Università Commerciale "Luigi Bocconi"
Ph.D. School

Ph.D. Program in: Economics and Finance
Cycle: XXXIV
Disciplinary Field: SECS-P/01

Essays in the Economics of Innovation

Advisor
Mara P. Squicciarini

Author
Davide M. Coluccia
ID Number: 3081045

Year 2023

Essays in the Economics of Innovation

By

Davide M. Coluccia

Abstract

Innovation is a crucial determinant of long-run economic growth in advanced economies. This dissertation explores the economic and social determinants of the production and diffusion of innovation in the context of Europe and the United States in the late nineteenth and early twentieth century.

The first chapter (jointly authored with Gaia Dossi) documents how out-migration impacts innovation in the country of origin of migrants. During the Age of Mass Migration, nearly four million English migrants settled in the US. We construct a novel individual-level dataset linking English immigrants in the US to the UK census and complement it with the newly digitized universe of UK patents. Using a new shift-share instrument for bilateral migration flows and a triple-differences design, we document a positive, significant, and persistent effect of exposure to US technology through migrant ties on the direction of innovation in Britain in 1870–1940. The individual-level analysis suggests that physical return migration is not the main factor underlying this "return innovation" effect. Instead, we find that migration ties generate information flows that facilitate the cross-border diffusion of novel knowledge. Furthermore, our findings suggest that market integration fostered by migration linkages is a crucial driver of information flows.

The second chapter (jointly authored with Lorenzo Spadavecchia) interprets out-migration through the lenses of directed technical change and adoption theory. We study the impact of immigration restriction policies on technology adoption in countries sending migrants. Between 1920 and 1921, the number of Italian immigrants to the United States dropped by 85% after Congress passed the Emergency Quota Act, a severely restrictive immigration law. In a difference-in-differences setting, we exploit variation in exposure across Italian districts to this massive restriction against human mobility. Using novel individual-level data on Italian immigrants to the US and newly digitized historical censuses, we show that this policy substantially hampered technology adoption and capital investment. We interpret this as evidence of directed technical adoption: an increase in the labor supply dampens the incentive for firms to adopt labor-saving technologies. To validate this mechanism, we show that more exposed districts display a sizable increase in overall population and employment in manufacturing. We

provide evidence that “missing migrants,” whose migration was inhibited by the Act, drive this result.

The third chapter (jointly authored with Enrico Berkes, Gaia Dossi, and Mara P. Squicciarini) investigates how societies respond to adversity. After a negative shock, separate strands of research document either an increase in religiosity or a boost in innovation efforts. In this paper, we show that both reactions can occur at the same time, driven by different individuals within the society. The setting of our study is 1918–1919 influenza pandemic in the United States. To measure religiosity, we construct a novel indicator based on the naming patterns of newborns. We measure innovation through the universe of granted patents. Exploiting plausibly exogenous county-level variation in exposure to the pandemic, we provide evidence that more-affected counties become both more religious and more innovative. Looking within counties, we uncover heterogeneous responses: individuals from more religious backgrounds further embrace religion, while those from less religious backgrounds become more likely to choose a scientific occupation. Facing adversity widens the distance in religiosity between science-oriented individuals and the rest of the population, and it increases the polarization of religious beliefs.

Acknowledgements

Each line of this dissertation attests my gratitude to the unwavering personal and professional support of my supervisor, Mara Squicciarini, throughout these years. When I entered graduate school, my research interests focused on Macroeconomics. As I transitioned toward Applied and Historical Economics, she patiently guided my first clumsy research efforts and helped me approach a subject so distant. In retrospect, those months were probably the most engaging and rewarding of my career thus far. I will forever be grateful for her support during that period and ever since.

I am deeply thankful to Simon Görlach and Maristella Botticini. Simon has been a constant source of insightful and thought-provoking discussions. He taught me that learning requires methodological rigor and motivated me to (try to) pursue it. The support of Maristella, especially as I first approached research in Economic History, greatly inspired my understanding of the role of supervisors. Shall I be given the opportunity to teach, her influence will not go unnoticed. Finally, I am grateful to Erika Deserranno, who agreed to join my committee at a later stage. I am indebted to the entire Bocconi faculty for the many discussions, talks, and coffees. Among others, my gratitude goes to Jan Bakker, Alexia Delfino, Stefano Fiorin, Thomas Le Barbanchon, Nicola Limodio, Tommaso Monacelli, Carlo Schwartz, and Antonella Trigari. Likewise, I sincerely thank the Economics community at Northwestern University, particularly Walker Hanlon, Leander Heldring, and Joel Mokyr, for hosting me and for the valuable feedback I received while in Evanston.

I doubt that a single line of this dissertation would have been written had it not been for the care and support of my friends and family. A few lines would neither suffice nor be necessary to acknowledge your importance in my life. You all know it, I hope. I most certainly do. This is, in all honesty, reassuring, for I could never convey in writing what I feel toward you. A simple yet so intensely heartfelt thank you, then, to Alice, Davide, Federico, Gianfranco, Ginevra, Giovanni, and Natascia. And to the many more whose life has crossed mine, and to those whose still does.

To Sofia.

To Mom and Dad.

Contents

List of Figures

List of Tables

Chapter 1

Return Innovation[*]

Evidence from the British Migration to the US, 1860-1940

1.1 Introduction

The adoption of foreign innovation accounts for a sizable proportion of national productivity growth and, ultimately, long-run development (Eaton and Kortum, 1999).[2] However, evidence on the drivers of the cross-country diffusion of knowledge remains surprisingly limited. In this paper, we document that exposure to foreign technology through out-migration linkages contributes to the diffusion of innovation to emigration countries.[3] This result—which we label the "return innovation" effect—offers a more nuanced perspective on the effects of emigration on the economic development of sending countries compared to the traditional "brain drain" narrative (Docquier and Rapoport, 2012). Depending on the specialization of their settling locations, emigrants are exposed to different types of novel knowledge. We find that this heterogeneity, in turn, generates information flows that shape the trajectory of innovation of their origin locations, thus contributing to international technology diffusion.

[*]This chapter is jointly authored with Gaia Dossi. We are extremely grateful to Mara Squicciarini for her continued guidance and support. For insightful comments and discussions, we also thank Jan Bakker, Enrico Berkes, Maristella Botticini, Carola Frydman, Simon Görlach, Walker Hanlon, Leander Heldring, Matteo Leombroni, Nicola Limodio, Paolo Masella, Filippo Mezzanotti, Joel Mokyr, Alessandro Nuvolari, Laura Ogliari, Sebastian Ottinger, Marta Prato, Nancy Qian, Paola Sapienza, Carlo Schwartz, Marco Tabellini, Edoardo Teso, and John Van Reenen. We thank audiences at the EHA (Pittsburgh), EHS (Cambridge), and ES (Berlin) conferences, at the 10th CEPR Economic History Symposium, and seminar participants at Bocconi, Milan, Northwestern, Cologne, and Uppsala. Antonin Bergeaud, Enrico Berkes, and Walker Hanlon kindly shared data with us. We are grateful to Bocconi, Fondazione Invernizzi, MapTiler AG, and POID for financial support. Davide Coluccia thanks the Department of Economics at Northwestern University for hospitality. All errors are our own.

[2]Economic historians have long argued that the diffusion of knowledge is a key driver of productivity growth and catching up (e.g., see Gerschenkron, 1962; Rosenberg, 1982). However, endogenous growth models featuring cross-country diffusion dynamics have emerged only recently (Alvarez *et al.*, 2013; Buera and Oberfield, 2020; Benhabib *et al.*, 2021; Perla *et al.*, 2021; Van Patten, 2023). Eaton and Kortum (1999), for example, estimate that 87% of productivity growth in France during the 1980s was due to foreign research.

[3]A vast scholarship documents that immigrants actively contribute to several dimensions of economic development in their destination countries spanning entrepreneurship (Kerr and Kerr, 2020; Azoulay *et al.*, 2022), innovation (Ganguli, 2015; Bahar *et al.*, 2019; Burchardi *et al.*, 2020; Bernstein *et al.*, 2022) and science (Moser *et al.*, 2014, 2020), local specialization (Ottinger, 2020), and the formation of political preferences (Giuliano and Tabellini, 2020). As pointed out by Clemens (2011), emigration has generally generated far less attention than immigration.

The impact of emigration on innovation is *ex-ante* ambiguous. Classical "brain drain" arguments suggest that emigration countries suffer from a loss of human capital (for a review, see Gibson and McKenzie, 2011). Standard growth theory, in turn, predicts that this would negatively affect the subsequent ability to innovate (*e.g.*, Benhabib and Spiegel, 2005). On the other hand, recent scholarship suggests that exposure to innovation is a crucial determinant of innovation activity (Akcigit *et al.*, 2018; Bell *et al.*, 2019). Building on these insights, we argue that as migrants are exposed to innovation in the areas where they settle, they facilitate knowledge flows between those areas and their origin country. Since this "return innovation" effect and the "brain drain" channel operate in opposing directions, empirical evidence is necessary to quantify their magnitudes and assess the overall impact of out-migration on innovation dynamics in countries sending migrants.

The setting of this study is the English and Welsh migration to the United States during the Age of Mass Migration (1850–1920). Between 1850 and 1920, approximately four million British migrated to the United States (Berthoff, 1953).[4] The British constitute the single largest immigrant ethnic group in the contemporary United States and have been credited as one major influence in the development of American culture (Fischer, 1989). Besides its historical importance, this setting allows us to overcome three critical limitations of contemporary scenarios. First, our novel individual-level data allow us to track bilateral migration flows within Britain and the United States at a granular level of spatial aggregation. Second, we measure international knowledge flows using detailed historical patent data. This approach would be unfeasible with contemporary data due to intellectual property protection laws enacted in 1945. Finally, the near-complete absence of emigration and immigration regulations ensures that endogenous migration policy interventions do not confound the analysis.

To estimate the effect of out-migration on innovation, we observe that districts in the UK would be exposed to different technologies depending on the county where emigrants from those areas would settle.[5] Our research design thus leverages the joint variation in (i) county-level specialization across technology classes and (ii) district-county bilateral migration flows. For the sake of argument, consider two English districts, call them A and B, and two US counties, call them a and b. Districts A and B have the same population and emigration rates, but all emigrants from district A settle in county a, whereas all those originating in B move to b.

[4] This figure does not include the Irish. In the paper, we focus on the English and Welsh migration. We thus use, for the sake of brevity, the terms "British" and "English" as shortcuts to collectively refer to England and Wales, thus excluding Scotland.

[5] In most of the analysis, the 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 terms of population (approximately 40,000). Unlike counties, however, registration districts were statistical entities that did not enjoy political or budgetary autonomy.

Furthermore, suppose that county a specializes in one sector, call it s_a, whereas county b specializes in s_b. Then, in our baseline empirical setting, district A will be exposed to technology in sector s_a, whereas district B shall experience exposure to s_b.

Existing data would not be suitable for this exercise. First, we lack disaggregated data on the origin of English immigrants in the United States. We leverage confidential individual-level data from the US and the UK historical population censuses (Schurer and Higgs, 2020; Ruggles *et al.*, 2021) to overcome this difficulty. We link US English immigrants' records to their entry in the UK census using state-of-the-art linking procedures (Bailey *et al.*, 2020; Abramitzky *et al.*, 2021). The resulting novel dataset allows us to track individual-level out-migration and return migration between the US and the UK and to compute granular bilateral flows between the two countries. Moreover, since we observe census-measured characteristics before and after the migration spell, we can study selection and assimilation dynamics in detail. Second, historical patent data for England were available starting in the late 1890s. We thus digitize the universe of patents issued in England and Wales from 1853–1899 to construct the first comprehensive dataset covering innovation during the Second Industrial revolution in the United Kingdom.

Assortative matching is the primary factor that cautions against a causal interpretation of the estimated relationship between exposure to US knowledge and innovation. For example, suppose that out-migration correlates with a—possibly unobserved—factor that also predicts the location where British immigrants settle in the United States. Then, the coefficient of a naï regression between US knowledge exposure and innovation would reflect the spurious association between district- and county-level underlying covariates. Again, consider the example above, and further suppose that districts A and B also specialize in one sector each, s_a and s_b, respectively. Then, the estimated association between knowledge exposure and innovation would conflate the spurious correlation driven by the initial specialization patterns. In practice, however, the unobserved confounding variable would need to vary over time *and* across technology classes since we control for time-varying unobserved heterogeneity at the district level.

We develop two empirical approaches to ensure that there exists a causal link between knowledge exposure spurred by migration ties and innovation. First, we construct a shift-share instrumental variable that exploits conditional county-level variation in the connection timing to the railway network to randomize immigration choices across counties (Sequeira *et al.*, 2020). The instrument's validity requires that these county-level shocks only affect British immigration through railway connection (Borusyak *et al.*, 2022). This assumption is especially appealing in our setting because the shift-share shocks predict overall—and not British only—immigration. Second, we note that the return innovation effect would imply that shocks to innovation activity in the United States would reverberate in the United Kingdom in districts whose emigrants had settled in those areas where these shocks manifest. To test this, we implement two triple differences analyses that compare districts and technology classes by exposure to US county-level (i) synthetic innovation shocks and (ii) Great Influenza pandemic mortality, which exerted a positive impact on pharmaceutical innovation in comparatively more affected counties (Berkes *et al.*, 2022). The absence of statistically significant pre-treatment differences between treatment and control districts provides evidence in favor of the standard parallel trends identification assumption.

Our main result is that exposure to foreign knowledge through migration linkages shapes the direction of innovation in emigration countries. We document a positive and statistically significant association between various metrics of exposure to US innovation and patenting activity in the United Kingdom. Importantly this result, which we label the "return innovation" effect, is quantitatively larger for technologies that were already present in the UK. This suggests that exposure to foreign knowledge through migration ties nurtures existing industries rather than creating new ones. The instrumental variable analysis confirms the existence of a causal link between exposure to US knowledge and innovation in the UK. Furthermore, the triple differences analysis provides evidence that innovation shocks in the US diffuse into the United Kingdom through migration ties. We estimate that, on average, exposure to a synthetic innovation shock in the United States results in .4 yearly patents. This figure is quantitatively sizable, accounting for approximately one-third of the average annual number of patents by district technology class. Similarly, districts more exposed to the pharmaceutical innovation boom ushered by the Great Influenza pandemic in the United States display higher rates of pharmaceutical innovation after the pandemic. Crucially, the flexible triple-differences estimates in both cases provide evidence of the parallel trends assumption.

Several potentially concurrent mechanisms may explain the return innovation effect. We first distinguish between those requiring the physical return of British immigrants in the US and those that do not. Return migration may impact home innovation directly if return migrants engage in innovation activity or indirectly if they spread the knowledge they had acquired during their period abroad (Bahar *et al.*, 2022b). To do so, we leverage the individual-level nature of our migration and patenting data. Using a linked patent-census sample and geo-coded information on the universe of the UK population, we estimate the effect of geographical proximity to US emigrants on innovation activity by non-migrants in a difference-in-differences setting. We find that patenting activity increases for non-migrants after their neighbor migrates to the United States. Statistically significant pre-trends do not drive this estimate. Moreover, the estimated effect remains positive and significant when restricting the treatment to include only

US migrants that never return. The first finding excludes that innovation activity by return migrants alone explains the return innovation effect. The second result, in turn, suggests that it is unlikely that the indirect return migration channel is its primary determinant.

Our favored interpretation of the return innovation effect is that migrant linkages foster the flow and diffusion of information, which, in turn, nurture innovation. This "information diffusion" effect would operate plausibly independently from the physical return of emigrants. We leverage the introduction of the first transatlantic telegraph cable (1866) as a sudden and sizable negative information friction shock to shed light on this mechanism.[6] In a difference-in-differences setting, we show that districts with higher US emigration rates display higher patenting activity after the introduction of the telegraph. Moreover, increased innovation is not even across technology classes. Instead, we estimate that the gains in patenting activity manifest in those same technologies that districts had been more exposed to through migration ties. Reassuringly, we find that the transatlantic telegraph positively impacted innovation only in districts connected to the domestic network.

Even though we cannot explicitly disentangle the "meta-mechanism" that drives information diffusion, our preferred interpretation is that migration ties facilitate cross-border market integration, thus fostering the exchange of knowledge and information (Aleksynska and Peri, 2014; Ottaviano *et al.*, 2018). Exploiting bilateral trade as a measure of market integration, we leverage technology-level variation in import duties following the Smoot-Hawley Tariff Act (1930) in a difference-in-differences setting. We document that innovation activity decreased in districts with more US emigrants in technology classes akin to industries targeted by the tariff. While suggestive, this analysis cannot refute that hampered export-driven innovation and not market integration explains the observed decrease in patenting activity (Bustos, 2011; Atkin *et al.*, 2017).

Finally, we explore the scope of the information flows generated by migration ties. More specifically, we investigate whether these are restricted to novel knowledge and innovation or if they encompass a broader set of subjects. We collect coverage data of general US-related information from a comprehensive repository of historical British newspapers. We find that newspapers in areas with more US emigrants are relatively more likely to cover US-related news. Analogously, newspaper coverage of a given state (resp. county) is broader in districts with more emigrants to a given state (resp. county). This exercise suggests that the scope of information flows generated by migration ties is not limited to the diffusion of novel knowledge

[6]The telegraph represented a fundamental development in information and communication technology. Steinwender (2018) documents that the transatlantic cable allowed information to flow more rapidly and efficiently across the Atlantic Ocean, thus enabling trade and reducing international arbitrage opportunities.

and innovation. Hence, out-migration exerts potentially wide-encompassing consequences on emigration countries.

This paper provides new evidence on how out-migration impacts innovation in the countries of origin of emigrants. We find that exposure to foreign knowledge through migration ties influences the direction of innovation in emigration countries and contributes to the diffusion of novel knowledge. Despite obvious cautions on the external validity, our results bear relevant policy implications on how to think of the role of out-migration in economic development and technology catch-up.

Related Literature. This paper contributes to four streams of literature. First, we inform literature that studies the determinants of the direction of innovation and the allocation of research activity across technological sectors. Pioneering work on directed technical change by Habakkuk (1962) was formalized by Acemoglu (2002, 2010). More recently, this question has been studied both theoretically (Bryan and Lemus, 2017; Bryan *et al.*, 2022; Hopenhayn and Squintani, 2021) as well as empirically (Aghion *et al.*, 2016; Moscona, 2021; Moscona and Sastry, 2022; Einiö *et al.*, 2022; Gross and Sampat, 2022). We inform this literature by introducing one novel determinant of the direction of innovation, namely, international human mobility, through the return innovation effect.

Second, we inform the literature that studies the effects of out-migration on countries sending migrants. Emigration has been shown to impact wages (*e.g.*, Dustmann *et al.*, 2015), attitudes towards democracy and voting (Spilimbergo, 2009; Batista and Vicente, 2011; Ottinger and Rosenberger, 2023) and political change (Chauvet and Mercier, 2014; Kapur, 2014; Karadja and Prawitz, 2019), technology adoption (Coluccia and Spadavecchia, 2022), and social norms (Beine *et al.*, 2013; Bertoli and Marchetta, 2015; Tuccio and Wahba, 2018). Our paper is closest to Andersson *et al.* (2022) who show that out-migration in nineteenth-century Sweden triggered labor-saving innovation in the origin areas of emigrants. Relative to their contribution, this paper provides novel evidence that out-migration generates international technology transfer and knowledge flows that shape the direction of innovation in emigration countries. Moreover, we highlight that this effect is likely driven by information flows rather than physical return migration.

By its setting, this paper adds to the literature that studies technical change and diffusion of novel technologies during the Age of Mass Migration. A growing number of papers examines the short-run (Arkolakis *et al.*, 2020; Moser *et al.*, 2020; Diodato *et al.*, 2022) as well as the long-run (Akcigit *et al.*, 2017; Burchardi *et al.*, 2020; Sequeira *et al.*, 2020) implications of immigration on US innovation. Ottinger (2020) shows that European immigration influenced US industry specialization. Andersson *et al.* (2018, 2022) document that the mass migration of about a quarter of the Swedish population between 1850 and 1913 ignited labor-saving technological change. Compared to Andersson *et al.* (2022), in particular, we show that exposure to US technology altered the *direction* of innovation in the country of origin of emigrants. Moreover, we document that migration ties foster the diffusion of information on novel knowledge absent physical return migration. Furthermore, we provide two methodological contributions: we develop a new database of English and Welsh patents that spans the Second Industrial Revolution, and we construct a linked sample of emigrants from individual-level census data.

