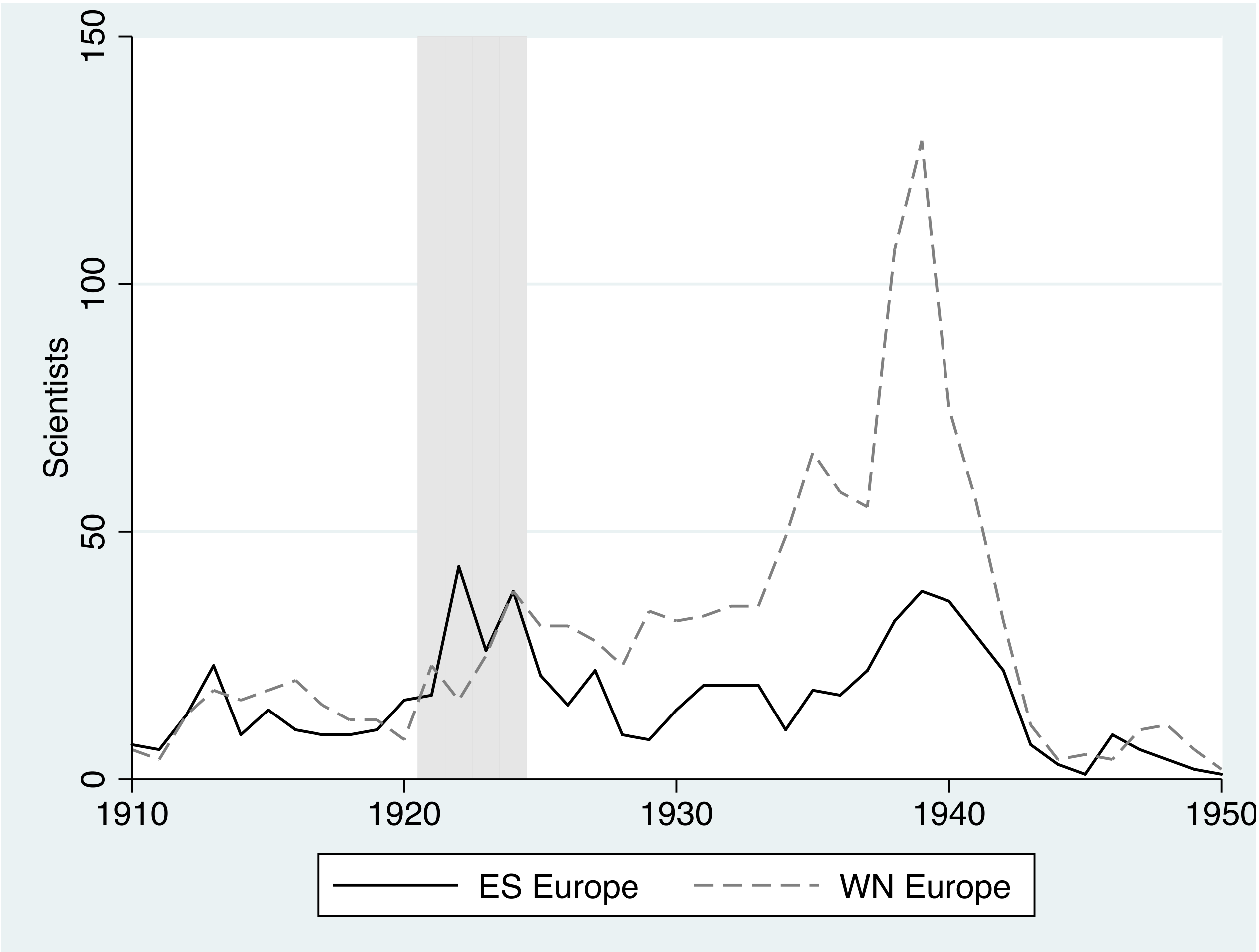
Notes: Average patents per scientist for scientists in the physical sciences who are x-year old at the year of the patent application. Patents matched on first, middle, and last names. The probability of a name is the calculated by multiplying the probability of the first name by the probability of the last name. Data on surname frequencies are from Census 2000 data contain surnames occurring 100 or more times. Data on first names frequencies are from the Social Security Administration for U.S. people born from 1880 to 2013 contain names occurring 5 or more times in each year.

