RETURN INNOVATION: THE KNOWLEDGE SPILLOVERS OF THE BRITISH MIGRATION TO THE UNITED STATES, 1870–1940 *

Davide M. Coluccia[†] Gaia Dossi[‡]

This Version: May, 2025 Latest version available here

Abstract

The cross-country diffusion of technology is a central driver of productivity growth and economic convergence. We document that international migration fosters technology transfer from the destination to the origin country of migrants. We construct an individual-level dataset linking four million British immigrants in the US over 1870–1940 to the UK census and digitize the universe of British patents granted in the 19th century. Double- and triple-difference designs and text analysis applied to patents reveal that migration ties enabled the diffusion of US technology to the UK. Migrants' social networks at home promoted technology diffusion even when migrants did not return.

Keywords: Age of Mass Migration, Innovation, Networks, Out-migration. **JEL Classification:** F22, N73, N74, O15, O31, O33.

* We are grateful to Mara Squicciarini and Joel Mokyr for their continued guidance and support. For helpful discussions, we thank Jan Bakker, Enrico Berkes, Maristella Botticini, Jeremiah Dittmar, Joe Ferrie, Sofia Franchini, Carola Frydman, Simon Görlach, Rachel Griffith, Walker Hanlon, Leander Heldring, Alessandro Iaria, Wolfgang Keller, Matteo Leombroni, Nicola Limodio, Paolo Masella, Filippo Mezzanotti, Nathan Nunn, Alessandro Nuvolari, Laura Ogliari, Sebastian Ottinger, Eleonora Patacchini, Marta Prato, Nancy Qian, Joshua Rosenbloom, Paola Sapienza, Carlo Schwartz, Marco Tabellini, John Van Reenen, Fabian Waldinger, and Yanos Zylberberg. We thank audiences at the EHA (Pittsburgh), EHS (Cambridge), and ES (Berlin) conferences, the 10th CEPR Economic History Symposium (Geneva), the 1st CEPR Applied Micro-Economic History Workshop (Mannheim), and seminar participants at Bocconi, Bristol, Cologne, Lugano, Milan, Northwestern, Southampton, Tilburg, Torino, and Uppsala. Antonin Bergeaud, Enrico Berkes, and Walker Hanlon kindly shared data. We acknowledge financial support from Bocconi, Fondazione Invernizzi, LSE, and POID. All errors are our own.

[†]School of Economics, University of Bristol. Email: davide.coluccia@bristol.ac.uk. [‡]Stone Centre, Department of Economics, University College London. Email: g.dossi@ucl.ac.uk.

I INTRODUCTION

The diffusion of new technologies across countries is a central driver of productivity growth and economic convergence (Comin and Hobijn, 2010).¹ Eaton and Kortum (1999), for example, estimate that 70% of productivity growth in Western Europe in the 1980s relied on technology developed in the United States and Japan. In this paper, we study out-migration as a novel channel for technology transfer from migrants' destination countries to their countries of origin.² We refer to this phenomenon as "return innovation."

Out-migration is a widespread feature of economic development (Clemens, 2020). Since migrants typically move to technologically advanced countries, they are ideally positioned to transfer technologies back to their *sending* regions, as documented qualitatively by Saxenian (2006). At the same time, the "brain drain" resulting from out-migration may deplete human capital and hinder innovation activities (Gibson and McKenzie, 2011). Despite this theoretical tension, existing studies have predominantly focused on the impact of migrants on innovation in *destination* countries.³ This gap is motivated primarily by data constraints. Moving beyond cross-country analyses requires information on where migrants settle within the receiving country and their origin area within the sending country, but such data are seldom available.

To overcome this limitation, we link English and Welsh migrants in the US Census between 1850 and 1940 to their records in the UK Census.⁴ This new dataset allows us to observe migrants' origin areas in the UK and their destinations in the US. Our main contribution is to show that higher US emigration rates promoted UK innovation and increased the similarity between British and American patents, especially in technological fields prevalent in the areas where migrants settled. About half of this "return innovation" effect is driven by returnees. Migrants' social networks, however, fostered cross-border technology flows even in the absence of return migration. Our results uncover that outmigration can act as a source of technology transfer to emigration countries.

¹Recent endogenous growth models include cross-country technology diffusion dynamics and quantify its implications for economic growth (e.g., Alvarez, Buera and Lucas, 2018; Buera and Oberfield, 2020; Perla, Tonetti and Waugh, 2021).

²This paper is therefore related to an important literature investigating the role of technology sourcing, trade, and foreign direct investments as drivers of cross-country technology transfer (Keller, 2004).

³Burchardi, Chaney, Hassan, Tarquinio and Terry (2020) and Bernstein, Diamond, Jiranaphawiboon, McQuade and Pousada (2022) provide causal evidence that immigration benefits US innovation. Moser, Parsa and San (2020) document that foregone immigration after the 1921 border closure hampered American innovation. Akcigit, Grigsby and Nicholas (2017a) and Arkolakis, Lee and Peters (2020) show that historical immigration enabled the United States' rise to the technology frontier.

⁴Throughout the paper, we focus on England and Wales. With a slight abuse of language, we use the terms "Britain" and "United Kingdom" (UK) as shortcuts to collectively refer to England and Wales.

Between 1850 and 1940, a period known as the "Age of Mass Migration," approximately 30 million European migrants settled across the Atlantic. Nearly four million came from England and Wales. Over this period, the US reached the technology frontier in many industries, from engines to agricul-tural machinery (David, 1966; Rosenberg, 1970).⁵ In this paper, we document that migration networks promoted the diffusion of these technologies back to Britain, the cradle of the Industrial Revolution.

We assemble two novel general-purpose datasets to overcome the limitations of existing sources. First, the available data do not contain information on the origin of British immigrants in the US beyond their country of origin. We thus use confidential individual-level UK and US census data to link the records of British immigrants in the US to the UK census (Schurer and Higgs, 2020; Ruggles, Fitch, Goeken, Hacker, Nelson, Roberts, Schouweiler and Sobek, 2021). This dataset enables us to track individual migration between the UK and the US, as well as return migration. Second, to reconstruct the geography of innovation in the UK in the second half of the nineteenth century, we digitized all 300,000 original patent documents issued in England and Wales between 1853 and 1899. This dataset allows us to cover the universe of patents granted over nearly a century, from 1853 to 1940.

We construct two measures of exposure to US technology through migration ties. First, we look at the number of US emigrants across UK districts.⁶ The simple rationale is that a larger emigrant community is more likely to be exposed to innovation and transmit it back to the UK. Second, we leverage joint variation in county-level specialization across technology classes in the United States and migration flows between UK districts and US counties to construct a technology-specific indicator of "knowledge exposure." In this case, exposure to US innovation is higher for UK districts with larger overseas emigrant communities who settle in counties that are more innovative in specific technologies.

Identifying the impact of migrants' exposure to US technology on innovation in the UK presents one key challenge: assortative matching. British migrants could sort across US counties in ways that are correlated with both the innovation of their district of origin and that of their county of destination.

To address this challenge, we isolate sudden increases in patenting activity—innovation "shocks"—

⁵By the 1890s, the American technological primacy was well established. Nelson and Wright (1992) note that, starting in the 1880s, American technology saw major advances in textiles, sewing machines, clocks, firearms, boots and shoes, locomotives, bicycles, and cigarettes. From the 1890s, mass production led to innovations in consumer products (canned goods, dairy, and grain products), light machinery (typewriters, cameras), electrical equipment, and industrial machinery such as boilers, pumps, and printing presses.

⁶Our core units of observation are UK registration districts and US counties. In 1901, there were 631 registration districts in England and Wales. Districts were comparable to US counties in population (approximately 40,000). Unlike counties, however, registration districts were statistical entities without political or budgetary autonomy.

in US counties and county-technology pairs. Using predetermined migration flows, we classify UK districts as "treated" if a relatively large number of their emigrants had settled in US areas experiencing these shocks. The key insight is that unexpected innovation shocks provide plausibly exogenous variation in UK districts' exposure to US technology via pre-existing migration linkages. Using double- and triple-difference regressions, we compare innovation outcomes across exposed and unexposed districts and technologies, before and after these shocks, to estimate how US technology shocks diffused into the UK through migration ties. The validity of this approach rests on a standard parallel trends assumption: in the absence of innovation shocks, treated and untreated units would have followed similar innovation trajectories. The event-study estimates provide consistent evidence supporting this assumption.

The main finding of this paper, which we label "return innovation," is that technological change in the UK is affected by exposure to US innovation through out-migration linkages. First, we document an increase in total patenting in districts with more US emigrants: moving from the 25th to the 75th percentile of US out-migration is associated with an 8% increase in the number of patents produced in the United Kingdom. Second, districts increase their innovation activity in the technological fields to which emigrants are exposed in the United States. Using our index based on joint variation in the location where emigrants settle and the innovation they are exposed to, we find that moving from the 25th to the 75th percentile of the "knowledge exposure" variable is associated with an 11% increase in innovation activity. Third, we adopt a text-analysis methodology that computes the textual similarity between UK and US patents to measure the transfer of innovation generated by migration ties. We find that patenting in districts with more US emigrants becomes more similar to US patents: moving from the 25th to the 75th percentile of emigration to the US is associated with UK patents that are twice as similar to American ones. This pattern also holds within technological fields.

In the double differences regression, innovation in districts exposed to a US innovation shock through migration ties increases by 9%. A back-of-the-envelope calculation suggests that migration ties generate a 15% pass-through rate of innovation shocks from the United States into the United Kingdom. In exposed districts, patents become substantially more similar in their textual content to American patents. We construct the text-based measure of patent "impact" developed by Kelly, Papanikolaou, Seru and Taddy (2021) for British patents and find that exposure to US technology further promoted the production of high-impact innovation in the years following the innovation shocks. These results also hold in the triple differences framework, which leverages shocks to US innovation activity across counties and technologies. A key advantage of the triple differences approach is that it allows us to include richer fixed effects that control for, among others, district-level time-varying unobserved heterogeneity and fixed district-technology specialization. In support of the standard parallel trends

assumption, we estimate a set of event-study regressions that indicate that treatment and control units were statistically similar before the exposure to the shock.

The effects of exposure to foreign innovation are heterogeneous across sectors. In particular, we uncover an inverted U-shaped relationship between return innovation and the specialization of the UK relative to the US. In response to emigrants' exposure to US innovation, patenting in the UK increases the most in sectors where the US and the UK have similar specialization rates. By contrast, the effect is considerably smaller in technologies where the US is more advanced, such as agriculture and metallurgy, or where the UK is more advanced, such as textiles. This pattern is consistent with the theoretical predictions of Van Patten (2023), who argues that the gains from international technology diffusion are highest when a country interacts with partners with higher but similar levels of development.

In the second part of the paper, we leverage the richness of our data to explore the mechanisms behind the return innovation effect. On the one hand, return innovation may require the physical return of migrants. On the other hand, migration ties may promote the diffusion of new technologies irrespective of physical return. We find that the physical return of migrants is an important driver of return innovation, but it accounts for only about half of the overall effect. The impact of exposure to US innovation through out-migration remains sizable and significant even in the absence of return migration. This finding suggests that migration ties contribute to the diffusion of knowledge independently of physical return

To provide more evidence in this direction, we study the role of interactions between emigrants and their communities of origin as a driver of the diffusion of innovation. We focus on two factors that could promote such interactions: family ties and geographical proximity. Our analysis builds on a large literature in development economics that links the diffusion of technology to network interactions (e.g. Bandiera and Rasul, 2006; Conley and Udry, 2010; Beaman, BenYishay, Magruder and Mobarak, 2021).

We document that family members—surname groups within small geographical areas—of emigrants increase their patenting activity after their relatives move to the US. It takes about three to five years for a British emigrant to contribute to innovation activity back home. Despite this delay, the magnitude of the effect is large. Importantly, we distinguish between emigrants who, at some point, return to the UK from those who do not. The impact of return emigrants is larger than those who never return. However, emigrants promote innovation within their families even if they never return. Since return emigrants account for approximately one-third of the entire migrant stock, the magnitudes of the two effects are, in aggregate, similar.

As an alternative proxy for local social networks, we leverage the geographical proximity between emigrants and their former neighbors. To estimate its impact on the innovation activity of stayers, we link the UK patent data to individual UK census records, and we geo-code the universe of the UK population. We find that patenting activity increases for non-migrants after their neighbors migrate to the United States. The estimated effect remains positive and significant even when we exclude the neighbors of return migrants, suggesting that cross-country interactions between emigrants and their origin communities are a key driver of return innovation.

This paper provides new evidence on how knowledge diffuses across countries. We document that exposure to foreign innovation through migration ties contributes to the diffusion of innovation from migrants' destination countries back to their countries of origin. Our results, therefore, suggest that out-migration can promote innovation by fostering the diffusion of knowledge into emigration countries.

Related Literature. This paper is related to four streams of literature. First, an important literature studies technology diffusion within and across countries (among others, see Jaffe, Trajtenberg and Henderson, 1993; Bahar, Hausmann and Hidalgo, 2014). Early contributions document the role of foreign direct investments (Aitken and Harrison, 1999), trade (Keller, 2002), and technology sourcing (Griffith, Harrison and Van Reenen, 2006) as drivers of international technology diffusion. Specifically, we relate to the papers exploring the relationship between human mobility and the diffusion of knowledge across countries (Kerr, 2008; Hornung, 2014; Bahar, Hauptmann, Özgüzel and Rapoport, 2019; Bahar, Choudhury, Sappenfield and Signorelli, 2022; Bahar, Choudhury, Miguelez and Signorelli, 2024; Prato, 2025). The primary focus of the literature is on receiving countries, whereas we study the sending country, where the effects of migration on innovation are theoretically ambiguous. Additionally, existing studies mainly look at the highly selected group of migrant inventors, whereas our dataset allows us to examine the universe of emigrants. This distinction is relevant because inventor migration does not fully capture the "brain drain" effect. Third, we find that migrants who are not inventors operate technology transfer. Hence, focusing on inventors likely understates the return innovation effect. Finally, the literature focuses on return migration as the key mechanism underlying technology diffusion, primarily because of data constraints. Our data, instead, allow us to document that migrants' social networks promote knowledge diffusion in the absence of return.

Second, we contribute to the literature on the determinants of the direction of innovation. Pioneering work on directed technical change was formalized by Acemoglu (2002, 2010). More recently, the question has been studied both theoretically (e.g., Bryan and Lemus, 2017; Hopenhayn and Squintani, 2021; Acemoglu, 2023) as well as empirically (e.g., Hanlon, 2015; Moscona, 2021; Moscona and Sastry, 2023; Einiö, Feng and Jaravel, 2023; Dossi, 2024; Truffa and Wong, 2025). We add to this literature by

introducing a novel determinant of the direction of innovation, namely, international human mobility, through the return innovation effect. We document that out-migration fosters innovation in sectors and technologies that are similar to those migrants are exposed to in the US, thereby shaping the direction of innovation in the United Kingdom.

Third, existing studies explore the effects of out-migration on sending countries. Emigration has been shown to affect wages (e.g., Dustmann, Frattini and Rosso, 2015), attitudes toward democracy (Spilimbergo, 2009; Batista and Vicente, 2011) and political change (Barsbai, Rapoport, Steinmayr and Trebesch, 2017; Karadja and Prawitz, 2019), technology adoption (Coluccia and Spadavecchia, 2024), entrepreneurship (Anelli, Basso, Ippedico and Peri, 2023), and social norms (Beine, Docquier and Schiff, 2013; Bertoli and Marchetta, 2015; Tuccio and Wahba, 2018). We provide new evidence that emigration shapes the rate and direction of innovation in the sending countries by increasing their exposure to knowledge developed abroad. Our findings add to the existing literature, which generally interprets out-migration as a depletion of the stock of human capital.

Finally, by its setting, this paper contributes to the literature on innovation and technology diffusion during the Age of Mass Migration. Several papers examine the short-run (e.g., Moser, Voena and Waldinger, 2014; Arkolakis et al., 2020; Moser et al., 2020) as well as the long-run (e.g., Akcigit, Grigsby and Nicholas, 2017b; Burchardi et al., 2020) implications of immigration for US innovation. Our work is closest to Andersson, Karadja and Prawitz (2022), who show that mass out-migration from Sweden triggered labor-saving innovation by increasing the relative cost of labor. In this paper, we document a different phenomenon–the "return innovation effect"—whereby emigrants generate knowledge flows from their destinations back to their areas of origin. Our findings show that the European mass migration to the US facilitated the diffusion of American technology into Europe.

Outline. The rest of the paper is structured as follows. In Section II, we describe the historical and institutional context. In Section III, we present the data. In Section IV, we discuss the research design. In Section V, we document the main findings. In Section VI, we discuss the mechanisms and possible alternative interpretations. In Section VII, we conclude.

II HISTORICAL BACKGROUND

In this section, we provide details on the English and Welsh migration to the United States (II.A) and on intellectual property protection in the United Kingdom and in the United States over this period (II.B).

II.A The English and Welsh Migration to the United States

Between 1850 and 1920—the so-called Age of Mass Migration—more than 30 million Europeans migrated to the United States (Abramitzky and Boustan, 2017). Migrants from England and Wales accounted for approximately 10% of the total (Willcox, 1928), and emigration rates in Britain were among the highest in Europe (Baines, 2002).⁷ In this section, we provide details on the English and Welsh migration to the US.

II.A.1 Migration Policy in the United Kingdom and the United States

Until 1917, the US applied minor restrictions on European immigration (Abramitzky and Boustan, 2017). As a result, the Age of Mass Migration was characterized by virtually no legal constraints on human mobility to the US. Until 1900, immigrants mostly originated from Northern Europe, particularly the United Kingdom, Ireland, Germany, Sweden, and Norway. This positive attitude towards immigration started to decline in the 1890s as flows from Eastern and Southern Europe increased. The restrictive immigration policies of the 1920s, however, allotted generous quotas to the United Kingdom, which were never filled (Abramitzky and Boustan, 2017).⁸

Like in other European countries, out-migration legislation in the UK encouraged emigration towards the Empire (Baines, 2002).⁹ Emigration to the United States was neither subsidized nor discouraged. Attitudes towards out-migration remained positive after World War I, and the perceived slowdown of emigrant flows after the War was viewed with concern by policymakers (Leak and Priday, 1933).

II.A.2 The Characteristics of British Emigrants

The British migration to the US presents two main distinctive features compared to the broader European phenomenon.¹⁰ First, unlike most continental countries, Britain was already highly urbanized

⁷Only Ireland, Italy, and Norway had higher emigration rates, although, in England, massive out-migration lasted longer than in the other countries above.

⁸The 1921 (resp. 1924) Act computed the quota for a given country as 3% (resp. 2%) of the population from that country that was recorded in the US census in 1910 (resp. 1880). This scheme favored first-wave immigration countries, such as the United Kingdom and Germany, at the expense of new ones, as recommended by the Dillingham Commission (Higham, 1955).

⁹Out-migration was encouraged in two ways: reduced and subsidized ticket fares and allotment of agricultural lands. Policy efforts were directed towards the Empire, particularly Canada and Australia, through the Committee of the Emigrants' Information Office. In general, however, these policies were not successful. Baines (2002) argues that less than 10% of emigrants traveled under government assistance during the entire 1814-1918 period, and Leak and Priday (1933) report similar figures for the post-War era.

¹⁰Throughout the period, the US was the most frequent destination for English and Welsh migrants. Between 1850 and 1930, more than 40% of total British and Welsh emigrants settled in the US. This figure compares to 25% in Canada, 20% in

and industrialized at the inception of the Mass Migration.

Second, the selection of British migrants differed from that in continental Europe (Erickson, 1957). Compared to the occupational structure of Great Britain, migrants were less likely to be employed in agriculture and more likely to be low and high-skilled industrial workers (Baines, 2002, p. 83). Until the 1880s, British emigrants generally came from rural areas and, consequently, the vast majority were farmers. However, as cities and smaller urban centers gained prominence, migrants started being increasingly employed in industrial manufacturing occupations. Starting in the 1880s, urban areas supplied most overseas migrants (Erickson, 1957, 1972; Thomas, 1954).

At the beginning of the 1860s, when the transatlantic migration was taking off, about 15% emigrants were employed in agriculture, and only five percent were white-collar workers. In the early 1900s, this composition had shifted as agricultural workers accounted for five percent of the overall emigrant stock, while those employed in white-collar occupations were 25%.

II.B Intellectual Property Protection in the United Kingdom and in the United States

During the period of our analysis, patents in the United Kingdom and the United States were granted by the respective patent offices. However, the UK and the US did not mutually recognize patents. In this section, we describe patent protection in the two countries, as well as international intellectual property protection.

II.B.1 Patent System in the United Kingdom

The United Kingdom established the world's oldest continuously operating patent system in 1623-1624. However, access to intellectual property protection was difficult until 1850 (Bottomley, 2014). Fees amounted to approximately four times the average annual income in 1860, and the application process was lengthy and uncertain (Dutton, 1984). A large literature documents the poor performance of this system during the Industrial Revolution (e.g., Macleod, 1988). The 1852 Patent Law Amendment Act sought to reform this process. The reform effort was inspired by the system in place in the United States and reduced application fees and shortened the bureaucratic process. A subsequent reform in 1883 further reduced fees, permitted applications by mail, and introduced both a litigation system and professional patent examiners. A technical examination of novelty was introduced only in 1902. Until 1907, patents were granted conditional on the invention being produced in Britain (Coulter, 1991).

Australia, and 15% in other destinations (Baines, 2002).

II.B.2 Patent System in the United States

The first article of the United States Constitution establishes that inventors should be granted exclusive rights over their discoveries. In 1836, US Congress passed the Patent Act, which formally instituted the US Patent Office (USPTO). The USPTO has been credited as the first modern patent system in the world (Khan and Sokoloff, 2004). Two features distinguished the American patent system from its British counterpart. First, professional examiners carried out a novelty examination to ascertain the originality of patent applications. Second, low application fees ensured that access to intellectual property protection was widespread.

II.B.3 International Intellectual Property Protection

As national patent systems were established in Europe and in the US during the 19th century, demands for international regulation increased. The Paris Convention—formally, the "Paris Convention for the Protection of Industrial Property"—of 1883 governed international patent protection (Penrose, 1951). This agreement emerged from a decade of multilateral confrontations that started with the World Exhibitions in Vienna (1873) and Paris (1878). The Paris Convention introduced two major principles. First, nationals and residents of subscribing countries were guaranteed equality of treatment with nationals. This concept, known as "national treatment," rejects the principle of "reciprocity," which maintains that nationals in subscribing countries would be granted the same protection as their origin country. The United States strongly advocated for reciprocity (Penrose, 1951). Second, upon applying for a patent in a member country, inventors were granted a "right of priority" of six months. Patents filed in foreign countries during the priority period would not invalidate the inventor's claim for protection in other member countries. However, patents obtained in one member state were *not* automatically recognized by other countries. To effectively claim protection, inventors had to submit different patent applications, which represented a substantial bureaucratic and financial burden. The UK joined the Convention in 1884, and the US joined in 1887.

While the Paris Convention—and its numerous amendments—are still in operation today, international patents were only established in 1970. Since the UK and the US did not mutually recognize patents, we use patenting patterns as a proxy for knowledge flows between the two countries.

III DATA

To conduct the analysis, we assemble two novel datasets. The first links British migrants in the United States to the UK census, allowing us to construct a matrix of bilateral migration flows at the UK district and US county level. This dataset provides individual-level information on the origin of migrants and allows us to build a dataset of return migrants. We describe this dataset in section III.A. The second

dataset consists of the universe of patents granted in England and Wales between the second half of the 19th century. We digitize these data, which to date were not available in disaggregated form, and link inventors to the UK census. We describe this dataset in section III.B. In section III.C, we describe the additional data sources we use in the analysis.¹¹

III.A A Novel Individual-level Dataset of British Immigrants in the United States

Our empirical analysis requires information on the detailed origin of English and Welsh immigrants *within* the United Kingdom. The available data, however, do not contain this information, as neither the US nor the UK collected disaggregated records on the origin of immigrants and the destination of emigrants. We address this limitation by developing a new dataset that links British immigrants in the US to the British census. In this way, we are able to observe an individual in the US and to track him to his UK census record before he emigrated.¹² This is the first dataset that reconstructs migration flows at this granular level of aggregation for a non-Scandinavian European country during the Age of Mass Migration and adds to work by Abramitzky, Boustan and Eriksson (2014) and Andersson et al. (2022) on Norway and Sweden.¹³

To construct the linked dataset, we leverage confidential individual-level data from population censuses in the United Kingdom (Schurer and Higgs, 2020) and the United States (Ruggles *et al.*, 2021). We extract the universe of British male immigrants from US censuses in 1900, 1910, 1920, and 1930.¹⁴ These records contain the name and surname, birth year, and immigration year of each migrant, which we use to match these individuals to the last British census where the emigrants appeared.¹⁵ We use state-of-the-art census-linking algorithms adapted from pioneering work by Abramitzky, Boustan, Eriksson, Feigenbaum and Pérez (2021), as described in Appendix Sections A.III.1 and A.III.2, which discuss in more detail the primary sources and the technical implementation of the algorithm.

¹¹Appendix A provides a more detailed discussion of the various data sources used in this paper.

¹²Throughout the paper, we use the masculine to refer to individuals in our data because, as we explain in detail later, we can only work with male individuals.

¹³England and Wales were larger in terms of the overall population and US immigrant population. The population of Sweden and Norway in 1890 was approximately 4.7 and 2 million, while it was 27 million in England and Wales.

¹⁴We cannot use information from the 1870 and 1880 censuses because the immigration year was not recorded. Individuallevel data from the 1890 census have not survived. Following the standard practice in the literature, we only link men because women usually changed their surname upon marriage.

¹⁵For example, we link an individual who immigrated to the US in 1905 to the 1901 UK census. Because no census was taken in 1870, we match those who migrated between 1870 and 1881 to the 1860 census. Moreover, since the last available UK census was in 1911, we match all those who emigrated after 1911 to that one. Thus, we have no information on migrants born after 1911. Since the median age of migrants is 30 and less than 10% of the distribution is younger than 19 in the rest of the sample, this limitation bears little quantitative implications for the matching rate in the later part of the sample.

In our preferred version of the data, we match approximately 65% of male British immigrants who appear in the United States census, as shown in Appendix Figure A.7. The linking rate is considerably higher than benchmark datasets linking individuals over time (e.g., Abramitzky *et al.*, 2021). There are at least three potential explanations for this difference. First, British immigrants were a positively selected group of urban skilled manufacturing workers. Their census records would, therefore, plausibly be more accurate than for other populations.¹⁶ Second, British immigrants spoke English. Linguistic factors do not threaten our linking algorithm, but they are challenging to resolve in intergenerational linking routines. Finally, the high quality of UK census records facilitates the linking procedure.

This approach nonetheless presents some important caveats.¹⁷ First, it may deliver spurious links if the matching variables are insufficient to restrict the pool of potential matches. Second, the matching probability may be correlated with individual characteristics. To address the first issue, we discard the matches that do not attain a high level of string similarity. Moreover, we weigh each migrant by the inverse of the number of matches he is paired with to minimize the weight placed on false-positive matches. The results remain, however, qualitatively unchanged if we only consider single matches. To address the second potential issue, in Appendix Table A.3, we report the correlation between emigrants' observable characteristics and the linking probability. Overall, this association is small and often statistically insignificant.

It is challenging to validate the data with external sources that contain disaggregated information on the origin of British emigrants to the US because, to the best of our knowledge, they do not exist. In Appendix Figure A.9, we correlate our data with county-level estimates of aggregate transatlantic emigration for the period 1880–1910 assembled by Baines (2002) and confirm that the two are positively correlated.¹⁸

In addition, we construct a dataset of return migrants. To assemble it, we follow a similar procedure. The only difference is that migrants are matched to the UK censuses taken in the decades *after* their immigration year. As an example, someone who migrated to the US in 1895 is matched to censuses in 1901 and 1911. To avoid double-counting, if an individual is matched to more than one census, we keep the match(es) in the first.

¹⁶This pattern is in line with evidence by Helgertz, Price, Wellington, Thompson, Ruggles and Fitch (2022) that individuals with relatively higher socio-economic status have higher linking rates in the US census.

¹⁷Appendix Section A.III.3 presents a more detailed discussion of the internal and external validation exercises we undertake to assess the plausibility of the linked emigrant sample.

¹⁸See Appendix A.III for a more detailed discussion.

Figure I reports in grey the number of English and Welsh immigrants in the United States by year of immigration, digitized from official statistics (Willcox, 1928). The blue line on the right *y*-axis tabulates the number of immigrants in our linked dataset. Our data co-moves with official statistics data. Panel IIa of Figure II reports the spatial distribution of emigration rates across districts in the final sample and highlights its cross-sectional spatial heterogeneity. In Appendix Figure A.10, we break down the map by decade and uncover substantial variation in the origin of US emigrants over time, as discussed in Appendix Section A.III. Table I, Panel A, provides descriptive statistics for variables tabulated from the 1880 individual census. Panel B lists the district-level number of emigrants and return migrants in the UK-US-linked migrants sample by decade.

III.B A Dataset Covering British Patents in the Second Half of the Nineteenth Century

We measure innovation activity using patents, as standard in the literature (Griliches, 1998).¹⁹ We use data on the universe of patents granted in the United States digitized by Berkes (2018). Leveraging information on the county of residence of inventors and the technology class (both reported on patent documents), we construct a balanced panel at the county-technology class-grant year level.²⁰

We collect all the patents granted in the United Kingdom in 1900-1939 from PATSTAT, which provides bulk access to documents from the European Patent Office. These data contain information on authors and CPC classes but do not report the geographic location of inventors. To retrieve the coordinates of the inventors, we merge them with data by Bergeaud and Verluise (2024) and map them to registration districts at their 1890 borders.

For the period between 1853 and 1898, UK patent data was not available in digitized format. To tackle this data limitation, we digitize the universe of patents granted in England and Wales between 1853 and 1899 from scans of the original documents we obtained from the UK Patent Office (UK IPO). We apply optical character recognition (OCR) techniques on more than 300,000 original documents to convert the scanned images into machine-readable text. Then, we use large language models to parse the information contained in the raw text into a structured dataset. Compared to standard text extraction algorithms, this method harnesses artificial intelligence to deal with OCR issues and non-

¹⁹Patents are an imperfect proxy for innovation, as non-patented innovation represents a meaningful share of overall technological progress (Moser, 2019). However, we believe this limitation is minor in our context. As discussed in Section II.B, both the US and the UK implemented reforms prior to our study period that lowered barriers to patenting and led to sharp increases in patent counts. This supports the use of patents as a meaningful indicator of technological activity. In addition, our results also hold *within* technologies. Hence, if the propensity to patent in a given sector is similar in the UK and in the US, different propensities to patent across sectors should not be a concern for our analysis.