Finally, we relate this paper to the literature studying the dynamics and determinants of knowledge flows and technology diffusion across countries (among others, see Jaffe *et al.*, 1993; Bahar *et al.*, 2014; Pauly *et al.*, 2021). In particular, we contribute to the papers documenting how human mobility fosters the diffusion of novel knowledge (Kerr, 2008; Bahar *et al.*, 2019; Fackler *et al.*, 2020; Bahar *et al.*, 2022a; Prato, 2021). We inform this literature from several perspectives. First, we enlarge the observation sample to include the universe of emigrants instead of a selected subgroup of highly skilled individuals. Second, we leverage recent insights by Akcigit *et al.* (2018) and Bell *et al.* (2019) and show that exposure to foreign technology is a major driver of technology transfers through migration ties. Finally, our setting allows us to uncover the long-run effects of emigration and the mechanisms through which it affects innovation in the home country of emigrants.

Outline. The rest of the paper is structured as follows. In section 1.2, we describe this study's historical and institutional context. Section 1.3 introduces the novel datasets we assemble and the rest of the data. We present the empirical research design in section 1.4 and discuss the main findings in section 1.5. Section 1.6 uncovers the possible mechanisms underlying the results and discusses possible alternative interpretations. Section 1.7 concludes.

1.2 Historical and Institutional Background

This section offers a concise overview of the historical and institutional features of our study setting. Throughout it, we highlight key aspects and details that are relevant to our empirical investigation.

1.2.1 The English and Welsh Migration to the United States

Between 1850 and 1930—during the so-called Age of Mass Migration—more than 30 million Europeans migrated to the United States (Abramitzky and Boustan, 2017). Migrants from Great

Britain—England and Wales in particular—accounted for approximately 10% of this flow (Willcox, 1928). Emigration rates in Britain were among the highest in Europe, except for the years 1890–1900. They steadily increased throughout the period (Baines, 2002).[7]

1.2.1.1 Migration Policy in the United Kingdom and the United States

The virtual absence of legal constraints to human mobility represents a major appealing feature of the Age of Mass Migration for economic research. Until 1917, the US applied minor restrictions on European immigration (Abramitzky and Boustan, 2017). Immigrants mostly originated from Northern Europe, particularly the United Kingdom, Germany, Sweden, and Norway. This positive attitude towards immigration ceased as flows from Eastern and Southern Europe increased in the 1890s (Goldin, 1994). The restrictive immigration policies of the 1920s, however, imposed 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 sought to help emigrants, if not explicitly to foster emigration (Baines, 2002, p. 72). Out-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. Emigration to the United States was neither subsidized nor discouraged. In general, however, these policies were not successful. Baines (2002) argues that less than 10% emigrants traveled under government assistance during the entire 1814-1918 period, and Leak and Priday (1933) report similar figures for the post-War era. Attitudes towards out-migration remained positive after the First World War. The perceived slowdown of emigrant flows after the War was viewed with concern by policymakers (Leak and Priday, 1933).

This overview suggests that institutional constraints to US out- and immigration were largely absent for English and Welsh migrants throughout the XIX and early XX century. Compared to contemporary scenarios, this historical setting thus allows us to abstract from confounding factors arising from endogenous migration legislation.

---

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

[8] The 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).

1.2.1.2 English and Welsh Emigrants: The Perspective of Great Britain

Compared to the broader European phenomenon, the British migration to the US presents two main distinctive features.[9] First, unlike continental countries, Britain was already highly urbanized and industrialized at the inception of the Mass Migration. Erickson (1957, 1972) and Thomas (1954) highlight the centrality of urban areas, starting in the 1880s, supplied the majority of overseas migrants. Baines (2002) provides some estimates on the origin of migrants based on birth certificates over the years 1850–1900. Emigration ratios were highest in Northern and South-Western England and lowest in Lancashire and neighboring areas. Second, the selection of British migrants radically differed from that in continental countries (Erickson, 1957; Abramitzky *et al.*, 2020). 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 were increasingly employed in industrial manufacturing occupations (Baines, 2002). During the 1860s, about 15% emigrants were employed in agriculture, and merely 5% were white-collar workers. In the 1900s, however, agriculture workers accounted for a mere 5% of the overall emigrant stock, while those employed in white-collar occupations were 25%.

Our newly constructed migration database allows us to assess the historical evidence quantitatively. In Appendix Table 1.D.4, we compare individual-level characteristics of US emigrants with the staying population. On average, emigrants are more likely to come from North West and South East England. Moreover, they are less likely to be farmers. By contrast, emigrants' share of high and low-skilled manufacturing workers is substantially larger than among stayers. Similar—although less marked—patterns are observed for return migrants. Appendix Figure 1.C.3 displays the origin of emigrants overtime at the district level. The data vividly show that rural areas in central and south-western England, which initially feature the highest emigration rates, are gradually replaced by urban industrial districts in the North and South. Taken together, this evidence confirms the qualitative historical knowledge.

1.2.1.3 English and Welsh Immigrants: The Perspective of the United States

British immigrants have been central throughout the economic and political history of the United States (Berthoff, 1953; Fischer, 1989). Several features distinguish the English from the

[9]Throughout the period, the US was the most relevant destination for English and Welsh migrants. Between 1850 and 1930, more than 40% emigrants settled in the US. This compares to 25% in Canada, 20% in Australia, and 15% in other non-European destinations (Baines, 2002, p. 63).

continental transatlantic migrations. First, English and Welsh immigrants were, especially after the 1880s, artisans and manufacturing workers, who settled where their skills were in highest demand (Berthoff, 1953).[10] Textile workers from Manchester typically settled in Massachusetts, whereas coal miners from Southern Wales mostly settled in the Midwest and Pennsylvania. In 1890, 63% of British-born were employed in American industry (Thistlethwaite, 1958). Second, English immigrants—unlike the Welsh—did not form ethnic clusters (Furer, 1972). Instead, they tended to be scattered around settlement areas in highly diverse ethnic communities. Finally, British immigrants were economically successful and assimilated relatively easily with the native population (Abramitzky *et al.*, 2020).

We quantitatively evaluate these observations in Table 1.D.5. First, we compare individual-level characteristics observed in the US census between the native and the British immigrants. The analysis suggests that British immigrants are substantially different from the average native. For example, they are richer, more literate, and more likely to live in urban centers. Consequently, they are less likely to be farmers and more likely to be employed in manufacturing occupations with high or low-skill content. In addition, English immigrants are comparatively more concentrated in North Atlantic and the West and less in Southern states. Similar patterns emerge for return migrants.

These results, coupled with Table 1.D.4, identify British immigrants in the US as part of an urban industrial class of skilled and semi-skilled workers. This is crucial in our analysis: it would have been much more difficult for illiterate farmers to facilitate knowledge flows across the Atlantic Ocean.

1.2.2 Intellectual Property Protection in the US and the UK

We measure innovation and knowledge flows using patent data. In this section, we briefly present the key features of the American and British patent systems and discuss the state of international intellectual property protection in the XIX and early XX centuries.

1.2.2.1 National Patent Systems

The first article of the United States Constitution establishes that inventors be granted exclusive rights over their discoveries. In 1836 the 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

[10] Thistlethwaite (1958) presents one instructive example. The pottery industry, a highly skilled and labor-intensive sector, was concentrated in the Five Towns of Staffordshire. As transatlantic migration ensued, ceramic workers located in just two localities: Trenton, New Jersey, and East Liverpool, Ohio.

system from its European counterparts. First, an examination of novelty was carried out by professional examiners to ascertain the originality of patent applications. Second, low application fees ensured that access to intellectual property protection was widespread (Sokoloff and Khan, 1990). Several scholars documented how effectively the US patent system fostered innovation well into the 20th century (Khan, 2020).

Britain established the world's oldest continuously operating patent system in 1623-1624. Until 1850, access to intellectual property protection was, however, difficult (Gomme, 1948; Bottomley, 2014). Fees amounted to approximately four times the average income in 1860, and the application process was lengthy and rife with uncertainty (Dutton, 1984). A large amount of literature documented the poor performance of this system during the Industrial Revolution (Macleod, 1988; Moser, 2012). The 1852 Patent Law Amendment Act sought to reform this process. The US system inspired the reform effort, which reduced application fees and tried to streamline bureaucratic procedures. One subsequent reform in 1883 further reduced fees, allowed applications by mail, designed a litigation system, and provided for the employment of professional patent examiners (Nicholas, 2011). 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).

1.2.2.2 International Intellectual Property Protection

As national patent systems spread across Europe and 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 throughout the period that we study (Penrose, 1951).

The Paris Convention emerged after a decade of multilateral confrontations spurred by World Exhibitions in Vienna (1873) and Paris (1878). The 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 had vigorously demanded reciprocity (Penrose, 1951, p. 66). Second, upon applying for a patent in one member country under Article 4, 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. The provisions contained in Article 4 were central within the broader legal apparatus (Penrose, 1951, p. 68). 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. This represented a substantial bureaucratic and financial burden. While the Paris Convention—and its numerous amendments—are still in operation today, international patents were established only in 1970.

The state of international intellectual property protection during our period is a major advantage of this historical setting. Since the UK and the US did not mutually recognize patents, we can use them as an informative proxy of knowledge flows between them. This approach would be impracticable in modern settings.

1.3 Data

This section presents our primary data sources and discusses the key methodology we adopt to assemble our final datasets. We provide a more detailed description of the data in Appendix sections 1.A, 1.B, and 1.C. Table 1.1 lists the main variables and provides descriptive statistics.

1.3.1 Migration Data

To conduct our analysis, we need information on the origin of English and Welsh immigrants in the United States *within* the United Kingdom. Currently available data, however, do not contain this information. Neither the US nor the UK collected disaggregated data on, respectively, the origin of immigrants and the destination of emigrants. We tackle this limitation of the data by developing a new dataset that links British immigrants in the US to the UK census. This allows us to observe an individual in the UK and to track him to his US census record after he emigrated.[11] This is the first dataset that reconstructs migration flows at this granular level of aggregation for a major European country in this period.[12]

To construct our linked dataset, we leverage non-anonymized individual-level data from the population censuses in the United Kingdom (Schurer and Higgs, 2020) and the United States (Ruggles *et al.*, 2021). We first extracted the universe of British immigrants from the censuses in 1900, 1910, 1920, and 1903. They list, among other variables, the name and surname, birth year, and immigration year of each migrant. We then match these records to the closest census when they appear. Hence, for example, we try to link an individual who immigrated to the US

[11] Throughout 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.

[12] Data assembled by Abramitzky *et al.* (2014) and Andersson *et al.* (2022) serve a similar purpose for, respectively, Norway and Sweden. England and Wales, however, were substantially larger in terms of the overall population and the US immigrant population. The population of Sweden and Norway in 1890 was approximately 4.7 and 2 million. In the same year, the population in England and Wales stood at 27 million.

in 1905 to the 1901 UK census.[13] The matching variables we consider are the name, surname, and reported birth year, using state-of-the-art census-linking algorithms adapted from pioneering work by Abramitzky *et al.* (2021). Appendix 1.C.1 lists in more detail the primary sources and the technical implementation of the algorithm. This class of linking algorithms relies on the observation that a simple exact matching routine would artificially discard many plausible links between the two sources because of minor coding errors by the census enumerators. Since human hand-checking is unfeasible, we implement an algorithm that returns a match whenever the string similarity between the US and the UK records is above a certain threshold, conditional on the birth year.

This approach presents some important caveats (Bailey *et al.*, 2020). 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. This would be the case if, for instance, the likelihood that names and surnames were correctly enumerated in the censuses correlated with education. We set high-quality thresholds to accept potential matches to address the first concern. Moreover, we only keep immigrants matched up to two records in the UK census. This ensures that we minimize the rate of false positives as much as possible. Finally, we provide evidence against the second issue in Table 1.C.1, which shows that the correlation between the number of matches and individual-level observable characteristics is seldom significant, and always very small in magnitude.

Finally, we construct a dataset of return migrants. To assemble it, we apply the exact previous logic, except that migrants are matched to the UK censuses taken in the decades *after* their immigration year. Hence, as an example, someone who migrated to the US in 1895 is matched to censuses in 1901 and 1911. To avoid double counting, if a migrant is matched to more than one census, we keep the match(es) in the first. Data on return migrants are generally scant historically and with modern data (Dustmann and Görlach, 2016). This exercise is thus a valuable feature of our methodology.

In Figure 1.1, we report in gray 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. We attain a matching rate of about 60% after dropping multiple matches and links with below-threshold matching quality. Note that we are forced to discard women whose surname was likely to change

[13] 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. This implies that 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 bears little quantitative implications for the matching rate in the later part of the sample.

after marriage. The matching rate aligns with the literature on census linking (Abramitzky *et al.*, 2021).[14] Moreover, reassuringly, our data co-moves with official statistics data. Figure 1.2 reports the spatial distribution of emigration rates across districts in the final sample and highlights its cross-sectional spatial heterogeneity. In Appendix Figure 1.C.3, we break down the map by decade and uncover substantial variation in the origin of US emigrants over time.

1.3.2 Patent Data

We measure innovation activity using patents, as is standard in the literature (Griliches, 1998).[15] Patents for the United States have been digitized from original documents by Berkes (2018). The data contain, among others, information on the authors' addresses, the filing date, and the CPC patent classification. We use these to construct a balanced panel dataset at the county-technology class-year level.[16]

We collect patents for the United Kingdom for the period 1895-1939 from PATSTAT, which in turn provides bulk access to data stored at the European Patent Office. These data contain information on authors and CPC classes but do not report the geographic location of inventors. We thus merge with data by Bergeaud and Verluise (2022) to retrieve the coordinates of inventors and map them to registration districts at their 1890 borders. Patent data for previous years, unfortunately, are not currently available. To tackle this limitation of the data, we digitized the universe of patents granted in England and Wales between 1853 and 1895. As a result, we assemble a unique patent-level database that leverages textual information from nearly 800,000 original patent documents.[17] We have information on the title, text, inventors' geo-references addresses, filing and issue date, and other variables not used in this paper. Next, we map patents to districts at 1890 borders. We then employ a simple machine learning classification algorithm, discussed in Appendix section 1.B.1, to assign technology classes leveraging information in titles.

[14] In Appendix section 1.C.2, we provide a more detailed discussion of the algorithm's performance.

[15] Previous research shows that patents are not a flawless measure of innovation because non-patented innovation represents a non-negligible share of overall technological progress (Moser, 2019). We nonetheless believe that this is a comparatively minor issue for our analysis. Before our study period, the US and the UK had enacted important reforms which decreased the cost of access to patent protection (Gomme, 1948). These drastically increased the number of patents in both countries, thus ensuring that patents convey an informative picture of the state of technology in both countries.

[16] We map patents to counties at 1900 borders using the inventors' coordinates. From the three-digit CPC class, we map patents to a coarser taxonomy of twenty sectors. Appendix 1.A.1 provides additional details.

[17] Appendix section 1.B.1 describes the primary sources and methodology we develop to extract and structure the data from the original documents. In section 1.B.2, we compare our series with two existing series and find that the three are highly consistent for the period of common support.

This newly developed dataset is the first with geographical and textual information on the universe of patents granted in England and Wales during the second half of the nineteenth century. Data by Hanlon (2016), for instance, do not list titles or texts and do not report geographic information. This dataset, which we plan to extend to 1617–1895, thus expands previous work by Nuvolari and Tartari (2011) and Nuvolari *et al.* (2021) and provides the first comprehensive assessment of innovation in Britain during the Second Industrial Revolution.

In some empirical applications, we link patent data to the census. This allows us to assign a unique label to single inventors appearing in multiple patents and to observe individual-level characteristics recorded in 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 1.A.3.

1.3.3 Other Variables

In this section, we provide a brief description of the additional heterogeneous data that we assemble. Appendix section 1.A.1 discusses each more diffusely.

1.3.3.1 Census Data

We assemble district-level statistics from population censuses at decade frequency between 1851 and 1911. Districts are the level of observation in most of the analysis. This is because they were statistical units with neither budgetary nor administrative authority. The average population was 40,000, which makes them roughly comparable to US counties. Districts undergo minor boundary changes during the analysis period. However, to ensure geographical consistency, we cross-walk all variables to districts in 1890 using the method described in Eckert *et al.* (2020). In particular, the census allows reconstructing the employment shares across sectors and other demographic information.

1.3.3.2 Newspapers

We use newspaper coverage of US-related topics as a measure of attention to the United States in public opinion. We collect the data from the British Newspaper Archive. Beach and Hanlon (2022) discuss this dataset in detail. We run three sets of queries. First, we search for the joint mention of the words “United States”; second, we search for mentions of each state; third, we search for mentions of each county. We collect these data at the newspaper level from 1850–1939. Additionally, we know each newspaper’s publishing address, which we geo-reference to 1890-border districts. Ultimately, we assemble three datasets at the district, district-state, and district-county levels, each at decade frequency. Figure 1.A.3 reports the distribution of newspapers.

1.3.3.3 Telegraph Network

We reconstruct the English and Welsh telegraph network from *Zeitschrift des Deutsch-Österreichischen Telegraphen-Vereins, Jahrgang*, volume IX, 1862. This directory lists all telegraph stations outside of London in 1862. To the best of our knowledge, it is the most comprehensive list before the establishment of the transatlantic telegraph cable connecting the UK and the US (1862). We georeference all the stations and assign them to 1890-border districts. Since, however, the source does not list stations in the London area, in the sample of the telegraph analysis, we conflate London urban districts into a single "London" unit, which we assume to be connected to the telegraph network. Figure 1.A.4 reports the distribution of the stations.

1.4 Empirical Strategy

This section describes our baseline empirical strategy. We discuss the potential caveats that hinder a causal interpretation of the resulting estimates. Then, we discuss two strategies to address these concerns.

1.4.1 Baseline Methodology

The central hypothesis of this paper is that exposure to foreign—in this case, American—knowledge through migrant linkages fosters the diffusion of novel knowledge—in this case, innovation. We thus develop a simple measure of exposure to US knowledge that leverages two sources of variation. First, local specialization across counties measures the knowledge that diffuses 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, and call them 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 σ_a. Emigrants from B settle in county b, which only innovates in sector σ_b. Then, we expect district A (resp. B) to innovate comparatively more in sector σ_a (resp. σ_b).

To implement this intuition, we define knowledge exposure as follows:

$$\text{Knowledge Exposure}_{ik,t} \equiv \sum_{j\in J} \left(\frac{\text{Patents}_{jk,t}}{\text{Patents}_{j,t}} \times \text{Emigrants}_{i\rightarrow j,t} \right) \tag{1.1}$$

where i, j, k, and t denote a (UK) district, a (US) county, a technology class, and a decade, respectively.[18] The set J denotes the universe of counties. The knowledge exposure term thus averages district-level exposure to county-level specialization across technology classes. The first term in the summation captures specialization, while the second term codes district-level exposure. One may argue, however, that the relative share of patents may inflate specialization in counties with a small number of granted patents. While this is unlikely to significantly bias our results as those countries would likely have low district-level exposure, we code an alternative knowledge exposure variable that measures specialization as the raw count of patents in a given technology class. One further challenge to measure (1.1) is that districts with larger bilateral linkages are probably larger and, hence, selected. To account for district-level time-varying confounding variables, we control non-parametrically for district-by-time fixed effects. However, we also 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 section 1.E.2.

We estimate variants of the following regression model:

$$\text{Patents}_{ik,t} = \alpha_{i\times t} + \alpha_{i\times k} + \beta \times \text{Knowledge Exposure}_{ik,t} + \varepsilon_{ik,t} \tag{1.2}$$

where the coefficient of interest (β) quantifies the correlation between innovation activity and exposure to foreign knowledge. The term $\alpha_{i\times t}$ denotes district-by-decade fixed effects whose inclusion allows to control non-parametrically for time-varying unobserved heterogeneity at the district level; the term $\alpha_{i\times k}$ denotes district-by-technology fixed effects and excludes variation arising, for example, from the possibility that district-level technology specialization and immigration location decisions may be correlated. We comment more on this second point in the next section. The error term is the $\varepsilon_{ik,t}$. Standard errors in this specification are clustered at the district level. We mainly estimate model (1.2) through ordinary least squares. Since the dependent variable presents a non-negligible share of zeros, we also report the estimates of the Poisson regression associated with the baseline model.[19]

[18] Throughout the paper, we refer to decade t to mean the ten years before the upper bound t. Hence, the decade indexed by 1890 refers to 1881–1890.

[19] In the innovation literature, it is common practice to apply a log transformation to the dependent variable. We do not follow this practice because Chen and Roth (2022) show that average treatment effects for transformations of the dependent variable defined in zero are arbitrarily scale-dependent. In Appendix section 1.D.4, we present alternative specifications with multiple transformations of the dependent variable.

1.4.2 Threats to Identification

The main factor that cautions against a causal interpretation of the estimates of model (1.2) is assortative matching, meaning that 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 across technology classes.

In section 1.2.1, we discussed that the historical and quantitative 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 the same 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 be against our result.