FIGURE A6 – ESTIMATING THE TYPE I ERROR RATE USING THE SCIENTIST'S AGE



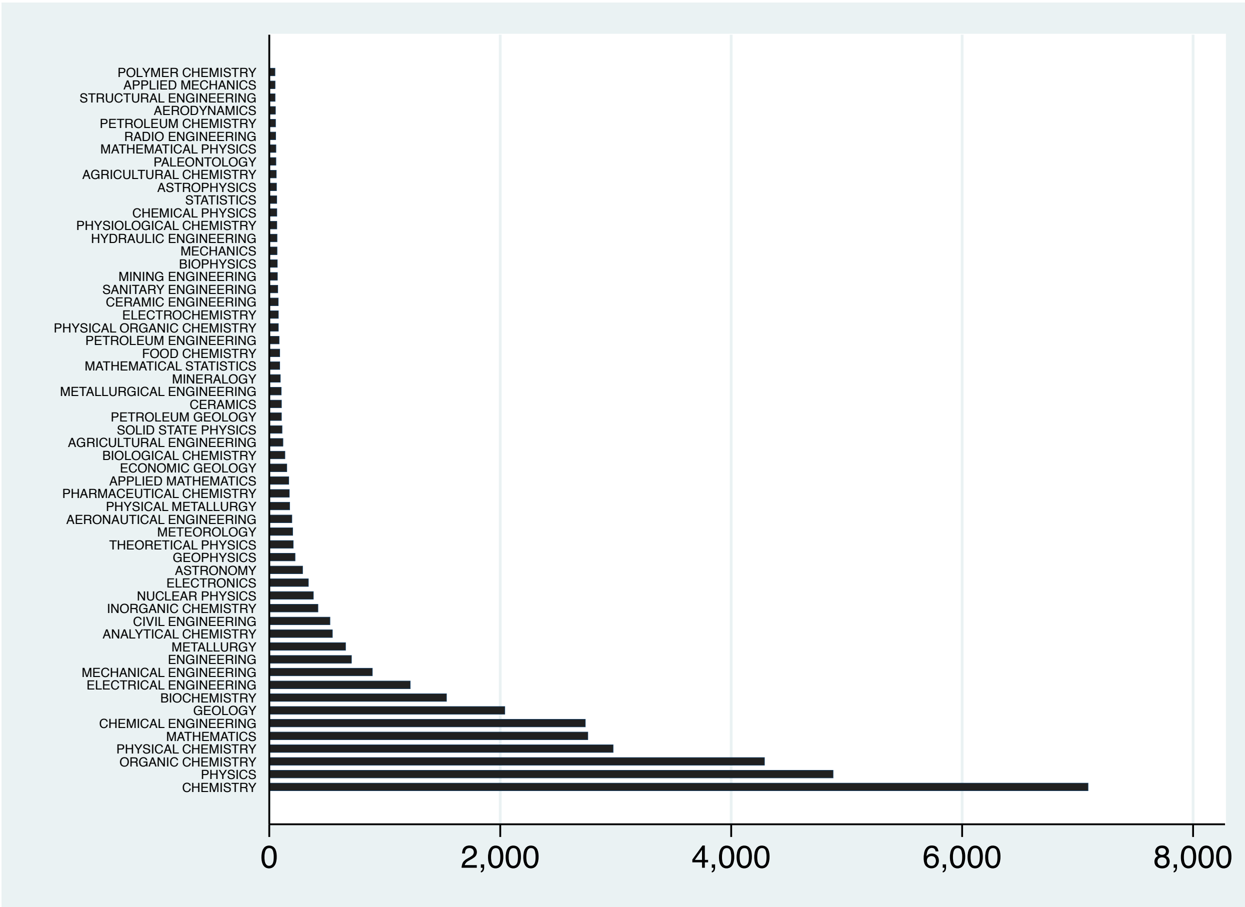
Notes: We use information on the age of the scientist in the year of the patent filing to filter out patents that are very likely to be false positive matches, because the scientist was between 0 and 18 years old in the year of the patent filing. Under the assumption that false positives are distributed uniformly across the life cycle of an inventor, we can use this information to calculate the false-positive (type I) error rate as the share of false-positives in the total number of matches between scientists and their patents.

FIGURE A7 – ARRIVALS OF NATURALIZED AMERICAN SCIENTISTS FROM ES VS. WN EUROPE



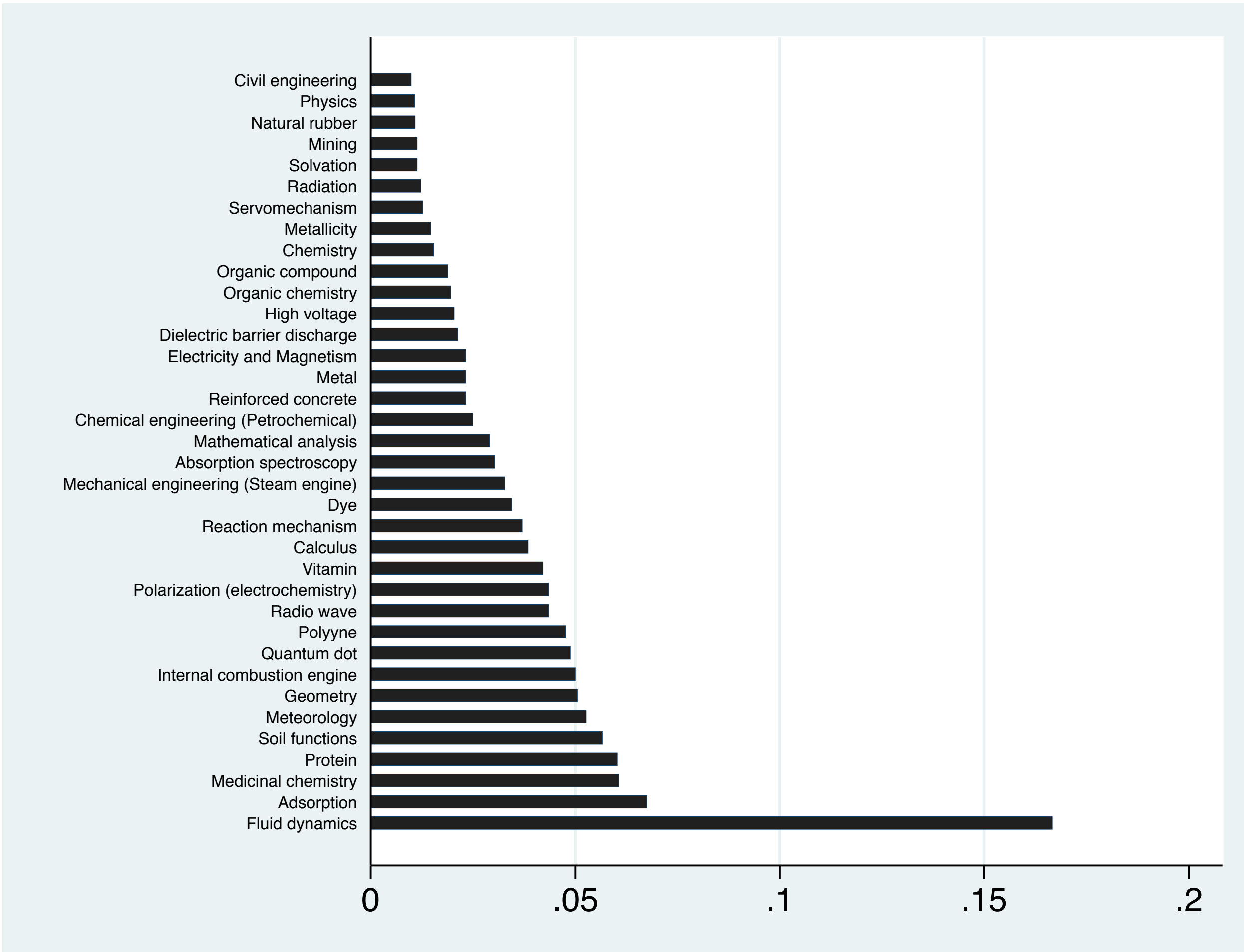
Notes: Arrivals per year of American citizens who are naturalized citizens, ESE-born scientists compared with scientists born with WN Europe. Years of arrivals are proxied by subtracting five years (the number of years required between immigration and naturalization) from the year when a scientist became a naturalized citizen. Data on the year of naturalization are available for 2,775 citizens of the United States.