²⁰We map patents to counties at 1900 borders. From the three-digit Cooperative Patent Classification (CPC) class, we map patents to a coarser taxonomy of twenty technological sectors.

standard formulations of patent texts.²¹ We parse information on the title, text, inventors' names, geo-referenced addresses, filing and issue dates, assignee status, and whether the patent was filed by a patent agent. To conduct the analysis, we map patents to districts at 1890 borders.

This dataset adds to the one developed by Hanlon (2016) (1855–1883) by compiling information on patent titles, texts, and geography and extending the period until 1899, and expands previous work by Nuvolari and Tartari (2011), which covers the period until 1853. Appendix Figure A.2 reports the number of patents granted in the UK by year (Panel A.2a). The blue dots report the dataset we assemble, and the red data are tabulated from PATSTAT. The two series are consistent and display a steady growth in the number of patents issued throughout the period. Panel A.2b shows the evolution in the number of patents granted by technological class.

Patents differ in terms of novelty and economic value. In contemporaneous settings, forward citations are the standard proxy for patent impact. However, these were not reported in historical documents (Andrews, 2021). To construct a measure of patent "impact," we apply the text-based methodology developed by Kelly *et al.* (2021) to the corpus of British patents. Additionally, we adapt this approach to measure the textual similarity between British and American patents and proxy for technology transfer between the US and the UK. Intuitively, a British and an American patent are similar if their texts feature semantically similar words. Appendix Section A.II.6 describes in detail the practical implementation of the algorithm.

Finally, to perform the neighborhood-level analysis, we link the patent data to the census. To perform this linking, we match inventors based on the string similarity between their name and surname and those recorded in the census, conditional on geographic proximity. We describe the precise implementation in Appendix Section A.II.7. Figure A.6 displays the geographic distribution of inventors across England and Wales using the geo-coded addresses of the inventors linked to the census.

In Table I, Panels C and D, we report descriptive statistics for this sample. Panel C reports the total number of patents granted across districts by year and breaks it down into the five most common fields in the sample. Panel D reports statistics for the sample of inventors linked to the population census. Figure Panel IIb reports the spatial variation in the number of patents normalized by district population in 1891.

²¹Appendix Section A.II provides more details on the implementation procedure of the OCR and parsing algorithms.

III.C Additional Variables

We assemble district-level statistics from UK population censuses at a decade frequency between 1851 and 1911. Districts are the level of observation in most of the analysis. These were statistical units with an average population of 40,000, which makes them broadly comparable to US counties.²² Districts undergo minor boundary changes during the analysis period. To ensure geographical consistency, we crosswalk all variables to districts at 1890 borders, as explained in Appendix Section A.I.5.

IV EMPIRICAL STRATEGY

In this section, we describe our empirical strategy. In section IV.A, we provide details on the construction of technology exposure in the United States. In section IV.B, we describe our baseline empirical specifications and discuss the potential endogeneity associated with them in subsection IV.C. In section IV.D, we outline our identification strategy and the associated double and triple differences designs.

IV.A Measuring Exposure to American Technology

We construct two measures of exposure to US technology. The first is the number of emigrants to the United States from any given UK district. The underlying hypothesis is that districts with relatively more emigrants to the United States have higher exposure to US technology. The second measure uses two margins of variation: along technological classes and along districts. First, local specialization across counties measures the knowledge diffusing from those counties. Second, the number of migrants that leave a given district and settle in a given county measures the intensity of the return knowledge channel. To fix ideas, consider two districts *A* and *B*. The same number of emigrants *n* leaves each district. Emigrants from *A* settle in county *a*, which only produces innovation in sector s_a . Emigrants from *B* settle in county *b*, which only innovates in sector s_b . Then, we expect district *A* (respectively *B*) to innovate comparatively more in sector s_a (respectively s_b).

We define knowledge exposure as:

Knowledge Exposure_{*ik,t*}
$$\equiv \sum_{j \in J} \left(\frac{\text{Patents}_{jk,t}}{\text{Patents}_{j,t}} \times \text{Emigrants}_{i \to j,t} \right)$$
, (1)

where i, j, k, and t denote a (UK) district, a (US) county, a technology class, and a decade, respectively.²³ The set J denotes the universe of counties. The knowledge exposure term averages district-

²²Differently from US counties, UK districts did not have budgetary or administrative authority.

 $^{^{23}}$ Throughout the paper, decade t refers to the ten years before the upper bound t. For example, the decade indexed by 1890

level exposure to county-level specialization across technology classes.

The first term in the summation captures specialization, while the second term captures district-level exposure. The relative share of patents may inflate the influence of specialization in counties with a small number of granted patents. While this bias is unlikely to be large, as those counties would likely have low district-level exposure, to address this concern we code an alternative knowledge exposure variable that measures specialization as the raw count of patents in a given technology class. An additional challenge is that districts with larger bilateral linkages are likely larger and, hence, selected. To account for district-level time-varying confounding variables, we control non-parametrically for district-by-time fixed effects. Additionally, we report results for an alternative knowledge exposure that measures exposure through relative emigrant shares. We discuss these alternative definitions in more detail in the Appendix Table C.I.2.

IV.B Empirical Specifications

We estimate two sets of regressions. First, we study the relationship between overall patenting activity in the UK and the number of transatlantic emigrants:

$$y_{i,t} = \alpha_i + \alpha_t + \beta \times \text{US Emigrants}_{i,t} + X'_{i,t}\Gamma + \varepsilon_{i,t}, \qquad (2)$$

where *i* and *t* denote, respectively, a district and a decade; $y_{i,t}$ is the outcome variable; α_i and α_t are district and decade fixed effects; US Emigrants_{*i*,*t*} is the number of emigrants to the United States from district *i* between years t - 9 and t; and $X_{i,t}$ is a set of district-level controls—population, employment shares across sectors, share of males—measured in 1880 and interacted with year dummies. Standard errors are clustered at the district level.

The second specification studies the relationship between patenting in the UK across technologies and exposure to US technology in (1):

$$y_{ik,t} = \alpha_{it} + \alpha_k + \beta \times \text{Knowledge Exposure}_{ik,t} + \varepsilon_{ik,t},$$
 (3)

where *k* denotes technology classes, the term α_{it} denotes district-by-decade fixed effects, whose inclusion allows us to control non-parametrically for time-varying unobserved heterogeneity at the district level; the term α_k denotes technology-fixed effects and excludes variation arising from factors that are technology-specific but common across areas and over time. Standard errors are clustered at

refers to 1881-1890.

the district level.

We consider three different sets of outcomes throughout the paper. First, we use the log(1+) number of patents to measure the volume of patenting activity. Second, to account for heterogeneous novelty across patents, we compute the log(1+) number of patents in the top 20% of the impact distribution of the measure proposed by Kelly *et al.* (2021) and constructed based on the corpus of UK patents. Third, to measure technology flows from the US into the UK, we compute the average text-based similarity between US and UK patents (Appendix Section A.II.6 provides additional details). In Appendix Section C.I.1, we present alternative specifications with different transformations and definitions of all the dependent variables and show that results remain qualitatively unchanged.

IV.C Potential Sources of Endogeneity

The main factor that cautions against a causal interpretation of the estimates of regression (2) is that out-migration is not random across districts. In particular, any unobserved variable correlating with out-migration and the volume of innovation *y* induces bias in the estimated β coefficient. To address this issue, in (3), we include district-by-time fixed effects to partial out time-varying heterogeneity within districts. The inclusion of these fixed effects allows us to exclusively leverage cross-technology variation.

A remaining potential issue is assortative matching, meaning there may be a—possibly unobserved variable that correlates with the location where emigrants settle in the United States and the composition of patenting activity at home across technology classes. In Section II.A, we discussed that the historical and empirical evidence suggests that, over time, emigrants originated from increasingly affluent and urbanized areas. Suppose emigrants also settled in comparatively more urban and affluent counties in the United States, and there was a correlation between patenting activity in specific fields and economic growth. In that case, the selection issue may bias the OLS estimates upward. We note that the bias arises only if (i) the correlation between patenting and the underlying confounding variable is heterogeneous across technology classes and (ii) the correlation is similar in the US and the UK. If (i) does not hold, then the omitted confounding variable would be absorbed by district-by-time fixed effects. If (ii) does not hold, the selection bias would work against our result.

Assortative matching may also arise if pre-existing differences in specialization across technology classes predicted the counties where emigrants chose to settle. For example, suppose that emigrants from a largely textile area were comparatively more likely to settle in counties with larger textile sectors. Then, the estimated β from equation (3) would reflect pre-existing innovation similarities between sending and settling areas rather than capture the effect of return innovation. Evidence by

Hanlon (2018) and Ottinger (2020), among others, suggests that location decisions may not be random. In Appendix Section B.III, we show that British migrants in this period do not display a similar pattern.²⁴ However, ultimately, we cannot rule out assortative matching across US counties. Assortative matching is part of the broader "reflection problem" whereby our estimates may reflect shocks that simultaneously influence innovation in migrants' destination county and origin district (Manski, 1993). In the next section, we propose an empirical strategy to address these issues.

IV.D The Effect of Exposure to US Technology on UK Innovation: Double and Triple Differences Analysis

To estimate the causal impact of immigrants' exposure to US technology on innovation in the United Kingdom, we leverage shocks to US innovation activity. Suppose there is a sudden increase in the number of patents granted in a particular technology class within a US county. Then, the "return innovation" effect would imply that UK districts with more emigrants who settled in that county experience increased innovation activity in that same class. In other words, innovation shocks in the United States should transmit to the United Kingdom through pre-existing migration linkages.

We define a "shock" to US innovation as any period of unusually high patenting activity. As in the previous analysis, shocks can either pool technologies together or be technology-specific. In the pooled case, we isolate shocks by residualizing the (log) number of patents on county and year fixed effects, and label as a shock any observation in the top 5% of the residualized distribution. For technologyspecific shocks, we residualize patenting on a fully saturated set of county, technology, and year fixed effects, and similarly identify shocks as the top 5% of the residualized distribution.²⁵

From the perspective of US counties, innovation shocks represent periods of exceptionally high patenting activity. Appendix Table C.14 and Appendix Figure C.14 report the difference-in-differences estimates of the increase in patenting associated with these shocks. Patenting increases on average by 58% (in the pooled case) and by 13% (in the technology-specific instance) in US districts after a shock.

²⁴Appendix Table B.3 presents evidence against this pattern.

²⁵While our algorithm identifies shocks to innovation in a data-driven way, their mapping to specific historical events is not straightforward. In Appendix Section B.IV, we use the 1918–1919 Great Influenza pandemic as a case study of a historically grounded technology shock. The influenza pandemic offers a rare opportunity to validate our approach against an episode of increased innovation. Consistent with Berkes, Coluccia, Dossi and Squicciarini (2023), we find that counties with higher influenza mortality experienced disproportionately greater innovation in pharmaceutical-related technologies. Our algorithm similarly detects a higher likelihood of innovation shocks in these areas, confirming that our method captures the response to a major exogenous event affecting innovation. From the "reflection problem" perspective, a strategy that leverages the resulting US innovation shocks would yield invalid causal inference only if the severity of the pandemic was correlated with pre-existing migration linkages between US counties and UK districts, which appears implausible.

Let $\sigma_{j,t}$ be an indicator variable equal to one if county *j* undergoes an innovation shock in year *t*. To measure the exposure of British districts to these shocks (occurring in the US), we compute the number of emigrants from district *i* who are exposed to a shock as $\sum_{j \in J} (\text{Emigrants}_{i \to j} \times \sigma_{j,t})$, where Emigrants_{*i*→*j*} is the number of emigrants who moved from district *i* to county *j* during the decade 1870–1879, and $\sigma_{j,t}$ measures the intensity of the shock in county *j* at time *t*. We then residualize the resulting series by district and year. Based on these residuals, we define a binary variable that identifies a district as exposed to an innovation shock if its residual falls within the top 5% of the overall distribution. This indicator variable, ξ_i , equals one in the first period when district *i* is exposed to a US innovation shock, and zero otherwise. The technology-specific case follows the same logic and leads to the construction of a set of indicator variables ξ_{ik} which take value one in the first year a district-technology pair *ik* is exposed to a US innovation shock is shown in Appendix Figure C.2, and patenting in districts exposed to US innovation shocks in Appendix Figure C.3.

Given a set of district-level shocks ξ_i and district-technology-level shocks ξ_{ik} , we estimate double and triple differences regressions to study how exposure to US innovation shocks impacts innovation in the UK. In the case of district-level shocks, we estimate the following regression:

$$y_{i,t} = \alpha_i + \alpha_t + \sum_{h=-a}^{b} \beta_h \times I\left[D_{i,t} = h\right] + \varepsilon_{i,t},$$
(4)

where *t* now denotes years and $D_{i,t} \equiv (t - \xi_i)$ is the number of periods since district *i* is first exposed to a US innovation shock ξ_i . $I[\cdot]$ is an indicator variable which is equal to one when the argument is true and zero otherwise. In the case of technology-specific shocks, the regression is a triple-differences model:

$$y_{ik,t} = \alpha_{ik} + \alpha_{it} + \alpha_{kt} + \sum_{h=-a}^{b} \beta_h \times I \left[D_{ik,t} = h \right] + \varepsilon_{ik,t},$$
(5)

where α_{ik} , α_{it} , and α_{kt} denote, respectively, district-by-technology, district-by-year, and technologyby-year fixed effects, and $D_{ik,t} \equiv (t - \xi_{ik})$ denotes the number of years since the district-technology couple *ik* is first exposed to a US innovation shock ξ_{ik} .

Under a standard parallel trends assumption, the coefficients $\beta_{h\geq 0}$ in equations (4)–(5) estimate the

²⁶Specifically, we define $\sigma_{jk,t}$ as an indicator equal to one if the county-technology pair *ik* undergoes a shock in year *t* and zero otherwise. We then compute the number of emigrants from district *i* that, in year *t*, are exposed to US innovation shocks in technology *k* as $\sum_{j \in J} (\text{Emigrants}_{i \to j} \times \sigma_{jk,t})$. We residualize the resulting series against district, technology, and year-fixed effects, and we label an observation as a shock if the residualized value is in the top 5% of the overall distribution of exposure. Then, we define as ξ_{ik} the indicator variable equal to one in the first year when the district-technology pair *ik* is exposed to a US innovation shock and zero otherwise.

dynamic treatment effects of exposure to US innovation shocks on outcome *y*. The parallel trends assumption requires that, in the absence of the shocks, the outcome variable *y* would not have differed across treated and untreated units. While this assumption is not testable, we show that, in all regressions, the estimates of the pre-treatment coefficients β_h for h < 0 are never statistically different from zero. This pattern provides evidence in favor of the plausibility of the parallel trends assumption.

The roll-out of shocks across units is staggered: different districts and district-technology pairs may be exposed to US innovation shocks at different points in time. Goodman-Bacon (2021) shows that, in these settings, the two-way fixed-effects estimator fails to estimate the average treatment effect when treatment effects are heterogeneous. In robustness checks, we employ the estimator proposed by Sun and Abraham (2021) to deal with the staggered nature of the research design and confirm all the baseline results. Moreover, we show that the results are not sensitive to alternative thresholds to define the shocks to US innovation or the district-level exposure to such shocks. We discuss robustness checks in more detail in Section V.

V MAIN RESULTS

In this section, we document the "return innovation" effect. Section V.A presents descriptive evidence on the association between exposure to US innovation through out-migration and innovation in the United Kingdom. Section V.B provides causal evidence of the return innovation effect. Section V.C shows that the effect is stronger in sectors where the United States and the United Kingdom have similar relative specialization. Section V.D documents that the knowledge spillovers from exposure to US technology are largest in UK sectors most similar to those that emigrants were exposed to. Finally, Section V.E discusses robustness checks.

V.A Exposure to Innovation in the United States and the Volume and Direction of Innovation in the United Kingdom

We begin by studying the association between emigration and patenting activity in the United Kingdom. In Table II, column (1), we report the estimates of equation (2). We find that moving from the 25th to the 75th percentile of the distribution of emigrants to the US is correlated with an 8% increase in patenting activity in the UK district of origin. The magnitude of the coefficient halves and remains statistically significant after including district-level controls interacted with time fixed effects (column 2). In column (3), we show that out-migration is associated with an increase in the average similarity between British and American patents. Moving from the 25th to the 75th percentile of emigration is linked to a doubling of average similarity. These results provide the first piece of evidence in support of the hypothesis that emigrants foster the transfer of technologies from the US into the UK. In column (4), we report the results of estimating equation (2) on the number of patents in the top quintile of the impact distribution, which we refer to as "breakthrough" innovations. More impactful patents are those whose text is more distant from existing patents at the time of issue, but more similar to those issued afterward.²⁷ We find that moving from the 25th to the 75th percentile of out-migration is associated with an 11% increase in the number of breakthrough patents. This pattern suggests that return innovation is not driven solely by low-impact inventions, but also by novel and influential technologies.

In columns (5–8), we report results similar to columns (1–4), but at the technology level. Column (5) reports the estimates of equation (3). We document a positive and statistically significant correlation between technology-specific exposure and innovation in that technology. Moving from the 25th to the 75th percentile of technology-specific exposure is associated with an 11% increase in patenting activity in that technology class in the district of origin. In column (6), we add more granular fixed effects that control for district-by-year and district-by-technology variation. The coefficient remains positive and statistically significant but its magnitude decreases, becoming equal to one-fourth that of column (5).

In column (7), we show that exposure to US technologies is associated with an increase in the similarity between British and American patents. Quantitatively, the association is large, as moving from the 25th to the 75th percentile of the main regressor is associated with a three-fold increase in the average similarity between UK and US patents. In column (8), we document that the association between exposure to US innovation increases and patenting remains positive, statistically significant, and quantitatively unchanged when the dependent variable only includes breakthrough patents.

Finally, we explore the long-run association between exposure to US technology through emigration and innovation. Using data on patenting activity between 1940 and 2015, we compute the association between knowledge exposure in the 1930s and innovation in the later twentieth century. In Appendix Figure B.1, we report the correlation between knowledge exposure and decade dummies between the 1940s and the 2010s, which serves as the baseline category. The association between knowledge exposure and innovation remained stable until the 1960s. It began to decline in the 1970s and, more substantially, in the 1980s. Appendix Section B.II provides a detailed discussion of this analysis.

²⁷We adopt the methodology introduced by Kelly *et al.* (2021) and apply it to the texts of British patents. We use a 5-year window to compute this metric, but results are not sensitive to alternative bandwidths.

V.B The Impact of Innovation in the United States on Innovation in the United Kingdom

In this section, we describe the results of estimating the double and triple differences research design described in Section IV.D. This analysis enables us to draw a causal link between exposure to US innovation via migration ties and the production of innovation in the UK. As discussed in Section IV.C, the estimates in the previous section may not have a causal interpretation if there is an omitted variable that predicts where British immigrants settled across the United States and is correlated with the direction of innovation in their areas of origin—an issue we refer to as "assortative matching."

We begin by studying the impact of US innovation shocks on the overall production of innovation in the UK. Column (1) of Table III reports the baseline estimate of regression (4), where we aggregate pre- and post-treatment periods into two bins, and the treatment is equal to one after a given district is exposed to a US innovation shock and zero otherwise. After a district is exposed to a shock, patenting increases by approximately 9%. In the United States, an innovation shock is associated with a 58% increase in patenting activity in the county where it occurs (see Appendix Table C.14). Hence, on average, emigration ties generate a 15% pass-through rate of innovation shocks from the US to the UK. Following a US innovation shock, the similarity between patents granted in exposed districts and American patents doubles (column 2). In addition, in column (3), we show that the number of patents in the top 20% of the impact distribution also increases by 8%, implying that emigrants facilitate the transfer of high-quality patents from the US into the UK.

In columns (4–6), we report the estimates associated with regression (5), where the treatment is an indicator equal to one after a district-technology class pair is exposed to a US innovation shock and zero otherwise. In this case, the regression includes the interacted fixed effects. Hence, identification hinges on within-district technology exposure to US innovation shocks over time. The results confirm the previous findings. Innovation increases by 6% in districts exposed to an innovation shock. Emigrant ties generate a 50% pass-through rate of the increased patenting activity in the US into the UK. Patents in the same technologies and districts that are exposed to US innovation shocks also become more similar in their textual content to American patents (column 5). Finally, in column (6), we confirm that the results are quantitatively similar if we only look at high-impact innovation, as the number of breakthrough patents increases by 5% when units are exposed to the treatment.

Figure III reports the event-study estimates associated with equation (4) (Panel IIIa) and (5) (Panel IIIb). In each figure, we test for the joint significance of all pre-treatment coefficients and find no statistical support for the alternative hypothesis that there is at least one significant pre-treatment coefficient. Overall, we interpret these patterns as evidence in favor of the parallel trends assumption stated in IV.D. After the treatment, the number of patents increases. The post-treatment coefficients are

highly statistically—and jointly—significant.²⁸ The exposure to US innovation shocks has a persistent impact on UK patenting. As shown in Appendix Figure C.7, patenting activity in the US increases sharply and remains significantly above the pre-shock mean after the shock period. The persistent effect observed in the UK is consistent with this pattern.

In Figure IV, we re-estimate equations (4) (Panel IVa) and (5) (Panel IVb) using the similarity between British and American patents as the outcome variable. As before, we estimate no statistically significant pre-treatment coefficient, except a ten-year lead in IVa. The tests for the joint significance of all pre-treatment coefficients provide further evidence in favor of the validity of the pre-trend assumption. After the shock, British patents in exposed units become more similar to American patents, and the effect persists until ten years after the shock occurs.

V.C Return Innovation and Relative Specialization

In this section, we study how return innovation varies depending on the sector-level comparative advantage of the UK and the US. The historical evidence suggests the US was relatively more specialized than the UK in agricultural and heavy machinery technologies, even early in the sample period, whereas the UK was more specialized in chemical and textile innovation (David, 1966; Rosenberg, 1970). To measure relative specialization across technologies, we build on the international trade literature and construct a measure of revealed comparative advantage (RCA). This index takes a value above one for technologies where the UK was more innovative than the US in terms of share of total patents, and below one otherwise (Balassa, 1965).²⁹ To study the heterogeneous responses to emigrants' exposure to US technology, we estimate equation (3), interacting knowledge exposure with a set of technology dummies.

Figure V reports the results of estimating the triple-differences regression (5) interacting the treatment with technology-specific dummy variables. Each dot reports the coefficient of one technology class

$$\text{RCA}_{ik} \equiv \frac{\text{Patents}_{ik} / \sum_{k' \in K} \text{Patents}_{ik'}}{\sum_{i' \in I} \text{Patents}_{i'k} / \sum_{i' \in I, k' \in K} \text{Patents}_{i'k}}$$

²⁸The timing of the effect, which starts showing up already two years after the shock in the US, is in line with those found in the literature. For example, Furman, Nagler and Watzinger (2021) and Berkes and Nencka (2024) find that patent repositories and Carnegie libraries, respectively, have a positive and statistically significant impact on local patenting activity within one year after their establishment.

²⁹In the international trade literature, revealed comparative advantage is a metric that hinges on the observation that a country's comparative advantage is revealed by the country's relative exports (Balassa, 1965). In our setting, we define the revealed comparative advantage as

where *i* and *k* denote countries and sectors within sets *I* and *K*. Specifically, $I = \{UK, US\}$. Then, the UK is relatively more specialized in sectors with RCA_{UK,k} above one.

on the *y*-axis, and the UK revealed comparative advantage for that class on the *x*-axis. We uncover an inverted U-shaped relationship between the size of return innovation and specialization. Exposure to US innovation generates larger gains in technologies where the US and the UK have similar rates of specialization, such as electricity, scientific instruments, and engines. By contrast, the gains are not statistically significantly different from zero in sectors where the US is more advanced, such as agriculture, and where the UK is at the frontier, such as textiles and chemistry. The figure reports the coefficients of a regression between the coefficients, the RCA, and its squared values, and confirms the existence of such a hump-shaped relationship. This finding is consistent with the theoretical argument of Van Patten (2023), who develops a model of endogenous technology adoption where the gains from international technology diffusion are highest when a country trades with partners with higher—but similar—levels of development.

V.D Return Innovation: Spillovers on Adjacent Technology Classes

We have shown that return innovation is driven by exposure to technologies in which the US and the UK share similar patterns of comparative specialization. We now explore how the impact of technology-specific shocks in the US varies across different technologies in the UK. Specifically, we examine how exposure to US innovation in sector k affects innovation output in sector k itself and in other sectors.

In Figure VI, the red dot reports the baseline estimate from the triple-differences specification. To compute the blue dots, we first measure the technological distance between two classes as the probability that a patent is assigned to both classes. We then divide the resulting distribution into quintiles: the first quintile contains the most similar technology classes, while the fifth quintile contains the least similar. We estimate a series of triple-difference models, each pairing a technology with US innovation shocks occurring in technology classes in a given quintile of similarity, excluding the technology itself. The resulting coefficients are plotted as blue dots.

As expected, the effect of return innovation is largest in the same technology to which emigrants were exposed, and it decays with distance from that technology. However, we find that return innovation also spills over to closely related technologies. The coefficient for the first quintile—which captures how UK innovation responds to US innovation shocks in the most similar set of technology classes— is approximately half the size of that for the same technology and is highly statistically significant. The coefficients for the more distant quintiles are negative and not statistically different from zero, except for the fifth quintile—the most distant—which is negative and statistically significant, with a magnitude equal to one-third that of the own-technology effect. This pattern suggests that exposure to US innovation shocks partially crowds out innovation in the most distant technology classes.

However, the crowding-out effect is more than offset by an increase in innovation within the exposed class and in technologically close sectors. Hence, this result is consistent with our finding that return innovation not only reallocates innovation across technologies but also leads to an overall increase in innovation levels.

V.E Robustness Checks

To conclude, we perform a series of checks to test the robustness of our results. More details are provided in Appendix Section C.

We begin by testing the robustness of our results to different definitions of our dependent variable, patenting activity. In the baseline analysis, our preferred measure is the log of the number of patents, to which we add one to avoid dropping the zeros. In Appendix Tables C.1 and C.9, we show robustness to alternative transformations of the dependent variable for the panel (2)-(3) and the doubleand triple- differences regressions (4)–(5) and find that all results remain qualitatively and quantitatively unchanged using the raw count, the inverse hyperbolic sine, or adding 0.1 instead of one to the log of the count. Moreover, in Appendix Table C.5 and Appendix Figure C.5, we replicate the baseline panel regressions and event studies (4)–(5) as Poisson quasi-maximum likelihood regressions, as patents are a count variable.³⁰ The estimates confirm the baseline results: we find no evidence of statistically significant pre-treatment coefficients, and innovation increases after UK districts are exposed to shocks to US innovation. A possible concern with patents as a proxy for innovation is that they do not adequately reflect economically relevant innovations, i.e., technologies that are concretely adopted by firms. To address this concern, in Appendix Table C.2, we replicate the baseline findings of regressions (2)–(3) using as outcome variable only patents with a firm assignee, which provide an arguably more direct measure of firm innovation.³¹ All results hold when applying this sample restriction. A plausible concern is that our results are driven, at least in part, by the same inventors patenting the same innovation in the US and in the UK. To quantify the importance of this phenomenon, we replicate the baseline analysis, excluding from the dependent variable all British patents whose inventor also appears on a US patent. Appendix Tables C.7–C.12 report the panel and double- and triple-differences results. The estimates using this alternative sample confirm the baseline findings, indicating that even if "double filing" was practically feasible, it is not a key driver of our results.

³⁰We prefer a log-linear to a Poisson specification as the baseline because it allows us to test for robustness to potential concerns such as spatial autocorrelation, which would be hard to address with a Poisson regression.

³¹By law, the inventor of a patent in the UK must be an individual. However, many patents indicate when the inventor is employed in a firm or whether they are the owner.

Second, we gauge the robustness of our findings to alternative definitions of our text-based dependent variables, similarity and originality. In the baseline analysis, these are computed over a five-year window around the issue year of the patent. In practice, this window implies that, when computing the originality of a given British patent, we compare it with those issued in the preceding and following five years. When computing the similarity between a British patent and other US patents, we consider those issued in the preceding five years. This is an arbitrary choice, hence in Appendix Tables C.1 and C.9, we consider two additional windows—one and ten years—and find that all the baseline results remain qualitatively unchanged.

Third, we test the robustness of our results to alternative ways of defining exposure. In Appendix Table C.3, we show that the baseline association between patenting and exposure to US technology remains qualitatively unchanged when using several alternative definitions of exposure. In Appendix Table C.4, we also find that the association is stronger when exposure is measured using only the most impactful patents. This pattern is consistent with the hypothesis that migrants had a greater incentive to transmit knowledge back to the UK when the underlying patents had higher economic returns. Further details on the construction of the alternative exposure measures are provided in Appendix Section C.I.2. Additionally, we show that our results are robust to excluding from the computation of exposure all US counties in the top 10% and 20% of the distribution of British-born inventors. The rationale is that in these areas, the contribution of British inventors to local innovation is disproportionately large and may reflect origin-country specialization rather than American innovation. Our results remain qualitatively unchanged under these sample restrictions. Together with the double- and triple-difference estimates-which show that episodes of unusually intense innovation diffused into the UK—this finding provides further reassurance that our results are not simply driven by British migrants sorting into US counties specialized in the same technologies as their regions of origin. The results are shown in Appendix Tables C.6–C.11.

Fourth, we show that the double and triple differences estimates are robust to alternative definitions of "innovation shocks." Specifically, in the main analysis, we define a unit—either a district or a district-technology pair—as exposed to a US innovation shock in a given period if the number of migrants who are exposed to a US innovation shock from that district is in the top 5% of the overall distribution of exposed emigrants. The 5% threshold is arbitrary. In Appendix Table C.10, we adopt two alternative definitions of the shocks (top 10% and top 1%). Intuitively, more restrictive thresholds should deliver larger treatment effects because they require more emigrants to be exposed to a US innovation shock to activate the treatment at the district level. The estimates confirm this insight: while most of the coefficients remain statistically significant irrespective of the threshold, the magnitude of the estimated treatment effect increases with more restrictive thresholds.

Fifth, we verify that our results hold when restricting the linked UK–US sample to exact census matches only. Our analysis relies on an algorithm that links British immigrants in the US to individuals in the UK census, which can return multiple potential matches per immigrant. In the baseline analysis, we follow standard practice and weight each match by the inverse of the number of matches to avoid multiple counting. The results remain unchanged when we repeat the panel, double-difference, and triple-difference analyses using only unique (single) matches, which comprise approximately 40% of the sample (Appendix Tables C.8–C.13).