Assortative matching also arises 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, say Lancashire, were comparatively more likely to settle in counties with larger textile sectors. Then, the estimated β of model (1.2) 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, suggest that non-random location decisions may represent a severe threat in this context. We attempt to quantify this issue in Appendix section 1.D.4.2. We measure the similarity of innovation portfolios between districts and counties and check whether this measure of specialization proximity correlates with observed bilateral migration flows. Table 1.D.2 reports the results. We find no significant association between innovation similarity and migration choices. This suggests that assortative matching is a plausibly minor concern for our analysis. Moreover, in the baseline estimation equation (1.2), we include district-by-technology fixed effects. Hence, for assortative matching to bias our estimates, the underlying confounding variable would need to vary over time across district-technology pairs.

While we present evidence against the presence of assortative matching, we ultimately cannot rule it out. We thus develop two strategies that, we argue, ameliorate residual endogeneity concerns.

1.4.3 Shift-Share Instrumental Variable Strategy

We design a shift-share instrument that leverages recent advancements in the econometric literature to deal with selection and assortative matching. Identification critically hinges on the observation that instrument validity can be obtained from the quasi-random assignment of shocks (Borusyak *et al.*, 2022). We construct county-specific immigration shocks by interacting aggregate immigration flows in the US with the gradual expansion of the railway network along the lines of Sequeira *et al.* (2020). These generate exogenous shocks in a quasi-experimental shift-share design à la Borusyak *et al.* (2022).

To construct the shocks, we predict the county-level immigrant share, which is not specific to British immigrants, from a regression between the actual immigrant shares and an interaction between the timing of connection to the railway network and the aggregate inflow of immigrants. Importantly, we control for county-level unobserved time-invariant heterogeneity and several other potential confounding variables at the county level.[20] In our context, shocks are conditionally exogenous if the settlement decisions of British immigrants did not influence the direction of the enlargement of the US railway network. In other words, instrument validity requires that shocks randomly assign British emigrants across counties. Under this assumption, the instrument breaks issues of assortative matching. This may fail if, for instance, British immigrants settled in counties more similar to their area of origin among the counties connected to the network in a given period. Since county-level shocks yield the overall predicted immigrant shares—and not those of the British only—we believe this is a relatively minor concern to rule out by assumption. Following Borusyak *et al.* (2022), we show that shocks are uncorrelated with county-level confounding variables and that the instrument does not systematically predict district-level characteristics. Appendix Figure 1.E.3 shows that while immigrant shares correlate with district-level observable characteristics (Panel A), predicted immigration shares do not (Panel B). Similarly, in Appendix Figure 1.E.4, we confirm that while out-migration correlated with most district variables, the instrument displays smaller and insignificant correlations with the same variables. These exercises provide evidence in favor of the validity of our research design.

Let $\omega_{j,t}$ be the immigrant share in county j in decade t, and let $\hat{\omega}_{j,t}$ its prediction. We thus define the instrument as

$$\widehat{\text{Emigrants}}_{i\to j,t} \equiv \hat{\omega}_{j,t} \times \sum_{j\in J} \left(\hat{\omega}_{j,t} \times \text{Emigrants}_{i\to j,1880}\right) \tag{1.3}$$

[20] In Appendix section 1.E.2, we describe in more detail the practical computation of the immigration shocks.

where $\text{Emigrants}_{i \to j,1880}$ denotes the number of emigrants leaving district i and settling in county j at the beginning of the sample period. Importantly, this exposure term is allowed to be endogenous by design. Identification stems from the quasi-exogeneity of the shocks $\{\hat{\omega}_{j,t}\}$. Given a predicted set of bilateral flows, we construct the instrument for knowledge exposure as in (1.1), except that the predicted flows replace the observed ones.

Even though we present evidence otherwise, the conditional exogeneity of the timing of railway connection is ultimately an untestable assumption. To validate the results obtained with the instrument (1.3), we construct an additional series of county-level shocks $\{\hat{\omega}_{j,t}\}$ that leverages a different source of variation. Specifically, we compute "leave-out" predicted county-level immigrant shares by interacting start-of-period immigrant shares with aggregate inflows by nationality. Importantly, we exclude British immigrants when calculating these shocks. This ensures that the "leave-out" shares do not reflect the settling decisions of the British. We describe the procedure in more detail in Appendix 1.E.2.2. This alternative instrument yields results that are highly consistent with the railway-based approach.

1.4.4 Shock Propagation Difference-in-Differences Strategy

To provide additional causal evidence on the effect of exposure to foreign knowledge through migration ties on domestic innovation, we devise a research design that leverages geographically clustered innovation shocks in the United States in a triple-differences setting. We start by observing a logical corollary of the return innovation argument. Suppose we observe a sudden increase in the number of patents granted in some counties. Then, one would expect that districts whose emigrants had settled more extensively in those counties would display increased innovation activity. In other words, innovation shocks in the United States should "reverberate" in the United Kingdom through pre-existing migration linkages.

We test this prediction using two sets of innovation shocks. First, as we describe in more detail in Appendix 1.E.3, we construct a set of county-technology class synthetic innovation shocks at yearly frequency. We regress the number of patents against fixed effects to obtain the residualized innovation activity. Then, we flag an innovation shock $\xi_{jk,t}$ whenever the residualized number of patents in a given county j, technology class k, and year t is in the top 0.1% of the overall distribution.[21] Appendix Table 1.D.8 documents that shocks are relevant, as one such shock is associated with an average of forty patents more in the given county. Second, we leverage recent evidence by Berkes *et al.* (2022), who document that the Great Influenza pandemic (1918–1919) significantly and positively affected pharmaceutical innovation in counties

[21] In Appendix Table 1.E.3.3, we show that the results remain consistent when imposing different values to flag innovation shocks.

that were more exposed to the pandemic. We thus claim that districts that were comparatively more exposed to affected counties should feature increased pharmaceutical innovation. We provide additional details on the construction of county-level exposure to the pandemic in Appendix 1.E.3.[22] We code county-level exposure to the pandemic as a dummy φ_j that returns value one if the ratio between deaths during the pandemic (1918–1919) and deaths in the preceding three years (1915–1917) is in the top 25%, and zero otherwise.

We measure district-level exposure to the county-level shocks in terms of the emigrants that had left the given district to settle in the given county *before* the period of analysis.[23] Formally, we compute exposure to synthetic shocks in technology class k as

$$\text{Synthetic Shock Emigrants}_{ik,t} = \sum_{j \in J} \left(\text{Emigrants}_{i \to j,1900} \times \xi_{jk,t} \right) \tag{1.4}$$

and analogously, we define exposure to counties affected by the pandemic as

$$\text{Influenza Emigrants}_{ik} = \sum_{j \in J} \left(\text{Emigrants}_{i \to j,1900} \times \varphi_j \right) \tag{1.5}$$

To avoid issues of continuous treatment described by Callaway *et al.* (2021), we recast each exposure metric in terms of a dummy variable that returns value one if the associated continuous measure is in the top 25%, and zero otherwise.[24]

To estimate the effect of US synthetic shocks on UK innovation activity, we estimate the following triple differences specification:

$$\text{Patents}_{ik,t} = \alpha_{i\times k} + \alpha_{k\times t} + \alpha_{i\times t} + \sum_{h=-a}^{b} \beta^h \times \mathrm{I}\left[D_{ik,t} = h\right] + \varepsilon_{ik,t} \tag{1.6}$$

where $\alpha_{i\times k}$, $\alpha_{k\times t}$, and $\alpha_{i\times t}$ denote, respectively, district-by-technology class, technology class-by-year, and district-by-year fixed effects.[25] The term $\left(D_{ik,t} \equiv t - \mathrm{I}\left[\text{Synthetic Shock Emigrants}_{ik,t}\right]\right)$ denotes the number of years since the district-technology class ik was exposed to a synthetic

[22]Since the technology taxonomy used in this paper is different from Berkes *et al.* (2022), in Appendix Table 1.D.7 we confirm that their result holds in our data. Figure 1.D.2 reports the associated flexible triple differences specification. Moreover, in Figure 1.E.6a, we confirm that the pandemic affected innovation activity only in the pharmaceutical sector.

[23]This part of the analysis restricts the outcome variable to 1900–1930, so we can leverage migrant flows in the preceding decade (1890–1899) to construct fixed exposure shares.

[24]In Appendix Table 1.E.7 we consider alternative thresholds to code the exposure variable (1.4). In Appendix Table 1.E.8, we report the results using the continuous measure (1.5).

[25]When we estimate regression (1.6) using variation in exposure to the pandemic shock, we normalize the dependent variable by the average number of patents granted before the pandemic to ensure that the estimated coefficients' size are comparable.

innovation shock ξ. The roll-out of the treatment is staggered across units. Different district-class pairs may be exposed to the exposure treatment at different points in time.[26] Goodman-Bacon (2021) showed that the standard two-way fixed effects estimator shown in (1.6) fails to estimate the ATE if treatment effects are not constant over time. Several estimators have been proposed to deal with this difficulty. In the main results, we report estimates obtained using the imputation procedure presented in Borusyak *et al.* (2021). Other estimators yield quantitatively similar results, as shown in Appendix Figure 1.E.7.

We follow a similar approach to estimate the effect of US exposure to the Great Influenza pandemic on UK innovation. In particular, the model is entirely similar to (1.6), except that the treatment variable is defined as $\left(D_{ik,t} \equiv t - I\left[\text{Influenza Emigrants}_{ik}\right]\right)$ as it codes the number of years since the influenza, and it is interacted with a dummy variable returning value one for the pharmaceutical technology class, and zero otherwise.[27]

The primary estimation strategy in this section is thus a triple difference estimator (Olden and Møen, 2022). A causal interpretation of the resulting estimates requires that the difference between the within-group differences are not statistically different from zero. Several papers highlight that, compared to the standard difference-in-differences estimator, the parallel trends assumption in this setting is relatively weak because it only requires that no contemporaneous shock affects the relative outcome of the treatment and the control group (Gruber, 1994). Throughout the paper, we present flexible triple difference estimates to provide evidence supporting the parallel trends assumption.

1.5 Empirical Results

In this section, we present the main return innovation result. Then, we document that shocks to US innovation diffuse into the UK through migration ties. We interpret these results as evidence that migration flows contribute to the diffusion of innovative knowledge to countries sending migrants.

[26]Notice that the treatment is also potentially repeated, for the same unit can be treated multiple times. This is, however, not the case in the baseline case, where we define synthetic shocks in the top 0.1% of the overall residualized innovation shock distribution.

[27]This specification focuses on the ATE on pharmaceuticals compared to other technology classes. In Appendix Table 1.E.8, we report the double differences estimates associated with model (1.6). Then, in Figure 1.E.6b, we show that, as in the United States, the influenza had a major effect on pharmaceutical innovation only.

1.5.1 Exposure to US Innovation Shapes Innovation in the UK

The primary finding of this paper is that exposure to foreign technology through migration ties shapes the dynamics of innovation in the emigrants' country of origin.[28] We label this novel finding "return innovation". We first estimate regression (1.2) through a simple OLS linear probability model to document it. We report the results in columns (1–3) of panel A of Table 1.2. There is a positive, significant, and quantitatively large correlation between the baseline measure of exposure to foreign knowledge and the number of patents at the district-technology class level. Moreover, the correlation persists over time, as the estimates remain statistically significant after two decades. In columns (1–3) of panel B we repeat this exercise, but we normalize the number of patents by the district-level population at the beginning of the analysis sample decade (1880). We confirm the positive association between knowledge exposure and per-capita patents.

As discussed in section 1.4.2, at least two factors hinder a causal interpretation of the estimates presented in panel A. First, out-migration is not random across districts. Second, there may be some latent determinant of the settlement location decisions of emigrants that correlates with innovation activity in their origin areas. To ensure that our estimates do not reflect spurious correlation arising from omitted variable bias issues, we estimate model (1.2) using the instrument (1.31). In columns (4–6) of panels A and B, we report the reduced-form association between the instrument and the dependent variable. Figure 1.3 visually represents the same regression. We confirm the positive and statistically significant effect of knowledge exposure on innovation. The effect persists until one decade, as opposed to two from the OLS estimates. Columns (7–9) report the two-stage least-squares (TSLS) estimation results. First, the instrument is relevant.[29] Second, the TSLS estimates confirm knowledge exposure's positive, large, and statistically significant effect on innovation. The magnitude of the TSLS estimates is roughly similar to the OLS, although the latter appears to be slightly upward biased. The OLS estimates possibly reflect the upward bias introduced by assortative matching across district-county pairs.

The evidence in Table 1.2 is at the district-technology level. To explore the heterogeneity of the return innovation effect across industries, however, we estimate model (1.2) at the district level separately for each technology class. We report the resulting reduced-form coefficients

[28] A recent literature produced compelling evidence that exposure to innovation is a key determinant of subsequent innovation activity (Akcigit *et al.*, 2018; Bell *et al.*, 2019). Our results can thus be interpreted as evidence in favor of this thesis.

[29] We report the complete first-stage estimates in Appendix Table 1.E.4. The instruments are always relevant and capture a substantial share of the variation of the endogenous variables.

of the knowledge exposure instrument—one for each regression—in Figure 1.4. We estimate the largest treatment effects for industries like electricity and chemistry at the forefront of the Second Industrial Revolution (Mokyr, 1998, e.g.,). We employ the UK-revealed comparative advantage to measure the relative sector-level innovation specialization.[30] We find that the return innovation effect is larger in sectors where the UK retained an advantage at the beginning of the period (the 1880s). Rather than igniting the emergence of entirely new sectors, our results suggest that exposure to US knowledge through migration ties nurtured already-existing industries.

The setting of this study allows for gauging the persistence of the association between exposure to foreign knowledge and innovation.[31] In Appendix Figure 1.D.3, we report the coefficients of a regression between the number of patents and an interaction term between knowledge exposure in the period 1900–1930 and biennial time dummies from 1940 to 2014. The estimates suggest that the positive effect of knowledge exposure on innovation persists for almost four decades, albeit the magnitude decreases over time. Starting in the mid-1970s, the association gradually becomes small and statistically insignificant. In Appendix Table 1.D.9, we repeat the exercise by technology class and find consistent results across sectors. Migration ties thus generate enduring knowledge flows that shape innovation activity over the long run.

The analysis presented thus far focuses on how out-migration shaped the *direction* of innovation.[32] A natural question is, however, whether it also impacted the *volume* of patents. Our data are not well-suited to answer this question because we lack disaggregated data on outright emigration. Nevertheless, if emigration to countries other than the United States correlated with US emigration, we can present some suggesting evidence. In Table 1.D.6, we estimate the effect of out-migration on innovation, measured as the number of patents granted. The OLS and TSLS estimates show that out-migration has a negative short-term impact on innovation, but this reverses in the medium run (after one decade). Our findings thus appear to reconcile evidence of "brain drain", in the short term, with "brain drain" arguments (Docquier and Rapoport, 2012). The effect of out-migration on the volume of innovation has been the focus of

[30]In the international trade literature, the revealed comparative advantage is a widely-employed 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

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

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

[31]We discuss the technical details of the long-run analysis in Appendix section 1.D.3.

[32]Appendix section 1.D.4.1 explores this aspect in more detail and provides the technical details of the analysis.

many of the existing studies (Agrawal *et al.*, 2011; Andersson *et al.*, 2022). This paper, instead, provides evidence that emigration is a fundamental driver of the direction of innovation.[33] From this perspective, our results thus inform the recent literature studying the determinants of the direction of innovation (Bell *et al.*, 2019; Einiö *et al.*, 2022).

We perform several robustness exercises to gauge the robustness of our results. First, we report them in the Appendix and discuss them in section 1.E.1. Second, we consider alternative dependent variable transformations in Table 1.E.1. Third, Table 1.E.2 reports the results using five different definitions of knowledge exposure that hold fixed various margins of variation. The baseline specification of model (1.2) includes district-by-decade and technology class fixed effects. In Table 1.E.3, we show that the results are robust to alternative, demanding specifications. The standard errors are clustered at the district level in the baseline specification. In Figure 1.E.1, we adopt various estimators and confirm that they all preserve the statistical significance of the main results. The instrument used in Table 1.2 leverages variation in the connection timing to the railway network to randomize immigration across counties. In Table 1.E.5, we report the results using an alternative "leave-out" instrument, described in section 1.E.2. Importantly, we can also use both instruments simultaneously and provide over-identification tests. In Table 1.E.6, we confirm that the leave-out instrument results are robust to various alternative definitions of the county-level shocks.

1.5.2 Innovation Shocks in the US Diffuse to the UK

The return innovation result indicates that migration ties shape the direction of innovation in the origin areas of emigrants. We claim that this finding implies that fluctuations in patenting activity in the United States would reverberate in the United Kingdom through migration linkages. We estimate model (1.6) using two different sources of such fluctuations—which we label innovation shocks—to test this hypothesis.

Table 1.3 reports the results of this exercise. Columns (1–4) refer to the synthetic shocks series we construct by residualizing the observed patenting activity against fixed effects and flagging large increases in the resulting series as "innovation shocks". As a preliminary robustness test, we report the full-sample estimate in column (1), while columns (2–4) exclude districts in the top three areas in terms of patents granted. We estimate a positive, large, and statistically significant effect of US synthetic innovation shocks on innovation activity in the UK.

[33] Our results resonate with evidence by Fackler *et al.* (2020). While their study essentially leverages cross-country variation in emigration destinations, our analysis is based on within-country disaggregated data on the origin and destination of migrants. This allows us to credibly estimate the causal effect of out-migration and investigate possible underlying mechanisms.

We estimate an average of 0.4 patents per year in the treated technology class after the shock in exposed districts. This is a quantitatively sizable effect since the average number of patents per district-class pair is 1.3. Moreover, the relative size of the effect remains consistent through the regression samples. Next, we explore heterogeneous treatment effects over time in Figure 1.5a. Reassuringly, the figure provides evidence that supports the parallel trends assumption. The effect of the innovation shock is the largest and most significant after two years since the shock initially manifested in the United States. This time lag seems plausible, especially since our data shows an average of 1.1 years delay between the application and issue date at the UK patent office. The effect persists up until six years following the synthetic shock. We estimate the effect of synthetic shocks sector by sector in the appendix Figure 1.E.5. As in 1.4, we find the largest treatment effect for electricity.

Next, we investigate how exposure to the Great Influenza pandemic across US counties impacted UK innovation. The logic behind this exercise is that exposure to the pandemic fostered innovation in the pharmaceutical sector (see Table 1.D.7 and Berkes *et al.*, 2022). We thus expect districts whose emigrants had settled in counties more exposed to the pandemic to display higher patenting rates in pharmaceuticals. We report our findings in columns (5–8) of Table 1.3. We estimate the pandemic shock's effect on British innovation to be positive and sizable. On average, two patents per year are granted in the pharmaceutical sector in districts more exposed to counties severely affected by the influenza. We estimate the associated dynamic treatment effects in Figure 1.5b. We find only one marginally significant and very small coefficient in the pre-treatment period. By comparison, the post-treatment coefficients are large and highly significant. The effect of the pandemic materialized six-seven years after the shock in the United States. As noted before, this delay is partly due to the shift between patent application and issue by the patent office, except that we now have to compound delays at the US and UK offices. Moreover, the effect of the pandemic shock on US innovation in pharmaceuticals was not immediate, as shown in Appendix Figure 1.D.2. Taken together, it is plausible that the shock propagation into the UK is observed with some delay. We estimate statistically significant treatment effect coefficients for more than a decade thereafter.

The pandemic shock only impacted innovation in pharmaceuticals in the US (Figure 1.E.6a). We thus expect to retrieve a similar effect in Britain. Figure 1.E.6b shows that, although the point estimates are not as sharp as in the US case, the pharmaceutical sector is the one that benefits the most from the influenza shock. The point estimate for pharmaceuticals is nearly three times larger than the second-largest estimate. The estimated effect in some sectors may be negative because of crowding-out out of those fields into pharmaceuticals, although we cannot entirely disentangle the underlying reason. We interpret this exercise as a falsification check: Figure

1.E.6 provides convincing evidence that the pandemic shock affected the same sector in the US and the UK.

We assess the robustness of these results through several robustness checks. First, we consider alternative thresholds to (i) flag synthetic shocks and (ii) flag district exposure to synthetic shocks. Table 1.E.7. We estimate larger treatment effects for smaller thresholds. This is reasonable since smaller thresholds impute, on average, larger innovation shocks. The synthetic shock triple differences model is a staggered design since shocks generally occur in different periods across technology classes and districts. The baseline estimates are obtained from the imputation estimator developed by Borusyak *et al.* (2021). In Figure 1.E.7, the estimated treatment effect remains consistent across various estimators. In particular, the one developed by De Chaisemartin and D'Haultfœuille (2022) allows repeated treatments and yields consistent results. Finally, in Table 1.E.8, we report several specifications to gauge the robustness of the pandemic shock results. First, in columns (1–2), we report the double differences estimates that compare pharmaceutical innovation across districts by exposure to counties affected by the pandemic. Then, in columns (3–7), we report various triple differences specifications that exclude districts in areas with very high patenting activity. The results remain consistent throughout.