FIGURE A8 – COUNT OF SCIENTISTS IN THE PHYSICAL SCIENCES IN 1956 BY ORIGINAL FIELD



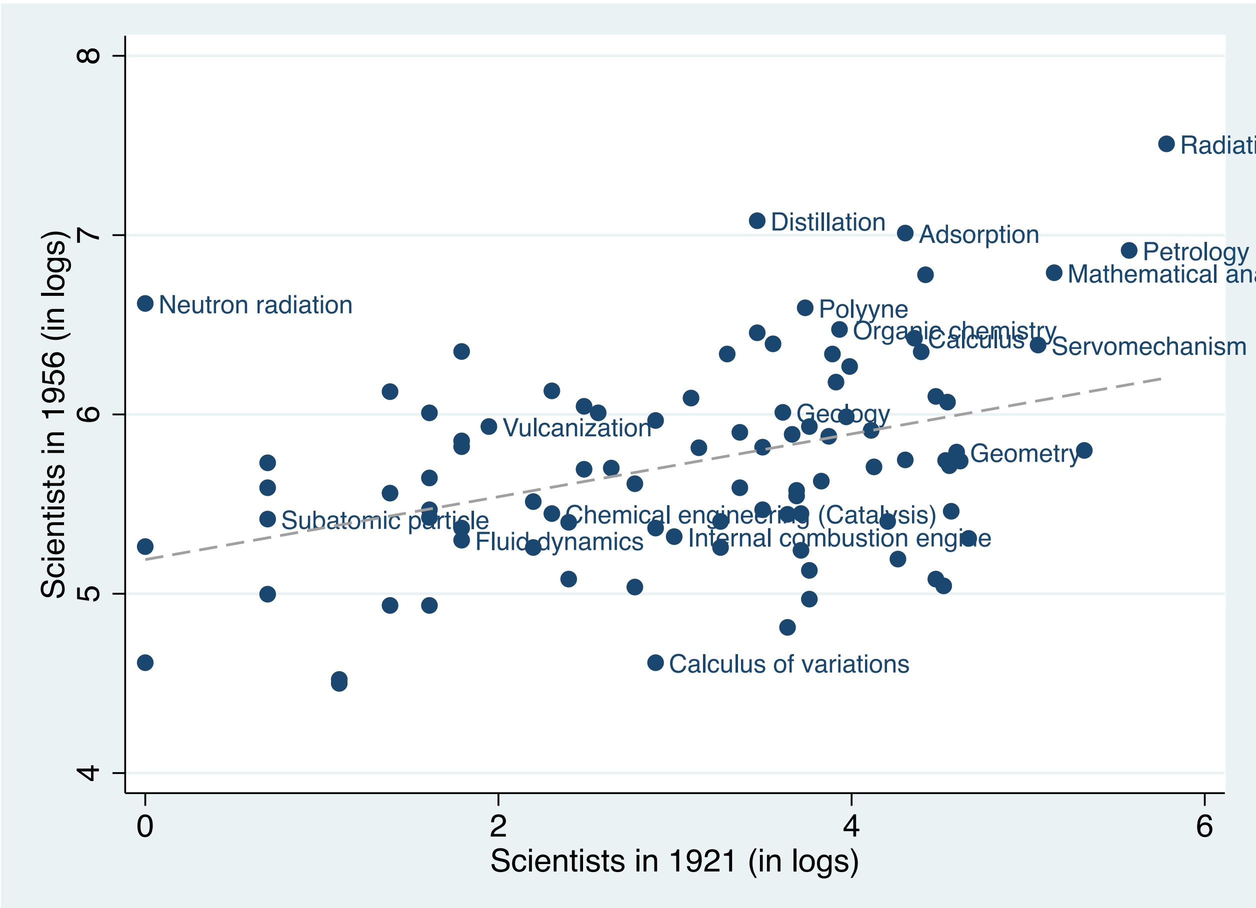
Note: Original MoS 1956 fields with 50 or more scientists. Additional 724 small fields are not presented in this figure.