Sixth, we show that the results from our double- and triple-difference designs are robust to alternative estimators. In our setting, the roll-out of innovation shocks is staggered, as they occur at different times across districts and technologies. Following the recent literature, we replicate our baseline results using the two-way fixed effects estimator developed by Sun and Abraham (2021), which accounts for staggered treatment timing (Appendix Figure C.4). The results are very similar when using the log(1+) number of patents (Panels C.4a–C.4b) and the similarity between British and American patents (Panels C.4c–C.4d) as outcome variables.

Finally, we verify that our results hold under different levels of clustering. In all our main specifications, standard errors are clustered at the district level. In Appendix Figures C.1 and C.6, we use a set of alternative standard errors for both the panel and the double and triple differences regressions. In particular, we confirm that the results remain statistically significant when accounting for spatial autocorrelation as suggested by Conley (1999).

VI MECHANISMS

After documenting the "return innovation" effect, we study its drivers. In section VI.A, we show that return migration accounts for approximately half of the overall effect. In section VI.B, we document that, even in the absence of return migration, emigrants impact the innovation activity of their social networks in the United Kingdom. In section VI.C, we discuss potential complementary mechanisms that cannot be tested empirically.

VI.A The Role of Return Migration

Return migration is a primary candidate for explaining our findings through two channels. First, return migrants may engage in innovation activities in the fields they were exposed to abroad. Second, return migrants may facilitate access to US knowledge without directly undertaking innovation activities. The literature does not offer conclusive evidence on the effect of return migration on inno-

vation.32

The baseline linked sample of British emigrants traces them back to the UK census before they migrated. To measure return migration, we perform one additional step. We link British migrants who migrated to the US to UK censuses completed after they had settled in the US. Then, we aggregate return migration flows at the district-by-county level and decade frequency. Analogously to equation (1), we compute a measure of "return knowledge exposure" as follows:

Return Knowledge Exposure_{*ik,t*}
$$\equiv \sum_{j \in J} \left(\frac{\text{Patents}_{jk,t}}{\text{Patents}_{j,t}} \times \text{Return Migrants}_{j \to i,t} \right)$$
 (6)

where Return Migrants_{$j\to i,t$} is the number of migrants that return from county j to district i in decade t. Because UK censuses are available only until 1911, our return migration dataset spans 1870–1910.

In Table IV we estimate equations (2) and (3) including returning US emigrants (columns 1–3) and the return knowledge exposure indicator (columns 4–6). In columns (1) and (4), the dependent variable is the total number of patents. In columns (2) and (5), the dependent variable is the similarity to US patents, while in columns (3) and (6), it is the number of breakthrough patents. Throughout all specifications, the coefficients of US emigrants and the baseline knowledge exposure metric remain statistically significant, and the magnitudes are similar to those of Table III. The association between return migration and patenting is less strong. The coefficient is generally positive—albeit not always statistically significant—and at least one-half smaller compared to the coefficient of out-migration.³³

As expected, we find a positive correlation between return migration and patenting, but there is substantial variation in innovation associated with out-migration that is not explained by return migration. In the rest of this section, we provide evidence of the role of interactions between emigrants and local communities in Britain in fostering knowledge exchange even in the absence of physical return migration.

³²On the one hand, several studies estimate modest effects for recruiting programs of high-skilled nationals working abroad (Ash, Cai, Draka and Liu, 2022; Shi, Liu and Wang, 2023). On the other hand, Giorcelli (2019) shows that individuals exposed to (managerial) foreign knowledge change their behavior once back in their origin country.

³³The association with US patent similarity is not statistically significant for overall patenting (column 2), and is negative (though small) for patenting by technology (column 5). This pattern suggests that return migrants—who likely resettled in the UK several years before being observed in our analysis—are less likely to foster the diffusion of novel innovations in the US in the recent past.

VI.B The Role of Social Networks

In this section, we study how emigrants impact the innovation activity of their social networks in Britain. We focus on two factors promoting interactions between the emigrants and the non-migrant population: family ties and geographical proximity. This analysis builds heavily on experimental evidence from developing countries, which links the diffusion of technology to interactions over social networks (e.g., Bandiera and Rasul, 2006; Conley and Udry, 2010; Beaman et al., 2021).

VI.B.1 Transatlantic Emigration Within Families

To reconstruct extended families from the census data, which only contains information on the members of the same household, we leverage the geographical distribution of surnames. Specifically, we assume that individuals with the same surname who live in geographical proximity—in the same county—are related. This assumption is reasonable as long as surnames are not too common: in this analysis, we therefore drop the top 5% most common surnames. Results are robust to alternative thresholds (excluding the 1% or 10% most common surnames).

We study how US emigration shapes the innovation activity of the relatives of the migrants who do not move in a difference-in-differences setting. The treatment leverages variation in the surname of US emigrants by county. Formally, we estimate variations on the following regression:

$$y_{f,t} = \alpha_f + \alpha_t + \sum_{h=-a}^{b} \beta_h \times I[D_{f,t} = h] + \varepsilon_{f,t},$$
(7)

where *f* denotes a family—i.e., a surname-by-county pair—and *t* is a year. The term $D_{f,t}$ denotes the number of periods since at least one member of family *f* moved to the United States. Since families are identified at the county level, standard errors are clustered by county. Under the standard parallel trends assumption, the coefficients $\beta_{h\geq 0}$ estimate the impact of emigration on patenting activity carried of the emigrant's family.

First, in Panel A of Table V we report the estimates of (7) using a pre-post indicator for treatment status. Emigration within the family is associated with an increase in patenting (column 1), which remains unchanged if we include county-by-year fixed effects (column 2) and is larger the more migrants from the family move to the United States (column 3). Since patenting is a skewed outcome, in column (4), we use a binary indicator that confirms the results. Importantly, the patents granted to the family members of those who migrate to the US become more textually similar to US patents (column 5). These findings provide further evidence pointing to information flows between migrants and their relatives in the UK. Finally, in column (6), using the text-based impact measure, we confirm

that the increased patenting activity generated by within-family migration is not restricted to lowimpact innovations. Figure VII–Panel VIIa provides the associated event-study estimates. We find no evidence of statistically significant pre-treatment coefficients and a large increase in patenting, which builds up over the three years following the emigration of the family member.

We then distinguish between emigrants who return and those who do not. Emigrants could interact with their families upon returning, but they could also maintain ties while abroad. In Panel B, we restrict the treatment to emigrants who never return to the UK. The estimates confirm that even when emigrants do not return, they nonetheless promote innovation within their families in the UK. Comparing the coefficients in Panels A and B, we find that the treatment effect of non-return migrants is about 70% of the average treatment effect, irrespective of return status. This difference is consistent with the descriptive evidence discussed in section VI.A and confirms that migrants promote innovation in their origin communities even when they do not return.

VI.B.2 Transatlantic Emigration Within Neighborhoods

Geographical proximity between the origin of emigrants and stayers is an alternative proxy for local social networks. In this section, we show an increase in the innovation activity of stayers whose former neighbors migrate to the US.

We leverage the granular nature of our data and perform an individual-level analysis. First, we extract all men aged between 18 and 50 in 1900 who do *not* emigrate from the 1891 census. We create a yearly balanced panel that reports the number of patents granted to each individual between 1880 and 1900 using the linked inventor-census data described in Appendix Section A.II.7.³⁴ We complement this data with information on the geographical proximity between these stayers and migrants to the US. Specifically, we define a variable $D_{p,t}$ that codes the periods since at least one individual living within *k* kilometers from a non-migrant *p* migrates to the US and zero otherwise. In the baseline analysis, we consider a threshold *k* of five kilometers.³⁵

The regression specification mirrors (7), except that the variable $D_{p,t}$ substitutes $D_{f,t}$. All regressions include individual and year-fixed effects. In various specifications, we include district-by-time and parish-by-time fixed effects to control for time-varying heterogeneity at fine levels of spatial aggre-

³⁴In Table C.15, we show that our results are robust to excluding Wales from the estimation sample. The probability of linking patent and census records for inventors residing in Wales is, in fact, lower than in England (see Appendix Table A.3).

³⁵In Appendix Figure C.8, we show that the results are robust to alternative thresholds. In particular, the estimated treatment effect remains positive as in the baseline specification. Exceedingly restrictive thresholds, however, exclude a large number of treated individuals, and consequently, the estimated treatment effect is biased toward zero. Slack definitions of neighborhoods, by contrast, introduce measurement error, which reduces the precision of the estimates. We discuss these modeling choices in Appendix Section C.III.

gation. Standard errors are clustered at the parish level because neighborhoods form subdivisions of parishes.

Table VI (Panel A) reports the results. Within-neighborhood migration to the United States has a positive and statistically significant impact on the innovation activity of stayers (column 1). Importantly, the effect remains when including parish-by-time fixed effects (column 3). In this case, the identifying variation consists of within-parish neighborhood-level emigration and ensures that we compare similar areas. In column (4), we show that, conditional on having at least one emigrant in the neighborhood, having more does not scale up the positive effect of emigration on innovation. In column (5), we use a binary indicator instead of the (log) number of patents, which confirms the baseline results. Importantly, in column (6), we find that individuals exposed to US emigrants through neighborhood ties are granted patents that are more similar to those issued in the United States. This pattern confirms that neighborhood migrants plausibly transfer knowledge to their communities of origin. Finally, column (7) shows that neighborhood out-migration also promotes high-impact patenting. Figure VIIb replicates these estimates in an event-study framework. We do not find statistically significant pretreatment coefficients, even though these are not precisely estimated. The effect of neighborhood out-migration is largest four to five years after the treatment is first active and decreases thereafter.

In analogy to the family analysis, in Panel B, we restrict the attention to non-return migrants. The framework mirrors the previous specification, except that we now only include non-return migrants in the definition of the treatment $D_{f,t}$. We find that emigration within the neighborhood increases innovation even if the emigrant never comes back to the United Kingdom. In this case, the estimates in Panels A and B largely overlap; hence, the contribution of return migrants appears modest.

In Appendix Figure B.2, we estimate the baseline regression and interact the treatment variable with dummies for the occupational categories of the inventor. We find substantial heterogeneity in response to neighborhood out-migration: treatment effects are larger for positively selected inventors working as entrepreneurs, utility workers—a catch-all term which comprises independent inventors—, chemistry and metallurgy workers, and engineers. By contrast, inventors employed in low-skilled occupations, such as agriculture and construction, do not react to exposure to within-neighborhood out-migration.

Taken together, the results presented in this section highlight the importance of social networks in driving return innovation. As evidenced by the text-based measure of similarity between British and US patents, emigrants generate knowledge flows that benefit those communities. Importantly, these effects operate even in the absence of physical return migration.

VI.C Complementary Mechanisms

To conclude this section, we discuss two alternative channels—temporary migrations and monetary remittances—that may play a complementary role in shaping innovation outcomes, although they are unlikely to be the primary drivers of the return innovation effect.

Temporary Migrations In our study of the mechanisms behind the return innovation effect, we find that physical return is an important determinant but not its exclusive driver. It is possible, however, that short-term temporary migrations influence the dynamics of innovation in the UK. Since we construct migration flows from population censuses, which are conducted every ten years, our data cannot identify temporary migrants. For the same reasons, we cannot quantify the importance of industrial espionage.

However, the quantitative effects of temporary migrations and industrial espionage are likely to be modest. First, the notion of a "temporary migrant" in XIX-century transatlantic migration is not welldefined. Piore (1980) refers to Southern and Eastern European migrants as temporary because they planned to return to their origin countries at some point. This spell could take, however, decades. For example, a one-way cabin travel ticket from New York to Liverpool, at roughly 100\$, would cost as much as 20% of the average annual US income. The extent of short-term stays must have been, therefore, relatively limited, especially for UK immigrants whose return rates were substantially lower than for Southern and Eastern migrants (Bandiera, Rasul and Viarengo, 2013) and, therefore, were more likely to settle permanently in the United States. Moreover, Piore (1980) notes that "temporary" migrants were relatively low-skilled and, therefore, less likely to operate technology transfer. Industrial espionage does not appear to be quantitatively sufficient to generate the return innovation effect that we estimate. Additionally, our research designs speak against the temporary migration and the industrial espionage mechanisms. For temporary migration or industrial espionage to explain the double and triple differences result, in fact, one would need such flows to be correlated with the county-level innovation shocks. This channel seems unlikely, although it cannot be directly tested due to the absence of data.

Monetary Remittances While the contribution of remittances to economic development appears to be modest, it is possible that this inflow of capital may have sustained increased innovation, perhaps by relaxing constraints in access to credit (Gorodnichenko and Schnitzer, 2013). This capital inflow, however, could not explain why out-migration influences the direction of innovation unless knowledge *and* monetary remittances go hand in hand. Disaggregated data on financial remittances, unfortunately, do not exist. Hence, this is a possibility we cannot investigate empirically. It nonetheless highlights that, if anything, financial and "knowledge" remittances are likely to shape innovation in a complementary way rather than in a mutually exclusive one.

VII CONCLUSIONS

Previous literature has shown that the diffusion of innovation across countries is a major driver of long-run growth and economic convergence. However, evidence on the mechanisms behind this diffusion remains limited. In this paper, we provide the first causal evidence that out-migration enables the diffusion of innovation from migrants' countries of destination to their countries of origin.

Our analysis focuses on the mass migration from England and Wales to the United States between 1870 and 1940. We link the individual-level census records of British immigrants in the US to the UK population census to measure spatially disaggregated emigration flows. We complement these data with newly digitized patents, which, combined with existing data, allow us to cover the universe of patents granted in England and Wales over the nineteenth and twentieth centuries.

Using a difference-in-differences and triple-differences research design, we document that exposure to foreign technology through migration ties contributes to its diffusion in migrants' countries of origin. First, we show that the volume of innovation in the UK increases in response to greater exposure to US innovation through emigration. Second, innovation activity in the UK shifts toward technologies to which emigrants are most exposed in the US. By analyzing the textual similarity between UK and US patents, we find that exposure to US knowledge stimulates technology transfer. Additionally, we show that this diffusion involves both low- and high-impact UK innovations. We define this phenomenon as "return innovation." We then turn to the drivers of return innovation. The physical return of migrants explains approximately half of the effect. However, we find that interactions between emigrants and their communities of origin represent another important channel of technology diffusion, even in the absence of return migration.

While the setting of this paper is historical, our results likely have implications for contemporary contexts. In particular, our findings echo qualitative evidence from Saxenian (1999, 2006) and are consistent with contemporary evidence on the importance of in-person interactions for innovation and diffusion (Boudreau, Brady, Ganguli, Gaule, Guinan, Hollenberg and Lakhani, 2017; Atkin, Khandelwal and Osman, 2017). Taken together, our results show that emigration does not necessarily lead to underdevelopment or stagnation, as suggested by the "brain drain" hypothesis, but can instead foster the diffusion of innovation, a key driver of economic growth. The return innovation and brain drain effects are likely to coexist, with the balance between them depending on the type of migration (skilled or unskilled) and the relative positions of the origin and destination countries on the technology frontier. Beyond the relative size of these two effects, our results suggest that fostering economic integration between migrants' countries of origin and destination is likely to promote knowledge exchange.

REFERENCES

- Abramitzky, Ran and Leah Boustan, "Immigration in American Economic History," *Journal of Economic Literature*, 2017, 55 (4), 1311–45.
- _ , _ , Katherine Eriksson, James Feigenbaum, and Santiago Pérez, "Automated Linking of Historical Data," *Journal of Economic Literature*, 2021, 59 (3), 865–918.
- __, Leah Platt Boustan, and Katherine Eriksson, "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, 2014, 122 (3), 467–506.

Acemoglu, Daron, "Directed Technical Change," The Review of Economic Studies, 2002, 69 (4), 781–809.

- _, "When Does Labor Scarcity Encourage Innovation?," *Journal of Political Economy*, 2010, 118 (6), 1037–1078.
- ____, "Distorted Innovation: Does the Market Get the Direction of Technology Right?," AEA Papers and Proceedings, 2023, 113, 1–28.
- Aitken, Brian J and Ann E Harrison, "Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela," *American Economic Review*, 1999, *89* (3), 605–618.
- Akcigit, Ufuk, John Grigsby, and Tom Nicholas, "Immigration and the Rise of American Ingenuity," *American Economic Review*, May 2017, 107 (5), 327–31.
- _ , _ , and _ , "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age," NBER Working Paper, 2017, (No. w23047).
- **Alvarez, Fernando E, Francisco J Buera, and Robert E Lucas**, "Idea Flows, Economic Growth, and Trade," *Annual Review of Economics*, 2018, *10*, 315–345.
- Andersson, David E, Mounir Karadja, and Erik Prawitz, "Mass Migration and Technological Change," *Journal of the European Economic Association*, 2022, 20 (5), 1859–1896.
- Andrews, Michael J, "Historical Patent Data: A Practitioner's Guide," Journal of Economics & Management Strategy, 2021, 30 (2), 368–397.
- Anelli, Massimo, Gaetano Basso, Giuseppe Ippedico, and Giovanni Peri, "Emigration and Entrepreneurial Drain," American Economic Journal: Applied Economics, 2023, 15 (2), 218–252.
- Arkolakis, Costas, Sun Kyoung Lee, and Michael Peters, "European Immigrants and the United States' Rise to the Technological Frontier," *Working Paper*, 2020.
- Ash, Elliott, David Cai, Mirko Draka, and Shaoyu Liu, "Bootstrapping Science? The Impact of a "Return Human Capital" Programme on Chinese Research Productivity," *Working Paper*, 2022.
- Atkin, David, Amit K Khandelwal, and Adam Osman, "Exporting and Firm Performance: Evidence from a Randomized Experiment," *The Quarterly Journal of Economics*, 2017, 132 (2), 551–615.

- Bahar, Dany, Andreas Hauptmann, Cem Özgüzel, and Hillel Rapoport, "Migration and Knowledge Diffusion: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia," *The Review of Economics and Statistics*, 2019, pp. 1–50.
- _ , Prithwiraj Choudhury, Ernest Miguelez, and Sara Signorelli, "Global Mobile Inventors," Journal of Development Economics, 2024, 171, 103357.
- __, __, James Sappenfield, and Sara Signorelli, "Talent Flows and the Geography of Knowledge Production: Causal Evidence from Multinational Firms," *Working Paper*, 2022.
- _, Ricardo Hausmann, and Cesar A Hidalgo, "Neighbors and the Evolution of the Comparative Advantage of Nations: Evidence of International Knowledge Diffusion?," *Journal of International Economics*, 2014, 92 (1), 111–123.
- **Baines, Dudley**, *Migration in a mature economy: emigration and internal migration in England and Wales* 1861-1900, Cambridge University Press, 2002.
- **Balassa, Bela**, "Trade Liberalisation and "Revealed" Comparative Advantage," *The Manchester School*, 1965, 33 (2), 99–123.
- **Bandiera, Oriana and Imran Rasul**, "Social Networks and Technology Adoption in Northern Mozambique," *The Economic Journal*, 2006, *116* (514), 869–902.
- ____, ___, and Martina Viarengo, "The Making of Modern America: Migratory Flows in the Age of Mass Migration," *Journal of Development Economics*, 2013, 102, 23–47.
- Barsbai, Toman, Hillel Rapoport, Andreas Steinmayr, and Christoph Trebesch, "The Effect of Labor Migration on the Diffusion of Democracy: Evidence from a Former Soviet Republic," American Economic Journal: Applied Economics, 2017, 9 (3), 36–69.
- **Batista, Catia and Pedro C Vicente**, "Do Migrants Improve Governance at Home? Evidence from a Voting Experiment," *The World Bank Economic Review*, 2011, 25 (1), 77–104.
- Beaman, Lori, Ariel BenYishay, Jeremy Magruder, and Ahmed Mushfiq Mobarak, "Can Network Theory-Based Targeting Increase Technology Adoption?," *American Economic Review*, 2021, 111 (6), 1918–1943.
- Beine, Michel, Frédéric Docquier, and Maurice Schiff, "International Migration, Transfer of Norms and Home Country Fertility," *Canadian Journal of Economics/Revue canadienne d'économique*, 2013, 46 (4), 1406–1430.
- **Bergeaud, Antonin and Cyril Verluise**, "A New Dataset to Study a Century of Innovation in Europe and in the US," *Research Policy*, 2024, 53 (1), 104903.
- **Berkes, Enrico**, "Comprehensive Universe of US patents (CUSP): Data and Facts," *Working Paper*, 2018.
- and Peter Nencka, "Knowledge Access: The Effects of Carnegie Libraries on Innovation," *Review of Economics and Statistics*, 2024, pp. 1–45.

- _ , Davide M Coluccia, Gaia Dossi, and Mara P Squicciarini, "Dealing with Adversity: Religiosity or Science? Evidence from the Great Influenza Pandemic," Working Paper, 2023.
- Bernstein, Shai, Rebecca Diamond, Abhisit Jiranaphawiboon, Timothy McQuade, and Beatriz Pousada, "The Contribution of High-Skilled Immigrants to Innovation in the United States," *NBER Working Paper*, 2022, (No. w30797).
- **Bertoli, Simone and Francesca Marchetta**, "Bringing It All Back Home–Return Migration and Fertility Choices," *World Development*, 2015, *65*, 27–40.
- **Bottomley, Sean**, *The British Patent System during the Industrial Revolution* 1700–1852: *From Privilege to Property*, Cambridge (UK): Cambridge University Press, 2014.
- Boudreau, Kevin J, Tom Brady, Ina Ganguli, Patrick Gaule, Eva Guinan, Anthony Hollenberg, and Karim R Lakhani, "A Field Experiment on Search Costs and the Formation of Scientific Collaborations," *Review of Economics and Statistics*, 2017, 99 (4), 565–576.
- Bryan, Kevin A and Jorge Lemus, "The Direction of Innovation," *Journal of Economic Theory*, 2017, 172, 247–272.
- **Buera, Francisco J and Ezra Oberfield**, "The Global Diffusion of Ideas," *Econometrica*, 2020, *88* (1), 83–114.
- Burchardi, Konrad B, Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J Terry, "Immigration, Innovation, and Growth," *NBER Working Paper*, 2020, (No. w27075).
- **Clemens, Michael A**, "The Emigration Life Cycle: How Development Shapes Emigration from Poor Countries," *IZA Discussion Paper*, 2020.
- **Coluccia, Davide M and Lorenzo Spadavecchia**, "Emigration Restrictions and Economic Development: Evidence from the Italian Mass Migration to the United States," *Working Paper*, 2024.
- Comin, Diego and Bart Hobijn, "An Exploration of Technology Diffusion," American Economic Review, 2010, 100 (5), 2031–2059.
- **Conley, Timothy G**, "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, 92 (1), 1–45.
- _ and Christopher R Udry, "Learning About a New Technology: Pineapple in Ghana," American Economic Review, 2010, 100 (1), 35–69.
- **Coulter, Moureen**, *Property in Ideas: The Patent Question in Mid-Victorian England*, Kirksville (MO): Thomas Jefferson Press, 1991.
- David, Paul A, "The Mechanization of Reaping in the Ante-Bellum Midwest," in "Industrialization in Two Systems: Essays in Honor of Alexander Gershenkron," Harvard University Press Cambridge, MA, 1966.

Dossi, Gaia, "Race and Science," Working Paper, 2024.

Dustmann, Christian, Tommaso Frattini, and Anna Rosso, "The Effect of Emigration from Poland
on Polish wages," The Scandinavian Journal of Economics, 2015, 117 (2), 522–564.

- **Dutton, H I**, *The Patent System and Inventive Activity during the Industrial Revolution,* 1750-1852, Manchester (UK): Manchester University Press, 1984.
- Eaton, Jonathan and Samuel Kortum, "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 1999, 40 (3), 537–570.
- **Einiö, Elias, Josh Feng, and Xavier Jaravel**, "Social Push and the Direction of Innovation," *Working Paper*, 2023.
- Erickson, Charlotte, American industry and the European immigrant, 1860–1885, Cambridge (MA): Harvard University Press, 1957.
- ____, "Who were the English and Scots emigrants in the 1880s?," in "Population and Social Change," Arnold, 1972.
- Furman, Jeffrey L, Markus Nagler, and Martin Watzinger, "Disclosure and Subsequent Innovation: Evidence from the Patent Depository Library Program," *American Economic Journal: Economic Policy*, 2021, 13 (4), 239–270.
- Gibson, John and David McKenzie, "Eight Questions About Brain Drain," *Journal of Economic Perspectives*, 2011, 25 (3), 107–28.
- Giorcelli, Michela, "The Long-Term effects of Management and Technology Transfers," American Economic Review, 2019, 109 (1), 121–52.
- **Goodman-Bacon, Andrew**, "Difference-in-Differences with Variation in Treatment Timing," *Journal* of *Econometrics*, 2021, 225 (2), 254–277.
- **Gorodnichenko, Yuriy and Monika Schnitzer**, "Financial Constraints and Innovation: Why Poor Countries Don't Catch Up," *Journal of the European Economic Association*, 2013, *11* (5), 1115–1152.
- **Griffith, Rachel, Rupert Harrison, and John Van Reenen**, "How Special is the Special Relationship? Using the Impact of US R&D Spillovers on UK Firms as a Test of Technology Sourcing," *American Economic Review*, 2006, *96* (5), 1859–1875.
- **Griliches**, **Zvi**, "Patent Statistics as Economic Indicators: A Survey," in "R&D and productivity: the econometric evidence," University of Chicago Press, 1998, pp. 287–343.
- Hanlon, W Walker, "Necessity is the Mother of Invention: Input supplies and Directed Technical Change," *Econometrica*, 2015, 83 (1), 67–100.
- __, "British Patent Technology Classification Database: 1855-1882," 2016. Unpublished [data collection].
- _____, "Skilled Immigrants and American Industrialization: Lessons from Newport News Shipyard," Business History Review, 2018, 92 (4), 605–632.
- Helgertz, Jonas, Joseph Price, Jacob Wellington, Kelly J Thompson, Steven Ruggles, and Catherine A Fitch, "A New Strategy for Linking US Historical Censuses: A Case Study for the IPUMS

Multigenerational Longitudinal Panel," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2022, 55 (1), 12–29.

- **Higham, John**, *Strangers in the land: Patterns of American nativism, 1860-1925*, Rutgers University Press, 1955.
- Hopenhayn, Hugo and Francesco Squintani, "On the Direction of Innovation," *Journal of Political Economy*, 2021, 129 (7), 1991–2022.
- **Hornung, Erik**, "Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia," *American Economic Review*, 2014, 104 (1), 84–122.
- Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, 1993, *108* (3), 577–598.
- **Karadja, Mounir and Erik Prawitz**, "Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States," *Journal of Political Economy*, 2019, 127 (4), 1864–1925.
- **Keller, Wolfgang**, "Trade and the Transmission of Technology," *Journal of Economic Growth*, 2002, 7, 5–24.

_, "International Technology Diffusion," *Journal of Economic Literature*, 2004, 42 (3), 752–782.

- Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation Over the Long Run," *American Economic Review: Insights*, 2021, *3* (3), 303–320.
- **Kerr, William R**, "Ethnic Scientific Communities and International Technology Diffusion," *The Review* of Economics and Statistics, 2008, 90 (3), 518–537.
- Khan, B Zorina and Kenneth L Sokoloff, Institutions and Technological Innovation During the Early Economic Growth: Evidence from the Great Inventors of the United States, 1790-1930, National Bureau of Economic Research Cambridge, Mass., USA, 2004.
- Leak, H and T Priday, "Migration from and to the United Kingdom," *Journal of the Royal Statistical Society*, 1933, 96 (2), 183–239.
- **Macleod, Christine**, *Inventing the Industrial Revolution*, Cambridge (UK): Cambridge University Press, 1988.
- Manski, Charles F, "Identification of Endogenous Social Effects: The Reflection Problem," *The Review* of *Economic Studies*, 1993, 60 (3), 531–542.
- **Moscona**, Jacob, "Environmental Catastrophe and the Direction of Invention: Evidence from the American Dust Bowl," *Working Paper*, 2021.
- _ and Karthik A Sastry, "Does Directed Innovation Mitigate Climate Damage? Evidence from US Agriculture," *The Quarterly Journal of Economics*, 2023, *138* (2), 637–701.
- **Moser, Petra**, "Patents and Innovation in Economic History," in "Research Handbook on the Economics of Intellectual Property Law," Edward Elgar Publishing, 2019, pp. 462–481.

- _, Alessandra Voena, and Fabian Waldinger, "German Jewish Émigrés and US Invention," American Economic Review, 2014, 104 (10), 3222–55.
- _, Sahar Parsa, and Shmuel San, "Immigration, Science, and Invention. Evidence from the Quota Acts," Working Paper, 2020.
- **Nelson, Richard R and Gavin Wright**, "The Rise and Fall of American Technological Leadership: The Postwar Era in Historical Perspective," *Journal of Economic Literature*, 1992, 30 (4), 1931–1964.
- Nuvolari, Alessandro and Valentina Tartari, "Bennet Woodcroft and the Value of English Patents, 1617–1841," *Explorations in Economic History*, 2011, 48 (1), 97–115.
- Ottinger, Sebastian, "Immigrants, Industries, and Path Dependence," Working Paper, 2020.
- **Patten, Diana Van**, "International Diffusion of Technology: Accounting for Heterogeneous Learning Abilities," *Working Paper*, 2023.
- **Penrose, Edith**, *The Economics of the International Patent System*, Baltimore (MD): Johns Hopkins University Press, 1951.
- **Perla, Jesse, Christopher Tonetti, and Michael E Waugh**, "Equilibrium Technology Diffusion, Trade, and Growth," *American Economic Review*, 2021, *111* (1), 73–128.
- **Piore, Michael J**, *Birds of Passage: Migrant Labor and Industrial Societies*, Cambridge (UK): Cambridge University Press, 1980.
- Prato, Marta, "The Global Race for Talent: Brain Drain, Knowledge Transfer, and Growth," *The Quarterly Journal of Economics*, 2025, 140 (1), 165–238.
- Rosenberg, Nathan, "Economic Development and the Transfer of Technology: Some Historical Perspectives," *Technology and Culture*, 1970, *11* (4), 550–575.
- Ruggles, Steven, Catherine Fitch, Ronald Goeken, J Hacker, M Nelson, Evan Roberts, Megan Schouweiler, and Matthew Sobek, "IPUMS ancestry full count data: Version 3.0 [dataset]," Minneapolis, MN: IPUMS, 2021.
- Saxenian, AnnaLee, Silicon Valley's New Immigrant Entrepreneurs, Public Policy Institute of California, 1999.
- _, The New Argonauts: Regional Advantage in a Global Economy, Harvard University Press, 2006.
- Schurer, K and E Higgs, "Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access," 2020. [data collection] Second Edition, UKDS.
- Shi, Dongbo, Weichen Liu, and Yanbo Wang, "Has China's Young Thousand Talents program been successful in recruiting and nurturing top-caliber scientists?," *Science*, 2023, *379* (6627), 62–65.
- **Spilimbergo, Antonio**, "Democracy and Foreign Education," *American Economic Review*, 2009, 99 (1), 528–43.
- Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2021, 225 (2), 175–199.