1.6 Potential Mechanisms and Discussion

Several concurrent, not necessarily mutually exclusive mechanisms can explain the return innovation result. In this section, we present our analysis to disentangle some. First, we establish whether return innovation is solely a consequence of return migration. Then, we discuss some complementary and possibly quantitatively more substantial channels.

1.6.1 Is Return Innovation Return Migration?

Return migration is a primary candidate to explain 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 innovation. On the one hand, several studies estimate modest effects for recruiting programs of high-skilled nationals working abroad (Ash *et al.*, 2022; Shi *et al.*, 2023). On the other, Giorcelli (2019) shows, although from a different perspective, that those exposed to (managerial) foreign knowledge change their behavior once back in their origin country.[34] In this section, we investigate whether the direct or indirect return migration effects explain the return innovation result.

As a first step, we estimate model (1.2) controlling for return knowledge exposure.[35] Table 1.D.10 reports the results: in columns (1–3), we present specifications with various levels of fixed effects; columns (4) and (5) display the coefficients of lagged values of the independent variables; in column (6) we report the full lag model. Throughout the specifications, the coefficient of knowledge exposure is always significant and substantially larger than that of return knowledge exposure, which is seldom significant. The results thus suggest that the role of return knowledge exposure is probably not a major driver of the return innovation effect.

To sharpen the focus, 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 that do *not* emigrate from the 1911 census. We then create a yearly balanced panel dataset that reports the number of patents obtained by each individual between 1900 and 1920. To do so, we leverage the linked inventor-census data described in Appendix 1.A.3. Next, each individual is geo-referenced to precise coordinates as described in Appendix 1.A.2. We complement this with information on the geographical proximity between these "stayers" and migrants. More specifically, we define a dummy variable ($\text{US Migrant}^{k}_{p,t}$) that returns value one in all periods after the first time at least one individual living within k meters from individual p migrates to the US, and zero otherwise. In the baseline analysis, we consider $k = 100$ and recast ($\text{US Migrant}^{100}_{p,t}$) as simply ($\text{US Migrant}_{p,t}$) for brevity. We label this variable an indicator of "neighborhood migration". To estimate the effect of neighborhood migration on the probability of patenting, we thus estimate the following double difference regression:[36]

$$\text{Patents}_{p,t} = \alpha_p + \alpha_t + \beta \times \text{Neighborhood Migrant}_{p,t} + \varepsilon_{p,t} \tag{1.7}$$

where p and t denote, respectively, individuals and years, and α_p and α_t are the associated fixed effects. The term β yields, under a standard parallel trends assumption, the estimated causal

[34] Choudhury (2016) shows that R&D firms with returnee managers are disproportionately more likely to file patents in the United States. Bahar *et al.* (2022b) show that return migration can influence trade.

[35] Return knowledge exposure is defined as in (1.1), except that return migrant replaces out-migration linkages. We formally describe how we define it in Appendix section 1.D.4.3.

[36] To avoid an excessive computational burden, we estimate model (1.7) on a 10% random sample of the population. Moreover, the model is a staggered difference-in-differences design with (potentially) repeated treatments. These issues arise because individuals in different neighborhoods are typically treated at different, possibly multiple, points in time. We thus estimate regression (1.7) using the estimator proposed by Borusyak *et al.* (2021). In Appendix Figure 1.E.8, we show that results hold if the neighborhood-migrant treatment is activated whenever emigrants within 100 meters from the individual in the sample migrate.

effect of neighborhood migration on innovation.

The logic beneath equation (1.7) builds on Bell *et al.* (2019), who document the importance of geographical proximity to inventors as a driver of subsequent innovation activity. A positive and significant estimate of β would be evidence against the direct return innovation effect since the sample comprises only individuals that never emigrate. Then, we define a (Non-Return Neighborhood Migrant$^k_{pt}$) dummy entirely analogous to the previous treatment, except that we condition the neighborhood emigrant to not return to the UK. In this case, a positive estimate of β would suggest that indirect return migration is also unlikely to be a primary driver of the return innovation effect.

We report the estimates of equation (1.7) in Table 1.4. The dependent variable is the yearly number of patents. In columns (1–4), the sample includes individuals from all districts; in columns (5–7), we exclude individuals in the top three-producing patents areas (London, Lancashire, and the South-West). In panel A, the treatment is activated by any US neighborhood emigrant. In panel B, we restrict to neighborhood emigrants that never return in the sample period. We estimate a positive effect of neighborhood emigration on innovation by non-migrants. The effects hold in the baseline specification (columns 1 and 5), as well as including parish-by-time fixed effects (columns 2 and 6) and applying coarsened exact matching (CEM, columns 3 and 4).[37] Importantly, the estimated coefficient remains if we restrict the sample to exclude all non-inventors, thus reducing the sample size considerably (columns 4 and 8). Panels A and B show that overall and non-return neighborhood migration has a positive statistically significant effect on the probability of inventing regardless of the dependent variable, the fixed effects, and the matching scheme. In Figure 1.E.9, we report the associated flexible difference-in-differences estimates, which indicate the absence of statistically significant pre-trends.

Evidence presented in Table 1.4 indicates that return migration is, either directly or indirectly, unlikely to be the main driver of the return migration effect. Nevertheless, we do not want to over-emphasize these results. They do not imply that return migration bears no impact on innovation activity. Instead, we interpret them as suggesting that some other mechanism that does not directly hinge on the physical return of emigrants exerts a more substantial influence on domestic innovation in the UK. We devote the rest of this section to studying these potential additional channels.

[37]Parishes are very small geographical units with a population of approximately 2,500. Coarsened exact matching weights are calculated to balance individuals in terms of age, parish of residence, and occupation. Appendix Figure 1.E.10 reports the correlation between treatment status and pre-treatment individual-level observable characteristics for the baseline sample (panel A) and the CEM weighted sample (panel B).

1.6.2 Return Innovation Through Information Diffusion

The leading alternative channel that we explore is that migration ties foster the spread of innovative knowledge through information diffusion, absent physical return. We cannot unambiguously disentangle the precise "meta-mechanism" through which information diffuses. However, we provide evidence suggesting the importance of market integration facilitated by migration ties.

1.6.2.1 The Transatlantic Telegraph Increased Innovation In Emigration Districts

We exploit one historically relevant event to provide evidence on the information diffusion channel: the first transatlantic telegraphic cable that connected the US and UK domestic networks (1866). The telegraph represented a major revolution in communication technology that ushered unprecedented market integration (Steinwender, 2018; Juhász and Steinwender, 2018). Before 1866, mail steam was the cheapest way to communicate between the UK and the US. It took fifteen days to transmit information in this way. This delay was reduced to one day overnight between June 27 and 28. The connection timing was unanticipated and exogenous (Steinwender, 2018).[38] We leverage this information friction shock to identify the effect of knowledge flows generated by migration ties on UK innovation. After introducing the telegraph cable, we thus expect to find (i) increased innovation activity in districts with more US emigrants and (ii) increased innovation activity in the fields emigrants were exposed to in the US.

To test these hypotheses, we estimate the following difference-in-differences models:

$$\text{Patents}_{i,t} = \alpha_i + \alpha_t + \sum_{h=-a}^{b} \beta^h \left[\text{US Emigrants}_i \times \mathrm{I}(t - 1866 = h)\right] + \varepsilon_{i,t} \tag{1.8a}$$

$$\text{Patents}_{ik,t} = \alpha_{i \times k} + \alpha_t + \sum_{h=-a}^{b} \beta^h \left[\text{Knowledge Exposure}_{ik} \times \mathrm{I}(t - 1866 = h)\right] + \varepsilon_{ik,t} \tag{1.8b}$$

where i, k, and t denote a district, technology class, and yea, respectively. The term (US Emigrants_i) and $(\text{Knowledge Exposure}_{ik})$ code the number of US emigrants and exposure to US knowledge.[39] Lastly, the variable $\mathrm{I}(t - 1866 = h)$ denotes the number of years since the transatlantic

[38] The project for a transatlantic telegraphic cable had been underway for a long time before 1866. Previous attempts in 1857, 1858, and 1865 all failed due to logistic and technical challenges. The 1866 attempt was thus one among many, and its success had not been anticipated.

[39] The cable was laid down in 1866. Our migration data started in 1870. To construct district-level emigration, we can only use emigrants from 1870–1875. This would be problematic if the telegraph fostered out-migration, which, by available historical accounts, was not the case.

cable was laid down. In equation (1.8a), the treatment coefficients $\{\beta^h\}$ quantify the effect of the transatlantic cable by comparing districts by the number of US emigrants; in equation (1.8b), we also leverage variation across sectors and exposure to US innovation.

We report the static versions that conflate pre- and post-treatment years in two periods in columns (1) and (4) of Table 1.5. We estimate a positive and significant effect of the transatlantic telegraph on innovation. To provide more convincing evidence on the plausibility of this result, we expect the transatlantic cable to affect innovation only in districts connected to the domestic network.[40] We thus reconstruct the entire telegraph network before the introduction of the transatlantic cable. The exact location of each station is displayed in Appendix Figure 1.A.4. We refer to districts with at least one station as "connected". In columns (2) and (5), we show the estimated effect of the telegraph on connected districts. By comparison, columns (3) and (6) report the estimates for non-connected districts. The results of this exercise are sharp. We estimate a positive effect of the transatlantic cable only for districts connected to the domestic UK network, as expected. We fail to detect any significant effect on non-connected districts.

Because the location of telegraph stations was not random, one may argue that this exercise only reflects pre-existing differences between connected and not connected districts. However, identification in this setting requires that patenting in connected and unconnected districts were on the same trend before the introduction of the telegraph and that it would not have differed had the cable not been laid down. In Figure 1.6, we thus report the flexible double-differences estimates of model (1.8a), which we estimate separately on connected and unconnected districts. We find that connected and unconnected districts were on the same trend before 1866. We estimate positive and significant treatment effects only for the former and after 1866, whereas the patenting in the latter does not respond to the shock. In 1873 and 1874, the second and third cables became operational. Our estimates suggest positive treatment effects for those.

Building on Steinwender (2018), we interpret these results as evidence that lower frictions to information diffusion enabled knowledge flows through pre-existing migration ties. In other words, return innovation manifests absent physical return of emigrants through knowledge flows ushered by migration ties. However, the precise mechanism that generates these flows is difficult to disentangle. Our favored interpretation is that the telegraph fostered the integration of the US and the UK markets and that this effect was more intense where migrant ties existed. Later, we provide more evidence in favor of the market integration effect and discuss potential additional mechanisms in section 1.6.3.

[40] We do not claim that there were no cross-district spillover effects even if districts were not connected to the domestic UK network. We nonetheless believe the effect on connected districts would arguably be more significant.

1.6.2.2 Newspaper Mentions of United States Topics in Emigration Districts

Thus far, we have restricted the focus of the analysis to information flows that pertain to innovative knowledge (patents). This section provides evidence that migration ties between the UK and the US generated more general-purpose information flows. We exploit the vast British Newspaper Archive that contains the digitized contents of thousands of historical British newspapers (for a detailed description of the data, see Appendix section 1.3.3.2 and Beach and Hanlon, 2022). Ideally, we would like to measure the intensity of US-related information flows into the United Kingdom. We tackle the absence of direct hard data by measuring how frequently US-related news appeared in historical newspapers.

We estimate three sets of regressions:

$$\text{US Mentions}_{i,t} = \alpha_i + \alpha_t + \beta_1 \times \text{US Emigrants}_{i,t} + \varepsilon_{i,t} \tag{1.9a}$$

$$\text{US State Mentions}_{is,t} = \alpha_i + \alpha_{s\times t} + \beta_2 \times \text{US Emigrants}_{i\to s,t} + \varepsilon_{is,t} \tag{1.9b}$$

$$\text{US County Mentions}_{ij,t} = \alpha_i + \alpha_{j\times t} + \beta_3 \times \text{US Emigrants}_{i\to j,t} + \varepsilon_{ij,t} \tag{1.9c}$$

where i, j, s, and t denote a UK district, a US county, a US state, and a decade, respectively. Regression (1.9a) is run at the district level and leverages the variation of the overall US emigration rate; in regressions (1.9b) (resp. (1.9c)), instead, we look at district-by-state (resp. district-by-county) migration flows. We estimate regressions (1.9) using actual out-migration and the shift-share instrument described in section 1.4.3.

Table 1.6 reports the results. Panels A, B, and C respectively display the estimated β coefficients of models (1.9a), (1.9b), and (1.9c). In columns (1–3), we report the correlation between measured out-migration flows and newspaper coverage; columns (4–5) report the OLS reduced-form association with the instrument; columns (7–9) display the two-stage least-square estimates. In columns (3), (6), and (9), we restrict the sample to districts with at least one newspaper. We find a strong and positive effect of out-migration on newspaper coverage of general-interest US-related news. Importantly, we always control for time-varying confounding factors at the level of the receiving place, whether it be the country, single states, or single counties. This ensures that the estimates do not reflect shocks in those areas.

We interpret this result as evidence that out-migration generates general—not only innovation—information flows between the areas where emigrants settle and where they originate. We cannot disentangle—-and this goes beyond the scope of this paper—the precise underlying mechanism. For example, increased coverage of US-related news may be demand-driven because the local population may demand information covering the areas where their loved ones settled. On the other hand, US emigrants could have sponsored local newspapers to cover news in the areas where they had located. In this sense, our estimates may reflect a supply-side factor. What is crucial for this paper is that, notwithstanding the precise underlying mechanism, out-migration ignites cross-country information flows. The return innovation effect is thus one of the possibly many effects of out-migration on countries sending migrants.[41]

1.6.2.3 Trade-Induced Technology Transfer

Our favored interpretation of the telegraph analysis is that market integration, fostered by migration ties, is a major driver of the return innovation result. Here we provide one additional piece of evidence to support this interpretation.[42] In particular, we focus on international trade as a measure of market integration. In 1930, the United States passed a tariff—the Smoot-Hawley Act —which sharply increased import duties and hampered trade (Eichengreen, 1986). We leverage variation in the tariff increase across technology classes in a difference-in-differences setting. We find that patenting decreases in districts more exposed to technologies that the Act comparatively more heavily targeted.

Since the tariff reform was one-sided, it is unlikely that depressed import competition or access to intermediate inputs drive this result (e.g., see Bloom *et al.*, 2016; Juhász and Steinwender, 2018; Autor *et al.*, 2020). We are unable to conclusively disentangle the impact of export opportunities (e.g., see Bustos, 2011; Atkin *et al.*, 2017) from the information access effect of migration ties (Aleksynska and Peri, 2014; Ottaviano *et al.*, 2018). However, we believe this exercise provides evidence in favor of the market integration interpretation of the knowledge flows generated by migration ties.

1.6.3 Potential Additional Mechanisms

In this section, we discuss some potential additional mechanisms that may explain the return innovation result. It is worth stressing that these may operate on top, and not instead, of the information and market integration mechanism.

[41] A recent literature has already documented the disparate effects of out-migration on attitudes towards democracy (Spilimbergo, 2009), demand for political change (Karadja and Prawitz, 2019), wages (Dustmann *et al.*, 2015), technology adoption and innovation (Andersson *et al.*, 2022; Coluccia and Spadavecchia, 2022), social norms (Tuccio and Wahba, 2018) We thus provide additional evidence of the wide-encompassing effects of out-migration on emigration areas.

[42] We discuss the literature and the technical implementation of the empirical analysis in Appendix section 1.D.1.

1.6.3.1 Temporary Migrations

When disentangling the possible mechanisms behind the return innovation effect, we contrasted those requiring physical return migration with those not. We concluded that physical return is unlikely to be a major driver of return innovation. It may nonetheless be possible that (unobserved) short-term temporary migrations influence the dynamics of innovation in the UK. We cannot observe temporary migrants because we construct migration flows from census data. Censuses are, in turn, only administered to the residing population. Our data would thus fail to reflect such temporary migration movements.

Temporary migrations would confound our estimates if such migrations were correlated with observed migration patterns. We believe that it is unlikely that this factor bears relevant quantitative implications. First, the notion of a "temporary migrant" in XIX-century transatlantic migration is unclear. Piore (1980) refers to Southern and Eastern European migrants as temporary because they planned to return to their origin countries at some point. This 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 (Dupont *et al.*, 2017). This suggests that the extent of short-term stays must have been relatively limited. Moreover, Piore (1980) notes that "temporary" migrants were relatively low-skilled and, thus, less likely to operate technology transfer.

Furthermore, our research designs largely rule the temporary migration mechanism out. First, our instrumental variable research design largely rules this mechanism out. Suppose that measured out-migration and unobserved temporary migrations were correlated across origin districts and destination counties. Our pull instrumental variable randomizes county-level immigration shocks leveraging (conditional) variation in the decade counties were connected to the railway network. While we show that the resulting instrument predicts actual out-migration, it is likely that the source of pull variation is not as active for temporary "business" migrants. Second, for temporary migration to explain the double and triple differences result, one would need such temporary flows to be correlated with the county-level innovation shocks. This seems unlikely, although we cannot directly test and rule it out.

1.6.3.2 Monetary Remittances

Along with classical "brain drain" arguments, monetary remittances have been a major subject of empirical investigations in the migration literature (Clemens, 2011). Remittances have been found to contribute only modestly to the economic development of emigration countries. This notwithstanding, it is possible that the inflow of capital through remittances may have sustained increased innovation activity, perhaps by relaxing financial constraints or access to credit (Gorodnichenko and Schnitzer, 2013). It would be more difficult, however, that it would have impacted the *direction* of innovation and, most importantly, that this effect would have been correlated with variation in knowledge exposure.

Disaggregated data on financial remittances, unfortunately, do not exist. We thus remain silent on the possibility that the documented positive effect of out-migration on innovation depends on financial remittances. This capital inflow, however, cannot explain why out-migration influences the direction of innovation unless knowledge *and* monetary remittances go hand in hand. This is a possibility that we cannot explore. It nonetheless highlights that financial and innovation remittances shape innovation in a complementary, rather than mutually exclusive, fashion.

1.6.4 Discussion

Our results bear potentially far-reaching implications for policy-makers. We show that emigration does not necessarily further underdevelopment or stagnation, as the "brain drain" literature seems to suggest (Docquier and Rapoport, 2012). Instead, out-migration can foster innovation, technology adoption, and diffusion and thus empower long-run economic growth. Instead of focusing on blocking the emigration of skilled individuals, our central recommendation to policy-makers in emigration countries would be to foster cooperation and exchanges between them and the staying population. Our results and more recent albeit narrative evidence by Saxenian (1999) suggest that this approach can yield important and lasting benefits on the economic development of emigration countries.

In 2020, for example, fifty-five Italian researchers were awarded a European Research Council (ERC) starting grant, possibly the most prestigious award for early-career scholars working in the European Union. Only nineteen (≈ 35%) of them worked in Italian institutions. This paper sheds new light on the economic contribution of the remaining thirty-six (≈ 65%) on science, innovation, and, ultimately, growth in Italy.[43] More generally, we study how the entire stock of emigrants influences the dynamics of innovation in their sending countries. In doing so, we vastly enrich our understanding of the consequences of emigration, compared to studies that focus on much narrower sub-samples of super-skilled emigrants (*e.g.*, see Prato, 2021).

Concerns over the external validity of these results are natural, given the setting we analyze. We nonetheless think that History can inform the scholarly debate and policy-making for two main reasons. First, as previously mentioned, Saxenian (1999) qualitatively documents

[43]These figures are the result of authors' calculations over data released by the ERC, available at this link.

similar return innovation effects with respect to the Taiwanese and Indian emigration to the Silicon Valley area. Second, we provide evidence that the UK emigration to the US in the XIX century largely resembles, *mutatis mutandis*, migration between European countries and the United States during the XXI. Compared to the rest of the English population, migrants were positively selected. They were similarly more likely to be employed in skilled occupations than the average native and to live in urban centers. These patterns suggest that a cautious comparison between historical and contemporary migration episodes can yield important insights for policy-makers and scholars (Abramitzky and Boustan, 2017).

1.7 Conclusions

The diffusion of innovation across countries is a major factor shaping long-run development trajectories and the economic catch-up of developing countries. In this paper, we argue that international migrations generate knowledge flows that shape the volume and direction of innovation in emigration countries. This result—which we label "return innovation"—offers a more nuanced view of emigration compared to the traditional "brain drain" hypothesis, which interprets out-migration as a depletion of the country's human capital sending migrants. Moreover, as the number of international migrants has been steadily rising over the past decades, the role of human mobility as a driver of knowledge and information diffusion across countries in a globalized world economy bears quantitatively relevant implications.