FIGURE A9 - SHARE OF ESE SCIENTISTS PER FIELD IN 1921



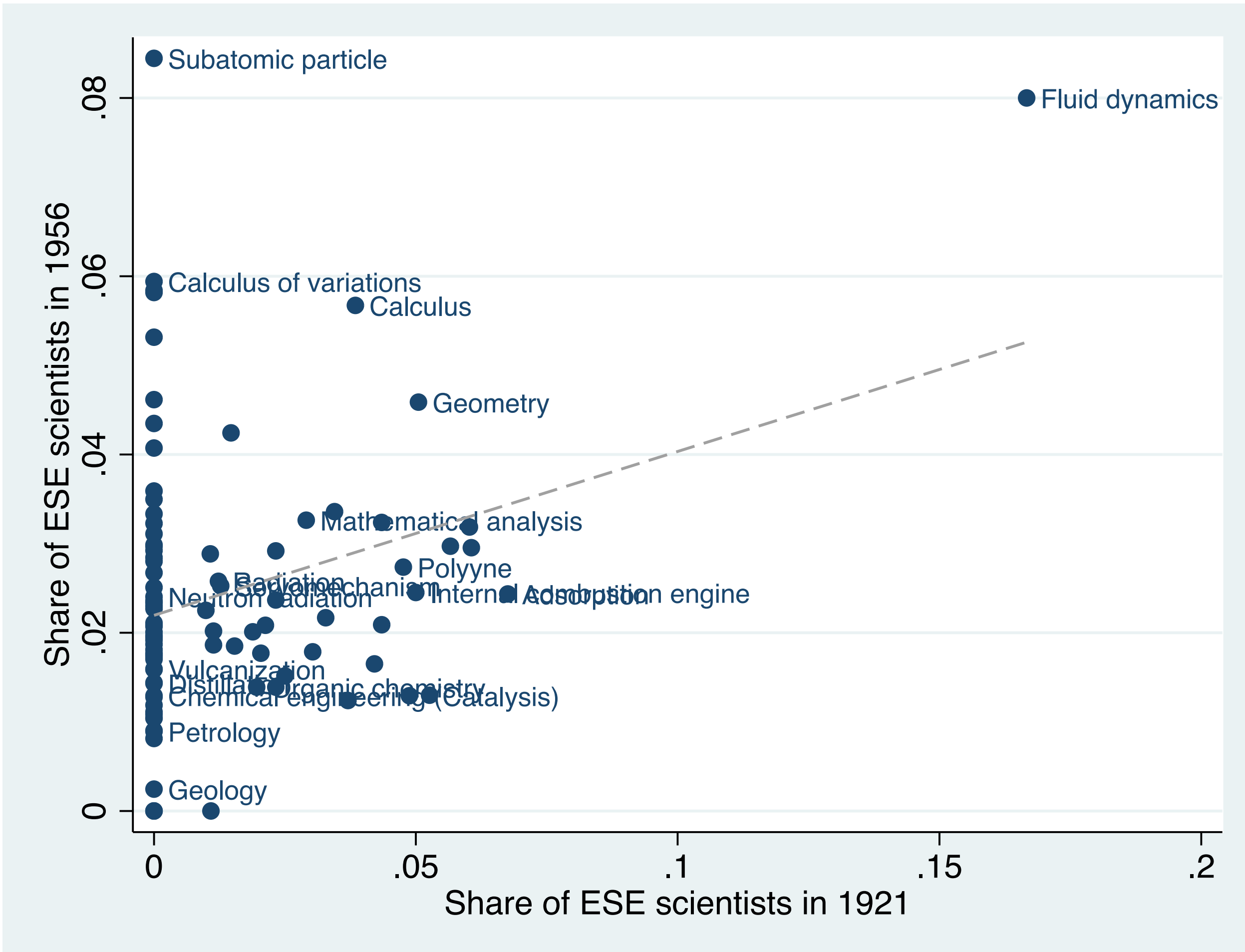
Note: The share of ESE scientists in 1921 for 36 fields with at least one ESE scientists in 1921. We call these fields "ESE fields" and use them as the treatment group in the main analysis. The control group consist of 59 fields that exist but have no ESE scientist in 1921. Five "new" fields have no scientists in 1921: "Solid-state chemistry", "Electronic engineering", "Aircraft", "Polymer" and "Nylon". We exclude them fields in the main specifications and include them in robustness checks.

FIGURE A10 - COUNTS OF SCIENTISTS PER FIELD IN 1921 AND 1956



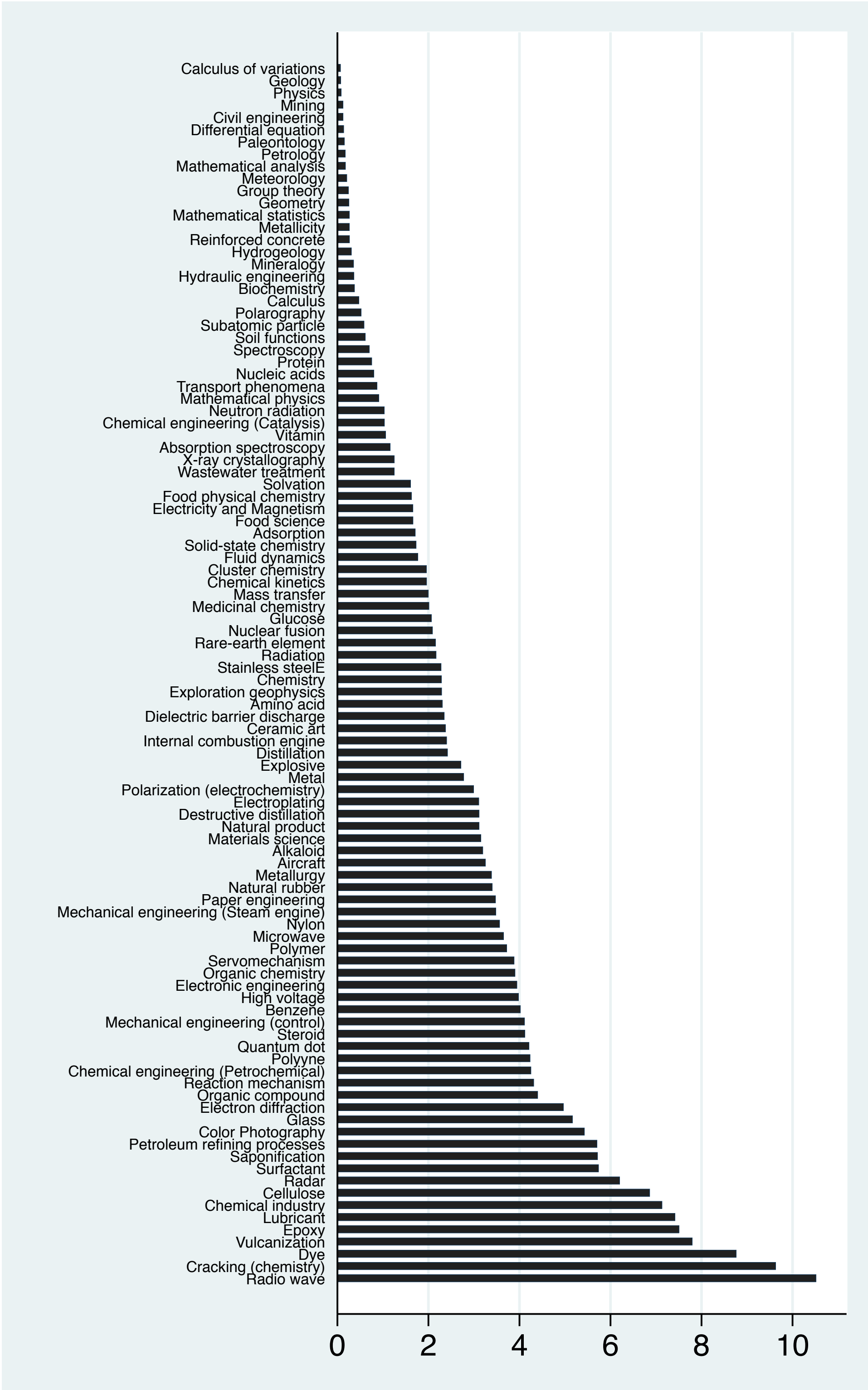
Note: Count of scientists in 1921 and 1956 in logs, excluding the residual cluster.