- **Thomas, B**, *Migration and economic growth. A study of Great Britain and the Atlantic Economy*, Cambridge (MA): NIESR and Cambridge University Press, 1954.
- Truffa, Francesca and Ashley Wong, "Undergraduate Gender Diversity and the Direction of Scientific Research," *American Economic Review*, 2025, *Forthcoming*.
- **Tuccio, Michele and Jackline Wahba**, "Return Migration and the Transfer of Gender Norms: Evidence from the Middle East," *Journal of Comparative Economics*, 2018, *46* (4), 1006–1029.
- Willcox, William F, International Migrations, Volume I: Statistics, Cambridge (MA): National Bureau of Economic Research, 1928.

TABLES

Table I. Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Units	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Demographics (in 1880)						
Population, (1,000s)	39.988	46.183	0.184	310.706	621	621
Share of Men_i (%)	49.238	2.124	38.662	56.616	621	621
Share of Employment _{<i>i</i>} (%)	61.091	7.240	0.000	74.427	621	621
in Agriculture _{<i>i</i>} (%)	14.406	6.892	1.989	31.117	621	621
in Construction _{<i>i</i>} (%)	3.747	1.597	1.420	20.618	621	621
in Engineering _i (%)	3.818	1.483	0.848	11.843	621	621
in Public and Liberal Occupations _i (%)	2.573	1.674	0.819	16.045	621	621
in Other Manufacturing $_i$ (%)	3.289	2.772	0.618	19.684	621	621
in Textiles _i (%)	4.996	5.135	0.736	34.390	621	621
in Trade _i (%)	1.612	1.003	0.372	9.783	621	621
in Transportation $_i$ (%)	2.382	1.340	0.438	13.857	621	621
Panel B. Emigration to the United States						
US Emigrants _{id}	105.783	148.023	0.000	1752.157	621	3726
Return US Emigrants _{id}	11.845	17.108	0.000	185.591	621	2484
Panel C. Patents						
Total Patents _{it}	21.355	68.158	0.000	2088.000	621	31671
Patents by Technology Class:						
in Electricity _{it}	1.855	14.724	0.000	1084.000	621	31671
in Instruments _{it}	1.777	7.349	0.000	217.000	621	31671
in Lighting and Heating _{it}	1.486	5.487	0.000	145.000	621	31671
in Personal Goods _{it}	1.810	6.573	0.000	170.000	621	31671
in Transportation _{it}	3.499	10.761	0.000	244.000	621	31671
Panel D. Linked Inventor-Census Sample						
Number of Patents _{it}	0.121	0.842	0.000	274.000	110532	2210640
Age _i	37.104	13.729	9.000	79.000	110532	110532
Employed in:						
Agriculture _i	0.133	0.339	0.000	1.000	110532	110532
Construction _i	0.140	0.347	0.000	1.000	110532	110532
Engineering _i	0.164	0.371	0.000	1.000	110532	110532
Trade _i	0.089	0.285	0.000	1.000	110532	110532
Transportation _i	0.092	0.289	0.000	1.000	110532	110532
Neighborhood Emigrants _i	2.294	4.099	0.000	53.424	110532	110532
Non-Return Neighborhood Emigrants _i	2.256	4.029	0.000	52.677	110532	110532

Notes. This Table presents descriptive statistics for the variables used in the main analysis at the district (Panel A), district-decade (Panel B), district-year (Panel C), and individual level (Panel D). Demographic data in Panels A and D are tabulated from the population census. Panel A refers to the 1880 census. Emigration data in Panels A and D are computed from the linked migrant sample. Demographic data are cross-walked to consistent 1891 district borders. The subscripts report the aggregation level: Panel A by district *i*; Panel B by district-decade *id*; Panel C by district-year *it*; Panel D by individuals *i*. Referenced on page(s) 11, 13.

	Patents by District				Patents by District-Technology			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number	Number	US Patents Similarity	High Impact	Number	Number	US Patents Similarity	High Impact
Dependent Variable Mean	3.518	3.518	162.555	2.042	1.006	1.006	74.488	0.388
US Emigrants	0.771***	0.482***	18.677***	0.976***				
	(0.132)	(0.178)	(5.348)	(0.164)				
Knowledge Exposure					0.023***	0.006***	0.231***	0.022***
					(0.002)	(0.001)	(0.058)	(0.001)
District FE	Yes	Yes	Yes	Yes	_	_	_	_
Decade FE	Yes	-	Yes	Yes	-	-	-	-
Controls \times Time	No	Yes	No	No	-	-	-	-
County-Year FE	No	Yes	No	No	-	-	-	-
District-Year FE	-	-	-	-	Yes	Yes	Yes	Yes
Technology FE	-	-	_	-	Yes	_	Yes	Yes
District-Technology FE	-	-	_	-	No	Yes	No	No
Technology-Year FE	-	_	-	-	No	Yes	No	No
Number of Districts	621	620	621	621	621	621	621	621
Observations	3,726	3,720	3,726	3,726	70,433	70,433	70,433	70,433

Table II. Exposure to Innovation in the United States and Innovation in the United Kingdom

Notes. This Table reports the association between emigration to the United States and innovation in the United Kingdom. In columns (1–4) (resp. 5–8), the unit of observation is a district (resp. district-technology pair) observed at a decade frequency between 1870 and 1930. In columns (1–2) and (5–6), the dependent variable is the log(1+) number of patents. In columns (3) and (7), the dependent variable is the average text-based similarity between UK and US patents issued in the previous five years. In columns (4) and (8), the dependent variable is the log(1+) of the number of patents in the top 20% of the impact distribution ("Breakthrough" patents). In columns (1–4), "US Emigrants (1,000s)" is the number of migrants to the United States from the given district and decade, in thousands. In columns (5–8), "Knowledge Exposure" is defined in equation (1). Columns (1) and (3–4) include district and decade fixed effects. Column (2) includes district fixed effects, district-level controls measured in 1880 and interacted with decade indicators, and county-by-decade fixed effects. Columns (5) and (7–8) include district-by-year and technology fixed effects. Column (6) includes district level. Referenced on page(s) 19. *: p < 0.10, **: p < 0.05, ***: p < 0.01

	Pat	tents by Distri	ct	Patents by District-Technology			
	(1) Number	(2) US Patents Similarity	(3) High Impact	(4) Number	(5) US Patents Similarity	(6) High Impact	
Dependent Variable Mean	1.687	121.561	0.736	0.292	30.124	0.082	
Post \times US Innovation Shock	0.091** (0.045)	18.523*** (3.473)	0.082* (0.043)	0.064*** (0.009)	8.760*** (1.145)	0.020*** (0.006)	
District FE	Yes	Yes	Yes	_	_	_	
Year FE	Yes	Yes	Yes	_	_	_	
District-Year FE	_	_	_	Yes	Yes	Yes	
Technology-Year FE	_	_	_	Yes	Yes	Yes	
District-Technology FE	_	_	_	Yes	Yes	Yes	
Number of Districts Observations	621 31,671	621 31,671	621 31,671	621 601,749	621 601,749	621 601,749	

Table III. The Effect of Shocks to Innovation in the United States on Innovation in the United Kingdom

Notes. This Table reports the effect of shocks to US innovation activity on innovation in the United Kingdom. In columns (1–3), the unit of observation is a district observed at a yearly frequency between 1870 and 1930. In columns (4–6), the unit of observation is a district-technology pair observed over the same period. In columns (1) and (4), the dependent variable is the log(1+) number of patents. In columns (2) and (5), the dependent variable is the average text-based similarity between UK patents and US patents issued in the previous five years. In columns (3) and (6), the dependent variable is the log(1+) number of patents in the top 20% of the impact distribution ("Breakthrough" patents). "Post" is an indicator equal to one for all years after the observation unit is exposed to a shock to US innovation activity, equal to zero otherwise. "US Innovation Shock" is defined in Section IV.D. Columns (1–3) include district and year fixed effects. Columns (4–6) include district-by-year, technology-by-year, and district-by-technology fixed effects. Standard errors, reported in parentheses, are clustered by district. Referenced on page(s) 21, 27. *: p < 0.10, **: p < 0.05, ***: p < 0.01

	Pa	tents by Distr	ict	Patents by District-Technology		
	(1) Number	(2) US Patents Similarity	(3) High Impact	(4) Number	(5) US Patents Similarity	(6) High Impact
Dependent Variable Mean	3.145	120.882	1.780	0.828	50.971	0.316
US Emigrants (\hat{eta}_1)	0.817*** (0.138)	19.982*** (7.360)	0.870*** (0.188)			
Return US Emigrants ($\hat{\beta}_2$)	3.059*** (1.136)	73.163 (71.053)	0.455 (1.300)			
Knowledge Exposure ($\hat{\beta}_1$)				0.020*** (0.001)	0.428*** (0.059)	0.016*** (0.001)
Return Knowledge Exposure ($\hat{\beta}_2$)				0.404*** (0.066)	-16.161*** (2.910)	0.629*** (0.076)
District FE	Yes	Yes	Yes	_	_	_
Decade FE	Yes	Yes	Yes	-	-	-
District-Year FE	-	_	-	Yes	Yes	Yes
Technology FE	_	_	_	Yes	Yes	Yes
Number of Districts Observations	621 2,484	621 2,484	621 2,484	621 46,835	621 46,835	621 46,835
Standardized $\hat{\beta}_1$ Standardized $\hat{\beta}_2$	0.075 0.029	0.054 0.021	0.091 0.005	0.195 0.046	0.068 -0.029	0.273 0.120

Table IV. Return Migration and Innovation in the United Kingdom

Notes. This Table compares the effect of out-migration and return-migration on innovation in the UK. The unit of observation is a district (columns 1–3) and a district-technology (columns 4–6). Units are observed at decade frequency between 1880 and 1910. Different sample periods depend on the lag structure. In columns (1) and (4), the dependent variable is the log(1+) number of patents. In columns (2) and (5), the dependent variable is the average text-based similarity between UK and US patents issued in the previous five years. In columns (3) and (6), the dependent variable is the log(1+) of the number of patents in the top 20% of the impact distribution ("Breakthrough" patents). In columns (1-3), "US Emigrants (1,000s)" and "Return US migrants (1,000s)" are the number of migrants and return migrants to the United States from the given district and decade, in thousands. In columns (4–6), "Knowledge Exposure" is defined in equation (1) and "Return Knowledge Exposure" is defined similarly to Knowledge Exposure, but only looks at flows of emigrants who eventually return to the UK. The coefficient of out-migration flows (US Emigrants and Knowledge Exposure) are labeled as β_1 , and the coefficients of return flows (Return US Emigrants and Return Knowledge Exposure) are labeled as β_2 . We report the standardized coefficients at the bottom of the table. These are calculated by dividing each estimated beta by the standard deviation of the corresponding right-hand side variable. Columns (1-3) include district and decade fixed effects. Columns (54-6) include district-by-year, technology-by-year, and district-by-technology fixed effects. Standard errors, reported in parentheses, are clustered at the district level. Referenced on page(s) **27.** *: p < 0.10, **: p < 0.05, ***: p < 0.01

Table V. Within-Family Migration to the United States and Innovation in the United Kingdom

		Numbe		Text-Based Measures		
	(1)	(2)	(3)	(4)	(5)	(6)
	Number	Number	Number	I(Patents > 0)	US Patents Similarity	High Impact
Panel A. Family Emigration						
Dependent Variable Mean	4.919	4.919	4.919	5.835	6.612	1.229
US Emigrant in Family \times Post	1.375***	1.505***	-0.312	1.303***	3.873***	0.452***
	(0.363)	(0.314)	(0.218)	(0.402)	(0.567)	(0.118)
US Emigrant in Family \times Post \times N. of Emigrants			1.299*** (0.316)			
Number of Families	46,864	46,864	46,864	46,864	46,864	46,864
Observations	2,858,704	2,858,704	2,858,704	2,858,704	2,858,704	2,858,704
Panel B. Family Non-Return Emigration						
Dependent Variable Mean	4.919	4.919	4.919	5.835	6.612	1.229
Non-Return US Emigrant in Family $ imes$ Post	1.030***	1.152***	-1.181**	0.970**	3.065***	0.306***
	(0.333)	(0.281)	(0.448)	(0.373)	(0.497)	(0.092)
Non-Return US Emigrant in Family \times Post \times N. of Emigrants			1.813***			
			(0.626)			
Number of Families	44,879	44,879	44,879	44,879	44,879	44,879
Observations	2,737,619	2,737,619	2,737,619	2,737,619	2,737,619	2,737,619
Family FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	-	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This Table reports the impact of migration to the United States on the innovation activity of migrants' relatives who remain in the UK. The unit of observation is a family defined by all individuals with the same surname in the same county. Units are observed at yearly frequency between 1870 and 1930. In columns (1–3), the dependent variable is the log(1+) number of patents issued to family members in the UK. In column (4), the dependent variable is an indicator equal to one if family members in the UK were issued at least one patent, equal to zero otherwise. In column (5), the dependent variable is the average text-based similarity between UK patents and US patents issued in the previous five years. In column (6), the dependent variable is the log(1+)number of patents in the top 20% impact distribution ("Breakthrough" patents). "US Emigrant in Family \times Post" is an indicator equal to one in the year the first emigrant in the family moves to the US and in later years, and equal to zero in the previous years. "US Emigrant in Family \times Post \times N. of Emigrants" is an interaction between "US Emigrant in Family \times Post" and the number of emigrants from that family. "Non-Return US Emigrant in Family \times Post" is an indicator equal to one in the year the first emigrant in the family moves to the US (without returning to the UK) and in later years and equal to zero in the previous years. "Non-Return US Emigrant in Family \times Post \times N. of Emigrants" is an interaction between "Non-Return US Emigrant in Family \times Post" and the number of non-return emigrants from that family. All columns include family (i.e., surnamecounty) and year fixed effects. Column (3) also includes county-by-year fixed effects. Standard errors, reported in parentheses, are clustered at the county level. The dependent variables in columns (1-4) are multiplied by 100. Referenced on page(s) 28. *: p < 0.10, **: p < 0.05, ***: p < 0.01

Table VI. Within-Neighbohood Emigration to the United States and Innovation in the United Kingdom

	Number of Patents					Text-Based Measures	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Number	Number	Number	Number	I(Patents > 0)	US Patents Similarity	High Impact
Panel A. Neighborhood Emigration							
Dependent Variable Mean	6.721	6.721	6.741	6.775	8.483	6.792	1.550
Post \times Emigrant in Neighborhood	0.440***	0.236**	0.254**	0.205	0.552***	0.477***	0.116**
	(0.130)	(0.110)	(0.116)	(0.154)	(0.160)	(0.132)	(0.052)
Post \times N. Emigrants in Neighborhood				0.210 (0.209)			
Number of Individuals	136,880	136,872	134,936	117,649	136,880	136,880	136,880
Observations	2,737,600	2,737,440	2,698,720	2,352,980	2,737,600	2,737,600	2,737,600
Panel B. Neighborhood Non-Return Emigration							
Dependent Variable Mean	6.721	6.721	6.741	6.775	8.483	6.792	1.550
Post $ imes$ Non-Return Emigrant in Neighborhood	0.409***	0.252**	0.263**	0.430***	0.517***	0.453***	0.109**
	(0.130)	(0.113)	(0.120)	(0.163)	(0.159)	(0.133)	(0.054)
Post \times N. Non-Return Emigrants in Neighborhood				0.064 (0.216)			
Number of Individuals	120,575	120,567	118,628	102,714	120,575	120,575	120,575
Observations	2,411,500	2,411,340	2,372,560	2,054,280	2,411,500	2,411,500	2,411,500
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parish-Year FE	No	No	Yes	No	No	No	No
District-Year FE	No	Yes	-	No	No	No	No
Year FE	Yes	-	-	Yes	Yes	Yes	Yes

Notes. This Table reports the impact of emigrants to the US on the innovation activity of their former neighbors in the UK. The unit of observation is an individual inventor observed yearly between 1880 and 1900. The sample includes the universe of British inventors linked to the 1891 population census, as detailed in the main text. In columns (1-4), the dependent variable is the log(1+) number of patents granted to members of the emigrant's family. In column (5), the dependent variable is an indicator equal to one if members of the emigrant's family are granted at least one patent, equal to zero otherwise. In column (6), the dependent variable is the average text-based similarity between UK patents and US patents issued in the previous five years. In column (7), the dependent variable is the log(1+) number of patents in the top 20% of the impact distribution. "Post \times Emigrant in Neighborhood" is an indicator equal to one in the year the first neighbor of the inventor moves to the US and in subsequent years, and equal to zero in the previous years. "Post \times N. Emigrants in Neighborhood" is an interaction between "Post \times Emigrant in Neighborhood" and the number of emigrants to the US from the neighborhood. "Post \times Non-Return Emigrant in Neighborhood" is an indicator equal to one in the year the first neighbor of the inventor moves to the US (without returning to the UK) and in subsequent years and equal to zero in the previous years. "Post \times N. Non-Return Emigrants in Neighborhood" is an interaction between "Post \times Non-Return Emigrant in Neighborhood" and the number of emigrants to the US from the neighborhood. Columns (1) and (4-7) include inventor and year fixed effects. Column (2) includes inventor and district-by-year fixed effects. Column (3) includes inventor and parish-by-year fixed effects. Standard errors, reported in parentheses, are clustered at the parish level. The dependent variables in columns (1-4) are multiplied by 100. Referenced on page(s) 30. *: p < 0.10, **: p < 0.05, ***: p < 0.01

FIGURES



Figure I. British Emigration to the United States: Official Statistics and Linked Sample

Notes. This Figure compares the total number of English and Welsh immigrants in the United States as recorded in official statistics from Willcox (1928) with the linked emigrants' sample assembled in this paper. The gray bars report, on the left *y*-axis, the number of English and Welsh immigrants in the official statistics by their recorded immigration year in the United States. The blue line, whose values are reported on the right *y*-axis, reports the total number of English and Welsh immigrants in the US that appear in our matched sample. By construction, we can only match men who appear at least once in one British census. Numbers are in thousand units. The black two-sided arrow marks the period covered by the linked sample, i.e., the years between 1870 and 1930. Referenced on page(s) 11.



Figure II. Patenting and Emigration Rates in the United Kingdom

Notes. This Figure reports the spatial distribution of emigrants across English and Welsh districts over the period 1870–1930 (Panel IIa) and the number of patents issued over the same period (Panel IIb). Both series are normalized by district population in 1891 and are reported in percentage points. Districts are displayed at 1891 historical borders, and the emigrant population is cross-walked to consistent borders. Lighter to darker tones of color indicate increasing quantiles of the variable. The London area is displayed separately. Referenced on page(s) 11.



Figure III. Dynamic Effect of Exposure to Innovation Shocks in the United States

Notes. This Figure reports the impact of shocks to US innovation activity on innovation in the UK. In Panel IIIa, the unit of observation is a district observed at yearly frequency between 1870 and 1930; in Panel IIIb, the unit of observation is a district-technology pair observed over the same period. The dependent variable is the log(1+) number of patents. Each dot reports the coefficient of an indicator variable, which codes the time since the observation unit is exposed to a shock in US innovation activity through emigration ties. The black dashed line indicates the treatment period. The last period before the shock serves as the baseline category. In Panel IIIa, the regression includes district and year fixed effects; in Panel IIIb, the regression includes district level. The bands report 95% confidence intervals. Each Figure reports separate *F*-statistics for the joint significance of the pre-and post-treatment coefficients and associated *p*-values. Referenced on page(s) 21.





Notes. This Figure reports the impact of shocks to US innovation activity on innovation in the UK. In Panel IVa, the unit of observation is a district observed at yearly frequency between 1870 and 1930; in Panel IVb, the unit of observation is a district-technology pair observed over the same period. The dependent variable is the average text-based similarity of UK patents issued in a given period to US patents issued in the previous five years. Each dot reports the coefficient of an indicator variable, which codes the time since the observation unit is exposed to a shock in US innovation activity through emigration ties. The black dashed line indicates the treatment period. The last period before the shock serves as the baseline category. In Panel IVa, the regression includes district-by-technology fixed effects. Standard errors are clustered at the district level. The bands report 95% confidence intervals. Each Figure reports separate *F*-statistics for the joint significance of the pre-and post-treatment coefficients and associated *p*-values. Referenced on page(s) 22.



Figure V. Heterogeneous Effects of Exposure to US Innovation Across Technologies

Notes. This Figure reports the heterogeneous effects of exposure to US technology on UK innovation across technology classes. The unit of observation is a district-technology couple observed at a yearly frequency over the same period. Each dot reports the coefficient of an interaction term between technology dummies and an indicator variable that is equal to one for periods after the unit is exposed to a US innovation shock for the first time and zero otherwise. The dependent variable is the log(1+) number of patents. Each regression includes district-by-year, technology-by-year, and district-by-technology fixed effects. The estimated coefficients are plotted against the measure of revealed comparative advantage discussed in the main text. The UK is relatively more active than the US in technologies to the right of one, and it is less active than the US otherwise. The plot superimposes a polynomial of degree two and reports the estimated coefficients of the measure of revealed comparative advantage discuss the measure of revealed comparative advantage of the estimated coefficients of the measure of revealed comparative advantage discussed in the text. The UK is relatively more active than the US in technologies to the right of one, and it is less active than the US otherwise. The plot superimposes a polynomial of degree two and reports the estimated coefficients of the measure of revealed comparative advantage of orders one and two, along with their robust standard errors. The size of each dot reflects the number of patents issued in the UK over the sample period in each technology class. Referenced on page(s) 22.



Figure VI. Knowledge Spillovers of US Innovation Shocks on British Innovation

Notes. This Figure reports the cross-technology knowledge spillovers of exposure to US technology shocks on UK innovation across technology classes. The unit of observation is a district-technology couple observed at yearly frequency between 1870 and 1930. The red square reports the baseline estimated effect from the triple differences specification. To compute the blue dots, we compute the distance between two technology classes as the probability that a patent is assigned to both classes. We divide the resulting measure into quintiles, where the first quintile collects the most similar classes, and the last quintile collects the least similar. There are three or four technology classes in each quantile. Then, we estimate a set of triple differences models where each technology is paired with a shock to US innovation in technology in the *k*-quintile of the similarity distribution. The dots report the resulting triple differences coefficients. Thus, the coefficient on the first quintile, for example, reports how innovation in the UK reacts to exposure to a US innovation shock in the most similar quintile of technology classes (excluding the technology itself). The bands report 95% confidence intervals. The graph reports the point estimate of each coefficient. Referenced on page(s) 23.



Figure VII. Dynamic Effects of Family and Neighborhood Emigration

Notes. This Figure reports the impact of emigration to the United States on the innovation activity of migrants' relatives in the United Kingdom (Panel VIIa) and of their neighbors (Panel VIIb). In Panel VIIa, the unit of observation is an extended family defined as comprising all those who share the same surname in a given county. Families are observed at yearly frequency between 1870 and 1930. In Panel VIIb, the unit of observation is an individual observed at yearly frequency between 1880 and 1900. The dots report the coefficients of an indicator variable that codes the period since the first emigrant in the family (Panel VIIa) or in the neighborhood (Panel VIIb). The black dashed line indicates the first treatment period; the previous period serves as the baseline category. The dependent variable is the log(1+) number of patents. All regression specifications control for unit (family or individual) and year fixed effects. Standard errors are clustered at the family (Panel VIIa) and district (Panel VIIb) level. The bands report 95% confidence intervals. Each Figure reports separate *F*-statistics for the joint significance of the pre- and post-treatment coefficients and associated *p*-values. Referenced on page(s) 28.

ONLINE APPENDIX

Return Innovation: The Knowledge Spillovers of the British Migration to the United States, 1870-1940

Davide M. Coluccia & Gaia Dossi May, 2025

TABLE OF CONTENTS

A DATA APPENDIX A.I Summary of Data Sources A.II Patent Data A.III Linked Migrants Sample Tables Figures B ADDITIONAL RESULTS B.I Selection of British Emigrants B.II Long-Run Association between Emigration and Innovation B.III Assortative Matching of Emigrants in the United States

B.IV A Case Study of a US Technology Shock: The Great Influenza Pandemic

B.V Anecdotal Evidence of Return Innovation

Tables

Figures

C ROBUSTNESS ANALYSES

C.I Panel Analysis

C.II Double and Triple Differences Analysis

C.III Neighborhood Analysis

Tables

Figures

B29

C42

A DATA APPENDIX

This section describes the data sources and the construction of the final sample (section A.I), the construction of the newly digitized patent data and the approach to link patents to the census (section A.II), and the methodology we developed to link the records of British immigrants in the United States to the UK census (section A.III).

A.I Summary of Data Sources

A.I.1 Patent Data

US patent data are from Berkes (2018), who digitizes the universe of patents granted between 1836, when the US patent and trademark office was established, and 2010. In this paper, we are interested in the CPC technology class, the issue year, and the coordinates of residence of each inventor. We then assign each patent to US counties at 1900 borders. Depending on the number of inventors, a single patent may be assigned to multiple counties. In the case of patents with numerous inventors, we weigh each by the inverse of the number of inventors to avoid multiple counting. English and Welsh patents after 1900 are available at the European Patent Office. To construct our dataset, we leverage bulk access to the PATSTAT dataset. Information contained in PATSTAT includes the CPC class and the issue year. To retrieve the location of each inventor, we merge the PATSTAT data with the PatCity repository, which contains geo-coded information on the universe of English and Welsh patents during this period (Bergeaud and Verluise, 2024). Data before 1900 are not available. In section A.II, we describe how we digitize the universe of patent documents issued over the period 1853–1899 to fill this substantial gap.

Importantly, we map 3-digit CPC classes to a coarser taxonomy of classes. To do that, we reduce them to functional units using the CPC classification scheme. The scheme is publicly available at the following link. To accommodate the historical context, we divide the transporting categories into two classes: "Transporting", which includes carriages, railways, and cars, and "Ships and Aeronautics". Moreover, we conflate the "Weapons and Blasting" and the "Mining" classes into the "Metallurgy" category because few patents were observed in those industries. We further augment patent data by defining a measure of "quality" or "innovativeness" following Kelly *et al.* (2021). This metric flags as influential those patents that introduce terms not used before they were granted and become common after that. We evaluate this metric on the abstracts of patents granted after 1900. We apply this sample restriction for consistency: in 1853–1899, we observe the full text of patents instead of their abstract.

A.I.2 Migration Data

Disaggregated data on the origin of English and Welsh immigrants—and, more generally, of all other nationalities—do not exist. Neither US authorities nor the sending ones in the UK collected them. We thus lack precise information on where British immigrants in the US came from *within* the UK. We fill this gap and link the individual-level UK and US censuses, as described in A.III. Ideally, we observe the universe of British emigrants to the United States between 1870 and 1930. For those individuals, we know all the information contained in the US Census and those detailed in the UK one. Most notably, we know where they came from. As we discuss more in detail later, we also link return migrants. Since the last publicly available UK census dates to 1911, we can only construct return migration flows over the period 1870–1910.

A.I.3 Data Constructed from the Population Censuses

The main data sources we leverage are the individual-level, non-anonymized UK and US population censuses. The US census features prominently in the economic history literature as a major source of detailed microdata, and we thus avoid discussing it any further (Ruggles *et al.*, 2021). The UK census is relatively less well-known (Schurer and Higgs, 2020). Although not as rich as its US counterpart, the UK population census covers individuals who have resided in the UK since 1850. The first census was run in 1841, but only 1851, 1861, 1881, 1891, 1901, and 1911 are entirely digitized.¹ Data in the census include the name and surname, birth year, division, county, district, parish, precise address of residence, the specific occupation detailed through HISCO codes, and other variables that we do not use in the paper, such as the type of dwelling and fertility information. We augment these variables by geo-coding the universe of addresses that appear in the census to precise geographical coordinates.

A.I.4 Miscellaneous Data

We construct US county-level exposure to the Great Influenza pandemic using mortality statistics collected by the US Bureau of Census. These data are available for a subset of counties representing approximately 60% of the US population in 1900.

A.I.5 GIS Shapefiles & Boundary Harmonization

Patents and telegraph stations are mapped to 1900 registration district borders using historical GIS files and their coordinates.² However, all data from the population censuses appear at historical bor-

¹The 1921 census is currently being digitized and is partially available. We do not use it because its coverage is still not complete and because it is not available in bulk. All censuses after 1921 are subject to privacy restrictions.

²GIS data for the US are provided by NHGIS, whereas district boundaries have been digitized by the Great Britain Historical GIS Project.

ders. Registration districts have not undergone major boundary changes over the period that we studied. However, we adapt the method presented by Eckert, Gvirtz, Liang and Peters (2020) to UK districts to ensure that we work with consistent geographical units.

To construct geographical crosswalks using their method, one needs to assume that variables are evenly distributed over the area of geographical units. The crosswalk is then obtained by overlapping geographical units over time. Suppose unit x in decade d is split, and 80% of its territory is assigned to itself, while 20% is assigned to another district y. To construct a cross-walk relative to period $d + t_2$ for a generic variable between decades $d - t_1$ and $d + t_2$, for $t_1, t_2 > 0$, one needs to multiply the variable measured in district x in $d - t_1$ by 4/5 and add 1/5 of the variable in x to that measured in y in the same decade. We map registration districts to their boundaries in 1901. Less than 5% of the overall area of England and Wales is re-assigned in this way. We adopt the same methodology to map counties to their 1900 borders.

A.I.6 Geo-referencing the Historical British Census

A notable feature of the UK census is that it contains precise information on the residential address of the universe of the British population. This information is extremely valuable because, in principle, it assigns the finest possible location to each individual. In practice, however, it is highly non-standardized and challenging to use. In this paper, we expand earlier work by Lan and Longley (2019), who adopt a different strategy, only analyze the 1901 census, and geo-reference the 1851-1911 censuses. The geo-coded census sample is used in the neighborhood individual-level analysis; all other exercises do not rely on these data.

There are two ways to geo-reference historical addresses. One approach is manually digitizing historical locations, either streets or enumeration units, from historical maps. However, this method does not scale up and becomes rapidly unfeasible as the data grows. A second automated approach is to run text-based address matching between historical data sources and address databases that have already been geo-referenced. We follow this latter method since we need to geo-reference 5,464,578 unique addresses.