We study this question in the context of the English and Welsh mass migration to the United States between 1870 and 1940. We leverage detailed US and UK population census data and assemble a novel individual-level dataset that allows us to observe the universe of English and Welsh immigrants in the US before and after migrating. We complement this with newly digitized patent data covering the universe of patents in England and Wales. On top of these unique, high-quality data, the absence of stringent international intellectual property protection and active migration policies represents a prominent appealing feature of this historical setting compared to contemporary scenarios.

We document that migration linkages increase UK innovation activity in the technologies that emigrants are exposed to in the US. To address endogeneity concerns arising from the assortative matching of British immigrants in the US, we develop a new shift-share instrument that exploits the conditional timing of connection to the railway network to randomize emigration across counties. Additionally, we implement a double and triple differences estimator that leverages variation across counties and technology classes. Thus, we can document a causal link between exposure to foreign knowledge through migration ties and innovation activity.

Exploiting the granular nature of our data, we find that physical return migration is not the primary determinant of the return innovation effect. Instead, the results suggest that migration ties ignite the cross-border diffusion of information and novel knowledge, thus influencing the direction of technological advancement. We provide evidence consistent with market integration ushered by migration linkages being a crucial driver of the information diffusion effect. Moreover, historical newspaper coverage of US-related news indicates that migration ties foster general-purpose information flows. From this perspective, the return innovation effect represents one of the potentially disparate effects of out-migration on the origin areas of emigrants.

Even though our results may not be immediately generalizable to contemporary scenarios, the historical evidence suggests that the British mass migration to the United States may be comparable to present-day cross-border movements between developed countries. History can thus inform the scholarly literature and policymakers on the complex relationship between out-migration, innovation, and, ultimately, long-run economic growth.

Tables

TABLE 1.1: Descriptive Statistics of Selected Variables

	Observations	Mean	Std. Dev.	Min.	Max.
	(1)	(2)	(3)	(4)	(5)
Panel A. Innovation					
Total Patents	5489	225.826	826.432	1	19789
Electricity	5489	23.565	200.755	0	9430
Instruments	5489	17.69	80.2	0	1850
Personal Articles & Furniture	5489	20.136	73.812	0	1548
Ships & Aeronautics	5489	16.463	55.062	0	1152
Transportation	5489	20.024	74.479	0	1923
Panel B. Emigration					
N. of US Emigrants	3779	61.765	91.36	0.303	1073.998
N. of Return US Emigrants	2494	35.342	52.202	0.064	730
Panel C. Census Tracts					
Population (1,000s)	3773	42.165	54.973	0.092	703.559
Share of Males (%)	3773	47.645	2.586	36.112	62.686
Share of Manufacture Empl. (%)	3773	13.213	6.306	2.569	42.723
Share of Agriculture Empl. (%)	3773	14.43	6.889	1.454	32.914
Share of Transportation Empl. (%)	3773	2.578	1.272	0	13.857
Share of Liberal Professions (%)	3773	1.679	0.65	0.43	6.873
Share of Public Servants (%)	3773	0.897	1.427	0	24.498
Panel D. Individual-Level Panel					
Share of Inventors	471013	0.009	0.094	0	1
N. of Patents	471013	0.018	0.356	0	87
N. of Patents if Inventor	4210	1.993	3.205	1	87
N. of Neighborhood Emigrants	471013	13.62	43.338	0	756
N. of Non-Return Neighborhood Emigrants	471013	12.979	40.888	0	512

Notes. This table displays summary descriptive statistics for a subset of the variables in the dataset. In Panels A, B, and C, variables are observed at the district level and at a decade frequency. In Panel D, the statistics are computed for individuals observed for twenty years around the 1891 and 1911 census years. An individual is labeled an inventor if they obtain at least one patent over this period. Panel A reports statistics for the top five most frequent technological classes. In Panels B and C, the underlying data are cross-walked to 1900 district borders.

TABLE 1.2: Effect of Exposure to US Technology on Innovation in Great Britain

	Ordinary Least Squares			Reduced Form			Two-Stages Least-Squares		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Dependent variable: Number of patents									
Knowledge Exposure$_t$	1.342***			0.037***			1.224***		
	(0.143)			(0.007)			(0.195)		
Knowledge Exposure$_{t-1}$		0.909***			0.015***			0.488**	
		(0.145)			(0.005)			(0.190)	
Knowledge Exposure$_{t-2}$			0.379***			-0.012			-0.398
			(0.112)			(0.014)			(0.478)
Mean Dep. Var.	10.392	13.345	15.256	8.706	12.045	15.314	8.708	12.049	15.319
Std. Beta Coef.	0.299	0.148	0.050	0.075	0.022	-0.013	0.296	0.088	-0.053
R^2	0.772	0.701	0.737	0.800	0.799	0.736	0.035	0.011	-0.003
K-P F-stat							109.826	109.826	109.826
Panel B. Dependent Variable: Patents per capita ($\times$ 10,000)									
Knowledge Exposure$_t$	0.178***			0.004***			0.146***		
	(0.020)			(0.001)			(0.027)		
Knowledge Exposure$_{t-1}$		0.092***			0.002***			0.078***	
		(0.018)			(0.001)			(0.024)	
Knowledge Exposure$_{t-2}$			0.049***			0.000			0.001
			(0.015)			(0.001)			(0.043)
Mean Dep. Var.	2.066	2.629	2.973	1.748	2.345	2.980	1.747	2.346	2.980
Std. Beta Coef.	0.124	0.054	0.023	0.023	0.011	0.000	0.093	0.046	0.000
R^2	0.426	0.443	0.476	0.417	0.452	0.476	0.001	0.000	0.000
K-P F-stat							107.825	107.825	107.825
District-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Technology Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of District-Class	11268	11268	11268	11214	11214	11214	11214	11214	11214
N. of Observations	67549	67549	56295	56070	56070	56070	56047	56047	56047

Notes. This table displays the association between innovation and exposure to US knowledge. The unit of observation is a district-technology class pair, observed at a decade frequency between 1880 and 1939. The main explanatory variable is knowledge exposure. In Panel A, the dependent variable is the number of patents; in Panel B, the dependent variable is the number of patents normalized by district-level population in 1880 and multiplied by 10,000 for readability. In columns (1–3), we estimate the OLS correlation with the observed measure of knowledge exposure; in columns (4–6), we estimate the reduced-form association with the railway-based instrument of knowledge exposure through OLS; columns (7–9) report the two-stage least-squares estimate. Each model includes district-by-decade and district-by-technology class fixed effects. Standard errors are reported in parentheses and are clustered at the district level.

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

TABLE 1.3: Triple Differences Effect of Exposure to US Shocks on UK Innovation

	Synthetic Shocks				Great Influenza Pandemic Shock			
	(1) Full Sample	(2) No London	(3) No Lancs	(4) No S/W	(5) Full Sample	(6) No London	(7) No Lancs	(8) No S/W
Synth. Shock × Post × Emigrants	0.434***	0.277***	0.578***	0.420***				
	(0.121)	(0.082)	(0.125)	(0.127)				
Pharma × Post × Emigrants					0.613***	0.417***	0.678***	0.461***
					(0.164)	(0.140)	(0.172)	(0.156)
District-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-by-Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Units	10029	9697	8760	9547	10727	10217	9384	10047
Number of Observations	393046	382153	343450	375850	429080	408680	375360	401880
Mean Dep. Var.	1.361	1.029	1.263	1.276	0.725	0.532	0.682	0.668

Notes. This table displays the effect of US innovation shocks on innovation activity in the UK. The unit of observation is a district-technology class pair observed at a yearly frequency between 1900 and 1939. The dependent variable is the number of patents. In columns (1–4), the independent variable is an indicator that, for a given district–technology, returns value one after a synthetic innovation shock in that technology class is observed in at least one county where the district has above-average out-migration. A synthetic innovation shock is observed whenever the residualized number of patents observed in the country is in the top 0.5% of the overall distribution. In columns (5–8), the independent variable is an indicator that returns value one for pharmaceutical patents only and only if emigration from the observed district to counties in the top quartile of the influenza mortality distribution is in the top quartile across districts. Both models should thus be interpreted as triple-difference designs. Since models in columns (1–4) are staggered designs, we estimate them using the imputation estimator developed by Borusyak *et al.* (2021). In columns (2) and (6), we drop districts in the London area; in columns (3) and (7), we exclude districts in the Lancashire area; in columns (4) and (8), we drop districts in the South-West area. Excluded regions are the first three in terms of patents granted. All models include district-by-year, district-by-technology class, and technology class-by-year fixed effects; standard errors, clustered two-way by district and technology class, are shown in parentheses.

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

TABLE 1.4: Double Differences Effect of Neighborhood Out-Migration on Innovation

	Baseline Sample				Dropping Individuals in...		
	(1)	(2)	(3)	(4)	(5) London	(6) Lancashire	(7) South-West
Panel A. All Emigrants							
Neighborhood Emigrant × Post	0.167***	0.180***	0.170***	16.208***	0.130**	0.180***	0.198***
	(0.053)	(0.055)	(0.056)	(5.758)	(0.061)	(0.055)	(0.061)
Std. Beta Coef.	0.022	0.024	0.023	0.211	0.018	0.025	0.025
Panel B. Only Non-Return Emigrants							
Non-Return Neighborhood Emigrant × Post	0.165***	0.189***	0.167***	15.749***	0.108*	0.183***	0.211***
	(0.054)	(0.056)	(0.057)	(5.987)	(0.060)	(0.056)	(0.062)
Std. Beta Coef.	0.021	0.024	0.021	0.196	0.014	0.024	0.026
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	–	Yes	Yes	Yes	Yes	Yes
Parish × Year FE	No	Yes	No	No	No	No	No
Matching	No	No	Yes	No	No	No	No
Sample	Full	Full	Full	Inventors	Full	Full	Full
N. of Individuals	473112	473112	469585	4224	410327	422230	352064
N. of Observations	9462240	9419787	9391700	84480	8206540	8444600	7041280
Mean Dep. Var.	0.890	0.892	0.893	99.716	0.794	0.836	0.893
S.D. Dep. Var.	40.291	40.324	40.351	414.695	37.439	39.126	41.333

Notes. This table reports the effect of neighborhood out-migration on innovation. The units of observation are individuals who are observed yearly between 1900 and 1920. In columns (1–3) and (5–7), the sample consists of the universe of males who did not emigrate over the period and that were at least 18 years old in 1900; in columns (4) and (8), we restrict the sample to inventors. The dependent variable is the number of patents obtained annually. In columns (1–4), the sample consists of individuals residing in all England and Wales divisions; in columns (5–7), we exclude the top tree-patents producing areas: London, Lancashire, and the South-West. In Panel A, the independent variable is an indicator that, for a given individual, returns value one after at least one person that was living in the same neighborhood as the individual migrates to the United States; in Panel B, we restrict to emigrants that never return in the period of observation. In this context, "neighborhood" refers to the same street, square, or similar. We explore an alternative distance-based definition in Appendix Table 1.E.9. Each model includes individual and—at least—year fixed effects; in column (2), we include parish-by-year fixed effects; in column (3), individuals are weighted by their coarsened exact matching weight. The estimates are obtained using the method discussed in Borusyak *et al.* (2021) to account for the staggered roll-out of the treatment across individuals. Standard errors, clustered at the district level, are reported in parentheses.

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

TABLE 1.5: Double Differences Effect of Transatlantic Telegraph on Innovation

	Double Differences			Triple Differences		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Connect	N. Connect	All	Connect	N. Connect
Post × Emigrants	1.345***	1.639***	-0.083			
	(0.451)	(0.559)	(0.097)			
Post × Knowledge Exposure				0.027**	0.027**	-0.003
				(0.010)	(0.011)	(0.005)
District FE	Yes	Yes	Yes	–	–	–
Class FE	Yes	Yes	Yes	–	–	–
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District-Class FE	–	–	–	Yes	Yes	Yes
N. of District-Class	631	463	168	631	463	168
N. of Observations	10096	7408	2688	181728	133344	48384
R^2	0.918	0.917	0.817	0.790	0.794	0.498
Mean Dep. Var.	5.610	7.241	1.115	0.312	0.402	0.062
Std. Beta Coef.	0.114	0.125	-0.039	0.035	0.034	-0.007

Notes. This table displays the estimated effect of the connection of the US and UK telegraph lines on innovation in the UK. The units of observation are districts in columns (1–3) and district-technology class pairs in columns (4–6). Units are observed yearly between 1860 and 1875. The dependent variable is the total number of patents granted. In columns (1–3), the independent variable is an interaction between the—time-invariant—number of US emigrants in the 1870s and an indicator variable that returns value one after the transatlantic cable successfully connected the UK and the US in 1866, and zero otherwise; in columns (4–6) the treatment interacts—time-invariant—knowledge exposure in the 1870s with the same posttreatment indicator. In columns (1) and (4), the sample includes all districts; in columns (2) and (5) (resp. 3 and 6), we restrict the estimation to districts that were (resp. were not) connected to the domestic UK telegraph system. Models (3) and (6) should be interpreted as placebo exercises. Regressions include fixed effects for district and year in columns (1–3) and district-by-class and year in columns (4–5). Standard errors, clustered at the district level, are reported in parentheses.

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

TABLE 1.6: Effect of US Emigration on Newspaper Coverage of US-related News

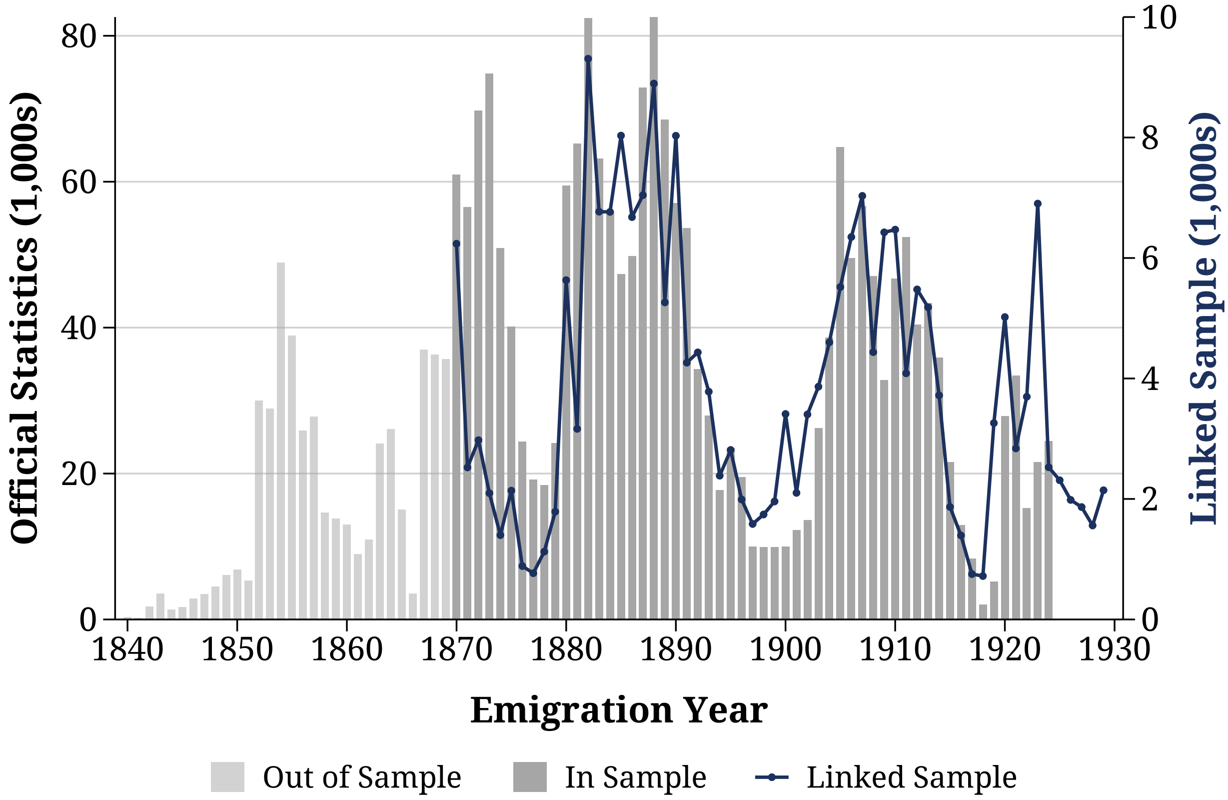
	Dependent Variable: Number of Newspaper Mentions								
	OLS			Reduced Form			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. US-Wide Coverage									
US Emigrants	6.753***	6.632***	7.207***				24.396***	24.228***	25.061***
	(0.958)	(1.006)	(0.600)				(1.570)	(1.611)	(0.912)
US $\widehat{\text{Emigrants}}$				1.451***	1.440***	1.501***			
				(0.121)	(0.124)	(0.078)			
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var.	1017.635	1017.635	2276.989	1017.635	1017.635	2276.989	1017.635	1017.635	2276.989
Panel B. State-Wide Coverage									
US Emigrants	1.050***	1.049***	1.103***				10.060***	10.061***	10.091***
	(0.095)	(0.096)	(0.052)				(0.428)	(0.430)	(0.458)
US $\widehat{\text{Emigrants}}$				0.038***	0.038***	0.039***			
				(0.001)	(0.001)	(0.001)			
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var.	57.486	57.486	127.984	64.672	64.672	136.660	64.672	64.672	136.660
Panel C. County-Wide Coverage									
US Emigrants	1.120***	1.120***	1.217***				4.861***	4.863***	5.130***
	(0.148)	(0.148)	(0.079)				(0.471)	(0.460)	(0.291)
US $\widehat{\text{Emigrants}}$				0.055***	0.055***	0.058***			
				(0.006)	(0.005)	(0.004)			
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var.	0.003	0.003	0.007	0.004	0.004	0.008	0.004	0.004	0.008
N. of Newspapers	No	Yes	No	No	Yes	No	No	Yes	No
Districts in Sample	All	All	w/News	All	All	w/News	All	All	w/News
N. of Districts	602	602	321	602	602	321	602	602	321

Notes. This table displays the effect of out-migration on newspaper coverage of emigrants' destinations. The observation unit is: in Panel A, a district; in Panel B, a district-US state pair; in Panel C, a district-US county pair. Units are observed at a decade frequency between 1880 and 1930. The dependent variable is the number of articles mentioned: in Panel A, "United States"; in Panel B, US states; in Panel C, US counties. The independent variable is: in Panel A, the number of US emigrants; in Panel B, the district-state emigrants; in Panel C, the district-county emigrants. Models (1–3) estimate the model through OLS; models (4–5) report the reduced-form association between mentions and the out-migration instrument; models (7–9) report the two-stage least squares estimates. Regressions include district fixed effects and: in Panel A, decade fixed effects; in Panel B: state-by-decade fixed effects; in Panel C: county-by-decade fixed effects. Standard errors, clustered at the district level, are reported in parentheses.