FIGURE A11 - SHARE OF ESE SCIENTISTS PER FIELD IN 1921 AND 1956



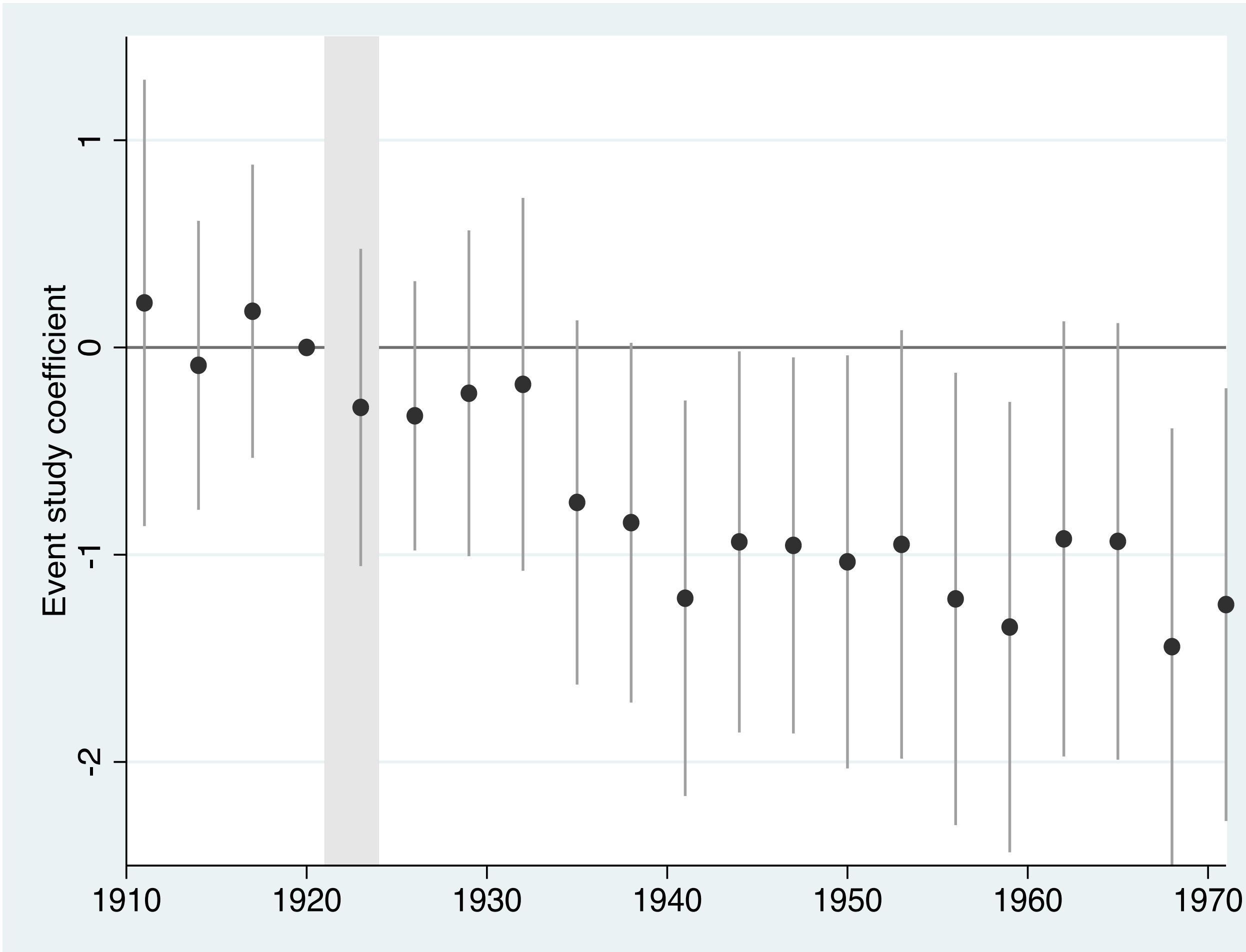
Notes: Share of ESE scientists per field in 1921 and 1956.

FIGURE A12 –PATENTS PER SCIENTIST BY FIELD



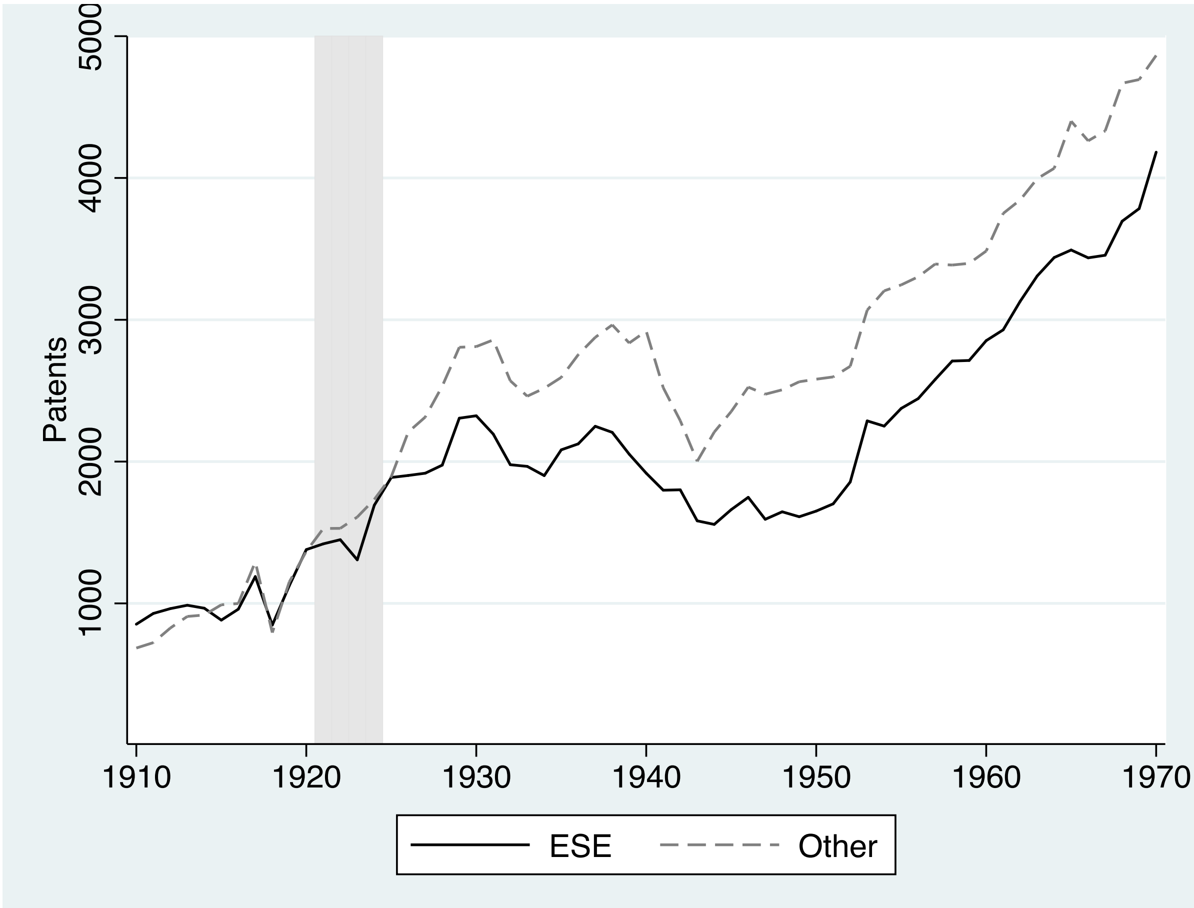
Notes: Overall number of patents per scientist by field for scientist in 1956.