To perform the actual matching, we first operate a preliminary manual trimming of addresses. First, we remove house numbers because they undergo many changes and re-sequencing over time. Second, we remove uninformative locations, such as "village", "farm", and "rectory". Then, we input the resulting addresses as queries into the geo-referencing engine. Crucially, we discard the match if the resulting coordinates are not within the parish's boundaries where the address is recorded. This consistency check is necessary because homonyms are frequent. Since observing two addresses with the same name within a given parish is extremely rare, this ensures that the algorithm matches are not spurious.

We use the geo-coded 1891 census in the neighborhood analysis. We geo-code 85% of the addresses listed in the census, accounting for 86% of the entire population. The geo-coding ratio is homogeneous across districts except in Wales, where many addresses are reported in Welsh. In one robustness exercise, we thus omit Wales from the analysis sample and confirm that the results hold.

A.II Patent Data

A.II.1 Motivation

Despite its historical significance, we lack comprehensive patent data for the Second Industrial Revolution period (1850–1900) in the United Kingdom. In particular, it is impossible to reconstruct the geographical distribution of innovation activity during this period. This data limitation sharply contrasts the effort undertaken to document patenting activity since the inception of the English patent law in 1617 up until the end of the First Industrial Revolution in the 1840s (Nuvolari and Tartari, 2011; Nuvolari, Tartari and Tranchero, 2021). We fill this gap by constructing the first dataset of English and Welsh patents that spans the period 1853- 1900 and contains detailed information on the text, geographical location, inventors' personal information, and date for the universe of patents.

A.II.2 Data Sources

The UK Intellectual Property Office allowed us access to restricted full-page scans of original patent documents. These are the universe of patents granted in England and Wales between 1617 and 1899. This paper focuses on the period 1853-1899 for two main reasons. First, Nuvolari and Tartari (2011) already digitized patents before 1853 from Bennet Woodcroft's index, although patent documents contain additional information compared to the index. Second, in 1853, a reform dramatically lowered patent application prices. This makes it challenging to compare patents before and after the reform. Patent documents contain a wealth of unstructured information. We provide two examples in Figure A.1: in panel A.1a, we show the patent granted to Henry Bessemer for the eponymous process to produce steel, and in panel A.1b we display the patent granted to John Starley for the first modern safety bicycle. Both patents are in our dataset. The rectangles identify the location of the textual data that we extract. These comprise (i) a short title, (ii) a long title, (iii) the author(s)'s name(s), (iv) the author(s)'s address(es), (v) the author(s)'s profession(s), (vi) the filing date, (vii) the issue date, (viii) the type of protection, (ix) an indicator of whether the application was filed by an agent on behalf of someone living abroad, and (x) the full text of the patent. Not all (i-x) are available throughout the sample. In particular, (i), (vi), and (viii) are available only until 1873. After that date, a short title is no longer reported, the filing date is reported only sporadically, and the type of protection becomes immaterial, for only granted patents are included in the sample.

A.II.3 Digitization

We individually perform optical character recognition (OCR) on each patent to structure the data in a machine-readable dataset. To ensure state-of-the-art performance, we OCR the first page of each document, where all the (i–ix) variables are located, using Amazon's commercial textract engine. To retrieve the rest of the text not used in this paper, we use the open-source engine tesseract. An OCR-ed document is a text file. To extract the relevant variables, we harness the flexibility of state-ofthe-art large language models—specifically, GPT 3.5—to parse all variables (i–ix). The text of patent grants is relatively standardized, but inconsistencies arise relatively frequently due to idiosyncratic writing, phrasing, and OCR errors. The flexibility of large language models allows us to consistently and precisely extract all the relevant information for a large (95%) share of all patent applications. Running the same model on the title of each patent, we assign the technological class to patent grants. This exercise resulted in a database of approximately 800,000 patents granted between 1853 and 1899.

A.II.4 Geo-Coding

To retrieve each patent's location, we geocode each inventor's listed address using the commercial geocoding engine provided by MapTiler AG. To geocode an address, if a coarse geographical unit is listed on the patent (e.g., the county), we condition the outcome coordinates to lie within that unit. In Figure IIb, we report the resulting distribution of patents per capita, whereas Figure A.3 reports the spatial distribution of patent grants across technologies. Reassuringly, these are consistent with underlying population and economic development indicators as well as with historical evidence (e.g., note the substantial clustering of textile patents in the Lancashire districts).

A.II.5 External Validation

To validate our data, we consider the only two series covering a portion of the 1853-1899 years. Hanlon (2016) digitized an index of patents issued between 1855 and 1883. His data list, for each patent, the inventor(s) and their profession(s), a technology class, and the issue year. On top of the longer time coverage, our data thus contain several additional information, including the geographical coordinates. The second dataset we use to compare is the "A Cradle of Invention" (COI) series, published by Finishing Publications (2018). These data, too, were digitized from indices and thus only list authors, issue year, and, often, titles. In principle, this series spans the years 1617–1895. However, after 1883, patent applications that were eventually denied protection were also listed. Absent a way to identify granted patents, we do not report figures after 1883 for the COI series.

In Table A.1, we report the aggregate number of patents issued according to our series (columns 2 and

6), COI (columns 3 and 7), and Hanlon (2016) (columns 4 and 8). Reassuringly, the three series are highly consistent. Our series is closest to Hanlon (2016), but the COI figures are not too far off either. Overall, the Table strongly suggests that our series is as complete as the Hanlon (2016) database. We cannot, however, externally validate it for the later part of the period because there is no data available.

A.II.6 Measuring the Similarity Between US and UK Patents and their Quality

This section describes how we construct the patent similarity metric to measure "copying" and "originality" of UK innovation activity. The approach borrows heavily on Kelly *et al.* (2021). We adapt their methodology to our context by leveraging text information in titles only. Even though we do not have access to full US patent texts, the title of a patent is usually very informative about its content. We previously showed that a title-based machine learning algorithm predicts the technological classification of the patent with nearly 90% accuracy. Titles for UK patents are embedded in the digitized text for 1870–1899 and collected from PATSTAT for the later years; titles for US patents are collected from PATSTAT throughout the sample period.

We define the backward inverse-document frequency associated with each word w. This expresses the inverse frequency with which the word w appears in US patents p issued until year t. Formally, we have

$$BIDF_{w,t} \equiv \log\left(\frac{\text{Number of Patents Issued Before }t}{1 + \text{Number of Patents Issued Before }t \text{ that contain word }w}\right)$$
(A.1)

Then, to each patent-word pair, we associate the term frequency variable that counts the number of instances word w appears in patent p, normalized by the length of the patent. With a slight abuse of notation, let p denote the patent's index and the set of words it contains. We shall have

$$TF_{wp} \equiv \frac{\sum_{c \in p} 1(c = w)}{\sum_{c \in p} 1(c)}$$
(A.2)

where the numerator returns how many times word w appears in patent p, and the denumerator is simply the number of words in patent p. Then, we define the TF-BIDF associated with word w, patent p at time t as the product between these two terms:

$$TF-BIDF_{wp,t} \equiv TF_{wp} \times BIDF_{w,t}$$
(A.3)

and, thus, the vector TF-BIDF_{*p*,*t*} collects the term frequency-backward inverse document frequency for all words w in p. For comparability, the vector TF-BIDF_{*p*,*t*} is normalized by its norm to have unit

length.

We compute the TF-BIDF_{*p*,*t*} vectors for US and UK patents, but the BIDF_{*w*,*t*} are computed on the corpus of US patents only. Then, we compute the cosine similarity $\rho_{i,j}$ between each UK patent *i* and each US patent *j*. This allows us to define two variables. First, we seek to measure the similarity between British innovation and previous American patents. This yields a measure of backward similarity that, for the sake of the paper's narrative, we define as "copying". Formally we define

Backward Similarity^{$$au$$}_i $\equiv \sum_{j \in \mathcal{F}_i^{-\tau}} \rho_{i,j}$ (A.4)

where the set $\mathcal{F}_i^{-\tau}$ denotes the set of US patents issued within τ years from the issue year of patent *i*. This measures the degree of similarity between a given patent in the UK and previous patents in the US. Second, we define a measure of "originality" of UK patents compared to previous US patents. This leverages the insight of Kelly *et al.* (2021), who suggest that innovative and influential patents are those that are most dissimilar from existing innovation while at the same time retaining semantic proximity with subsequent patents. Formally, we have

Excess Forward Similarity
$$\equiv \frac{\sum_{j \in \mathcal{F}_j^{+\tau}} \rho_{i,j}}{\sum_{j \in \mathcal{F}_i^{-\tau}} \rho_{i,j}}$$
 (A.5)

where $\mathcal{F}_i^{+\tau}$ denotes the set of US patents issued within τ years after the issue year of patent *i*. In the baseline analysis, we set a symmetric window of $\tau = 5$ years around each patent's issue date. In Tables C.1 and C.9, we report the results using an alternative threshold of ten years. Moreover, in the same Table, we report the results obtained by netting out year and technology class fixed effects at the patent level. As noted by Kelly *et al.* (2021), this ensures we do not conflate shifting terminology fashions in the similarity measures.

A.II.7 Census Linking

To perform the individual-level analysis on neighborhood emigration, we link the inventors listed on patent documents issued between 1881 and 1900 to the 1891 population census. The availability of the census data determines the sample restriction. First, we seek to follow the inventing activity of individuals over time. Hence, we want to link patent records to one single census. When linking inventors to the census, we exploit the location listed on the document to restrict the pool of potential matches in the census. This practice, however, assumes that the inventor did not move between the time when the patent was granted and the closest census year. This assumption is sensible for patents issued not too far from the closest census year. Since the records of the 1871 census are not available, **Matching Algorithm** Given a patent *p*, define the set of inventors as $\mathcal{A}_p = \{A_1, \ldots, A_{n_p}\}$. Most patents are solo-authored in this period, meaning $|\mathcal{A}_p| = 1$. Call $\mathcal{L}_p = \{\ell_1, \ldots, \ell_{m_p}\}$ the set of locations patent *p* is associated to. Each ℓ is a couple of latitude-longitude coordinates. Let $\mathcal{L}_p^{\text{parish}}$ be the set of parishes associated with each coordinate. Analogously, let $\mathcal{L}_p^{\text{neighbor parishes}}$, $\mathcal{L}_p^{\text{district}}$, $\mathcal{L}_p^{\text{neighbor districts}}$, and $\mathcal{L}_p^{\text{county}}$ be the set of, respectively, neighboring parishes, districts, neighboring districts, and counties where each coordinate locates. Notice that these are progressively coarser units: parishes are contained in districts, which form counties. Unfortunately, we do not know the inventor-location pair. To match the generic A_p , we thus perform the following operations:

- 1. With a slight abuse of notation, let $\mathcal{L}_p^{\text{parish}}$ —and, analogously, $\mathcal{L}_p^{\text{district}}$ and $\mathcal{L}_p^{\text{county}}$ —denote the set of census records in each parish, district, and county within the respective sets.
- 2. Take all entries *i* within the set of parishes $\mathcal{L}_p^{\text{parish}}$ that are at least 18 when the patent *p* is filed. Let year_{*i*} and *t*_{*p*} respectively denote the birth year of *i* and the issue date:

$$\mathcal{M}_{A_p}^{\text{parish}} = \left\{ i \in \mathcal{L}_p^{\text{parish}} \middle| t_p - \text{year}_i \ge 18 \right\}$$
(A.6)

3. For each $i \in \mathcal{M}_{A_p}^{\text{parish}}$, compute the distance between the name and surname of *i*, and that of A_p :

Similarity^{$$A_p$$}_{*i*} = $\alpha \times$ Name Similarity ^{A_p} _{*i*} + (1 - α) × Surname Similarity ^{A_p} _{*i*} (A.7)

for some $\alpha \in [0, 1]$. In our baseline setting, we pick $\alpha = .3$ to assign a larger weight to the surname.

4. Define the set of acceptable matches as those with the highest similarity with the given A_p :

$$\overline{\mathcal{M}}_{A_p}^{\text{parish}} = \left\{ i \in \mathcal{M}_{A_p}^{\text{parish}} \middle| \operatorname{Similarity}_{i}^{A_p} = \max_{i' \in \mathcal{M}_{A_p}^{\text{parish}}} \operatorname{Similarity}_{i'}^{A_p} \right\}$$
(A.8)

and define Similarity^{A_p} as the similarity between all elements in $\overline{\mathcal{M}}_{A_p}^{\text{parish}}$ and A_p . Notice that this is the same across all $i \in \overline{\mathcal{M}}_{A_p}^{\text{parish}}$.

- 5. Set a threshold τ such that if Similarity^{A_p} < τ , $\overline{\mathcal{M}}_{A_p}^{\text{parish}} = \emptyset$, otherwise pass.
- 6. If $\overline{\mathcal{M}}_{A_p}^{\text{parish}}$ is not empty, then inventor A_p is matched to all records in $\overline{\mathcal{M}}_{A_p}^{\text{parish}}$. If empty, repeat steps 2–4 conditioning on records in the coarser matching set.

Patent data have the clear advantage that we have geographical information on the location of inventors. Inventors are mobile, however, and there may be a considerable time between the moment the patent is granted and the 1891 census. For these reasons, we incrementally exploit geographical information on the inventor's location. First, we look for high-quality matches within the same parish where the patent is filed. We then progressively expand the set of records by coarsening their geographic location to their neighboring parishes, districts, neighboring districts, and counties.

Evaluation of the Matching In Figure A.4, we report the matching rate of this exercise. We match approximately 75% of the inventors in the sample. The share remains constant throughout the period. Almost 50% of the matches are attained at the parish level—the smallest geographical layer we consider. About half of the remaining matches are obtained within the neighboring parishes of the location indicated in the patent document. These figures indicate that mobility remained relatively low within the relatively narrow time frame around the census we consider.

A plausible concern is that the probability of obtaining a link is not random. This may be the case if, for instance, more successful inventors were more educated and, hence, more likely to report their names correctly in the census. On the other hand, if successful inventors were relatively more mobile, we may fail at linking them because we may need to go national to obtain a match, which would most likely be dropped because of the multiple-match issue. These hypotheses are difficult to test. In Figure A.5, however, we report the overall distribution of the number of matches in the sample. Approximately 60% of the inventors listed in the patent grants are linked to one single census record, and the share of inventors linked to more than 10 records is negligible. In fact, we can restrict the analysis to single matches and obtain very similar results. In Table A.2, we compute the correlation between the number of matches in our sample and a set of individual observed characteristics. In Panel A, we have the age and the residence divisions while in Panel B we list the occupation reported in the census. The number of matches is correlated with some of these characteristics, but the magnitude of this association is small, except for the fact that inventors living in Wales have substantially more matches in the sample. To ensure that Wales does not drive our results, we replicate the main results, exclude them from the sample, and confirm that the results hold.

A.III Linked Migrants Sample

A.III.1 Data Sources

We rely on two sources of externally compiled data.³ For the US, we have access to the IPUMS fullcount non-anonymized census (Ruggles *et al.*, 2021). A census was taken in the US every ten years starting in 1790, except for 1890. Until 1840, the census was run at the household level. From 1850 on, instead, we have detailed *individual* information on the universe of the US population.⁴ For confidentiality, these data are available up until 1940. Our dataset, therefore, contains snapshots of the entire US population at any given decade between 1850 and 1940, although for the sake of this paper, we restrict to the years 1870-1930. Crucially, we have access to the non-anonymized version of the IPUMS data. Hence, we also know each individual's recorded name and surname besides publicly available information.

In the UK, the I-CeM data mirrors the IPUMS (Schurer and Higgs, 2020) content. More precisely, it contains information on the universe of people living in England, Scotland, and Wales. Similarly to the US and virtually every other census, it was run at a decade frequency from 1851 until 1911. No census was taken in 1871. As with the IPUMS data, we can access the full-count, non-anonymized version of the dataset. Besides publicly available information, this contains full names and addresses of the universe of individuals living in the UK at any given decade.

A.III.2 Linking Algorithm

Our methodology relies on Abramitzky *et al.* (2021). This dataset tackles the problem that neither the US nor the UK—nor any other European countries—recorded where British immigrants came from *within* the UK. Thus, we try to match British immigrants residing in the US with their entry in the UK census, which records where they come from at a granular geographical level.⁵ More precisely, we take the stock of British residing in the US in a given census year—say, 1900—and match them with their entry in the preceding UK census—in this case, 1891.⁶ This implies that we measure the *flow* of British immigrants over time rather than their stock.

We use three variables to link individuals: first name, surname, and birth year. The baseline sample

³We are deeply thankful to IPUMS and I-CeM for allowing us access to their confidential data. Without their help, this paper would not have been possible.

⁴By US population, we refer to the universe of individuals who *lived* in the US at a given point in time.

⁵Since women usually change their name upon marriage, we cannot match them. This is a common problem in linking algorithms (Abramitzky *et al.*, 2021).

⁶Since no census was taken in the UK in 1871, we link the 1880 US census to the 1861 UK one. This is not overly problematic because we can still match all those aged ten or older in 1871.

we link consists of individuals who report, in the US census, either England or Wales—or analogous denominations, such as Great Britain—as their country of origin. In the 1900 census, we take all those who immigrated between 1870 and 1899. In the subsequent censuses, until 1930, we retrieve stock of those who immigrated in the preceding decade. Then, to match each unit in the sample—call the generic one A—to an entry in the UK census, we perform this sequence of operations:

- 1. Take the census that precedes the immigration year of *A*. Hence, for instance, we match all those who immigrated in 1896 to the 1891 census.
- 2. Select all records in that census with the same reported birth year as *A*—call the resulting sample $\mathcal{M}^A = \{m_1^A, \dots, m_N^A\}.$
- 3. Compute a string-similarity measure between the name and surname of *A* and that of all elements of \mathcal{M}^A . In other words, for every $m_i^A \in \mathcal{M}^A$, compute⁷

Similarity^{*A*}_{*i*} =
$$\alpha$$
 × Name Similarity^{*A*}_{*i*} + (1 - α) × Surname Similarity^{*A*}_{*i*} (A.9)

for some $\alpha \in [0, 1]$. In our baseline setting, we set $\alpha = 0.3$ to give higher weight to the surname.

4. The set of matches is defined as

$$\overline{\mathcal{M}}^{A} = \left\{ m_{i}^{A} \in \mathcal{M}^{A} \mid \text{Similarity}_{i}^{A} = \max_{m_{i'}^{A} \in \mathcal{M}^{A}} \text{Similarity}_{i'}^{A} \right\}$$
(A.10)

which means that we restrict the set of possible matches to include only those whose similarity score with the entry in the US census *A* is the largest.

5. Finally, for a given threshold $\tau > 0$, we select only the possible matches whose similarity score is above τ . The set of effective matches thus boils down to:

$$\widetilde{\mathcal{M}}_{\tau}^{A} = \left\{ m_{i}^{A} \in \overline{\mathcal{M}}^{A} \mid \text{Similarity}_{i}^{A} \ge \tau \right\}$$
(A.11)

Clearly, $\widetilde{\mathcal{M}}^A$ can ideally be empty, meaning that *A* has no effective matches. It can have one element, in which case we refer to it as a "perfect match," or it can have multiple matches. In our baseline exercise, we set $\tau = 0.7$ as we see a clear elbow in the distribution of similarities there.

⁷We cannot simply match on exact same name and surname because coding errors are commonplace in historical census data (Abramitzky *et al.*, 2021).

We evaluate the distance between two strings *i* and *j* in terms of their Jaro-Winkler similarity d_{ij} :

$$d_{ij} \equiv \widehat{d_{ij}} + \ell p(1 - \widehat{d_{ij}}) \tag{A.12}$$

where

$$\widehat{d_{ij}} \equiv \begin{cases} 0 & \text{if } m = 0\\ \frac{1}{3} \left(\frac{m}{|i|} + \frac{m}{|j|} + \frac{m-t}{m} \right) & \text{else} \end{cases}$$
(A.13)

where *m* is the number of matching characters, |i| is the length of string *i*, *t* is half the number of transpositions, ℓ is the length of common an eventual common prefix no longer than four characters between *i* and *j*, and p = 0.1 is a constant scaling factor. Two characters are matching only if they are the same and are not farther than $\left\lfloor \frac{\max(|i|,|j|)}{2} \right\rfloor - 1$. Half the number of matching characters in different sequence order is the number of transpositions.⁸

The Jaro-Winker distance has been shown to perform well in linking routines (Abramitzky *et al.*, 2021). In our particular case, however, this metric outperforms more standard string dissimilarity metrics, such as the cosine or the Levenshtein distances, because the Jaro-Winkler assigns a "bonus" score to strings starting with closer initial substrings. In addition, coding errors are far more frequent at the end of names and surnames than at the beginning. A manual assessment confirmed that the Jaro-Winkler metric outperforms other measures in our setting.

A.III.3 Internal and External Validation

Matching Rate In Figure A.7, we report the key matching rate statistics of the linked sample. In particular, Figure A.7a displays the crude matching rate over time, i.e., the share of English immigrants recorded in the US census linked to at least one record in the UK census. Since it is impossible to link British immigrants born after the last census before they migrated (i.e., those who migrate when younger than ten), we distinguish between the full sample in blue and the "matchable" sample in red. The matching rate in the matchable sample is approximately 65%. It declines from 75% in the 1870s to 60-65% and remains constant after that. The large jump in 1880 in the crude matching rate is because the 1871 census is unavailable; hence, between 1870 and 1881, we cannot match anyone younger than 20—as opposed to 10 in the rest of the sample—when they migrated. The matching rate is relatively high compared to other algorithms, such as Abramitzky *et al.* (2021). In the next section, we thus asked whether matched and un-matched immigrants are balanced on observables. Panel A.7b replicates the previous plot but reports the absolute number of emigrants in the sample and those linked instead of the percentage shares.

⁸The Jaro-Winkler distance has been recently employed in the economic history literature for intergenerational linking purposes by, among others, Abramitzky *et al.* (2021)

In Panel A.7c, we display the distribution of the number of matches for linked immigrants. A large fraction ($\approx 45\%$) of immigrants are paired with a single entry in the UK census. Approximately 60% are matched to one or two entries. Beyond this threshold, multiple matches are more common, and the average number of matches is 9.4. To ensure that multiple matches do not confound our analysis, we always weigh emigrants by the inverse of the number of matches. In unreported analyses, however, we confirm that all our results remain qualitatively unchanged even if we exclude instances with more than two matches from the analysis sample.

In Figure A.8, we report the distribution of the name and surname similarities between the records of British immigrants in the US census and their links in the UK census. The dashed black lines plot the threshold below which we reject the link (.9). In the sample; we find that if an immigrant is linked to at least one record in the UK census, then the quality of the match is high, as more than 90% of the sample lies right to the .9 line. However, it is important to note that for a link to be accepted in the final sample, we require that *both* the name and the surname similarities be above 0.9.

Balance of Linked Emigrants In Table A.3, we report the correlation between the probability that a British immigrant in the US is linked to an entry in the UK census and a set of individual-level variables observed in the US census. Column (1) reports the sample value of each variable for unmatched immigrants, and column (2) refers to matched immigrants. Column (3) reports the difference between the two groups. As one may expect, the linking probability does not correlate with either literacy since educated immigrants could present systematically more precise census records—or income, another proxy for educational standing. Matched immigrants are less likely to work as professionals, in clerical occupations, and as sales workers. These differences, however, are small in magnitude and affect a small portion of the population. Manufacturing workers (skilled "craftsmen" and unskilled "operatives"), instead, account for almost 60% of the population and do not appear to be selected in terms of their linking probability. In terms of their region of residence, emigrants living in the Northeast are more likely to be linked, although the difference is minimal in magnitude (less than 2%). They are also less likely to reside in the South, but again, this difference affects only approximately 1% of the sample. Overall, Table A.3 provides reassuring evidence that linked emigrants are primarily not selected on observable characteristics and, importantly, the statistically significant differences are minor in magnitude.

Validation of the Data It is challenging to validate our sample with external data because, as we note in the main text, we do not have primary geographically disaggregated data on the origin of British emigrants. Baines (2002) attempts to estimate decade-level data for British counties between 1880 and 1910. To the best of our knowledge, these estimates are the only alternative source we can use to gauge the plausibility of our data. They nonetheless present important issues. First, they are based on

interpolations from population data tabulated from the population census. First, the author tries to account for internal migration, which remains an important confounding factor. Second, the estimates refer to aggregate emigration outflows, whereas our data precisely captures emigration to the United States. Third, the estimates are computed at the county level. Counties are coarse geographical units. Hence, this dataset does not warrant any modern econometric exercise.

Keeping these shortcomings in mind, in Figure A.9, we aggregate our data at the county-decade level and plot it against the estimates produced by Baines (2002). Both series are taken in log terms to reduce the influence of extreme values. Moreover, we weight the data by county population. We do so to ensure that the resulting correlation reflects the actual sizable differences in population across counties. We estimate a positive and statistically significant correlation between our data and the Baines series. Our figures are generally lower than those provided by Baines, which plausibly reflects that our data do not cover emigration towards countries other than the United States. Overall, we view this graph as providing supporting evidence of the plausibility of our dataset.

Finally, in unreported results, we construct an intergenerational linked sample from the English and Welsh population census. This sample allows us to follow individuals over two consecutive censuses between 1851 and 1911. Using this sample, we find that individuals recorded in census t that also appear as emigrants between census t and t + 10 are 60% less likely to be linked to the census in t + 10 than those that do not appear in the linked migrant sample. Moreover, eliminating from the emigration data those linked to the census in t + 10 yields qualitatively and quantitatively similar results to those shown in the paper. The intergenerational linking algorithm presents important issues, as described in Abramitzky *et al.* (2021), but provides additional evidence supporting the informativeness of our intergenerational linking exercise.

A.III.4 Return Migration Data

Following the same linking algorithm described before, we construct a linked sample of return migrants. This identifies English and Welsh immigrants in the US in year t and looks for possible matches in the UK census in year t + 10, using a minor variation on the algorithm described previously. Since the last UK census is the 1911 one, we face a hard upper bound for the coverage of return migration, as we can only construct return migrants linked samples spanning the period 1870–1910.

Previous research suggests that return migration rates during the Age of Mass Migration were substantial (Bandiera *et al.*, 2013), although probably less so in the UK than in second-wave countries such as Italy. Using our linked sample methodology, we find an approximately 30% return migration rate, broadly consistent with previous estimates.

	Years 185	3–1876		Years 1877–1899					
Year	New Data	Hanlon	COI	Year	New Data	Hanlon	COI		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
1853	2926		3042	1877	4937	4940	5133		
1854	2737		2715	1878	5338	5333	5258		
1855	2868	2955	2883	1879	5312	5325	5442		
1856	3204	3102	3003	1880	5531	5509	5211		
1857	3175	3197	3108	1881	5762	5745	5760		
1858	3071	2999	2988	1882	6244	6233	6187		
1859	3031	2998	3008	1883	6118	5981	6075		
1860	3231	3190	3161	1884	9809				
1861	3279	3272	3303	1885	8695				
1862	3479	3486	3485	1886	8913				
1863	3312	3308	3411	1887	9070)			
1864	3245	3257	3286	1888	9283	283			
1865	3391	3378	3436	1889	10315				
1866	3422	3452	3481	1890	10376				
1867	3720	3720	3723	1891	10768				
1868	3976	3984	3955	1892	11454				
1869	3837	3781	3781	1893	11986				
1870	3459	3405	3323	1894	11664				
1871	3574	3525	3542	1895	12243				
1872	3951	3967	4013	1896	13619				
1873	4261	4282	4336	1897	14304				
1874	4419	4491	4533	1898	13105				
1875	4576	4557	4537	1899	13267				
1876	5048	5064	5085						

 Table A.1. External Validation of Newly Digitized Patent Data

Notes. This Table reports the total number of patents in England and Wales between 1853 and 1899. Columns (2) and (6) report the series constructed from our novel dataset; columns (3) and (7) tabulate data from *A Cradle of Inventions* (Finishing Publications, 2018); columns (4) and (8) report data from Hanlon (2016). The *A Cradle of Inventions* series potentially stretches until 1899. However, after 1883 there is no way to distinguish between patents granted and applications. Hence we do not report figures for these later years (Nicholas, 2014). Data from Hanlon (2016) only cover the years 1855–1883. Referenced on page(s) A6.

	Uncond	ditional	Year FE		
	Mean	Std. Err.	Mean	Std. Err.	
	(1)	(2)	(3)	(4)	
Panel A. Age and Place	e of Resider	nce			
Age	-0.002	(0.002)	-0.002	(0.001)	
London	-0.197	(0.144)	-0.177	(0.128)	
South East	-0.362***	(0.132)	-0.342***	(0.114)	
East	-0.246*	(0.149)	-0.242*	(0.136)	
Yorkshire	-0.324**	(0.137)	-0.308**	(0.122)	
South West	-0.412***	(0.124)	-0.404***	(0.114)	
West Midlands	0.022	(0.199)	0.022	(0.186)	
East Midlands	-0.246*	(0.135)	-0.230*	(0.119)	
North East	-0.260**	(0.133)	-0.234**	(0.110)	
North West	-0.165	(0.159)	-0.169	(0.154)	
Wales	1.809***	(0.227)	1.752***	(0.186)	
Panel B. Occupation					
Agriculture	0.528	(0.365)	0.494	(0.320)	
Chemicals	-0.209**	(0.105)	-0.189**	(0.083)	
Construction	-0.061	(0.076)	-0.055	(0.068)	
Engineering	-0.070	(0.062)	-0.063	(0.052)	
Liberal Professions	-0.186***	(0.071)	-0.180***	(0.064)	
Metallurgy	0.114	(0.081)	0.109	(0.075)	
Public Administration	-0.150	(0.094)	-0.148*	(0.089)	
Textiles	-0.119	(0.093)	-0.109	(0.083)	
Trade	-0.164*	(0.084)	-0.155**	(0.072)	
Transports	0.016	(0.027)	0.018	(0.024)	
Utilities	-0.178**	(0.080)	-0.170**	(0.071)	

Table A.2. Balance of Linked Inventor Sample

Notes. This Table reports the correlation between the number of matches in the linked inventor sample and a set of individual-level co-variates observed in the census. In columns (1–2), we display the unconditional correlation and the associated standard error. In columns (3–4), we repeat the exercise but control for the issue year of the patent. Standard errors are clustered at the county level and are displayed in parentheses. Referenced on page(s) A10, C47.