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

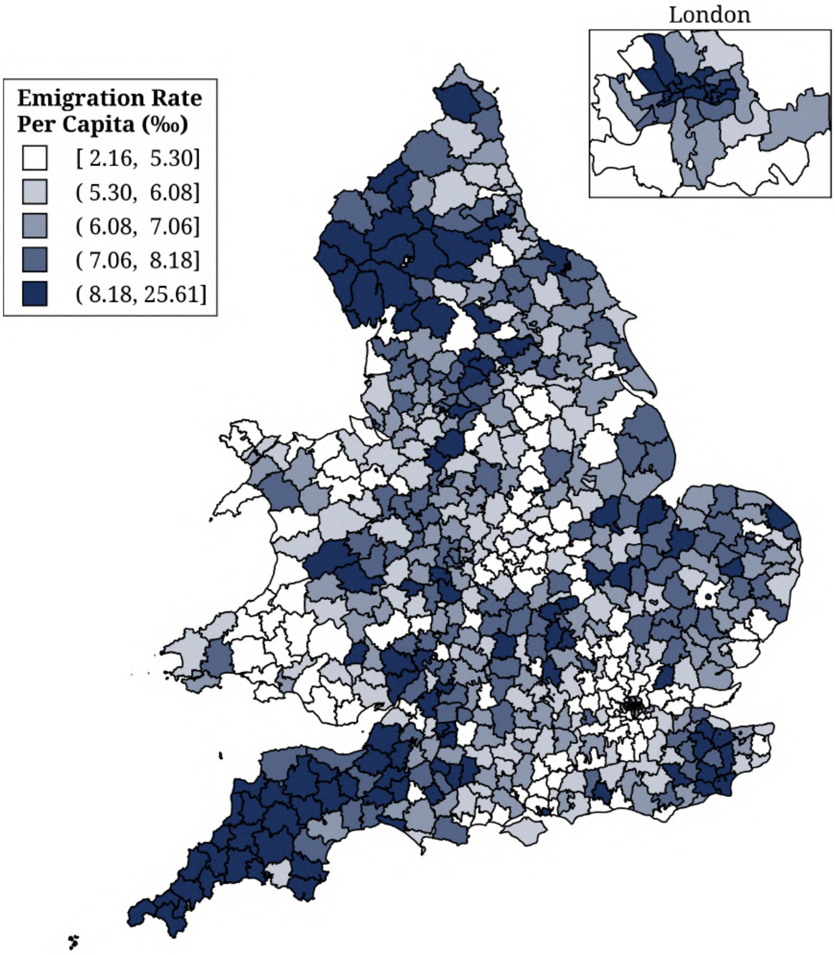
Figures

FIGURE 1.1: Immigrants in the US from the UK and Linked US-UK Migrants, 1840–1930



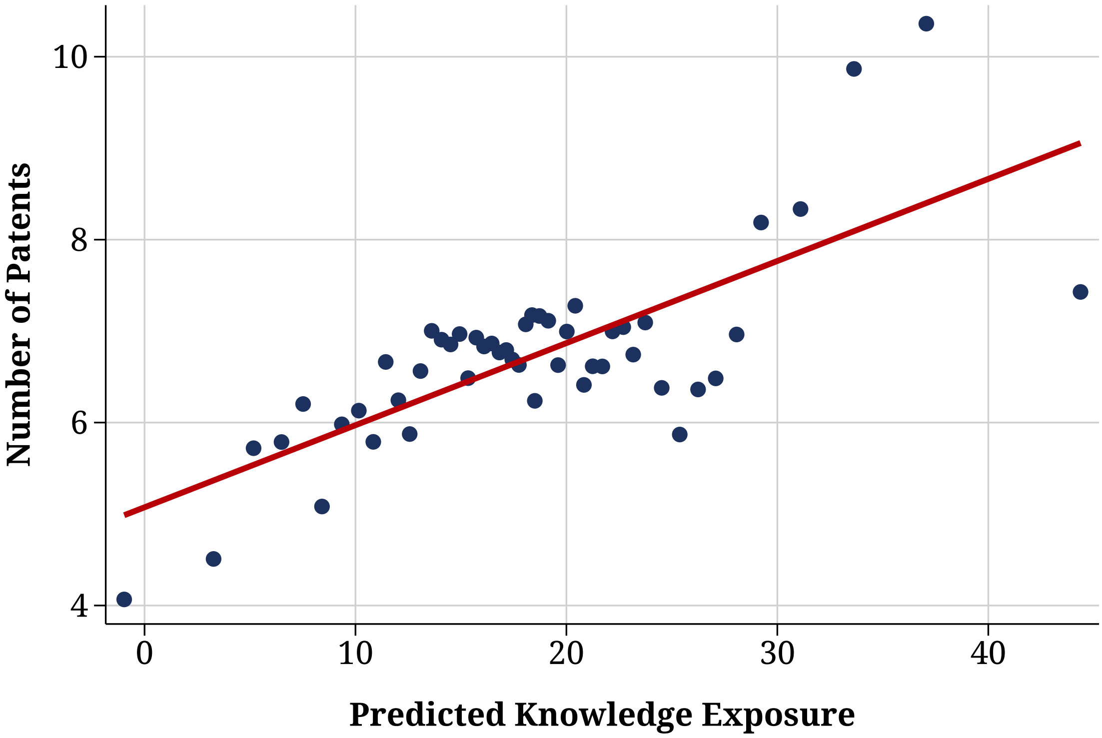
Notes. This figure compares the total number of English and Welsh immigrants in the United States as recorded in official statistics from Willcox (1928) with the linked emigrants' sample developed in this paper. The light gray bars display the total inflow of English and Welsh immigrants in the US over the period 1840–1870, i.e., out of the period we study. The darker gray bars report the same figure for 1870–1924. 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. Figures are in thousand units.

FIGURE 1.2: Spatial Distribution of US Migrants Across British Districts



Notes. This figure reports the spatial distribution of emigrants across English and Welsh districts over the period 1870–1930. Data are from the matched emigrants' sample. The total number of emigrants over the period is normalized by district population in 1900 and is reported in ‰ units. Districts are displayed at 1900 historical borders, and the emigrant population is cross-walked to consistent borders as described in 1.A.1. Lighter to darker blues indicate higher emigration rates.

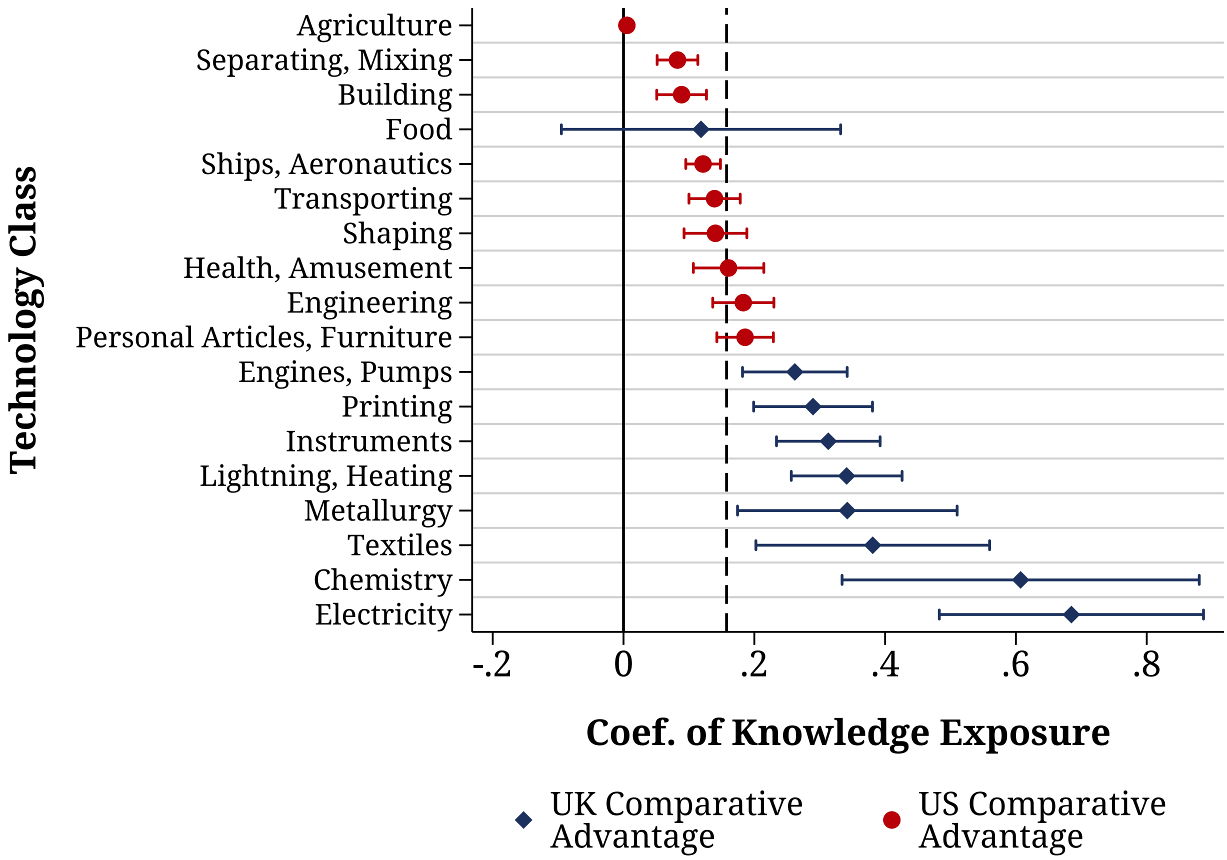
FIGURE 1.3: Reduced-Form Effect of Knowledge Exposure on Innovation



Notes. Coefficient = 0.037 (Clust. Std. Err. = 0.007). R^2 = 0.731.

Notes. This figure is a binned scatter plot of the reduced-form effect of instrumented knowledge exposure on innovation. The unit of observation is a district-technology class observed at a yearly frequency between 1880 and 1939. The graph partials out district-by-decade and district-by-technology class fixed effects. We report in note the regression coefficient along with its standard error, clustered at the district level, and the R^2.

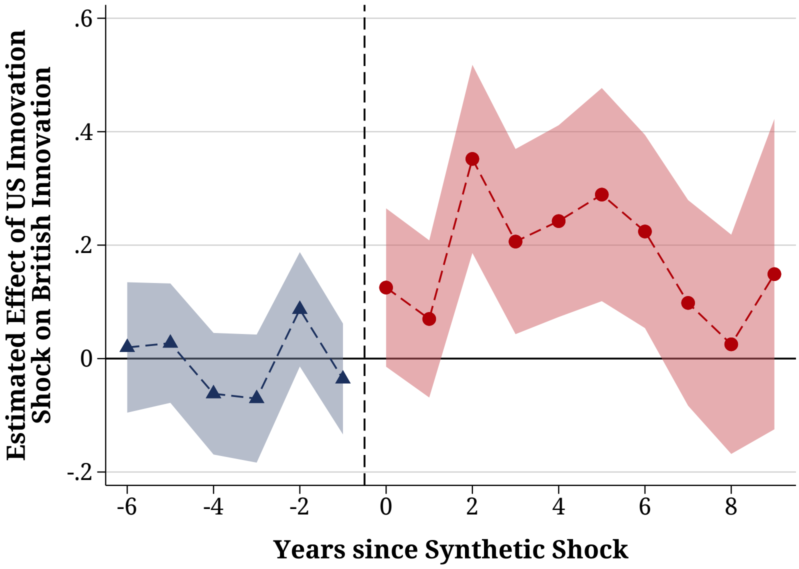
FIGURE 1.4: Heterogeneous Effects of Return Innovation Across Technology Classes



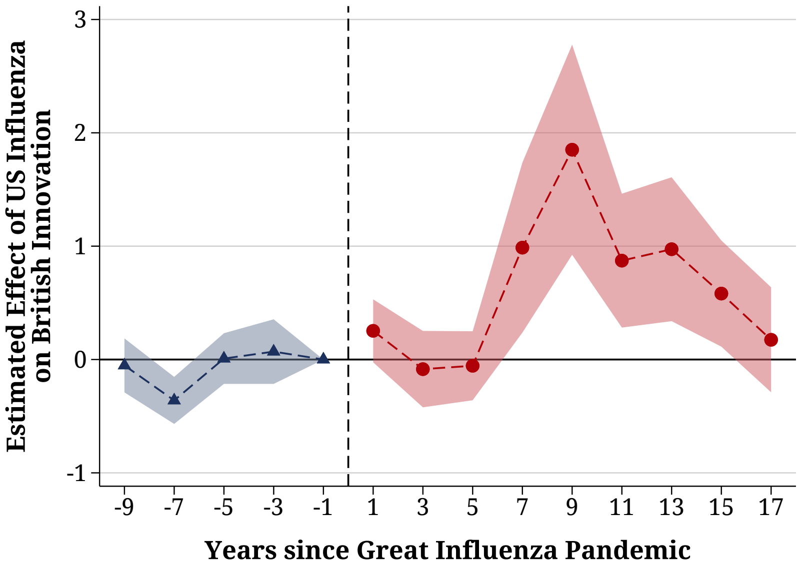
Notes. This figure reports the reduced-form effect of instrumented knowledge exposure on innovation by technology class. Each dot reports the coefficient of a regression between the total number of patents and the instrument for knowledge exposure in a given technology class. The unit of observation in each regression is a district, observed at a decade frequency between 1880 and 1939. Regressions include district and decade fixed effects. Bands report 95% confidence intervals. Standard errors are clustered at the district level. The dashed black line reports the average unconditional association between the instrument and the patents across classes. Blue (resp. red) dots display the regression coefficients for the UK (resp US) revealed comparative advantage sectors.

FIGURE 1.5: Flexible Triple Differences Effect of UK Exposure to US Shocks

(A) Synthetic Shocks

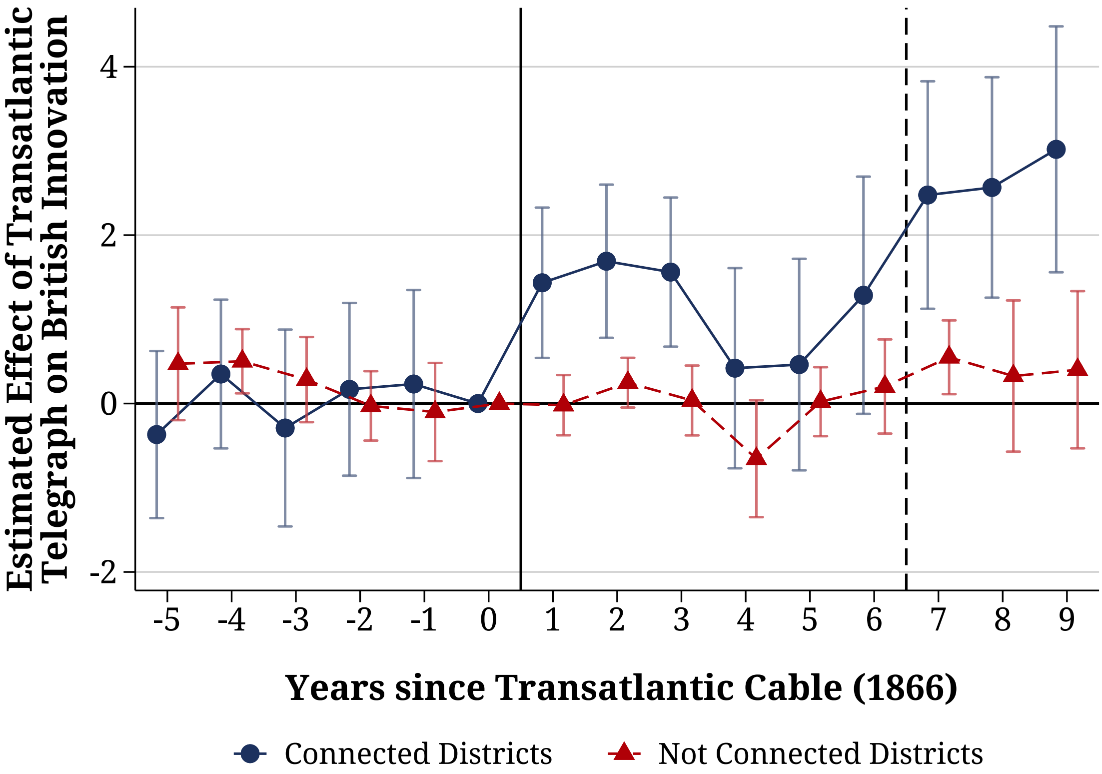


(B) Great Influenza Pandemic Shock



Notes. These figures report the dynamic treatment effects of synthetic shocks (Panel 1.5a) and the Great Influenza Pandemic shock (Panel 1.5b) on innovation. The units of observation are district-technology class pairs; units are observed at a yearly frequency in Panel 1.5a and at a biennial frequency in Panel 1.5b between 1900 and 1939. The dependent variable is the number of patents. The treatment is an indicator equal to one if: in Panel 1.5a, a synthetic shock is observed in a given technology in at least one county where the district has above-median out-migration; in Panel 1.5b, for pharmaceutical patents, emigration from a given district to counties in the top quartile of the mortality distribution is in the top quartile across districts. The black dashed line indicates the timing of the treatment. Standard errors are two-way clustered by district and technology class; bands report 95% confidence intervals.

FIGURE 1.6: Flexible Double Differences Effect of Transatlantic Telegraph on Innovation



Notes. The figure displays the estimated dynamic treatment effect of the connection of the US and UK telegraph lines on innovation in the UK. The units of observation are districts observed at a yearly frequency between 1860 and 1875. The dependent variable is the number of patents. The independent variable is an interaction between the—time-invariant—number of emigrants in the 1870s and a posttreatment indicator that equals one after the transatlantic telegraph cable. Blue dots report the dynamic treatment effects on the sample of districts connected to the domestic UK telegraph network in 1862; red dots report those for the districts not connected to the network. The black solid vertical bar indicates the year the first cable was laid down (1866); the dashed black vertical line flags the year when the second and third cables were laid (1873-1874). Regressions include district and year fixed effects. Standard errors are clustered at the district level; bands report 95% confidence bands.

Bibliography

ABRAMITZKY, R. and L. BOUSTAN (2017). "Immigration in American Economic History." *Journal of Economic Literature*, 55(4): 1311–45.

ABRAMITZKY, R., L. BOUSTAN and K. ERIKSSON (2020). "Do Immigrants Assimilate More Slowly Today Than in the Past?" *American Economic Review: Insights*, 2(1): 125–41.

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.

ACEMOGLU, D. (2002). "Directed Technical Change." *The Review of Economic Studies*, 69(4): 781–809.

——— (2010). "When Does Labor Scarcity Encourage Innovation?" *Journal of Political Economy*, 118(6): 1037–1078.

AGHION, P., A. DECHEZLEPRÊTRE, D. HEMOUS, R. MARTIN and J. VAN REENEN (2016). "Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry." *Journal of Political Economy*, 124(1): 1–51.

AGRAWAL, A., D. KAPUR, J. MCHALE and A. OETTL (2011). "Brain Drain or Brain Bank? The Impact of Skilled Emigration on Poor-Country Innovation." *Journal of Urban Economics*, 69(1): 43–55.

AKCIGIT, U., S. CAICEDO, E. MIGUELEZ, S. STANTCHEVA and V. STERZI (2018). "Dancing With the Stars: Innovation Through Interactions." *NBER Working Paper*, (No. w24466).

AKCIGIT, U., J. GRIGSBY and T. NICHOLAS (2017). "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age." *NBER Working Paper*, (No. w23047).

ALEKSYNSKA, M. and G. PERI (2014). "Isolating the Network Effect of Immigrants on Trade." *The World Economy*, 37(3): 434–455.

ALVAREZ, F. E., F. J. BUERA, R. E. LUCAS *et al.* (2013). "Idea Flows, Economic Growth, and Trade." *NBER Working Paper*, (No. w19667).

ANDERSSON, D., M. KARADJA and E. PRAWITZ (2022). "Mass Migration and Technological Change." *Journal of the European Economic Association*, 20(5): 1859–1896.

ANDERSSON, D. E., M. KARADJA and E. PRAWITZ (2018). "Mass Migration, Cheap Labor, and Innovation." *Journal of Economic History*, 78(2): 623–623.

ARKOLAKIS, C., S. K. LEE and M. PETERS (2020). "European Immigrants and the United States' Rise to the Technological Frontier." *Working Paper*.

ASH, E., D. CAI, M. DRAKA and S. LIU (2022). "Bootstrapping Science? The Impact of a "Return Human Capital" Programme on Chinese Research Productivity." *Working Paper*.

ATKIN, D., A. K. KHANDELWAL and A. OSMAN (2017). "Exporting and Firm Performance: Evidence from a Randomized Experiment." *The Quarterly Journal of Economics*, 132(2): 551–615.

AUTOR, D., D. DORN, G. H. HANSON, G. PISANO and P. SHU (2020). "Foreign Competition and Domestic Innovation: Evidence from US Patents." *American Economic Review: Insights*, 2(3): 357–74.

AZOULAY, P., B. F. JONES, J. D. KIM and J. MIRANDA (2022). "Immigration and Entrepreneurship in the United States." *American Economic Review: Insights*, 4(1): 71–88.

BAHAR, D., P. CHOUDHURY, J. SAPPENFIELD and S. SIGNORELLI (2022a). "Human Mobility and the Globalization of Knowledge Production: Causal Evidence from Multinational Enterprises." *Working Paper*.

BAHAR, D., A. HAUPTMANN, C. ÖZGÜZEL and H. RAPOPORT (2019). "Migration and Knowledge Diffusion: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia." *The Review of Economics and Statistics*, pp. 1–50.

——— (2022b). "Migration and Knowledge Diffusion: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia." *The Review of Economics and Statistics*, pp. 1–50.

BAHAR, D., R. HAUSMANN and C. A. HIDALGO (2014). "Neighbors and the Evolution of the Comparative Advantage of Nations: Evidence of International Knowledge Diffusion?" *Journal of International Economics*, 92(1): 111–123.

BAILEY, M. J., C. COLE, M. HENDERSON and C. MASSEY (2020). "How Well Do Automated Linking Methods Perform? Lessons from US Historical Data." *Journal of Economic Literature*, 58(4): 997–1044.

BAINES, D. (2002). *Migration in a mature economy: emigration and internal migration in England and Wales 1861-1900*. Cambridge University Press.

BALASSA, B. (1965). "Trade Liberalisation and "Revealed" Comparative Advantage." *The Manchester School*, 33(2): 99–123.

BATISTA, C. and P. C. VICENTE (2011). "Do Migrants Improve Governance at Home? Evidence from a Voting Experiment." *The World Bank Economic Review*, 25(1): 77–104.

BEACH, B. and W. W. HANLON (2022). "Historical Newspaper Data: A Researcher's Guide and Toolkit."