FIGURE A13– TIME-VARYING EFFECTS OF THE QUOTAS ON AMERICAN INNOVATION, CONTROLLING FOR PRE-TRENDS



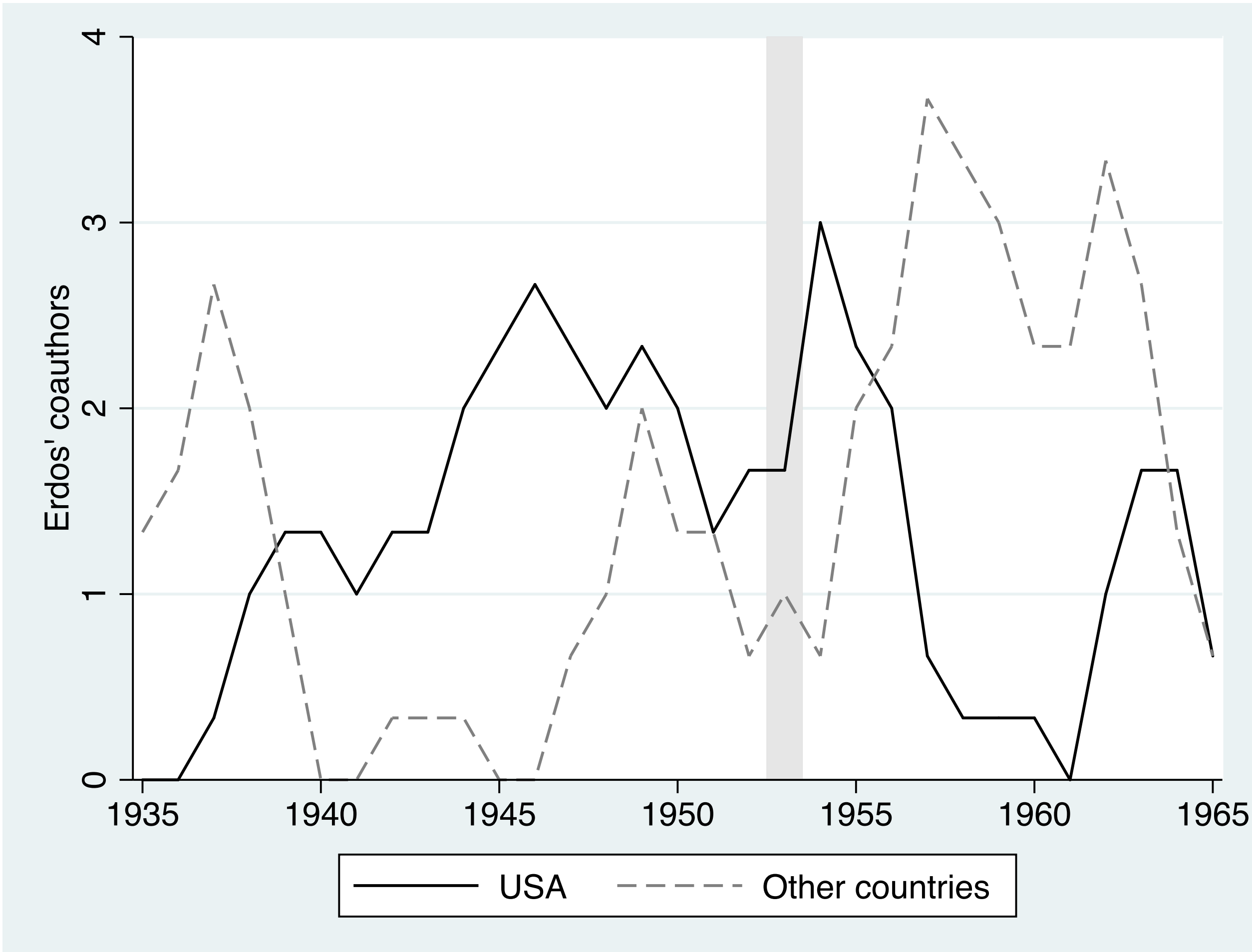
Notes: : OLS estimates of $\ln(y_{it}) = \beta_t ESE_i + t \cdot \eta_i \cdot \mathbb{I}(t \leq 1924) + \gamma_i + \delta_t + \epsilon_{it}$.