*: p < 0.10, **: p < 0.05, ***: p < 0.01

	Correlation with Matching Status								
	= 1 if Not matched	= 1 if Matched	Diff.	Std. Err.					
	(1)	(2)	(3)	(4)					
Panel A. Individual Characteristics									
Literacy	0.885	0.965	0.079	(0.069)					
Income	3.147	3.141	-0.006	(0.011)					
Panel B. Occupation									
Professional	0.053	0.041	-0.012***	(0.002)					
Farmer	0.085	0.091	0.006	(0.010)					
Manager	0.066	0.059	-0.006	(0.004)					
Clerical	0.058	0.045	-0.013***	(0.004)					
Sales	0.062	0.047	-0.014***	(0.001)					
Craftsman	0.260	0.281	0.020*	(0.010)					
Operative	0.310	0.330	0.020	(0.014)					
Service	0.050	0.049	-0.001	(0.003)					
Laborer	0.057	0.056	-0.000	(0.004)					
Panel C. Region of Residence									
North East	0.519	0.539	0.019***	(0.005)					
Midwest	0.277	0.271	-0.006	(0.005)					
South	0.057	0.046	-0.011***	(0.001)					
West	0.146	0.145	-0.002	(0.008)					

Table A.3. Balance of Linked Emigrants Sample

Notes. This Table reports the correlation between the matching probability in the migrants-linked sample and a set of individual-level characteristics observed in the US census. The dependent variable equals one if the immigrant is linked and zero otherwise. We report the average value of the row variable for unmatched (column 1) and matched immigrants (column 2), as well as the difference between the two groups (column 3) along with its standard error clustered at the census year level (column 4). All row variables are indicators except income, which is the log of the occupational income score. Standard errors are clustered at the census year level and are displayed in parentheses. Referenced on page(s) 11, 29, A14, A14.

*: p < 0.10, **: p < 0.05, ***: p < 0.01
Figures

Figure A.1. Sample Annotated Patent Documents: the Bessemer Process and the First Modern Safety Bicycle

(a) Henry Bessemer's 1856 Patent



(b) John K. Starley's 1885 Bicycle Patent



Notes. This figure displays two sample patent documents in our dataset. Panel A.1a was granted to Henry Bessemer in 1856 to invent the famous eponymous process for the mass production of steel from molten pig iron. Panel A.1b was granted to John Starley in 1885 to invent the first modern bicycle, which would soon revolutionize mobility in Europe and the US. Colors mark different variables that we structure in the dataset: (i) in brown, the short title; (ii) in red, the complete title (iii) in green, the type of protection granted; (iv) in blue, the author(s) name(s); (v) in yellow, the author(s)'s address(es); (vi) in light blue, the application date; (vii) in purple, the issue date; (viii) in black, the patent text that continues in the rest of the patent document; (ix) in dark purple, the author(s) profession(s). Not all (i–ix) data are available on every patent and in each year. Referenced on page(s) A5.



Figure A.2. Number and Composition of Patents Over Time

Notes. This Figure reports time-series information on the innovation activity in Britain between 1853 and 1939. In Panel A.2a, we report the total number of patents granted in the UK over the period. The blue dots report the newly digitized data that we assembled for this paper; the red dots report tabulations from the Patstat repository. In Panel A.2b, we plot the share of patents granted over time across technology classes. Referenced on page(s) 13.

(a) Time Series of Patents



Figure A.3. Geographic Distribution of Patents Across Technologies

Notes. This Figure reports the intensity of patenting activity across districts over 1880–1939 for selected technology classes. Districts are displayed at 1891 borders. To assign patents to districts, we geo-reference the address of each author listed in the patent document and assign districts based on historical district borders. Black edges display county borders. The London area is displayed separately. Darker shades of blue indicate increasing quantiles of the patenting rate, defined as the percentage ratio between the number of patents in a given technology class and the overall number of patents produced. Referenced on page(s) A6.



Figure A.4. Matching Rate of Linked Inventors Sample

Notes. This Figure reports the matching rate of the linked sample of inventors. The records of the inventors who obtained a patent between 1880 and 1899 are linked to the 1891 population census, as detailed in the main text. The matching rate, i.e., the share of inventors successfully matched to the census, is reported on the *y*-axis in percentage points. The matching rate is broken down by the geographic layer of aggregation where the match is attained. Hence, we match over 75% inventors to the census throughout the period, and among those, the census record of slightly more than 50% of them is found in the same parish where the inventor is recorded living on the patent document. Referenced on page(s) A10.



Figure A.5. Number of Matches in Linked Inventor Sample

Notes. This Figure reports the number of census entries each inventor in the linked sample is matched to. Panel A.5a reports the overall distribution, while Panels A.5b–A.5e report the distributions broken down by geographic layers. In each graph, we separately report the average number of matches and its standard deviation. Referenced on page(s) A10.



Figure A.6. Geographic Distribution of Linked Inventors

Notes. This Figure displays the spatial distribution of inventors across districts between 1880 and 1900. Each marker reports one inventor, defined as an individual who obtains at least one patent over the sample period. To retrieve the coordinates of the inventors, we first link population censuses, whose entries are, in turn, georeferenced. The background map displays black counties and gray districts at historical borders in 1891. We highlight the ten largest urban centers at the time. Referenced on page(s) 13.



Figure A.7. Matching Rate and Number of Matches

Notes. This Figure reports information on the matching performance of the emigrants' linked sample. In Panel A.7a, we display the matching rate, i.e., the share of emigrants recorded in the US census that are successfully linked to at least one entry in the UK census. The blue dots refer to the overall sample, and the red dots refer to the emigrants who could be recorded in the UK census (see main text for more precise information). In Panel A.7b, we plot the overall number of immigrants in the US census (in green), the number of those that are matched (in red), and the number of matches that we accept (in blue). Panel A.7c reports the distribution of the number of matches, where the last bin collects all those with more than 20 matches. Referenced on page(s) 10, A13.

Figure A.8. Quality of Matches



(a) Name Similarity

Notes. This Figure reports the distribution of the match quality in terms of name and surname similarity for the set of records with no more than two matches in the baseline sample. The similarity measure we use to construct the links is the Jaro-Winkler. This string metric measures the edit distance between the name and surname of the British immigrant recorded in the US census and their match(es) in the UK census. Panel A.8a reports the distribution of the name similarity; Panel A.8b refers to surnames. The vertical lines mark the quality thresholds we impose for a match to be part of the final linked sample. Referenced on page(s) A14.



Figure A.9. External Validation with Out-Migration Estimates

Notes. This Figure reports the correlation between out-migration in our linked inter-census sample and the estimates produced by Baines (2002). These estimates are at the county level and span 1880–1910. We thus aggregate our data by county. Furthermore, we exclude the London area because we cannot map our data to Baines' geographical divisions. Both variables are expressed in log terms. The gray dots display county-decade observations. Counties are weighted by their population. The blue dots report binned means. The red line overlays a linear fit between the variables. The Figure also reports the regression coefficient, its robust standard error, and the R^2 of the regression. Referenced on page(s) 11, A15.



Figure A.10. Geographic Distribution of Emigration Rates Over Time

Notes. This Figure reports the distribution of US emigrants across districts in England and Wales over the period 1870–1940 by decade. Data are from the matched emigrants sample. The number of emigrants in each decade is normalized by population in 1891 and is expressed in percentage terms. Districts are displayed at their 1891 historical borders. Black edges also display historical county borders. Out-migration is cross-walked to consistent historical borders. Lighter to darker shades of blue indicate increasing quantiles of the emigration rate. The London area is displayed separately. Referenced on page(s) 11.

B ADDITIONAL RESULTS

This section presents in some detail several additional results that are mentioned in passing in the main text: the selection of British emigrants (Section B.I), the long-run association between emigration and innovation (Section B.II), and the possibility that immigrants sort into counties that are similar to their origin area (Section B.III).

B.I Selection of British Emigrants

The historical scholarship argues that the English and Welsh mass migration to the US starkly differed from that of other countries (Berthoff, 1953; Baines, 2002). Unlike other European countries, such as Germany, Sweden, or Italy, UK emigration to the US in the second half of the nineteenth century was not a low-skilled rural phenomenon. Especially after the 1880s, people started to leave urban, industrial areas. Importantly, emigrants did not represent the bottom of the human capital distribution, as was the case in Italy (Spitzer and Zimran, 2018) or Norway (Abramitzky *et al.*, 2014). This is crucial for our analysis, as it is unlikely that illiterate farmers would facilitate the flow of novel knowledge back to their origin areas. Even if this was the case, it would be equally unlikely that those rural areas would have the ability to reproduce US patents. While these considerations are helpful for our analysis, they largely rely on anecdotal evidence or analyses of incomplete census sources. In this section, we present evidence on the selection of English emigrants to the United States and returning migrants to the United Kingdom relative to the staying population in the UK. To construct these data, we exploit the linked US-UK migrants sample and the English population census.

Table B.1 presents. Column (1) refers to non-migrants, and columns (2) and (5) refer to emigrants and return migrants, respectively. In columns (3) and (6), we compute the difference between non-migrants and emigrants and non-migrants and return migrants, respectively. Emigrants are more likely than stayers to work as engineers and as manufacturing workers in metallurgy and textiles. Unsurprisingly, they are less likely to work in public administration and as liberal professionals, since those occupations could not be transferred overseas. Overall, these patterns confirm historical evidence by, among others, Baines (2002), who describes transatlantic emigrants as a positively selected group of entrepreneurial individuals well-versed in the manufacturing crafts, especially in the second half of the Nineteenth century. Return migrants appear somewhat different. They are more likely to work in construction, a lesser-skilled sector compared to emergent industrial jobs, and are more likely to work as professionals, public officers, in transports—which, in this period, would mainly comprise railway workers—and as utility workers. According to this sketched first inquiry, it is likely that the decision to return to the United Kingdom was relatively more common among those who were less successful in their American enterprise.

Individuals who migrated to the United States are more likely to originate from the North West including the industrial Lancashire districts—South West—especially the rural areas of Cornwall and Devon—and Wales. As mentioned in the main text, the origin of emigrants shifts over time from the mainly rural areas in Southern England to the industrial regions in the North and the Midlands. It appears that the probability of returning was not homogeneous across sending regions. Return migrants are less likely to reside in the East, in Wales, and in the West Midlands, while they are more common in the London area and, as are the emigrants, in South West. The different geographic distribution of emigrants and return migrants is crucial when disentangling their contributions to innovation activity in the UK.

B.II Long-Run Association between Emigration and Innovation

We now investigate the persistence of the effect of exposure to foreign knowledge through migration ties on the direction of patenting activity. While this exercise cannot be tasked with any claim of causality, it nonetheless suggests the possible far-reaching effects of out-migration on innovation.

We estimate the following regression:

$$Patents_{ik,t} = \alpha_{i \times k} + \alpha_t + \sum_{\tau \in \mathcal{T}} \beta^{\tau} \left[Knowledge \ Exposure_{ik} \times 1 \left(t = \tau \mid t = \tau + 1 \right) \right] + \varepsilon_{ik,t}$$
(B.1)

where *i*, *k*, and *t* denote a district, technology class, and year, respectively. In this setting, we have $t \in [1940, 2015]$. The term Knowledge Exposure_{*ik*} refers to knowledge exposure in the years 1900–1930, i.e., before the sample period. To reduce noise in the estimated β^{τ} coefficients, we conflate years in \mathcal{T} in biennial windows. The estimated set of β^{τ} expresses the conditional correlation between historical exposure to knowledge exposure and innovation activity in the two-year window indexed by τ .

In Figure B.1, we report the set of estimated β^{τ} over time. The correlation between historical knowledge exposure and patenting activity remained positive and significant until the early 1980s, although it—reassuringly— decreased over time. We interpret this as evidence that exposure to foreign knowledge through migration ties has a potentially long-lasting effect on the composition of innovation activity over time. In Table B.2, we re-estimate model (B.1), sector-by-sector, by decade. Compared to (B.1), we can thus only include district and decade-fixed effects. Columns report the estimated β^{τ} by decade. The estimated correlation between historical exposure and patenting decreases over time in all sectors and, by the 1990s, it is no longer significant in many.

B.III Assortative Matching of Emigrants in the United States

In this section, we lay down a simple framework to test whether British immigrants sort into US counties depending on the innovation similarity between the settlement location and their origin district. Let $\mathbf{P}_{j,t} = \{p_{1j,t}, \dots, p_{Nj,t}\}$ denote the patent portfolio of county *j* in decade *t*, whose generic entry p_{kjt} returns the number of patents in technology class *k*. Analogously, let $\mathbf{P}_{i,t}$ be the portfolio of district *i*. We define a metric of innovation similarity as follows:

Innovation Similarity_{*ij,t*}
$$\equiv \frac{\mathbf{P}_{i,t}^{\mathsf{T}} \mathbf{P}_{j,t}}{\|\mathbf{P}_{i,t}\| \cdot \|\mathbf{P}_{j,t}\|} = \frac{\sum_{k} p_{ki,t} p_{kj,t}}{\sqrt{\sum_{k} p_{ki,t}^2} \sqrt{\sum_{k} p_{kj,t}^2}} \le 1$$
 (B.2)

which is a simple cosine similarity. The similarity measure returns a value of one if the patent portfolios of district *i* and county *j* are equal, meaning their composition across classes is the same. The index is normalized between zero and one.

We then estimate variations on the following simple linear probability model:

$$\text{Emigrants}_{i \to j,t} = \alpha_{i \times j} + \alpha_t + \beta \times \text{Innovation Similarity}_{ij,t} + X_{ij,t}\Gamma + \varepsilon_{ij,t}$$
(B.3)

where the dependent variable is the flow of emigrants from district *it* to county *j* in decade *d*, and $\alpha_{i\times j}$ denotes county-by-district fixed effects. Controlling for the dyadic fixed effects $\alpha_{i\times j}$ is important because it implies that $\hat{\beta}$ relies on time variation in innovation similarity rather than pre-determined district-county factors that persist over time. The coefficient β thus yields the correlation between the similarity of innovation activity and migration flows. Standard errors are two-way clustered by district and county. Under sorting, one would expect $\hat{\beta} > 0$.

We test this prediction in Table B.3. We do not find evidence that the innovation similarity between origin and destination areas explains migration patterns. The sign of the regression coefficients is unstable and, once we control for time-varying shifters at the county or district level, negative and statistically significant. In columns (4–6), we estimate the same regressions on the smaller sample of county-district pairs with positive migration flows. The results remain largely similar. We do not wish to over-emphasize these results. Our measure of innovation similarity, while intuitive, may be subject to measurement error, which may reduce the precision of the estimates. Taken together, however, we interpret these results as evidence that it is unlikely that assortative matching played a pivotal role in determining the location choice of immigrants.

B.IV A Case Study of a US Technology Shock: The Great Influenza Pandemic

Our data-driven approach to isolate shocks to US innovation activity yields rich variation in the number and timing of innovation shocks. However, it is hard to trace those shocks to concrete periods of heightened innovation activity. In this section, we discuss one such example. Berkes *et al.* (2023) document that innovation activity in pharmaceuticals increased in response to higher mortality due to the Great Influenza pandemic that spread throughout the United States between 1918 and 1919. We thus ask whether our algorithm to identify innovation shocks reflects the heterogeneous responses to the pandemic across counties and technologies.

We follow the approach of the original paper and use the US mortality statistics to construct a measure of excess mortality due to the Flu. This "Excess Deaths" measure is defined as

Excess Deaths_j
$$\equiv \frac{\frac{1}{2}\sum_{t=1918}^{1919} \text{Deaths}_{jt}}{\frac{1}{3}\sum_{t=1915}^{1917} \text{Deaths}_{jt}}$$
, (B.4)

where *j* and *t* denote a county and year, and Deaths_{*jt*} is the number of individuals who died in county *j* and year *t*. Mortality data are available since 1915—hence, the running window of 1915–1917 to compute the pre-pandemic average mortality—and cover approximately 1,170 counties, which approximately account for 50% of the US population in 1920.

We estimate two specifications. The double-differences estimator compares counties by the severity of the influenza, before and after 1918:

$$y_{jkt} = \alpha_{j \times k} + \alpha_{k \times t} + \beta \times (\text{Excess Deaths}_j \times \text{Post}_t) + \varepsilon_{jkt}, \tag{B.5}$$

where *j*, *k*, and *t* denote a county, technology class, and year. The terms $\alpha_{j\times k}$ and $\alpha_{k\times t}$ denote countyby-technology and technology-by-time fixed effects, and Post_t is equal to one for $t \ge 1918$ and zero otherwise. In this specification, we leverage across-county variation in influenza mortality. The second specification is a triple-difference regression:

$$y_{jkt} = \alpha_{j \times k} + \alpha_{k \times t} + \alpha_{j \times t} + \beta \times (\text{Excess Deaths}_j \times \text{Post}_t \times \text{Pharma}_k) + \varepsilon_{jkt}, \tag{B.6}$$

where we further include county-by-year fixed effects ($\alpha_{j \times t}$) and the treatment includes an interaction with a term (Pharma_k) which is equal to one if *k* is a pharmaceutical-related technology ("health, amusement" and "chemistry") and zero otherwise. In this case, we further leverage cross-technology variation because we expect the influenza to only affect pharmaceutical-related innovation. Both models are estimated over a 10-year window around the influenza (1910–1930).

First, we replicate the baseline results of the original paper. In Table B.4 (columns 1–3), we use the (log 1+) number of patents as the dependent variable. Column (1) refers to the triple differences specification (B.5), whereas in columns (2) and (3), we estimate regression (B.6) separately on non-pharmaceutical and pharmaceutical patenting. Exposure to the influenza does not impact non-pharmaceutical patenting (column 2), whereas it presents a positive and large effect on pharmaceutical-related innovation (column 3). Moving from the 25th (1.03) to the 75th (1.20) percentile of excess deaths increases innovation by 25%. This heterogeneity is reflected in the triple-differences estimate, which similarly indicates that exposure to the influenza increased innovation in pharmaceuticals relative to all other technology classes.

Second, we explore the effect of the influenza on US innovation shocks. In this case, the dependent variable is an indicator equal to one if, in cell *jkt*, our algorithm identifies a technology shock, and zero otherwise. We report the estimates in columns (4–6). In column (5), we confirm that exposure to the pandemic had virtually no effect on the probability of observing a shock on non-pharmaceutical innovation. By contrast, column (6) displays that higher exposure to the pandemic results in a higher probability of observing a shock to US pharmaceutical innovation. Quantitatively, moving from the 25th to the 75th percentile of excess deaths results in a 2% increase in the probability of a US innovation shock in pharmaceuticals, which approximately corresponds to 22% of the pre-treatment mean. As in the previous analysis, this pattern implies that the triple differences estimates indicate the Great Influenza pandemic's positive and sizable effect on US innovation shocks in pharmaceutical technology relative to other sectors.

Our data-driven approach to measuring county-technology-level innovation shocks yields rich variation in US technology shock which, however, may lack historical interpretability. This exercise illustrates that, at least in one significant case, one can map these shocks to historically relevant episodes. We view this correspondence as instructive about the ability of our approach to capture appropriate variation in innovation activity.

B.V Anecdotal Evidence of Return Innovation

Who were the immigrants that contributed to the diffusion of US technology in Britain? History is rife with examples of skilled artisans, entrepreneurs, and factory workers who were exposed to some novel technology where they settled and promoted its diffusion, or in some cases appropriated it, in the UK.

In this section, we provide three instructive examples. All three are cases of return migration. Historical records typically focus on successful migrants who, upon returning, bring their technology to their origin areas and promote economic development there. The statistical analysis that we present, however, suggests that this was only part of the story. In fact, we find that emigrants interacted with their origin communities even without returning.

B.V.1 British Puddlers and the Kelly-Bessemer Process

An 1856 article published in Scientific American described a new patent granted in the UK to Henry Bessemer (Wagner, 2008). Bessemer had discovered a new process, the would-be eponymous Bessemer process, that, for the first time, allowed the production of inexpensive steel from molten pig iron.⁹ American inventor William Kelly complained:

"I have reason to believe my discovery was known in England three or four years ago, as a number of English puddlers visited this place to see my new process. Several of them have since returned to England and may have spoken of my invention there." (Wagner, 2008, p. 363)

The veracity of Kelly's allegations remains unverified. They nonetheless indicate three important elements. First, American inventors knew that British immigrants posed a threat to the secrecy of their inventions. Second, technology transfer did not necessitate the very upper tail of the human capital distribution. Skilled workers, such as puddlers, could be the agents of technology diffusion. Finally, the precise mechanism that emerges is return migration. Kelly expects British puddlers to speak of "his" invention upon returning to England.

B.V.2 Henry Marsden and the Industrialization of Leeds

Henry Rowland Marsden was born in Leeds to poor parents in 1823 (Curtis, 1875). At age twentyfive, he emigrated to the United States, first to New York and then to Connecticut. There, he took on apprenticeships in engineering and metal-working firms. He obtained several engineering patents chiefly related to steam engines and pumps, including a "stone-crusher" which is still in use today. In 1862, Marsden returned to Leeds, where he set up a flourishing business centered around his newly patented inventions. A wealthy man respected for his philanthropic endeavors, he was elected mayor of Leeds in 1873. He died in 1878 and is credited as one of the most prominent figures in the industrial development of Leeds.

⁹The Bessemer process was one of the most transformative technological developments of the nineteenth century (Rosenberg and Trajtenberg, 2004).

B.V.3 Migrants as Agents of Technology Transfer: Wellstood & Smith Ltd.

The case of Stephen Wellstood and John Smith illustrates how international migration spurs technology transfers across countries. At age 16, James Smith (1811–1886) left Bonnybridge, Scotland, and migrated to the US. There, he established himself selling cooking stoves and married. However, as his wife got ill, Smith returned to Bonnybridge and started re-selling imported stoves from the US. He soon realized, however, that he could manufacture stoves directly in Britain. He then partnered with his long-time friend Stephen Wellstood and opened a foundry. They patented the exact same cooking stove Smith had been selling in the US and started a business that remained active until 1983.

	Non-Migrants		Emigrants		I	Return Migra	nts
	Mean	Mean	Difference	Std. Err.	Mean	Difference	Std. Err.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Employment	by Sector (Depen	dent vari	able = 100 if	individual	employe	d in:)	
Agriculture	29.349	28.995	-0.100	(0.072)	26.255	-1.369***	(0.244)
Chemicals	0.949	0.916	-0.057***	(0.015)	1.006	-0.035	(0.055)
Construction	15.686	16.014	0.068	(0.059)	17.029	0.444**	(0.210)
Engineering	14.501	14.612	0.225***	(0.056)	14.280	0.185	(0.195)
Entrepreneur	0.007	0.003	-0.003***	(0.001)	0.004	-0.001	(0.004)
Liberal Professions	3.928	3.184	-0.792***	(0.029)	4.516	0.247**	(0.118)
Metallurgy	3.110	3.704	0.514***	(0.030)	3.247	-0.068	(0.097)
Public Administration	3.497	3.437	-0.124***	(0.029)	4.007	0.220**	(0.110)
Textiles	9.488	10.002	0.658***	(0.049)	8.135	-0.226	(0.154)
Trade	7.909	7.852	-0.068	(0.045)	8.534	-0.038	(0.158)
Transports	10.771	10.475	-0.317***	(0.048)	11.759	0.482***	(0.180)
Utilities	0.805	0.807	-0.003	(0.014)	1.226	0.159***	(0.061)
Panel B. Region of Orig	gin (Dependent v	ariable =	100 if indivi	dual reside	s in:)		
East	10.494	8.436	-2.061***	(0.033)	10.017	-0.328**	(0.149)
East Midlands	6.384	5.883	-0.519***	(0.028)	6.212	-0.062	(0.119)
Greater London	13.639	12.059	-1.712***	(0.040)	15.358	1.733***	(0.182)
North East	6.580	7.107	0.562***	(0.031)	6.436	-0.210*	(0.121)
North West	17.431	19.598	2.521***	(0.047)	17.287	-0.407**	(0.184)
South East	12.408	10.443	-1.918***	(0.037)	12.899	-0.139	(0.167)
South West	6.444	7.377	0.923***	(0.033)	6.521	0.932***	(0.126)
Wales	6.450	8.365	1.918***	(0.030)	6.379	-1.204***	(0.115)
West Midlands	11.189	11.396	0.182***	(0.038)	10.413	-0.565***	(0.149)
Yorkshire	8.596	8.900	0.057	(0.035)	7.874	0.098	(0.133)

Table B.1. Comparison between English Emigrants and Stayers

Notes. This Table displays the selection of emigrants and returning migrants to and from the United States relative to the rest of the British population. The unit of observation is an individual. In each row, the dependent variable is equal to 100 if the individual belongs to the given category (e.g., if they are employed in agriculture) and zero otherwise. In columns (1), (2), and (5) we report the averages for non-migrants, migrants, and return migrants. In columns (3) and (6), we display the difference between non-migrants and, respectively, migrants and return migrants. The associated robust standard errors are displayed in parentheses in columns (4) and (7). Referenced on page(s) B29.

Table B.2. Long-Run Correlation between Exposure to US Technology and Innovation in the UK

	(log) Number of Patents by Technology Class							
	1940s	1950s	1960s	1970s	1980s	1990s	2000s	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Agriculture	0.078***	0.044***	0.029***	0.029***	0.017*	0.016**	0.020***	
	(0.012)	(0.009)	(0.010)	(0.010)	(0.009)	(0.008)	(0.007)	
Building	0.186***	0.158***	0.129***	0.112***	0.055***	0.025**	0.013	
	(0.017)	(0.022)	(0.016)	(0.017)	(0.011)	(0.012)	(0.010)	
Chemistry	0.040***	0.037***	0.033***	0.019***	0.014***	0.007***	0.006***	
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	
Electricity	0.046***	0.045***	0.036***	0.032***	0.022***	0.007	0.005	
	(0.007)	(0.008)	(0.007)	(0.007)	(0.005)	(0.005)	(0.004)	
Engineering	0.112***	0.101***	0.110***	0.092***	0.062***	0.018	0.031***	
	(0.020)	(0.016)	(0.020)	(0.019)	(0.016)	(0.011)	(0.008)	
Engines, Pumps	0.053***	0.048***	0.050***	0.043***	0.023***	0.014***	0.002	
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	
Food	0.098***	0.098***	0.098***	0.082***	0.057***	0.058***	0.046***	
	(0.012)	(0.011)	(0.012)	(0.010)	(0.011)	(0.005)	(0.005)	
Health, Amusement	0.040***	0.037***	0.036***	0.029***	0.016***	0.006***	0.001	
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	
Instruments	0.109***	0.089***	0.106***	0.090***	0.071***	0.054***	0.033***	
	(0.009)	(0.011)	(0.008)	(0.011)	(0.008)	(0.006)	(0.006)	
Lighting, Heating	0.259***	0.191***	0.253***	0.180***	0.157***	0.060***	0.020	
	(0.037)	(0.040)	(0.031)	(0.029)	(0.038)	(0.021)	(0.013)	
Metallurgy	0.107***	0.105***	0.093***	0.087***	0.059***	0.030***	0.021***	
	(0.008)	(0.007)	(0.009)	(0.007)	(0.007)	(0.006)	(0.005)	
Mining	0.047***	0.039***	0.039***	0.030***	0.019***	0.016***	0.006*	
	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	
Personal Articles, Furniture	0.108***	0.124***	0.112***	0.112***	0.132***	0.011	-0.020	
	(0.029)	(0.028)	(0.027)	(0.028)	(0.035)	(0.019)	(0.018)	
Printing	0.063***	0.054***	0.050***	0.036***	0.020***	0.001	0.004	
	(0.007)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.004)	
Separating, Mixing	0.064***	0.061***	0.061***	0.048***	0.031***	0.010***	0.002	
	(0.006)	(0.005)	(0.005)	(0.006)	(0.004)	(0.004)	(0.004)	
Shaping	0.072***	0.060***	0.061***	0.040***	0.028***	0.018***	0.012***	
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	
Textiles	0.344***	0.125***	0.121***	0.103***	0.142***	0.084***	0.052**	
	(0.079)	(0.028)	(0.033)	(0.024)	(0.031)	(0.027)	(0.023)	
Transporting	0.047***	0.045***	0.041***	0.032***	0.019***	0.005	0.000	
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	
Weapons, Blasting	0.061***	0.046***	0.041***	0.033***	0.017***	-0.005*	-0.004	
	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	

Notes. This Table reports the long-run association between exposure to US technology and innovation in the UK. The unit of observation is a district-technology pair observed at a decade frequency between 1940 and 2010. The dependent variable is the (log) number of patents. The independent variable is an interaction term between exposure to US technology over the 1930s, decade dummies, and technology dummies. The 2010 decade dummy serves as the baseline category. The regression includes district-by-time, district-by-technology, and technology-by-time fixed effects. Standard errors are shown in parentheses and are clustered two-way by district and technology. Referenced on page(s) B30.

			Number of I	Emigrants		
		All Pairs		Posit	ive Migration	Ties
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable Mean	0.036	0.036	0.036	0.422	0.422	0.422
Innovation Similarity	2.875	-7.383***	-4.558*	-77.969**	-149.167***	14.269
	(1.848)	(2.831)	(2.664)	(35.306)	(46.664)	(51.243)
County-District FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	_	_	Yes	_	-
County-Decade FE	No	Yes	Yes	No	Yes	Yes
District-Decade FE	No	No	Yes	No	No	Yes
Number of Counties	1,799,936	1,799,936	1,799,936	255,523	254,968	254,968
Observations	10,799,616	10,799,616	10,799,616	869,632	868,418	868,418

Table B.3. Assortative Matching of British Immigrants in the United States

Notes. This Table tests the hypothesis that British immigrants settled in US counties that innovated in the same fields of their district of origin. The unit of observation is a district-county pair. Units are observed at a decade frequency between 1870 and 1930. The dependent variable is the number of migrants between the (UK) district and the (US) county. The independent variable is the similarity of the patent portfolios between the district and the county. The details of the similarity metric are explained in the main text. In columns (1–3), we include all district-county dyads; in columns (4–6), we include only the pairs with positive migration ties. All regressions include county-by-district and decade fixed effects; in columns (2) and (4), we include county-by-time fixed effects; columns (3) and (6) further include district-by-time fixed effects. Standard errors are clustered two-way by district and county and are displayed in parentheses. Referenced on page(s) 16, B31.