BEINE, M., F. DOCQUIER and M. SCHIFF (2013). "International Migration, Transfer of Norms and Home Country Fertility." *Canadian Journal of Economics/Revue canadienne d'économique*, 46(4): 1406–1430.

BELL, A., R. CHETTY, X. JARAVEL, N. PETKOVA and J. VAN REENEN (2019). "Who Becomes an Inventor in America? The Importance of Exposure to Innovation." *The Quarterly Journal of Economics*, 134(2): 647–713.

BENHABIB, J., J. PERLA and C. TONETTI (2021). "Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier." *Econometrica*, 89(5): 2261–2301.

BENHABIB, J. and M. M. SPIEGEL (2005). "Human Capital and Technology Diffusion." *Handbook of Economic Growth*, 1: 935–966.

BERGEAUD, A. and C. VERLUISE (2022). "A new Dataset to Study a Century of Innovation in Europe and in the US." *Working Paper*.

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 (2022). "Religiosity and Science: An Oxymoron? Evidence from the Great Influenza Pandemic." *Working Paper*.

BERNSTEIN, S., R. DIAMOND, A. JIRANAPHAWIBOON, T. MCQUADE and B. POUSADA (2022). "The Contribution of High-Skilled Immigrants to Innovation in the United States." *NBER Working Paper*, (No. w30797).

BERTHOFF, R. (1953). *British Immigrants in Industrial America 1790-1950*. Harvard University Press.

BERTOLI, S. and F. MARCHETTA (2015). "Bringing It All Back Home–Return Migration and Fertility Choices." *World Development*, 65: 27–40.

BLOOM, N., M. DRACA and J. VAN REENEN (2016). "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity." *The Review of Economic Studies*, 83(1): 87–117.

BORUSYAK, K., P. HULL and X. JARAVEL (2022). "Quasi-Experimental Shift-Share Research Designs." *The Review of Economic Studies*, 89(1): 181–213.

BORUSYAK, K., X. JARAVEL and J. SPIESS (2021). "Revisiting Event Study Designs: Robust and Efficient Estimation." *Working Paper*.

BOTTOMLEY, S. (2014). *The British Patent System during the Industrial Revolution 1700–1852: From Privilege to Property*. Cambridge University Press.

BRYAN, K. A. and J. LEMUS (2017). "The Direction of Innovation." *Journal of Economic Theory*, 172: 247–272.

BRYAN, K. A., J. LEMUS and G. MARSHALL (2022). "R&D Competition and the Direction of Innovation." *International Journal of Industrial Organization*, 82: 102841.

BUERA, F. J. and E. OBERFIELD (2020). "The Global Diffusion of Ideas." *Econometrica*, 88(1): 83–114.

BURCHARDI, K. B., T. CHANEY, T. A. HASSAN, L. TARQUINIO and S. J. TERRY (2020). "Immigration, Innovation, and Growth." *NBER Working Paper*, (No. w27075).

BUSTOS, P. (2011). "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms." *American Economic Review*, 101(1): 304–40.

CALLAWAY, B., A. GOODMAN-BACON and P. H. SANT'ANNA (2021). "Difference-in-differences with a continuous treatment."

CHAUVET, L. and M. MERCIER (2014). "Do Return Migrants Transfer Political Norms to Their Origin Country? Evidence from Mali." *Journal of Comparative Economics*, 42(3): 630–651.

CHEN, J. and J. ROTH (2022). "Log-like? ATEs Defined with Zero Outcomes are (Arbitrarily) Scale-dependent." *Working Paper*.

CHOUDHURY, P. (2016). "Return Migration and Geography of Innovation in MNEs: A Natural Experiment of Knowledge Production by Local Workers Reporting to Return Migrants." *Journal of Economic Geography*, 16(3): 585–610.

CLEMENS, M. A. (2011). "Economics and Emigration: Trillion-dollar Bills on the Sidewalk?" *Journal of Economic perspectives*, 25(3): 83–106.

COLUCCIA, D. M. and L. SPADAVECCHIA (2022). "The Economic Effects of Immigration Restriction Policies: Evidence from the Italian Migration to the United States." *Working Paper*.

COULTER, M. (1991). *Property in Ideas: The Patent Question in Mid-Victorian England.* Kirksville (MO): Thomas Jefferson Press.

DE CHAISEMARTIN, C. and X. D'HAULTFŒUILLE (2022). "Difference-in-Differences Estimators of Intertemporal Treatment Effects." *NBER Working Paper*, (No. w29873).

DIODATO, D., A. MORRISON and S. PETRALIA (2022). "Migration and Invention in the Age of Mass Migration." *Journal of Economic Geography*, 22(2): 477–498.

DOCQUIER, F. and H. RAPOPORT (2012). "Globalization, Brain Drain, and Development." *Journal of Economic Literature*, 50(3): 681–730.

DUPONT, B., D. KEELING and T. WEISS (2017). "First Cabin Fares from New York to the British Isles, 1826–1914." In "Research in Economic History," Emerald Publishing Limited.

DUSTMANN, C., T. FRATTINI and A. ROSSO (2015). "The Effect of Emigration from Poland on Polish wages." *The Scandinavian Journal of Economics*, 117(2): 522–564.

DUSTMANN, C. and J.-S. GÖRLACH (2016). "The Economics of Temporary Migrations." *Journal of Economic Literature*, 54(1): 98–136.

DUTTON, H. I. (1984). *The Patent System and Inventive Activity during the Industrial Revolution, 1750-1852*. Manchester (UK): Manchester University Press.

EATON, J. and S. KORTUM (1999). "International Technology Diffusion: Theory and Measurement." *International Economic Review*, 40(3): 537–570.

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).

EICHENGREEN, B. (1986). “The Political Economy of the Smoot-Hawley Tariff.” *NBER Working Paper*, (No. w2001).

EINIÖ, E., J. FENG and X. JARAVEL (2022). “Social Push and the Direction of Innovation.” *Working Paper*.

ERICKSON, C. (1957). *American industry and the European immigrant, 1860–1885*. Cambridge (MA): Harvard University Press.

——— (1972). *Who were the English and Scots emigrants in the 1880s?*, pp. 87–125. Arnold.

FACKLER, T. A., Y. GIESING and N. LAURENTSYEVA (2020). “Knowledge Remittances: Does Emigration Foster Innovation?” *Research Policy*, 49(9): 103863.

FISCHER, D. H. (1989). *Albion’s seed: Four British folkways in America*. Oxford University Press.

FURER, H. B. (1972). *The British in America: 1578-1970*. Oceana Publications.

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

GERSCHENKRON, A. (1962). “Economic Backwardness in Historical Perspective.” *Cambridge University Press*.

GIBSON, J. and D. MCKENZIE (2011). “Eight Questions About Brain Drain.” *Journal of Economic Perspectives*, 25(3): 107–28.

GIORCELLI, M. (2019). “The Long-Term effects of Management and Technology Transfers.” *American Economic Review*, 109(1): 121–52.

GIULIANO, P. and M. TABELLINI (2020). “The Seeds of Ideology: Historical Immigration and Political Preferences in the United States.” *NBER Working Paper*, (No. w27238).

GOLDIN, C. (1994). “The Political Economy of Immigration Restriction in the United States, 1890 to 1921.” In “The Regulated Economy: A Historical Approach to Political Economy,” pp. 223–258. University of Chicago Press.

GOMME, A. A. (1948). *Patents of invention: origin and growth of the patent system in Britain*. British council.

GOODMAN-BACON, A. (2021). “Difference-in-Differences with Variation in Treatment Timing.” *Journal of Econometrics*, 225(2): 254–277.

GORODNICHENKO, Y. and M. SCHNITZER (2013). “Financial Constraints and Innovation: Why Poor Countries Don’t Catch Up.” *Journal of the European Economic Association*, 11(5): 1115–1152.

GRILICHES, Z. (1998). “Patent Statistics as Economic Indicators: A Survey.” In “R&D and productivity: the econometric evidence,” pp. 287–343. University of Chicago Press.

GROSS, D. P. and B. N. SAMPAT (2022). “Crisis Innovation Policy From World War II to COVID-19.” *Entrepreneurship and Innovation Policy and the Economy*, 1(1): 135–181.

GRUBER, J. (1994). “The Incidence of Mandated Maternity Benefits.” *American Economic Review*, 84(3): 622–641.

Habakkuk, H. J. (1962). *American and British Technology in the Nineteenth Century: the Search for Labour Saving Inventions*. Cambridge (MA): Cambridge University Press.

Hanlon, W. W. (2016). "British Patent Technology Classification Database: 1855-1882." Unpublished [data collection].

——— (2018). "Skilled Immigrants and American Industrialization: Lessons from Newport News Shipyard." *Business History Review*, 92(4): 605–632.

Higham, J. (1955). *Strangers in the land: Patterns of American nativism, 1860-1925*. Rutgers University Press.

Hopenhayn, H. and F. Squintani (2021). "On the Direction of Innovation." *Journal of Political Economy*, 129(7): 1991–2022.

Jaffe, A. B., M. Trajtenberg and R. Henderson (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *The Quarterly Journal of Economics*, 108(3): 577–598.

Juhász, R. and C. Steinwender (2018). "Spinning the Web: The Impact of ICT on Trade in Intermediates and Technology Diffusion." *NBER Working Paper*, (No. w24590).

Kapur, D. (2014). "Political Effects of International Migration." *Annual Review of Political Science*, 17: 479–502.

Karadja, M. and E. Prawitz (2019). "Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States." *Journal of Political Economy*, 127(4): 1864–1925.

Kerr, S. P. and W. Kerr (2020). "Immigrant Entrepreneurship in America: Evidence from the Survey of Business Owners 2007 & 2012." *Research Policy*, 49(3): 103918.

Kerr, W. R. (2008). "Ethnic Scientific Communities and International Technology Diffusion." *The Review of Economics and Statistics*, 90(3): 518–537.

Khan, B. Z. (2020). *Inventing Ideas: Patents, Prizes, and the Knoweldge Economy*. Oxford University Press, USA.

Khan, B. Z. and K. L. Sokoloff (2004). *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.

Leak, H. and T. Priday (1933). "Migration from and to the united kingdom." *Journal of the Royal Statistical Society*, 96(2): 183–239.

Macleod, C. (1988). *Inventing the Industrial Revolution*. Cambridge (UK): Cambridge University Press.

Mokyr, J. (1998). "The Second Industrial Revolution, 1870-1914." In V. Castronovo, ed., "Storia dell'Economia Mondiale," pp. 219–245. Rome (Italy): Laterza.

Moscona, J. (2021). "Environmental Catastrophe and the Direction of Invention: Evidence from the American Dust Bowl." *Working Paper*.

MOSCONA, J. and K. A. SASTRY (2022). "Does Directed Innovation Mitigate Climate Damage? Evidence from US Agriculture." *The Quarterly Journal of Economics*.

MOSER, P. (2012). "Innovation without patents: Evidence from World's Fairs." *The Journal of Law and Economics*, 55(1): 43–74.

——— (2019). "Patents and Innovation in Economic History." In "Research Handbook on the Economics of Intellectual Property Law," pp. 462–481. Edward Elgar Publishing.

MOSER, P., S. PARSA and S. SAN (2020). "Immigration, Science, and Invention. Evidence from the Quota Acts." *Working Paper*.

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

NICHOLAS, T. (2011). "Cheaper patents." *Research Policy*, 40(2): 325–339.

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.

OLDEN, A. and J. MØEN (2022). "The Triple Difference Estimator." *The Econometrics Journal*, 25(3): 531–553.

OTTAVIANO, G. I. P., G. PERI and G. C. WRIGHT (2018). "Immigration, Trade and Productivity in Services: Evidence from UK Firms." *Journal of International Economics*, 112: 88–108.

OTTINGER, S. (2020). "Immigrants, Industries, and Path Dependence." *Working Paper*.

OTTINGER, S. and L. ROSENBERGER (2023). "The American Origin of the French Revolution." *Working Paper*.

PAULY, S., F. STIPANICIC *et al.* (2021). "The Creation and Diffusion of Knowledge: Evidence from the Jet Age."

PENROSE, E. (1951). *The Economics of the International Patent System*. Baltimore (MD): Johns Hopkins University Press.

PERLA, J., C. TONETTI and M. E. WAUGH (2021). "Equilibrium Technology Diffusion, Trade, and Growth." *American Economic Review*, 111(1): 73–128.

PIORE, M. J. (1980). *Birds of Passage: Migrant Labor and Industrial Societies*. Cambridge (UK): Cambridge University Press.

PRATO, M. (2021). "The Global Race for Talent: Brain Drain, Knowledge Transfer, and Economic Growth." *Working Paper*.

ROSENBERG, N. (1982). *Inside the Black Box: Technology and Economics*. New York: Cambridge University Press.

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*.

SAXENIAN, A. (1999). *Silicon Valley's New Immigrant Entrepreneurs*. Public Policy Institute of California.

SCHURER, K. and E. HIGGS (2020). "Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access." [data collection] Second Edition, UKDS.

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

SHI, D., W. LIU and Y. WANG (2023). "Has China's Young Thousand Talents program been successful in recruiting and nurturing top-caliber scientists?" *Science*, 379(6627): 62–65.

SOKOLOFF, K. L. and B. Z. KHAN (1990). "The Democratization of Invention During Early Industrialization: Evidence from the United States, 1790–1846." *The Journal of Economic History*, 50(2): 363–378.

SPILIMBERGO, A. (2009). "Democracy and Foreign Education." *American economic review*, 99(1): 528–43.

STEINWENDER, C. (2018). "Real Effects of Information Frictions: When the States and the Kingdom Became United." *American Economic Review*, 108(3): 657–96.

THISTLETHWAITE, F. (1958). "The Atlantic Migration of the Pottery Industry." *The Economic History Review*, 11(2): 264–278.

THOMAS, B. (1954). *Migration and economic growth. A study of Great Britain and the Atlantic Economy*. Cambridge (MA): NIESR and Cambridge University Press.

TUCCIO, M. and J. WAHBA (2018). "Return Migration and the Transfer of Gender Norms: Evidence from the Middle East." *Journal of Comparative Economics*, 46(4): 1006–1029.

VAN PATTEN, D. (2023). "International Diffusion of Technology: Accounting for Heterogeneous Learning Abilities." *Working Paper*.

WILLCOX, W. F. (1928). *International Migrations, Volume I: Statistics*. Cambridge (MA): National Bureau of Economic Research.

Appendix

1.A Data Sources and Methods

This section describes the sources and methods we adopted to assemble and merge the various datasets that underlie the empirical analysis. We defer a more detailed discussion of the novel patent data that we digitize, and the linked international migrants sample to sections 1.B and 1.C, respectively.

1.A.1 Summary of Data Sources

1.A.1.1 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 multiple 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 (2022). Data before 1900 are not available. In section 1.B, we describe how we digitize the universe of patent documents issued over the period 1853–1900 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.

1.A.1.2 Migration Data

Disaggregated data on the origin of English and Welsh immigrants—and, more generally, of all other nationalities—do not exist. These we collected neither by receiving US authorities nor by sending UK offices. 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 1.C. Ideally, we observe the universe of British emigrants to the United States between 1870 and 1930. For those individuals, we know all 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, however, we can only construct return migration flows over the period 1870–1910.

1.A.1.3 Population Census

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 completely digitized.[44] 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, as detailed in section 1.A.2.

1.A.1.4 Newspapers

We collect data on newspaper coverage of US-related news from the British Newspaper Archive.[45] Beach and Hanlon (2022) describe this dataset in detail. In this paper, we run a set of three queries. First, we search for the words "United States". Second, we perform fifty searches, one for each state. Finally, we perform approximately three thousand searches, one for each county. Each search spans the period 1850–1939. We collect the information at the article level. For each

[44] The 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 that were after 1921 are subject to privacy restrictions.

[45] A limited free-tier access to newspaper data is available at the following `link`.

entry in the database, we know the journal, day, month, and year of publication, whether it is an article or some other type of content—e.g., an obituary–, the page, and the word count. Importantly, we collect information on the universe of newspapers in the archive. Journal-level data contain the publishing address at the city level, the first and last day, month, and year of activity, and the publication frequency—e.g., quarterly, daily. We then geocode each newspaper to the coordinates of the city where it was published and map those to 1891 registration districts. We can thus construct a measure of newspaper coverage at the district-year level.[46] In Table 1.A.1, we provide a set of summary statistics on the resulting dataset. We collect information for a total of 2022 newspapers: of these, 1459 are based in England, and 93 are published in Wales. We exclude Scottish and Irish newspapers from the analysis. The average life of a newspaper in this period is 40 years. In Panel B, we report district-level statistics by decade. The number of newspapers decreases over time, as noted by Beach and Hanlon (2022), from an average of 2.3 newspapers per district in the 1870s to 0.7 in the 1930s. It is unclear whether this is due to incomplete coverage in the later period. In Panel C, we report the district-level statistics by division and find that, except for the London division, newspapers appear to be quite sparse across the country. Figure 1.A.3 displays the spatial distribution of the number of newspapers across districts over the period and confirms the impression that newspapers tend to evenly cover a substantial share of districts. London stands as a major outlier: we thus perform all exercises dropping London districts and find consistent results.

1.A.1.5 Miscellaneous

To construct the domestic UK telegraph network prior to the first transatlantic UK–US cable (1866), we digitize the list of telegraph stations reported in the *Zeitschrift des Deutsch-Österreichischen Telegraphen-Vereins, Jahrgang*, volume IX, 1862. This directory lists the universe of telegraph stations outside of London in 1862. To the best of our knowledge, it is the most complete directory prior to the introduction of the transatlantic cable. We geo-code each station to precise coordinates. The red dots in Figure 1.A.3 report each station. We then label each district with at least one telegraph station as "connected" to the domestic network and as "not connected" otherwise.

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.

[46]Unfortunately, for newspapers based in London, we only know their city, i.e., London. In the newspaper analysis, we are thus forced to conflate all urban London districts into a single "London" geographical unit.

To compute the railway-based instrument, we construct US-county level immigration shocks following the methodology described in Sequeira *et al.* (2020). We use the same data sources. Hence we defer the interested reader to their paper for a more detailed discussion.

We digitize import and export yearly data from the 1935 edition of the *Statistical Abstract of the United States*.[47] In particular, we collect the yearly tariff rate applied between 1925 and 1929, i.e., before the Smoot-Hawley Act, and between 1930 and 1935, i.e., after the Act. Tariff rates are available by sector. We then map each industry to a technology class, as listed in Table 1.A.3. In the baseline analysis, we consider an industry protected if its tariff rate increases by more than 50% between 1925-1929 and 1930-1935. We consider alternative thresholds as robustness checks.

1.A.1.6 GIS Shapefiles & Boundary Harmonization

Patents and telegraph stations are mapped to 1900 registration district borders using historical GIS files and their coordinates.[48] However, all data from the population censuses appears at historical borders. Registration districts do not undergo major boundary changes over the period that we study. However, we adapt the method presented by Eckert *et al.* (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.

1.A.2 Geo-referenced Census Records

A notable feature of the UK census is that it contains precise information on the residential address of the universe of 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 section, we discuss the methodology

[47] This publication is freely available at the following link.

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

that we apply to assign geographical coordinates to textual addresses. This dataset expands earlier work by Lan and Longley (2019), who adopt a different strategy and only analyze the 1901 census, whereas we geo-reference the entire 1851-1911 censuses. Furthermore, the geo-coded census sample is used in the individual-level analysis only. All other exercises do not rely on these data.

1.A.2.1 Methodology

There are two ways to geo-reference historical addresses. One approach is to manually digitize 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 implement the latter approach Lan and Longley (2019) exploit open-source address data from `OpenStreetMaps`. In this paper, instead, we take advantage of the commercial geo-referenced database developed by `MapTiler AG`. This has three key benefits compared to `OpenStreetMaps`-powered engines. First, the data has some historical "depth", meaning that historical names of locations are sometimes recorded. Second, `MapTiler AG` provides a flexible address-correction engine that matches the query to the closest address available in their dataset. Finally, this commercial database has better coverage than `OpenStreetMaps` in rural areas.

To perform the actual matching, we perform a preliminary simple 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.

1.A.2.2 Matching Performance

In Figure 1.A.2, we report the distribution of the share of geo-referenced addresses by district and census decade. The blue bars refer to the simple matching rate, defined as the share of geo-referenced addresses. The black-contoured bars, instead, adjust for the number of residents recorded in each address. In each figure, we report the average matching rates and their respective standard deviations. The average matching rate ranges between 76% in 1851 and 86% in 1911. All distributions display substantial right-skewness, meaning there are very few districts with a matching rate lower than 50%. The matching rate increases over time for two reasons. First, the quality of recorded addresses increases in more recent censuses. Second, the urban geography in 1911 is more similar to that in the `MapTiler AG` database than in 1851. This is due to street re-labeling and urban agglomerates' growth and consolidation. Figure 1.A.1 displays the spatial distribution of the average geo-referencing rates across censuses. Figure 1.A.1a reports the crude rate, whereas Figure 1.A.1b, we adjust by address-population. Except for Wales and some rural districts at the center of England, the geo-referencing rates are above 80% everywhere. It is particularly high—above 90%—in North-Western and South-Eastern England. More urbanized areas generally tend to feature larger geo-referencing rates because addresses tend to be more informative. This notwithstanding, differences are quantitatively small as the matching rate is remarkably homogeneous across registration districts. Wales is the single most relevant exception. The geo-referencing rate there is very low because addresses in the census until 1901-1911 tend to be reported in Welsh, especially in Western areas.