FIGURE A14 –- PATENTS PER YEAR IN ESE AND OTHER FIELDS, AGGREGATE US INVENTION



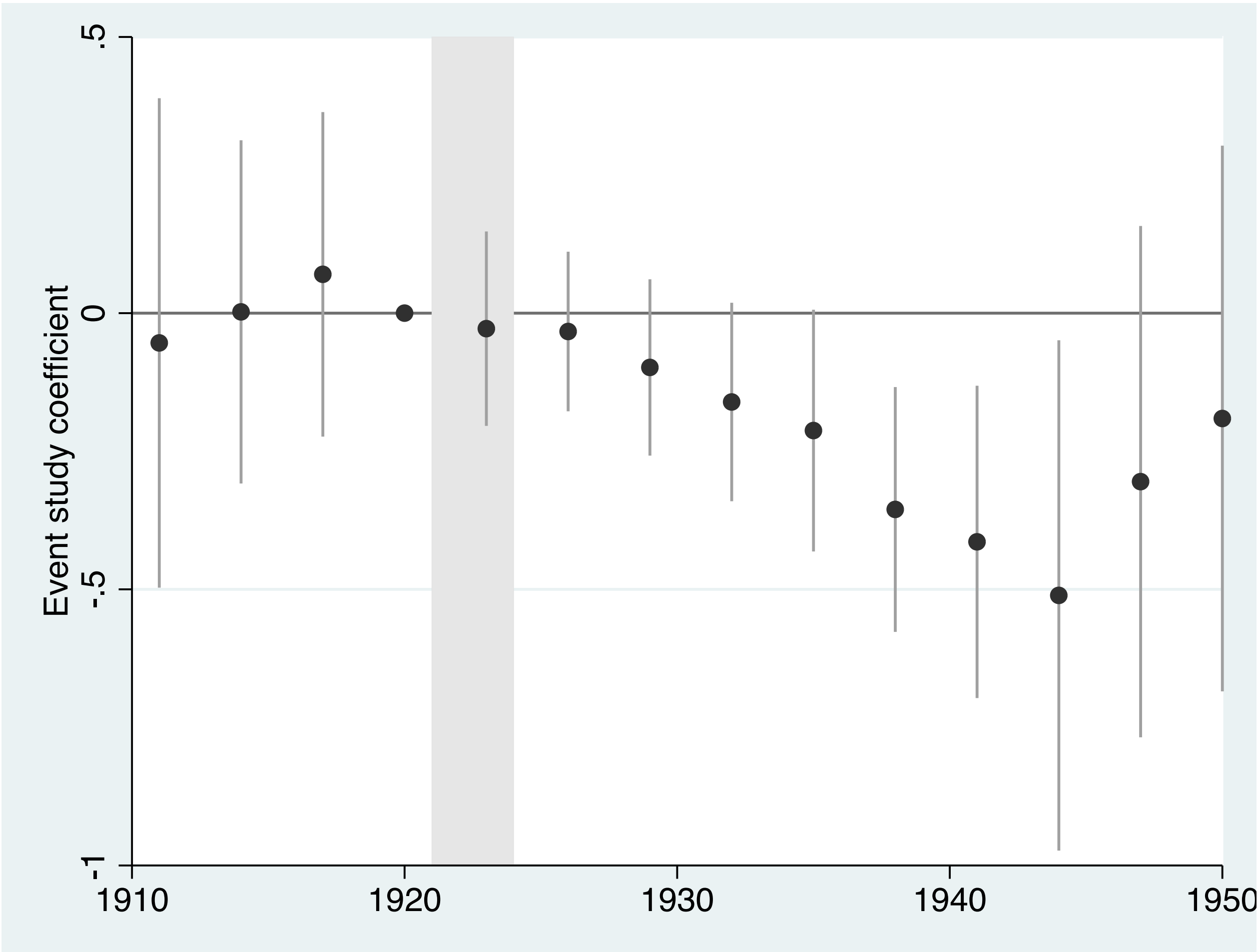
Notes: Classification to fields by titles of patents.

FIGURE A15– ERDOS' COAUTHORS BY YEAR AND COUNTRY OF FIRST JOINT PUBLICATION



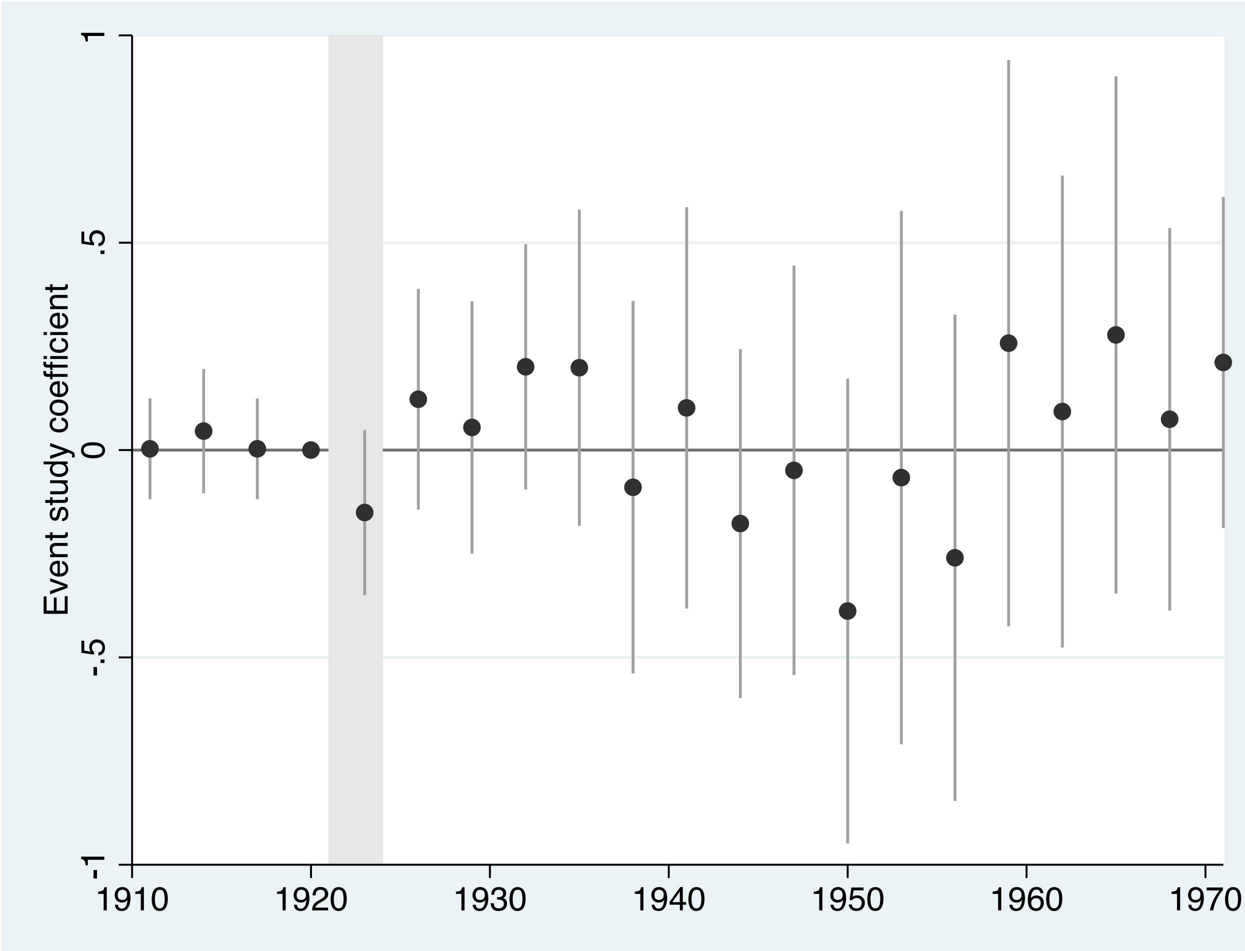
Notes: 3 years moving average.