	Ν	J. of Patent	ts	I(Tec	hnology Sl	nock)
	(1) All Patents	(2) Non Pharma	(3) Pharma	(4) All Patents	(5) Non Pharma	(6) Pharma
Dependent Variable Mean	0.274	0.281	0.215	0.090	0.091	0.080
Excess Deaths \times Post 1918 \times Pharma	0.109*** (0.030)			0.086*** (0.024)		
Excess Deaths \times Post 1918		0.018 (0.018)	0.127*** (0.034)		0.001 (0.007)	0.087*** (0.023)
County-Year FE	Yes	_	_	Yes	_	_
County-Technology FE	Yes	Yes	Yes	Yes	Yes	Yes
Technology-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Counties Observations	1,170 422,370	1,170 377,910	1,170 44,460	1,170 422,370	1,170 377,910	1,170 44,460

Table B.4. Case Study of US Innovation Shock: The Great Influenza Pandemic

Notes. This Table illustrates the effect of the Great Influenza Pandemic (1917–1918) impacted patents (columns 1–3) and innovation shocks (columns 4–6). The unit of observation is a county-technology pair observed at a yearly frequency between 1910 and 1930. In columns (1–3), the dependent variable is the (log) number of patents; in columns (4–6), it is a binary variable equal to one if in the given county (or county-technology pair) experiences an innovation shock. In columns (2–3) and (5–6), the independent variable is an interaction term between excess mortality in 1919–1918 relative to 1915–1917 and a post-Flu indicator equal to one after the Influenza pandemic (i.e., after 1917). In columns (1) and (4), the baseline treatment is further interacted with an indicator equal to one for the pharmaceutical—i.e., "health, amusement" and "chemistry"—technology classes. Regressions in columns (1) and (4) include county-by-year, technology-by-year, and county-by-technology fixed effects; regressions in columns (2–3) and (5–6) include county-by-technology and technology-by-year fixed effects. Standard errors are shown in parentheses and are clustered at the county level. Referenced on page(s) B33.

Figures





Notes. This Figure reports the long-run association between exposure to US technology and innovation in the UK. The unit of observation is a district-technology pair observed at a decade frequency between 1940 and 2010. The dependent variable is the (log) number of patents. The independent variable is an interaction term between exposure to US technology over the 1930s and decade dummies. The 2010 decade dummy serves as the baseline category. Each dot reports the coefficient of an interaction term by decade. The regression includes district-by-time, district-by-technology, and technology-by-time fixed effects. Standard errors are clustered two-way by district and technology. Bands report 95% confidence intervals. Referenced on page(s) 20, B30.

Figure B.2. Heterogeneous Effects of Within-Neighborhood Emigration on Innovation by Occupation of the Stayers



Notes. This Figure reports how emigrants to the United States impact the innovation activity fulfilled by their neighbors who remain in the UK. The unit of observation is an individual inventor observed at a yearly frequency between 1880 and 1900. The analysis sample is the universe of inventors linked to the 1891 population census, as detailed in the main text. The dependent variable is the (log) number of patents produced by the members of the family. Each coefficient refers to an interaction term between the baseline treatment—an indicator equal to one after the first neighbor of the inventor moves to the US, and zero otherwise—and a dummy variable that codes the occupation of the inventor. The dashed red line indicates the average treatment effect reported in the main text. All regressions include inventor and year-fixed effects. Standard errors are clustered at the county level. The bands report 95% confidence intervals. Referenced on page(s) 30.

C ROBUSTNESS ANALYSES

In this section, we describe in detail the exercises we perform to assess the robustness of the results presented in the paper on the OLS analysis (Section C.I), the double and triple differences regressions (Section C.II), and the neighborhood analysis (Section C.III).

C.I Panel Analysis

C.I.1 Alternative Dependent Variables

In the principal analysis, we use the (log 1+) number of patents at varying levels of aggregation as the dependent variable. Following Chen and Roth (2024), who note that this transformation makes the estimates of the average treatment effect scale-dependent, in Table C.1 (columns 1–5), we show that the results are robust using a battery of alternative transformations.

In the baseline analysis, we measure the originality of British patents and their similarity to US patents using a five-year window, following Kelly *et al.* (2021), as detailed in Appendix A.II. In columns (6–8) and (9–11), however, we show that the results remain robust when adopting different time windows to compute these text-based measures.

In addition, in Table C.2, we restrict the dependent variable to comprise only patents with at least one firm assignee. This restriction generates a series of patenting that can potentially reflect technology adoption and diffusion within firms rather than the activity of independent inventors. We confirm that the baseline results are confirmed using this stricter definition of patenting.

C.I.2 Alternative Definitions of Knowledge Exposure

In Table C.3, we employ four alternative measures of knowledge exposure. First, we take the log of the baseline. Second, we construct a measure that fixes bilateral emigrant flows:

Knowledge Exposure²_{*ik,t*} =
$$\sum_{j} \left(\frac{\text{Patents}_{jk,t}}{\text{Patents}_{j,t}} \times \text{Emigrants}_{i \to j,1880} \right)$$
, (C.1)

which, compared to the main measure, restricts assortative matching to the first decade of the analysis. Third, we define the mirror measure that holds fixed specialization patterns across counties:

Knowledge Exposure³_{*ik,t*} =
$$\sum_{j} \left(\frac{\text{Patents}_{jk,1880}}{\text{Patents}_{j,1880}} \times \text{Emigrants}_{i \to j,t} \right).$$
 (C.2)

Compared to the main measure, this ensures that knowledge exposure does not conflate variation in patenting activity across counties determined or influenced by English immigrants. Finally, we define an alternative measure that leverages the *stock*, instead of the *flow* of patents issued:

Knowledge Exposure⁴_{*ik*,*t*} =
$$\sum_{j} \left[\sum_{\tau \le t} \left(\frac{\text{Patents}_{jk,\tau}}{\text{Patents}_{j,\tau}} \right) \times \text{Emigrants}_{i \to j,t} \right].$$
 (C.3)

The idea behind (C.3) is that specialization can be defined in terms of the cumulative number of patents filed before the given period. Finally, we construct the baseline measure but using the level of patents instead of the share:

Knowledge Exposure⁵_{*ik*,*t*} =
$$\sum_{j} \left(\text{Patents}_{jk,t} \times \text{Emigrants}_{i \to j,t} \right).$$
 (C.4)

This last metric accounts for the fact that, in small counties with little patenting activity, the share of patents would misrepresent the actual composition of innovation relative to large counties with diversified innovation portfolios. In practice, however, the measure in (C.4) is not vastly different from the baseline metric because few immigrants settled in those small counties to begin with. In Table C.3, we show that all these measures yield quantitatively similar results.

Additionally, in Table C.4, we construct variations of the baseline knowledge exposure measure, which include only patents in the top ℓ % of the impact distribution. Specifically, the generic *k*-Knowledge Exposure variable is defined as

Knowledge Exposure⁶_{*ik,t*}(
$$\ell$$
) = $\sum_{j} \left(\frac{\text{Patents in Top } \ell\% \text{ Impact}_{jk,t}}{\text{Patents}_{j,t}} \times \text{Emigrants}_{i \to j,t} \right).$ (C.5)

We consider three ℓ thresholds: the top 20%, 10%, and 5%. The association between innovation and knowledge exposure is larger for more restrictive ℓ -Knowledge Exposure metrics. This pattern is consistent with the plausible hypothesis that migrants' incentive to transmit knowledge to the UK was larger for patents with a higher economic return.

C.I.3 Alternative Standard Errors

In Figure C.1, we estimate the baseline regressions using several estimators for the standard errors. In particular, we employ the estimator developed by Conley (1999) to account for the potential spatial autocorrelation in both emigration rates and exposure to US innovation across technology classes. All the results remain statistically significant at standard confidence levels using these alternative estimators.

C.I.4 Alternative Cut to Linked Sample

In the baseline analysis, we retain all sufficiently high-quality migrant matches. Some British immigrants recorded in the US census are paired with multiple UK census records, as detailed in Section A.III. In such cases, we follow the standard practice in the literature and weigh each such match by the inverse of the number of matches to avoid multiple counting.

In Table C.8, we take an alternative route and compute the number of US emigrants (columns 1– 4) and knowledge exposure (columns 5–8) using single matches. This approach ensures we do not need to apply the inverse matching weights. An advantage of only considering single matches is that one may be worried that the probability of obtaining multiple matches correlates with unobserved individual characteristics. In practice, however, the estimates remain qualitatively and quantitatively unchanged relative to the baseline.

C.I.5 Alternative Sample of US Counties

A possible concern for the baseline analysis is that our results are driven by British immigrants directly impacting innovation in the counties where they settle, thus generating a spurious association between innovation in those countries and in their areas of origin. To address this concern, we leverage a sample of US inventors linked to the census produced by Sarada, Andrews and Ziebarth (2019). We construct the share of British-born inventors for each US county based on patents filed between 1870 and 1879 and divide this variable into deciles.¹⁰ We use these deciles to construct two additional measures of US emigration and knowledge exposure, which exclude the migrants who settled in counties in the top 10% and 20% share of British-born inventors.

Table C.6 replicates the baseline analysis using these two modified treatments—displayed in Panels A and B. We find that the baseline results remain largely unchanged even when we exclude counties in the top 10% or top 20% share of British-born inventors to compute the number of US emigrants (Panel A) and the exposure to US innovation (Panel B).

C.I.6 Exclusion of Inventors Possibly Appearing on UK and US Patent Records

Our results may be driven by inventors applying for the same patent in the US and the UK. It is worth noticing that this phenomenon would impact our estimates only if there were a correlation between county-district migration and the inventors' location. Second, this behavior may still signify technology transfer operated by the inventors and, thus, contribute to return innovation.

To quantify its relevance, we compile the number of British patents-as well as their average simi-

¹⁰The deciles are constructed over the set of counties with a non-zero share of British-born inventors.

larity and the number of high-impact patents—excluding those filed by inventors who also appear in some US patents over a ten-year window around the patent issue year. In practice, this approach is very conservative because we do not know whether the inventor that appears in the British patent is effectively the same person that appears in the UK patent, so, in practice, this count excludes all patents produced by at least one inventor whose *name* also appears in US patents.

Table C.7 reports the results. We find that all the baseline results are confirmed using this approach, suggesting that it is unlikely that our results are driven by inventors "double filing" patents in the US and the UK.

C.II Double and Triple Differences Analysis

C.II.1 Alternative Dependent Variables

As in the previous analysis, the baseline results use the (log 1+) as the main dependent variable. In columns (1–4) of Table C.9, we adopt different transformations of the dependent variables and confirm that our results hold throughout. In the baseline analysis, we also adopt a five-year window to compute the text-based measures (originality and similarity). In columns (5–7) and (8–10), we show that the qualitative nature of the results remains unchanged when using different thresholds.

C.II.2 Alternative Definitions of the Shocks

In the baseline analysis, we consider a district in the UK to be exposed to an innovation shock in the United States if the number of emigrants from that district that are exposed to an innovation shock is in the top 5% of the overall distribution of exposed emigrants, net of district, year, and, when applicable, technology class fixed effects. In Table C.10, we show that results remain qualitatively unchanged when using the top 1% and 10% as alternative thresholds. Unsurprisingly, the magnitude of the estimates of the average treatment effect increases in the restrictiveness of the threshold. As we move from the top 10% to the top 1% we require that increasingly more emigrants are exposed to a shock in the United States. According to the logic explained in the main text, this shift would translate into a relatively more intense exposure to the shock from the perspective of the district.

C.II.3 Alternative Standard Errors

In Figure C.6, we show that the effect of US innovation shocks on the volume of patents, either pooled (Panel C.6a) or broken down by technology (Panel C.6b), remains statistically significant when using alternativeC.6a estimators for the standard errors. In Panel C.6a, we find that the statistical significance of the effect of the pooled shocks is diminished when adjusting for spatial autocorrelation. The

estimates remain, however, significant at the 10% level and, importantly, adjusting for spatial autocorrelation does not alter the statistical significance of the triple differences estimates shown in Panel C.6b.

C.II.4 Alternative Estimator

The roll-out of US innovation shocks across districts is staggered, because different districts—or districttechnology class pairs—can be exposed to a shock to US innovation at different times. Goodman-Bacon (2021) shows that, in this case, the standard two-way fixed-effects (TWFE) may fail to estimate the average treatment effect if the effect is heterogeneous across units and/or over time. In Figure C.4, we show that all the results shown in the main text remain unchanged when employing the estimator developed by Sun and Abraham (2021). In particular, we estimate similar pre- and post-treatment coefficients for the volume of patents (Panels C.4a–C.4b) and the similarity between US and UK patents (Panels C.4c–C.4d). We estimate standard TWFE models because they allow for simpler deviations from the baseline estimation strategy for the standard errors and the heterogeneity analyses, as previously discussed.

C.II.5 Innovation Shocks in the United States

The double and triple differences analyses implicitly rely on the fact that our methodology to flag innovation shocks in US counties can reliably isolate periods of intense patenting activity. We test this assumption in Table C.14. We consider the universe of US counties (in columns 1–2) and county-technology pairs (columns 3–4). The dependent variable is the (log) number of patents (in columns 1 and 3) and the (log) number of original patents (in columns 2 and 4), and the treatment returns a value of one for treated units, i.e. counties or county-technology pairs after the innovation shocks occur, and zero otherwise. The regressions are saturated with fixed effects. These double and triple differences regressions do not capture a causal effect but rather indicate the actual intensity of the shock to US innovation activity.

An innovation shock is associated with a 50% increase in the number of patents (column 1). In the triple-differences setting, patenting increases by 10% on average after the shock. In Figure C.7, we repeat the estimation in a flexible setting, which uncovers substantial heterogeneity over time. When counties undergo an innovation shock, they produce 125% more patents. The effect persists over time, but the spike is short-lived. Within technologies, a US innovation shock is associated with an 80% shock to innovation activity, which reverts to the pre-shock mean very rapidly. We thus conclude that our methodology successfully isolates sharp and large shocks to innovation activity in the United States. The salience of the shock is, in both cases, relatively short-lived, albeit substantial.

C.II.6 Alternative Cut to Linked Sample

In Table C.13, we show that the baseline double- and triple-difference results are largely confirmed when excluding all migrants with more than one match in the linked sample. The results are more robust for the triple differences model (columns 4–6) than for the double differences estimates (columns 1–3). However, it is not surprising that the precision of the estimates decreases since this cut excludes more than 60% migrants from the linked sample. Importantly, the sign and magnitude of each estimate remain largely unchanged.

C.II.7 Alternative Sample of US Counties

In Table C.11, we show that the baseline double- and triple-difference results remain quantitatively stable when we exclude all counties in the top 10% and 20% of the share of British-born inventors. This exercise addresses the concern that the shocks reflect the innovation activity of British migrants in the United States.

C.II.8 Exclusion of Inventors Possibly Appearing on UK and US Patent Records

The baseline results remain similarly unchanged when we look at British patents issued to inventors who do not also appear on American patent records, as shown in Table C.12. In this context, our results could have been driven by inventors in the United States applying for patent protection for the same invention in the US and the UK if the invention contributed to generating a US innovation shock. Since the estimated treatment effect remains stable when we exclude British patents obtained by inventors whose names also appear on US patent records, we conclude that this practice is unlikely to have a quantitatively sizable influence on our results.

C.III Neighborhood Analysis

We consider two very simple departures from the baseline scenario described in the paper. First, we exclude Wales from the estimation sample in Table C.15. We apply this sample cut because, as shown in Table A.2, the number of matches in the sample of inventors linked to the census is larger for inventors residing in Wales. By excluding them from the estimation sample, we thus ensure that this imbalance does not drive the results. The estimates presented in the Table suggest that this does not appear to be the case.

Second, in the paper, we consider an emigrant to be in the neighborhood of an inventor if, before migrating, he lived within five kilometers of the inventor. In Figure C.8, we consider five alternative threshold distances: one, two, three, ten, and twenty kilometers. We then estimate the baseline regression for these various thresholds and separately report the estimated average treatment effects.

The coefficients remain positive for all thresholds, but they are largest in magnitude—and statistically significant at the 1% level—for three- and five-kilometer neighborhoods. The estimate for the ten-kilometer definition is similar, albeit noisier. From an economic perspective, it is plausible that as the intensity of the social connections vanishes with distance, including larger neighborhoods in the treatment introduces noise which reduces the precision of the estimates. In turn, a very low threshold discards a large number of emigrants, thus artificially dampening the treatment effect. This notwithstanding, it is reassuring that the coefficients remain positive throughout.

		Nun	nber of Pate	ents		k	c-Originalit	y	k-Simila	rity with U	S Patents
	(1) Number	(2) $\ln(1+\cdot)$	(3) $\ln(\varepsilon + \cdot)$	(4) $\ln(\cdot)$	(5) Arcsinh	(6) 1 yr.	(7) 5 yrs.	(8) 10 yrs.	(9) 1 yr.	(10) 5 yrs.	(11) 10 yrs.
Panel A. Patents Pooled	Across Tech	nologies									
Dependent Variable Mean	161.499	3.512	3.254	3.568	4.105	2.083	2.037	2.021	33.833	162.046	305.021
US Emigrants (1,000s)	403.841** (166.325)	0.686*** (0.226)	0.572* (0.310)	0.685*** (0.236)	0.657*** (0.238)	0.712*** (0.252)	0.987*** (0.262)	0.915*** (0.255)	3.783** (1.807)	15.014* (8.577)	23.412 (15.728)
District FE Decade FE Observations	Yes Yes 3,720	Yes Yes 3,720	Yes Yes 3,720	Yes Yes 3,575	Yes Yes 3,720						
Panel B. Patents by Dist	rict-Technol	ogy									
Dependent Variable Mean	0.079	1.000	-1.389	1.678	1.236	0.403	0.382	0.376	15.432	73.914	139.440
Knowledge Exposure	0.014*** (0.001)	0.040*** (0.001)	0.027*** (0.003)	0.034*** (0.001)	0.042*** (0.002)	0.035*** (0.001)	0.034*** (0.001)	0.034*** (0.001)	0.137*** (0.017)	0.640*** (0.083)	1.201*** (0.156)
District-Year FE Technology-Year FE Observations	Yes Yes 70,433	Yes Yes 70,433	Yes Yes 70,433	Yes Yes 35,924	Yes Yes 70,433						
Number of Districts	620	620	620	619	620	620	620	620	620	620	620

Table C.1. Effect of US Emigration and Knowledge Exposure on Innovation: Alternative Dependent Variables

Notes. This Table displays the association between UK innovation and US emigration (Panel A) and knowledge exposure to US technology (Panel B). The unit of observation is a district (Panel A) and a district-technology class pair (Panel B) observed at a decade frequency between 1870 and 1930. The dependent variables are: in column (1), the number of patents; in column (2), the log(1+) number of patents; in column (3), the log(0.1+) number of patents; in column (4), the log number of patents, which excludes zeros; in column (5), the inverse hyperbolic sine of the number of patents; in columns (6–8), patents in the top 20% of the novelty distribution in the previous 1, 5, and 10 years; in columns (9–11), the dependent variable is the similarity of British patents with American patents issued in the previous 1, 5, and 10 years. In Panel A, the independent variable is the number of US emigrants; in Panel B, the independent variable is the baseline measure of exposure to US technology through migration ties. All regressions in Panel A include district and year-fixed effects; all regressions in Panel B include district-by-year and technology-by-year fixed effects. Standard errors in parentheses are clustered at the district level. Referenced on page(s) 24, 24, A8, C42.

	Pa	tents by Distr	ict	Patents b	y District-Tec	hnology
	(1) Number	(2) US Patents Similarity	(3) High Impact	(4) Number	(5) US Patents Similarity	(6) High Impact
Dependent Variable Mean	2.686	107.436	1.426	0.619	36.724	0.216
US Emigrants (1,000s)	0.728*** (0.189)	41.363*** (10.374)	0.765*** (0.210)			
Knowledge Exposure				0.038*** (0.001)	1.225*** (0.061)	0.024*** (0.001)
District FE	Yes	Yes	Yes	_	-	-
Decade FE	Yes	Yes	Yes	_	_	_
District-Decade FE	_	_	_	Yes	Yes	Yes
Technology FE	_	_	_	Yes	Yes	Yes
Number of Districts Observations	621 3,726	621 3,726	621 3,726	621 70,433	621 70,433	621 70,433

Table C.2. Emigration to the United States and Innovation: Patents with Firm Assignee

Notes. This Table displays the association between emigration and innovation in the United Kingdom. In columns (1–3) (resp. 4–6), the unit of observation is a district (resp. district-technology pair) observed at a decade frequency between 1870 and 1930. In columns (1) and (4), the dependent variable is the (log) number of patents; in columns (2) and (5), the dependent variable is the average similarity between UK and US patents; in columns (3) and (6), we include patents in the top 20% of the novelty distribution. We only include patents with a firm assignee. In columns (1–3) (resp. 4–6), the independent variable is the number of migrants (resp. exposure to US technology). In columns (1–3), regressions include district and decade-fixed effects; in columns (4–6), regressions include district-by-time and technology-fixed effects. Standard errors in parentheses are clustered at the district level. Referenced on page(s) 24, C42.

		log	(1 + Numb)	er of Paten	ts)	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable Mean	0.999	0.999	0.999	0.999	0.999	0.999
Knowledge Exposure	0.373*** (0.012)					
Log(1+Knowledge Exposure)		6.944*** (0.275)				
Fixed-Emigration Knowledge Exposure			0.655*** (0.034)			
Fixed-Innovation Knowledge Exposure				0.297*** (0.012)		
Cumulative Knowledge Exposure					0.007 (0.007)	
Level of Patents Knowledge Exposure						0.165*** (0.004)
District-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Technology FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts Observations	621 70,794	621 70,794	621 70,794	621 70,794	621 70,794	621 70,794

Table C.3. Correlation between Exposure to US Technology and Innovation: Alternative Measures of Knowledge Exposure

Notes. This Table displays the association between exposure to US technology through migration ties and innovation in the United Kingdom. The unit of observation is a district-technology pair observed at a decade frequency between 1870 and 1930. The dependent variable is the (log) number of patents. The independent variable is: in column (1), the baseline metric of knowledge exposure (divided by 10 for comparability); in column (2), the log of the baseline metric; in column (3), the baseline measure but keeping migration ties between districts and counties fixed in 1880; in column (4), the baseline metric but keeping the share of patents across classes fixed in 1880; in column (5), the baseline metric but considering patents as a stock instead of a flow and taking the cumulative number of patents over time; in column (6), the exposure metric interacts migration ties with the number of patents produced in each technology by the county, instead of the share. All regressions include district-by-year and technology-fixed effects. Standard errors in parentheses are clustered at the district level. Referenced on page(s) 25, C42, C43.

	Pate	nts by Dist	rict-Techno	ology
	(1)	(2)	(3)	(4)
Dependent Variable Mean	1.006	1.006	1.006	1.006
Knowledge Exposure	0.023*** (0.002)			
Knowledge Exposure (Top 20% Impact)		0.074*** (0.005)		
Knowledge Exposure (Top 10% Impact)			0.112*** (0.007)	
Knowledge Exposure (Top 5% Impact)				0.156*** (0.010)
District-Year FE	Yes	Yes	Yes	Yes
Technology FE	Yes	Yes	Yes	Yes
Number of Districts Observations	621 70,433	621 70,433	621 70,433	621 70,433

Table C.4. Correlation between Exposure to US Technology and Innovation: Exposure to High-Impact Patents

Notes. This Table displays the association between exposure to US technology through migration ties and innovation in the United Kingdom. The unit of observation is a district-technology pair observed at a decade frequency between 1870 and 1930. The dependent variable is the (log) number of patents. The independent variable in column (1) is the baseline metric of knowledge exposure. In columns (2) to (4), to construct the knowledge exposure metric, we only include high-impact patents to measure county specialization. In particular, in column (2), we only include patents in the top 20% of the impact distribution; in column (3), we only include patents in the top 10%; and in column (4), we only include patents in the top 5% of the distribution. All regressions include district-by-year and technology-fixed effects. Standard errors in parentheses are clustered at the district level. Referenced on page(s) 25, C43.

		Patents	by District		Patents by District-Technology			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number	Number	US Patents Similarity	High Impact	Number	Number	US Patents Similarity	High Impact
Dependent Variable Mean	182.270	182.484	41.365	38.521	10.006	11.816	2.275	2.506
US Emigrants (1,000s)	0.867***	0.385*	0.846***	0.917***				
	(0.212)	(0.230)	(0.214)	(0.282)				
Knowledge Exposure					0.437***	0.033	0.423***	0.754***
					(0.091)	(0.097)	(0.094)	(0.132)
District FE	Yes	Yes	Yes	Yes	_	_	-	-
Decade FE	Yes	-	Yes	Yes	-	-	-	-
Controls \times Time	No	Yes	No	No	-	-	-	-
County-Year FE	No	Yes	No	No	-	-	-	-
District-Year FE	-	-	_	-	Yes	Yes	Yes	Yes
Technology FE	-	-	_	-	Yes	-	Yes	Yes
District-Technology FE	-	-	_	-	No	Yes	No	No
Technology-Year FE	-	-	-	-	No	Yes	No	No
Number of Districts	621	620	621	610	621	620	621	610
Observations	3,726	3,720	3,726	3,660	67,754	57,371	67,697	56,164

Table C.5. Emigration, Exposure to US Knowledge, and Innovation in the United Kingdom: Poisson Quasi-Maximum Likelihood Estimates

Notes. This Table displays the association between emigration and innovation in the United Kingdom. In columns (1–4) (resp. 5–8), the unit of observation is a district (resp. district-technology pair) observed at a decade frequency between 1870 and 1930. Each regression is estimated using the Poisson quasi-maximum likelihood estimator described in Correia *et al.* (2020). The dependent variable is: in columns (1–2) and (5–6), the number of patents; in (3) and (7), the text-based similarity between British and American patents issued in the previous five years; in (4) and (8), the number of patents in the top 20% of the novelty distribution. In columns (1–4) (resp. 5–8), the independent variable is the number of migrants (resp. exposure to US technology). In columns (1) and (3–4), the model includes district and decade-fixed effects; column (2) includes district-level controls measured in 1880 and interacted with decade indicators and county-by-decade fixed effects. In columns (1) and (7–8), regressions include district-by-year and technology-fixed effects; column (6) also includes district-by-technology and technology-by-year fixed effects. Standard errors in parentheses are clustered at the district level. Referenced on page(s) 24. *: p < 0.10, **: p < 0.05, ***: p < 0.01

		Patents	by District		Pa	tents by Dis	strict-Technolo	ogy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number	Number	US Patents Similarity	High Impact	Number	Number	US Patents Similarity	High Impact
Panel A. Excluding US C	Counties in	top 10% of	Share of Brit	ish-Born I	nventors			
Dependent Variable Mean	3.520	3.521	163.036	2.043	1.006	1.006	74.488	0.388
US Emigrants (1,000s)	0.743*** (0.137)	0.499*** (0.184)	19.007*** (5.517)	0.882*** (0.168)				
Knowledge Exposure					0.023*** (0.002)	0.006*** (0.001)	0.231*** (0.058)	0.022*** (0.001)
Panel B. Excluding US C	Counties in	top 20% of	Share of Brit	ish-Born I	nventors			
Dependent Variable Mean	3.520	3.521	163.036	2.043	1.006	1.006	74.488	0.388
US Emigrants (1,000s)	0.766*** (0.143)	0.508*** (0.191)	19.562*** (5.641)	0.928*** (0.175)				
Knowledge Exposure					0.024*** (0.002)	0.007*** (0.001)	0.256*** (0.061)	0.023*** (0.001)
District FE	Yes	Yes	Yes	Yes	_	_	_	_
Decade FE	Yes	-	Yes	Yes	-	-	-	_
Controls \times Time	No	Yes	No	No	-	-	-	-
County-Year FE	No	Yes	No	No	-	_	-	_
District-Year FE	-	-	_	-	Yes	Yes	Yes	Yes
Technology FE	-	-	-	-	Yes	-	Yes	Yes
District-Technology FE	-	-	-	-	No	Yes	No	No
Technology-Year FE	-	-	-	-	No	Yes	No	No
Number of Districts	620	619	620	620	621	621	621	621
Observations	3,699	3.691	3,699	3.699	70.433	70.433	70.433	70.43

Table C.6. Emigration, Exposure to US Knowledge, and Innovation in the United Kingdom: Excluding US Counties with a High Share of British-born Inventors

Notes. This Table displays the association between emigration and innovation in the United Kingdom. To construct emigration and knowledge exposure, we exclude migrants that settle in US counties in the top 10% (Panel A) and 20% (Panel B) of the distribution of the share of British-born inventors. In columns (1–4) (resp. 5–8), the unit of observation is a district (resp. district-technology pair) observed at a decade frequency between 1870 and 1930. The dependent variable is: in columns (1–2) and (5–6), the number of patents; in (3) and (7), the text-based similarity between British and American patents issued in the previous five years; in (4) and (8), the number of patents in the top 20% of the novelty distribution. In columns (1–4) (resp. 5–8), the independent variable is the number of migrants (resp. exposure to US technology). In columns (1) and (3–4), the model includes district and decade-fixed effects; column (2) includes district-level controls measured in 1880 and interacted with decade indicators and county-by-decade fixed effects. In columns (1) and (7–8), regressions include district-by-year and technology-fixed effects; column (6) also includes district-by-technology and technology-by-year fixed effects. Standard errors in parentheses are clustered at the district level. Referenced on page(s) 25, C44.
Table C.7. Emigration,	Exposure to US k	Knowledge, an	nd Innovation	in the United I	Kingdom:
Excluding UK Patents	with Inventors W	hose Name A	Appears in US I	Patents	

		Patents	by District		Pa	Patents by District-Technology				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
	Number	Number	US Patents Similarity	High Impact	Number	Number	US Patents Similarity	High Impact		
Dependent Variable Mean	3.072	3.072	154.386	1.699	1.006	1.006	74.488	0.388		
US Emigrants (1,000s)	0.731***	0.345**	25.852***	0.956***						
	(0.132)	(0.174)	(6.106)	(0.170)						
Knowledge Exposure					0.023***	0.006***	0.231***	0.022***		
					(0.002)	(0.001)	(0.058)	(0.001)		
District FE	Yes	Yes	Yes	Yes	_	_	-	_		
Decade FE	Yes	-	Yes	Yes	-	-	-	-		
Controls \times Time	No	Yes	No	No	-	-	-	-		
County-Year FE	No	Yes	No	No	-	-	-	-		
District-Year FE	-	-	-	-	Yes	Yes	Yes	Yes		
Technology FE	-	-	-	-	Yes	-	Yes	Yes		
District-Technology FE	-	-	-	-	No	Yes	No	No		
Technology-Year FE	-	-	-	-	No	Yes	No	No		
Number of Districts	621	620	621	621	621	621	621	621		
Observations	3,726	3,720	3,726	3,726	70,433	70,433	70,433	70,433		

Notes. This Table displays the association between emigration and innovation in the United Kingdom. To construct the patent count, we exclude all UK patents with at least one inventor whose name appears in at least one US patent over a ten-year window around the patent filing date. In columns (1–4) (resp. 5–8), the unit of observation is a district (resp. district-technology pair) observed at a decade frequency between 1870 and 1930. The dependent variable is: in columns (1–2) and (5–6), the number of patents; in (3) and (7), the text-based similarity between British and American patents issued in the previous five years; in (4) and (8), the number of patents in the top 20% of the novelty distribution. In columns (1–4) (resp. 5–8), the independent variable is the number of migrants (resp. exposure to US technology). In columns (1) and (3–4), the model includes district and decade-fixed effects; column (2) includes district-level controls measured in 1880 and interacted with decade indicators and county-by-decade fixed effects. In columns (1) and (7–8), regressions include district-by-year and technology-fixed effects; column (6) also includes district-by-technology and technology-by-year fixed effects. Standard errors in parentheses are clustered at the district level. Referenced on page(s) 24, C45. *: p < 0.05, ***: p < 0.01

		Patents	by District		Pat	Patents by District-Technology				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
	Number	Number	US Patents Similarity	High Impact	Number	Number	US Patents Similarity	High Impact		
Dependent Variable Mean	3.518	3.518	162.555	2.042	1.006	1.006	74.527	0.388		
US Emigrants (1,000s)	1.485***	0.880**	46.065***	1.585***						
	(0.274)	(0.371)	(11.711)	(0.348)						
Knowledge Exposure					0.052***	0.011***	0.461***	0.052***		
					(0.004)	(0.002)	(0.134)	(0.003)		
District FE	Yes	Yes	Yes	Yes	_	_	-	-		
Decade FE	Yes	-	Yes	Yes	-	-	-	-		
Controls \times Time	No	Yes	No	No	-	-	-	-		
County-Year FE	No	Yes	No	No	-	-	-	-		
District-Year FE	-	-	-	-	Yes	Yes	Yes	Yes		
Technology FE	-	-	-	-	Yes	-	Yes	Yes		
District-Technology FE	-	-	-	-	No	Yes	No	No		
Technology-Year FE	_	_	-	-	No	Yes	No	No		
Number of Districts	621	620	621	621	621	621	621	621		
Observations	3,726	3,720	3,726	3,726	70,395	70,395	70,395	70,395		

Table C.8. Emigration, Exposure to US Knowledge, and Innovation in the United Kingdom: Census-Linking Algorithm Excludes Multiple Matches

Notes. This Table displays the association between emigration and innovation in the United Kingdom. To construct the treatment variables, we exclude all migrants with more than one match in the linked sample. In columns (1-4) (resp. 5–8), the unit of observation is a district (resp. district-technology pair) observed at a decade frequency between 1870 and 1930. The dependent variable is: in columns (1-2) and (5-6), the number of patents; in (3) and (7), the text-based similarity between British and American patents issued in the previous five years; in (4) and (8), the number of patents in the top 20% of the novelty distribution. In columns (1-4) (resp. 5–8), the independent variable is the number of migrants (resp. exposure to US technology). In columns (1) and (3–4), the model includes district and decade-fixed effects; column (2) includes district-level controls measured in 1880 and interacted with decade indicators and county-by-decade fixed effects. In columns (1) and (7–8), regressions include district-by-year and technology-fixed effects; column (6) also includes district-by-year fixed effects. Standard errors in parentheses are clustered at the district level. Referenced on page(s) 25, C44.