FIGURE A16 –TIME-VARYING EFFECTS ON THE NUMBER OF AMERICAN SCIENTISTS



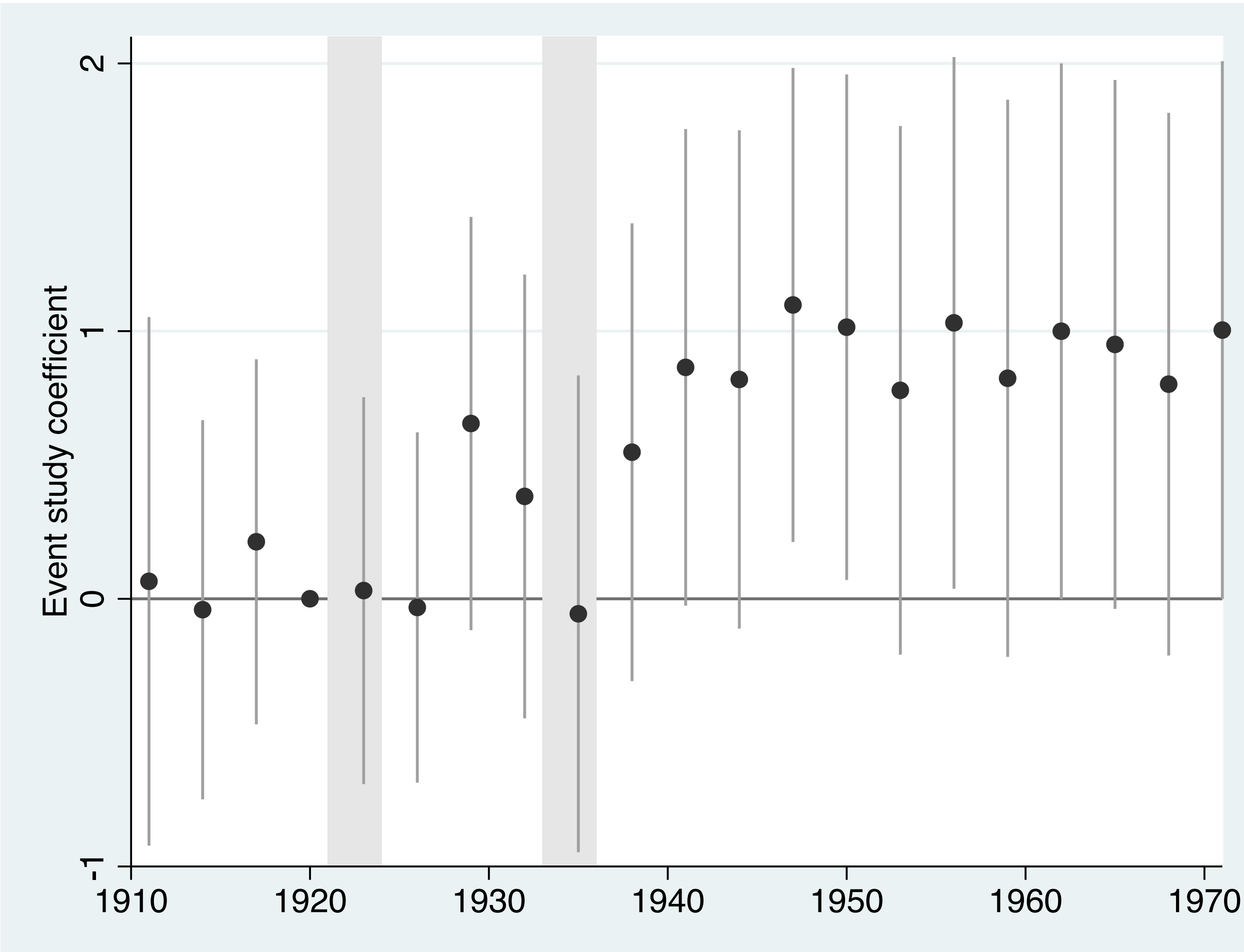
Notes: Time-varying estimates of β_t in the OLS regression $\ln(y_{it}) = \beta_t ESE_i + \gamma_i + \delta_t + \epsilon_{it}$ where $\ln(y_{it})$ is the natural logarithm of the number of scientists started to study at US institution at year t. The variable ESE_i indicates the research fields of ESE scientists in 1921, and γ_i and δ_t are field and year fixed effects, respectively, and 1918-1920 is the excluded period. Standard errors are clustered at the field level.

FIGURE A17 – TIME-VARYING EFFECTS OF THE QUOTAS ON CANADIAN INVENTION



Notes: OLS estimates of $\ln(y_{it}) = \beta_t ESE_i + \gamma_i + \delta_t + \epsilon_{it}$ by scientists worked in Canada in 1956.

FIGURE A18– CHANGES IN INVENTION BY AMERICAN SCIENTISTS IN FIELDS OF GERMAN-JEWISH ÉMIGRÉS



Notes: OLS estimates of the regression $\ln(y_{it}) = \beta_t German_i + \gamma_i + \delta_t + \epsilon_{it}$ where *German* is an indicator variable for fields who have an above median share of German or Austrian scientists in 1956.