		Number c	of Patents		k-	<i>k</i> -Originality <i>k</i> -Similarity			rity with US	S Patents
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number	$\ln(1+\cdot)$	$\ln(\epsilon + \cdot)$	Arcsinh	1 yr.	5 yrs.	10 yrs.	1 yr.	5 yrs.	10 yrs.
Panel A. Innovation by Distr	ict									
Dependent Variable Mean	1.687	21.355	0.206	2.055	0.759	0.736	0.728	2.549	12.156	23.010
Post \times US Innovation Shock	0.091**	20.777***	0.077	0.066	0.138***	0.082*	0.089*	0.364***	1.852***	3.569***
	(0.045)	(5.847)	(0.083)	(0.049)	(0.041)	(0.043)	(0.047)	(0.071)	(0.347)	(0.672)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,671	31,671	31,671	31,671	31,671	31,671	31,671	31,671	31,671	31,671
Panel B. Innovation by Distr	ict-Technol	ogy								
Dependent Variable Mean	0.292	1.124	-3.456	0.371	0.297	0.304	0.313	0.629	3.012	5.708
Post \times US Innovation Shock	0.064***	0.994**	0.093**	0.073***	0.057**	0.052**	0.047**	0.172***	0.876***	1.712***
	(0.019)	(0.361)	(0.042)	(0.023)	(0.020)	(0.020)	(0.021)	(0.047)	(0.234)	(0.465)
District-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Technology-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Technology FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	601,749	601,749	601,749	601,749	601,749	601,749	601,749	601,749	601,749	601,749

Table C.9. Double and Triple Differences: Alternative Dependent Variables

Notes. This Table displays the effect of shocks to US innovation activity on innovation produced in the United Kingdom. In Panel A, the unit of observation is a district observed at a yearly frequency between 1870 and 1930; in Panel B, the unit of observation is a district-technology pair observed over the same period. The dependent variables are: in column (1), the number of patents; in column (2), the log(1+) number of patents; in column (3), the log(0.1+) number of patents; in column (4), the inverse hyperbolic sine of the number of patents; in columns (5–7), the number of patents in the top 20% of distribution of originality in the previous 1, 5, and 10 years; in columns (8–10), the dependent variable is the similarity of British patents with American patents issued in the previous 1, 5, and 10 years. The independent variable is equal to one for all years after the observation unit is exposed to a shock to US innovation activity. The definition of exposure is provided in the main text. All regressions in Panel A include district and year fixed effects; all regressions in Panel B include district and are displayed in parentheses. Referenced on page(s) 24, 24, A8, C45. *: p < 0.10, **: p < 0.05, ***: p < 0.01

	Nur	nber of Pat	tents	High	High-Impact Patents			Similarity with US Patents		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Panel A. Innovation by District										
Dependent Variable Mean	1.687	1.687	1.687	1.738	1.738	1.738	12.156	12.156	12.156	
Post \times US Innovation Shock (Top 10%)	0.081* (0.043)			0.060 (0.043)			1.587*** (0.341)			
Post \times US Innovation Shock (Top 5%)		0.091** (0.045)			0.074* (0.045)			1.852*** (0.347)		
Post \times US Innovation Shock (Top 1%)			0.111* (0.063)			0.070 (0.063)			2.869*** (0.307)	
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	31,671	31,671	31,671	31,671	31,671	31,671	31,671	31,671	31,671	
Panel B. Innovation by District-Technol	logy									
Dependent Variable Mean	0.292	0.292	0.292	0.304	0.304	0.304	0.302	0.302	0.302	
Post \times US Innovation Shock (Top 10%)	0.035*** (0.011)			0.027** (0.011)			0.040** (0.014)			
Post \times US Innovation Shock (Top 5%)		0.062*** (0.019)			0.050** (0.020)			0.087*** (0.023)		
Post \times US Innovation Shock (Top 1%)			0.093*** (0.028)			0.078** (0.030)			0.144*** (0.037)	
District-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Technology-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
District-Technology FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	612,408	612,408	612,408	612,408	612,408	612,408	612,408	612,408	612,408	

Table C.10. Double and Triple Differences: Alternative Definitions of the Shocks

Notes. This Table displays the effect of shocks to US innovation activity on innovation produced in the United Kingdom. In Panel A, the unit of observation is a district observed at a yearly frequency between 1870 and 1930; in Panel B, the unit of observation is a district-technology pair observed over the same period. The dependent variable is in columns (1–3), the (log) number of patents; in columns (4–6), the number of patents in the top 20% of the originality distribution; in columns (7–9), the average text-based similarity between British patents and American patents produced in the previous five years. The independent variable is equal to one for all years after the observation unit is exposed to a shock to US innovation activity. A unit is exposed to a shock if the number of emigrants from that unit that are exposed to a US innovation shock is in the *k*-percentile of the distribution. We consider three alternative values for the *k* threshold, 10%, 5% (the baseline), and 1%. All regressions in Panel A include district and year fixed effects; all regressions in Panel B include district-by-technology fixed effects. Standard errors are clustered by district and are displayed in parentheses. Referenced on page(s) 25, C45.

Table C.11. Double and Triple Differences: Excluding US Counties with a High Share of British-born Inventors

	Pat	tents by Distri	ct	Patents b	Patents by District-Technology			
	(1) Number	(2) US Patents Similarity	(3) High Impact	(4) Number	(5) US Patents Similarity	(6) High Impact		
Panel A. Excluding US Coun	ties in top	10% of Share	of British	-Born Inve	ntors			
Dependent Variable Mean	1.687	121.561	0.736	0.292	30.124	0.082		
Post \times US Innovation Shock	0.091** (0.045)	18.523*** (3.473)	0.082* (0.043)	0.064*** (0.009)	8.760*** (1.145)	0.020*** (0.006)		
Panel B. Excluding US Counties in top 20% of Share of British-Born Inventors								
Dependent Variable Mean	1.687	121.561	0.736	0.292	30.124	0.082		
Post \times US Innovation Shock	0.070 (0.045)	16.619*** (3.578)	0.079* (0.044)	0.068*** (0.009)	9.457*** (1.139)	0.022*** (0.006)		
District FE	Yes	Yes	Yes	_	_	_		
Year FE	Yes	Yes	Yes	-	_	-		
District-Year FE	-	_	-	Yes	Yes	Yes		
Technology-Year FE	-	_	-	Yes	Yes	Yes		
District-Technology FE	-	-	-	Yes	Yes	Yes		
Number of Districts	621	621	621	621	621	621		
Observations	31,671	31,671	31,671	601,749	601,749	601,749		

Notes. This Table reports the effect of shocks to US innovation activity on innovation in the United Kingdom. To construct emigration and knowledge exposure, we exclude migrants that settle in US counties in the top 10% (Panel A) and 20% (Panel B) of the distribution of the share of British-born inventors. In columns (1–3), the unit of observation is a district observed at a yearly frequency between 1870 and 1930. In columns (4–6), the unit of observation is a district-technology pair observed over the same period. In columns (1) and (4), the dependent variable is the log(1+) number of patents. In columns (2) and (5), the dependent variable is the average text-based similarity between UK patents and US patents in the top 20% of the impact distribution ("Breakthrough" patents). "Post" is an indicator equal to one for all years after the observation unit is exposed to a shock to US innovation activity, equal to zero otherwise. "US Innovation Shock" is defined in Section IV.D. Columns (1–3) include district and year fixed effects. Columns (4–6) include district-by-year, technology-by-year, and district-by-technology fixed effects. Standard errors, reported in parentheses, are clustered by district. Referenced on page(s) 25, C47.

	Pa	tents by Distr	ict	Patents by District-Technology				
	(1) Number	(2) US Patents Similarity	(3) High Impact	(4) Number	(5) US Patents Similarity	(6) High Impact		
Dependent Variable Mean	1.380	105.767	0.569	0.292	30.124	0.082		
Post \times US Innovation Shock	0.135*** (0.047)	20.336*** (4.015)	0.120*** (0.043)	0.064*** (0.009)	8.760*** (1.145)	0.020*** (0.006)		
District FE	Yes	Yes	Yes	_	_	_		
Year FE	Yes	Yes	Yes	-	-	-		
District-Year FE	-	_	-	Yes	Yes	Yes		
Technology-Year FE	_	-	_	Yes	Yes	Yes		
District-Technology FE	-	-	_	Yes	Yes	Yes		
Number of Districts	621	621	621	621	621	621		
Observations	31,671	31,671	31,671	601,749	601,749	601,749		

Table C.12. Double and Triple Differences: Excluding UK Patents with Inventors Whose Name Appears in US Patents

Notes. This Table reports the effect of shocks to US innovation activity on innovation in the United Kingdom. To construct the patent count, we exclude all UK patents with at least one inventor whose name appears in at least one US patent over a ten-year window around the patent filing date. In columns (1–3), the unit of observation is a district observed at a yearly frequency between 1870 and 1930. In columns (4–6), the unit of observation is a district-technology pair observed over the same period. In columns (1) and (4), the dependent variable is the log(1+) number of patents. In columns (2) and (5), the dependent variable is the average text-based similarity between UK patents and US patents issued in the previous five years. In columns (3) and (6), the dependent variable is the log(1+) number of patents in the top 20% of the impact distribution ("Breakthrough" patents). "Post" is an indicator equal to one for all years after the observation unit is exposed to a shock to US innovation activity, equal to zero otherwise. "US Innovation Shock" is defined in Section IV.D. Columns (1–3) include district and year fixed effects. Columns (4–6) include district-by-year, technology-by-year, and district-by-technology fixed effects. Standard errors, reported in parentheses, are clustered by district. Referenced on page(s) 24, C47.

	Pat	ents by Distri	ct	Patents by District-Technology				
	(1)	(2)	(3) High	(4)	(5) LIS Patonto	(6) High		
	Number	Similarity	Impact	Number	Similarity	Impact		
Dependent Variable Mean	1.687	121.561	0.736	0.292	30.124	0.082		
Post \times US Innovation Shock	0.037	7.301**	0.068*	0.051***	6.691***	0.021***		
	(0.041)	(3.183)	(0.038)	(0.007)	(0.883)	(0.005)		
District FE	Yes	Yes	Yes	_	_	_		
Year FE	Yes	Yes	Yes	-	_	-		
District-Year FE	_	-	_	Yes	Yes	Yes		
Technology-Year FE	_	-	_	Yes	Yes	Yes		
District-Technology FE	_	_	_	Yes	Yes	Yes		
Number of Districts	621	621	621	621	621	621		
Observations	31,671	31,671	31,671	601,749	601,749	601,749		

Table C.13. Double and Triple Differences: Census-Linking Algorithm Excludes Multiple Matches

Notes. This Table reports the effect of shocks to US innovation activity on innovation in the United Kingdom. To construct the treatment variables, we exclude all migrants with more than one match in the linked sample. In columns (1–3), the unit of observation is a district observed at a yearly frequency between 1870 and 1930. In columns (4–6), the unit of observation is a district-technology pair observed over the same period. In columns (1) and (4), the dependent variable is the log(1+) number of patents. In columns (2) and (5), the dependent variable is the average text-based similarity between UK patents and US patents issued in the previous five years. In columns (3) and (6), the dependent variable is the log(1+) number of patents in the top 20% of the impact distribution ("Breakthrough" patents). "Post" is an indicator equal to one for all years after the observation unit is exposed to a shock to US innovation activity, equal to zero otherwise. "US Innovation Shock" is defined in Section IV.D. Columns (1–3) include district and year fixed effects. Columns (4–6) include district-by-year, technology-by-year, and district-by-technology fixed effects. Standard errors, reported in parentheses, are clustered by district. Referenced on page(s) 25, C46.

	Patents by County		Pater County-Te	nts by echnology
	(1)	(2)	(3)	(4)
	Number	Number High Impact		High Impact
Dependent Variable Mean	0.954	0.165	0.135	0.016
Post \times Innovation Shock	0.585***	0.057***	0.129***	0.030***
	(0.018)	(0.009)	(0.010)	(0.003)
County FE	Yes	Yes	_	_
Year FE	Yes	Yes	_	_
County-Year FE	_	_	Yes	Yes
Technology-Year FE	_	_	Yes	Yes
County-Technology FE	_	_	Yes	Yes
Number of Counties	2,848	2,848	2,848	2,848
Observations	202,208	196,953	3,841,952	3,832,217

Table C.14. Shocks to United States Innovation Activity

Notes. This Table reports the "first stage" of the shocks to US innovation, i.e., how much innovation in the US increases in the period following what we define as a shock. In columns (1–2) (resp. 3–4), the unit of observation is a county (resp. county-technology pair) observed at a yearly frequency between 1870 and 1930. In columns (1) and (3), the dependent variable is the (log) number of patents; in columns (2) and (4), it is the (log) number of patents in the top 20% of the novelty distribution. The independent variable is an indicator equal to one in the years after the observation units undergo an innovation shock. Regressions in columns (1–2) include county and year-fixed effects; regressions in columns (3–4) include county-by-year, technology-by-year, and county-by-technology fixed effects. Standard errors are shown in parentheses and are clustered at the county level. Referenced on page(s) 17, 17, 21, C46.

Table C.15. Within Neighborhood US Emigration and Innovation in the United Kingdom: Excluding Wales

	Number of Patents					Text-Based Measures	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
					I(Patents > 0)	US Patents Similarity	High Impact
Panel A. Neighborhood Emigration							
Dependent Variable Mean	6.818	6.818	6.839	6.855	8.579	6.891	1.585
Post \times Emigrant in Neighborhood	0.428***	0.220**	0.238**	0.176	0.540***	0.473***	0.111**
	(0.133)	(0.112)	(0.118)	(0.157)	(0.164)	(0.135)	(0.054)
Post \times N. Emigrants in Neighborhood				0.261			
				(0.214)			
Panel B. Neighborhood Non-Return Emigration							
Dependent Variable Mean	6.818	6.818	6.839	6.855	8.579	6.891	1.585
Post $ imes$ Non-Return Emigrant in Neighborhood	0.394***	0.236**	0.248**	0.409**	0.502***	0.448***	0.101*
	(0.133)	(0.116)	(0.123)	(0.166)	(0.163)	(0.136)	(0.055)
Post \times N. Non-Return Emigrants in Neighborhood				0.108			
				(0.222)			
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parish-Year FE	No	No	Yes	No	No	No	No
District-Year FE	No	Yes	-	No	No	No	No
Year FE	Yes	-	-	Yes	Yes	Yes	Yes
Number of Individuals	124,329	124,321	122,528	108,319	124,329	124,329	124,329
Observations	2,486,580	2,486,420	2,450,560	2,166,380	2,486,580	2,486,580	2,486,580

Notes. This Table reports how emigrants to the United States impact the innovation activity fulfilled by their neighbors who remain in the UK. The unit of observation is an individual inventor observed at a yearly frequency between 1880 and 1900. The analysis sample is the universe of inventors linked to the 1891 population census, as detailed in the main text, excluding inventors residing in Wales. The dependent variable is in columns (1–4), the (log) number of patents produced by the members of the family; in column (5), an indicator equal to one for the same variable; in column (6), the average text-based similarity between British patents and American patents produced in the previous five years; in column (7), the number of patents in the top 20% of the novelty distribution. In Panel A, the baseline treatment is an indicator equal to one after the first neighbor of the inventor moves to the US, and zero otherwise; in Panel B, we restrict the treatment to non-return emigrants. In column (4), we further include an interaction with the number of emigrants. All regressions include inventor and year fixed effects; in columns (2) and (3) we include, respectively, district-by-year and parish-by-year fixed effects. Standard errors are clustered at the county level and are displayed in parentheses. Referenced on page(s) 29, C47.

Figures

Figure C.1. Emigration, Exposure to US Knowledge, and Innovation: Alternative Standard Errors



(a) Patents Pooled Across Technologies (b) Patents by Technology

Notes. This Figure displays the association between emigration and innovation in the United Kingdom using various estimators for the standard errors. In Panel C.1a(resp. C.1b), the unit of observation is a district (resp. district-technology pair) observed at a decade frequency between 1870 and 1930. Each dot reports the coefficient of a regression between the (log) number of patents and an independent variable which, in Panel C.1a (resp. C.1b) is the number of migrants (resp. exposure to US technology). In Panel C.1a (resp. C.1b), regressions include district and decade-fixed effects (resp. district-by-technology and year-fixed effects). We consider various estimators for the standard errors: robust to heteroskedasticity; clustered by district, technology, and two-way by district and technology; robust to heteroskedasticity and autocorrelation at various bandwidths; and robust to spatial autocorrelation following (Conley, 1999) at various bandwidths. All confidence bands report 95% confidence intervals. Referenced on page(s) 26, C43.



Figure C.2. Distribution of Exposure to US Technology Shocks

Notes. This Figure reports the distribution of exposure to US technology shocks. The formal definition of exposure to a US technology shock—the number of migrants from each district between 1870 and 1880 that settle in counties with at least one innovation shock—is provided in Section IV.D. The gray bars report the distribution of the exposure metric. The blue line reports the cumulative distribution (shown on the right *y*-axis). The dashed black line marks the 95th percentile of the distribution, which is the threshold used in the main analysis to compile the double- and triple-differences shocks. Referenced on page(s) 17.



Figure C.3. Raw Mean Patenting in Districts Exposed to US Innovation Shocks

Notes. This Figure reports the average number of patents in UK districts before and after they are exposed to a US innovation shock. In Panel IIIa, the unit of observation is a district observed at yearly frequency between 1870 and 1930; in Panel IIIb, the unit of observation is a district-technology pair observed over the same period. The dependent variable is the log(1+) number of patents. Each dot reports the average number of patents produced in districts that, at some point, are exposed to an innovation shock, by relative time since the treatment period. The black dashed line indicates the treatment period. The last period before the shock serves as the baseline category. The bands report 95% confidence intervals. The black solid line reports, in each figure, the average number of patents computed through linear interpolation on the pre-treatment periods. Referenced on page(s) 17.

Figure C.4. The Dynamic Effect of Shocks to US Innovation on Innovation in the UK: Alternative Estimator



(c) Similarity to US Patents: Pooled





(b) Patents by Technology





Notes. This Figure displays how shocks to US innovation activity impact innovation in the UK. In Panels C.4a and C.4c, the unit of observation is a district observed at a yearly frequency between 1870 and 1930; in Panels C.4b and C.4d, the unit of observation is a district-technology pair, observed over the same time. The dependent variable is, in Panels C.4a–C.4b, the (log) number of patents, and in Panels C.4c–C.4d, the average text-based similarity of the UK patents issued in a given period to US patents issued in the previous five years. Each dot reports the coefficient of an indicator variable which codes the time since the observation unit is exposed to a shock to US innovation activity through emigration ties. We employ the estimator developed by Sun and Abraham (2021) to account for the staggered roll-out of the shocks across observation units. The black dashed line indicates the treatment period. The last period before the shock serves as the baseline category. In Panel C.4a, the regression includes district and year-fixed effects, and standard errors are clustered at the district level; in Panel C.4b, the regression includes district-by-year, technology-by-year, and district-by-technology fixed effects, and standard errors are clustered at the district level. The bands report 95% confidence intervals. Referenced on page(s) 26, C46.





Notes. This Figure displays how shocks to US innovation activity impact innovation in the UK. In Panel C.5a, the unit of observation is a district observed at a yearly frequency between 1870 and 1930; in Panel C.5b, the unit of observation is a district-technology pair observed over the same time. The dependent variable is the number of patents winsorized at the 1% level. Each regression is estimated using the Poisson quasi-maximum likelihood estimator described in Correia *et al.* (2020). Each dot reports the coefficient of an indicator variable, which codes the time since the observation unit is exposed to a shock in US innovation activity through emigration ties. The black dashed line indicates the treatment period. The last period before the shock serves as the baseline category. In Panel C.5a, the regression includes district-by-technology fixed effects; in Panel C.5a, the regressions control for the total number of patents in the pre-treatment period interacted with a post-treatment indicator variable. Standard errors are clustered at the district level. The bands report 95% confidence intervals. Referenced on page(s) 24.





(a) Patents Pooled Across Technologies

(b) Patents by Technology

Notes. This Figure displays the effect of shocks to US innovation activity on innovation produced in the United Kingdom using various estimators for the standard errors. In Panel C.6a, the unit of observation is a district observed at a yearly frequency between 1870 and 1930; in Panel C.6b, the unit of observation is a district-technology pair observed over the same period. The dependent variable is the (log) number of patents. Each dot reports the coefficient of the treatment variable, which is equal to one for all years after the observation unit is exposed to a shock to US innovation activity. The definition of exposure is provided in the main text. All regressions in Panel C.6a include district-by-technology fixed effects; all regressions in Panel C.6b include district-by-technology fixed effects. We consider various estimators for the standard errors: robust to heteroskedasticity; clustered by district, technology, and two-way by district and technology; robust to heteroskedasticity and autocorrelation at various bandwidths; and robust to spatial autocorrelation following (Conley, 1999) at various bandwidths. All confidence bands report 95% confidence intervals. Referenced on page(s) 26, C45.



Figure C.7. Dynamic Shocks to US Innovation Activity

Notes. This Figure reports the "first stage" of the shocks to US innovation, i.e., how much innovation in the US increases in the period following what we define as a shock. In Panel C.7a (resp. C.7b), the unit of observation is a county (resp. county-technology pair) observed at a yearly frequency between 1870 and 1930. The dependent variable is the (log) number of patents. Each dot reports the coefficient of an indicator variable that codes the number of periods since the observation units undergo an innovation shock. Regressions in Panel C.7a (resp. C.7b) include county and year-fixed effects (resp. county-by-year, technology-by-year, and county-by-technology fixed effects). Standard errors are clustered at the county level. Bands report 95% confidence intervals. Each graph reports separately the *F*-statistics of joint significance of the pre-and post-treatment coefficients, along with their *p*-values. Referenced on page(s) 21, C46.

Figure C.8. Within-Neighborhood US Emigration: Alternative Neighborhood Threshold Values



Notes. This Figure reports how emigrants to the United States impact the innovation activity fulfilled by their neighbors who remain in the UK. The unit of observation is an individual inventor observed at a yearly frequency between 1880 and 1900. The analysis sample is the universe of inventors linked to the 1891 population census, as detailed in the main text. The dependent variable is the (log) number of patents produced by the members of the family. Each dot reports the coefficient of a treatment variable equal to one after the first neighbor of the inventor moves to the US, and zero otherwise. A neighborhood is defined as a *k*-kilometer radius area around each inventor. Each dot refers to a different definition of neighborhood with *k* equal to 1, 2, 3, 5, 10, and 20 kilometers. The baseline case reported in the main text assumes k = 5. All regressions include inventor and year-fixed effects. Standard errors are clustered at the county level. Bands report 95% confidence intervals. Referenced on page(s) 29, C47.

APPENDIX REFERENCES

- ABRAMITZKY, R., L. BOUSTAN, K. ERIKSSON, J. FEIGENBAUM and S. PÉREZ (2021). "Automated Linking of Historical Data." *Journal of Economic Literature*, 59(3): 865–918.
- ABRAMITZKY, R., L. P. BOUSTAN and K. ERIKSSON (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy*, 122(3): 467–506.
- BAINES, D. (2002). *Migration in a mature economy: emigration and internal migration in England and Wales 1861-1900.* Cambridge University Press.
- BANDIERA, O., I. RASUL and M. VIARENGO (2013). "The Making of Modern America: Migratory Flows in the Age of Mass Migration." *Journal of Development Economics*, 102: 23–47.
- BERGEAUD, A. and C. VERLUISE (2024). "A New Dataset to Study a Century of Innovation in Europe and in the US." *Research Policy*, 53(1): 104903.
- BERKES, E. (2018). "Comprehensive Universe of US patents (CUSP): Data and Facts." Working Paper.
- BERKES, E., D. M. COLUCCIA, G. DOSSI and M. P. SQUICCIARINI (2023). "Dealing with Adversity: Religiosity or Science? Evidence from the Great Influenza Pandemic." *Working Paper*.
- BERTHOFF, R. (1953). British Immigrants in Industrial America 1790-1950. Harvard University Press.
- CHEN, J. and J. ROTH (2024). "Logs with Zeros? Some Problems and Solutions." *The Quarterly Journal of Economics*, 139(2): 891–936.
- CONLEY, T. G. (1999). "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45.
- CORREIA, S., P. GUIMARÃES and T. ZYLKIN (2020). "Fast Poisson Estimation with High-Dimensional Fixed Effects." *The Stata Journal*, 20(1): 95–115.
- CURTIS, S. J. (1875). The story of the Marsden mayoralty: with sketch of the mayor's life. Leeds (UK): Express Office.
- ECKERT, F., A. GVIRTZ, J. LIANG and M. PETERS (2020). "A Method to Construct Geographical Crosswalks with an Application to US Counties Since 1790." *NBER Working Paper*, (No. w26770).
- FINISHING PUBLICATIONS, L. (2018). "Early British Patents. A Cradle of Inventions. British Patents 1617 to 1895." MFIS [data collection].
- GOODMAN-BACON, A. (2021). "Difference-in-Differences with Variation in Treatment Timing." *Journal* of *Econometrics*, 225(2): 254–277.
- HANLON, W. W. (2016). "British Patent Technology Classification Database: 1855-1882." Unpublished [data collection].
- KELLY, B., D. PAPANIKOLAOU, A. SERU and M. TADDY (2021). "Measuring Technological Innovation Over the Long Run." *American Economic Review: Insights*, 3(3): 303–320.
- LAN, T. and P. LONGLEY (2019). "Geo-Referencing and Mapping 1901 Census Addresses for England and Wales." *ISPRS International Journal of Geo-Information*, 8(8): 320.

- NICHOLAS, T. (2014). "Technology, Innovation, and Economic Growth in Britain since 1870." In R. Floud, J. Humphries and P. Johnson, eds., "The Cambridge Economic History of Modern Britain," pp. 181–204. Cambridge (UK): Cambridge University Press.
- NUVOLARI, A. and V. TARTARI (2011). "Bennet Woodcroft and the Value of English Patents, 1617–1841." *Explorations in Economic History*, 48(1): 97–115.
- NUVOLARI, A., V. TARTARI and M. TRANCHERO (2021). "Patterns of Innovation During the Industrial Revolution: A Reappraisal Using a Composite Indicator of Patent Quality." *Explorations in Economic History*, 82: 101419.
- ROSENBERG, N. and M. TRAJTENBERG (2004). "A General-Purpose Technology at Work: The Corliss Steam Engine in the Late-Nineteenth-Century United States." *The Journal of Economic History*, 64(1): 61–99.
- RUGGLES, S., C. FITCH, R. GOEKEN, J. HACKER, M. NELSON, E. ROBERTS, M. SCHOUWEILER and M. SOBEK (2021). "IPUMS ancestry full count data: Version 3.0 [dataset]." *Minneapolis, MN: IPUMS*.
- SARADA, S., M. J. ANDREWS and N. L. ZIEBARTH (2019). "Changes in the demographics of American inventors, 1870–1940." *Explorations in Economic History*, 74: 101275.
- SCHURER, K. and E. HIGGS (2020). "Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access." [data collection] Second Edition, UKDS.
- SPITZER, Y. and A. ZIMRAN (2018). "Migrant Self-Selection: Anthropometric Evidence from the Mass Migration of Italians to the United States, 1907–1925." *Journal of Development Economics*, 134: 226– 247.
- SUN, L. and S. ABRAHAM (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.
- WAGNER, D. (2008). Science and Civilisation in China. Cambridge University Press (UK).