Taken together, the results of the geo-referencing algorithms are satisfactory. More than 80% addresses are successfully matched to precise geographical coordinates. This ratio is even higher in areas outside Wales, where innovation and migration activity are more extensive.

1.A.3 Linked Inventor Sample

This section presents the methodology we use to link patents to census records. The linked inventors-census sample is used in the individual-level analysis only. All other exercises do not rely on these data.

1.A.3.1 Methodology

We follow the logic of Berkes (2018), who links patents to census records in the US. We link patents between 1881 and 1899 to the 1891 census and those between 1901 and 1920 to the 1911 census. Relative to our baseline sample, we thus drop patents issued after 1920 because we cannot observe individuals born after 1911. While this is unlikely an issue for patents granted until 1930, it may induce some selection of linked inventors for later patents. Patent data contain the name and surname of inventors, their residence, and the issue year.

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 $\mathcal{L}_p^{\text{district}}$ and $\mathcal{L}_p^{\text{county}}$ be the set of, respectively, 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:

$$\mathcal{M}_{A_p}^{\text{parish}} = \left\{ i \in \mathcal{L}_p^{\text{parish}} \middle| t_p - \text{year}_i \geq 18 \right\} \tag{1.10}$$

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:

$$\text{Similarity}_i^{A_p} = \alpha \times \text{Name Similarity}_i^{A_p} + (1 - \alpha) \times \text{Surname Similarity}_i^{A_p} \tag{1.11}$$

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:

$$\overline{\mathcal{M}}_{A_p}^{\text{parish}} = \left\{ i \in \mathcal{M}_{A_p}^{\text{parish}} \middle| \text{Similarity}_i^{A_p} = \max_{i' \in \mathcal{M}_{A_p}^{\text{parish}}} \text{Similarity}_{i'}^{A_p} \right\} \tag{1.12}$$

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 $\text{Similarity}^{A_p} < \tau$, $\overline{\mathcal{M}}_{A_p}^{\text{parish}} = \emptyset$, 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 it is empty, repeat steps 2–4 conditioning on records in $\mathcal{L}_p^{\text{district}}$. If $\overline{\mathcal{M}}_{A_p}^{\text{district}}$ is empty, repeat steps 2–4 conditioning on records in $\mathcal{L}_p^{\text{county}}$. If $\overline{\mathcal{M}}_{A_p}^{\text{county}}$ is empty, repeat steps 2–4 without imposing geographical conditions on records i. In the baseline setting, we only accept county-level and country-level matches if the name and surname of the match(es) exactly match A_p's.

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 1911 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. Parishes are small, as their average population is less than 10,000. When a match at the parish level is feasible, it is usually unique. We then progressively expand the set of records by coarsening their geographic location. Districts are larger than parishes, and counties are, in turn, larger than districts. If we cannot find one match at the county level, we look for one within the entire population of England and Wales. Unlike the migrants sample, we do not have information on the birth year. To ensure that county- and country-level matches are reliable, we require that their name and surname are verbatim those recorded in the patent document.

1.A.3.2 Matching Statistics

In Figure 1.A.5, we report the matching rate of this exercise. We focus on two matching rates: a gross rate is the share of inventors that have at least one match, relative to the overall set of inventors; a net rate is the share of inventors with at least one *acceptable* match, relative to the overall group of inventors. In the analysis, a match is acceptable if (i) the similarity between name and surname is above 0.95 and (ii) a given inventor has no more than five matches. Panel 1.A.5a reports both margins over time. The gross matching rate remains consistently above 80 throughout the period. The net matching rate, however, rejects approximately 20% of the matches. This is mainly due to inventors linked to more than five census records. This notwithstanding, the share of acceptable matches is approximately constant and above 60% each year. Our algorithm delivers satisfactory performance compared to standard linking rates in the literature. In panel 1.A.5b, we break down the number of matches by geographical unit where the match is attained. Blue, red, green, and yellow bars report the matching rates at the parish, district, county, and national levels. The share of inventors matched with more than 20 census records is larger at the national level; there, we look for possible matches with no information on the residence. Multiple matches are somewhat common at the parish level as well. This is because we first try to match inventors at the parish level. Hence parish matches represent the large majority of the linked sample, while district-level matches are residual and, thus, more accurate. Figure 1.A.6 displays the spatial distribution of inventors, who are plotted using the geo-coded census coordinates described in the previous section.

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. While these hypotheses are ultimately challenging to test, in Table 1.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 age; in Panel B, we list the set of occupational categories; in Panel C, we list the residence divisions. We find no clear association between the number of matches and these variables in the overall sample (column 1) and across matches selected by geographical layer (columns 2–5). Overall, we interpret the Table as conveying reassuring evidence that the selection of inventors into the linked sample does not appear to favor particular groups systematically.

1.A.4 Historical Examples of Technology Transfer

Personal biographies of British emigrants to the United States often reveal instructive examples of return innovation. In this section, we briefly present three examples whose personal stories we recovered from historical newspapers.[49] The biographies of Marsden and Hughes, discussed respectively in sections 1.A.4.1 and 1.A.4.2, attest to how experience and exposure to novel knowledge in the US shaped the lives of emigrants. As emigrants returned, they retained these newly acquired skills and put them to productive use in Britain. In section 1.A.4.3, instead, we mention the story of Wellstood and Smith, who convey a vivid example of how overseas migrants operated technology transfer.

1.A.4.1 Henry R. Marsden and the Metal-Working Industry in Leeds

Henry Rowland Marsden was born in Leeds to poor parents in 1823 (Curtis, 1875). At age twenty-five, 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. In 1855 he developed a "stone-crusher", which is still used today and bears his name. In 1862 Marsden and his family returned to Leeds, where he set up a flourishing business centered around his newly patented invention and several other patents he had subsequently obtained in England. A wealthy man, respected for his philanthropic endeavors, he was elected mayor of Leeds in 1873 with the Liberals. He died in 1878 and is credited as one of the most prominent figures in the industrial development of Leeds.

[49] Unless otherwise specified, biographical information was collected from original newspapers; hence we omit the source.

1.A.4.2 David E. Hughes: The Inventor of the Microphone

Born in 1831, either in London or in Wales, David Edward Hughes was one of the most distinguished inventors of the Victorian Age (Encyclopædia Britannica, 2018). At age seven, his family emigrated to Kentucky. A musical talent from a young age, he eventually obtained a professorship in music at St. Joseph's College in Bardstown, Kentucky. In 1855, inspired by the mechanics of pianos, Hughes designed and patented the first printing telegraph. The printing telegraph essentially integrates a piano-style keyboard onto a standard telegraph, dramatically increasing the transmission speed. In 1857 Hughes moved to London to market his invention. Within a few years, the printing telegraph became the standard in Europe, and the then-small Western Union promptly brought it to commercial success. In later years, Hughes invented the microphone and may have detected radio waves as early as ten years before Heinrich Hertz.

1.A.4.3 Stephen Wellstood & James Smith: Migrants as Agents of Technology Transfer Overseas

Compared to Marsden and Hughes, the story of Stephen Wellstood and John Smith, at first glance, pales. It nonetheless highlights 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.

1.A.5 Tables

TABLE 1.A.1: Descriptive Statistics on Newspapers and Newspaper Coverage in the UK

	(1) Mean	(2) Std. Dev.	(3) Min.	(4) Max.	(5) Observations
Panel A. Journal-Level Statistics					
Number of Issues	2795.843	4959.740	1	46163	2022
First Publication Year	1869.746	44.171	1699	1996	2094
Last Publication Year	1910.692	49.470	1699	2009	2094
Publication Lifespan	40.946	40.490	0	273	2094
Publication Lifespan if English	40.993	41.921	0	273	1459
Publication Lifespan if Welsh	38.161	36.920	0	178	93
Publication Lifespan if Scottish	45.144	41.107	0	251	229
Publication Lifespan if Irish	41.336	34.809	0	170	241
Panel B. District-Level Statistics, by Decade					
1870s	2.309	14.860	0	285	637
1880s	1.885	11.610	0	233	636
1890s	1.494	8.587	0	160	634
1900s	1.166	5.893	0	114	634
1910s	0.942	3.845	0	83	633
1920s	0.809	2.381	0	50	633
1930s	0.714	1.274	0	24	633
Panel C. District-Level Statistics, by Division					
East	1.631	1.272	1	8	111
East Midlands	2.349	2.409	1	14	43
London	18.767	97.312	1	534	30
North East	2.079	1.761	1	8	38
North West	3.600	3.477	1	17	40
South East	1.800	1.271	1	6	100
South West	1.747	1.382	1	8	79
Wales	2.327	2.391	1	10	52
West Midlands	2.342	2.722	1	18	79
Yorkshire	2.186	2.201	1	10	59

Notes. This table reports descriptive statistics on newspapers active in the UK between 1850 and 1940. In Panel A, figures are computed at the newspaper level; Panel B computes district-level statistics on the number of newspapers by decade; Panel C computes district-level statistics on the number of newspapers by division. Panels B and C only restrict the observation sample to English and Welsh districts. Newspapers were geo-coded to their publishing address and assigned to districts based on their borders in 1900.

TABLE 1.A.2: Correlation Between Inventors' Characteristics and Number of Matches

	Overall Sample	Parish Matches	District Matches	County Matches	Nationwide Matches
	(1)	(2)	(3)	(4)	(5)
Panel A. Demographics					
Age	0.005	-0.018	-0.005	-0.014*	0.026***
	(0.010)	(0.021)	(0.012)	(0.007)	(0.005)
Dependent Variable – Dummy = 1 if Matched Inventor is in:					
Panel B. Occupation					
Agriculture	0.123*	0.272**	0.073***	0.000	0.028**
	(0.073)	(0.129)	(0.019)	(0.016)	(0.011)
Chemicals	-0.010***	-0.019***	-0.012*	-0.011***	-0.005***
	(0.004)	(0.005)	(0.006)	(0.004)	(0.001)
Construction	-0.015	-0.032	0.006	0.010	-0.001
	(0.016)	(0.031)	(0.008)	(0.017)	(0.002)
Engineering	-0.018	-0.042*	-0.017**	-0.007	0.004
	(0.012)	(0.025)	(0.007)	(0.007)	(0.003)
Liberal Professions	-0.014***	-0.016***	-0.003	-0.018***	-0.019***
	(0.005)	(0.006)	(0.010)	(0.002)	(0.005)
Metallurgy	-0.020	-0.031	0.018	-0.015	0.004**
	(0.013)	(0.019)	(0.019)	(0.019)	(0.002)
Other Manufacturing	-0.024*	-0.042**	-0.027	0.005	-0.007**
	(0.012)	(0.018)	(0.016)	(0.007)	(0.003)
Public Administration	-0.009	-0.017	-0.014	-0.015	-0.008**
	(0.008)	(0.014)	(0.010)	(0.011)	(0.003)
Textiles	-0.013	-0.042*	0.002	0.058***	0.003
	(0.012)	(0.026)	(0.024)	(0.013)	(0.005)
Trade	-0.031***	-0.044***	-0.038***	-0.021*	-0.025***
	(0.011)	(0.016)	(0.007)	(0.011)	(0.005)
Transport	-0.008	-0.025	0.001	0.006	0.003
	(0.014)	(0.026)	(0.008)	(0.011)	(0.003)
Utilities	-0.013***	-0.020***	-0.031***	-0.016***	-0.007*
	(0.004)	(0.005)	(0.005)	(0.006)	(0.004)
Panel C. Division of Residence					
East	-0.004	-0.059	-0.061	-0.074	-0.010
	(0.015)	(0.067)	(0.065)	(0.089)	(0.011)
East Midlands	0.004	-0.061	0.070	-0.055	0.012
	(0.012)	(0.070)	(0.102)	(0.066)	(0.013)
London	-0.048	-0.039	-0.053	-0.008	-0.025
	(0.069)	(0.173)	(0.053)	(0.079)	(0.022)
North East	0.028	-0.045	-0.041	-0.060	0.016
	(0.031)	(0.052)	(0.049)	(0.072)	(0.016)
North West	-0.057	-0.165	-0.069	0.195***	0.012
	(0.050)	(0.167)	(0.080)	(0.056)	(0.013)
South East	-0.024	-0.050	-0.106	-0.118	-0.028
	(0.030)	(0.057)	(0.106)	(0.136)	(0.027)
South West	0.001	-0.025	-0.049	-0.051	-0.019
	(0.008)	(0.029)	(0.054)	(0.062)	(0.020)
Wales	0.233	0.469**	0.540***	-0.026	0.046
	(0.187)	(0.212)	(0.137)	(0.032)	(0.046)
West Midlands	-0.049	-0.102	-0.104	-0.130	0.004
	(0.050)	(0.112)	(0.103)	(0.148)	(0.007)
Yorkshire	0.005	-0.021	-0.029	0.008	-0.003
	(0.013)	(0.025)	(0.032)	(0.021)	(0.007)

Notes. This table reports the correlation between inventor-level variables observed in the UK census and the number of matches in the linked sample. In column (1), the sample is the entire linked dataset. We restrict to matches at the parish (column 2), district (column 3), county (column 4), and national level (column 5). The Table reports standardized beta coefficients for comparability. Regressions include decade fixed effects. Standard errors are clustered at the division level and are reported in parentheses.

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

TABLE 1.A.3: List of Industries By Tariff Rate, 1925–1935

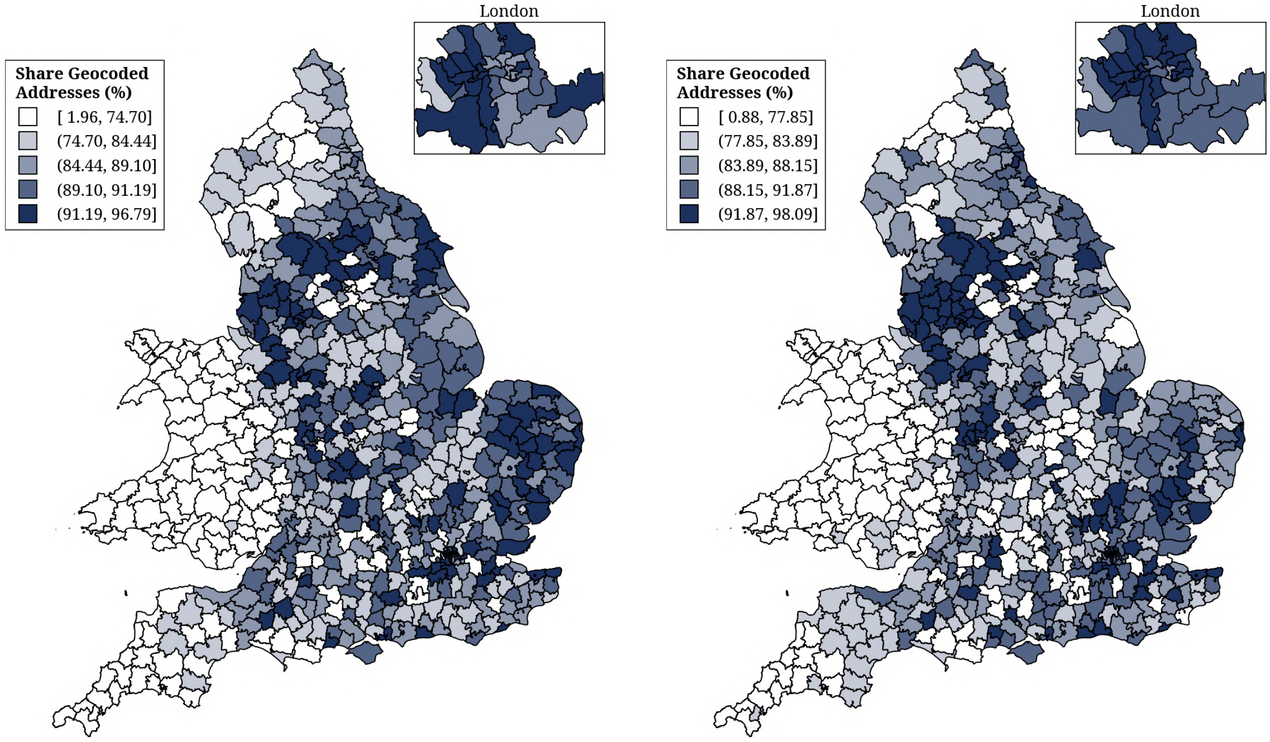
Sector	Technology Class	Tariff Rate Before S-H	Tariff Rate After S-H	Change in Tariff	Treated
(1)	(2)	(3)	(4)	(5)	(6)
Agricultural products and provisions	Agriculture	23.059	40.204	74.352	Yes
Chemicals, oils, and paints	Chemistry	29.577	40.195	35.900	No
Cotton Manufactures	Textiles	34.876	44.764	28.352	No
Earths, earthenware, and glassware	Personal Articles, Furniture	47.321	53.049	12.106	No
Flax, hemp, and jute, and manufacture thereof	Textiles	18.948	26.104	37.766	No
Metals, and manufacture thereof	Metallurgy	34.534	36.803	6.572	No
Pulp, paper, and books	Printing	25.652	25.591	-0.239	No
Silk and silk goods	Textiles	55.768	58.115	4.208	No
Spirits, wines, and other beverages	Food	37.298	59.007	58.226	Yes
Sugar, molasses, and manufactures thereof	Food	68.971	110.022	59.519	Yes
Sundries	Personal Articles, Furniture	38.149	36.587	-4.096	No
Tobacco, and manufactures thereof	Agriculture	58.176	81.636	40.326	No
Wood, and manufactures thereof	Building	23.727	20.672	-12.875	No
Wool, and manufactures thereof	Textiles	49.344	78.255	58.591	Yes

Notes. This table reports the US tariff rate applied to the categories listed in the *Statistical Abstracts of the United States*. Column (1) reports the listed sector; column (2) maps the sector to technology classes in our baseline taxonomy; columns (3) and (4) report the tariff rate applied, respectively, before and after the Smoot-Hawley Act (1930). Tariff rates before the Act are averages in the five years before the reform (1925–1929); tariff rates after the Act are averages in the five years posterior to the reform (1930–1935). Column (5) computes the change in the tariff rates. In column (6), we list the technology classes we considered targeted by the Act, namely, those whose tariff rate increase exceeded 50%. Data are digitized from the 1935 *Statistical Abstracts of the United States*.

1.A.6 Figures

FIGURE 1.A.1: Spatial Distribution of the Share of Geo-coded Addresses, 1851–1911

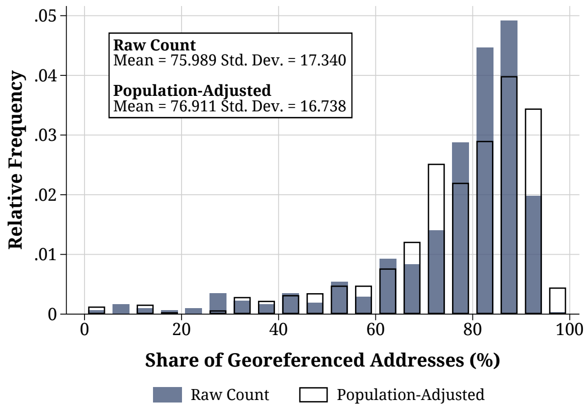
(A) Unweighted Geo-coding Ratio (B) Population-Weighted Geo-coding Ratio



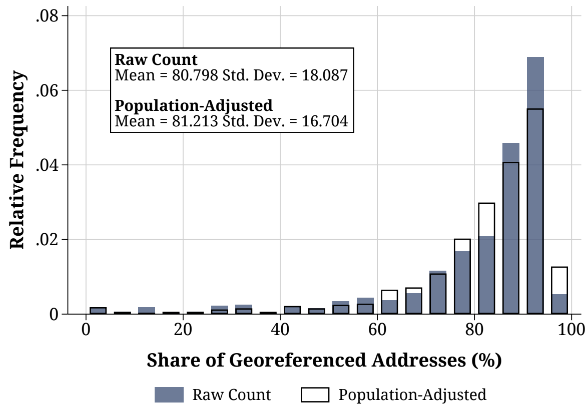
Notes. These figures report the spatial distribution of the share of geo-referenced addresses from the UK censuses, 1851–1911. For each census, we obtain a list of more than five million addresses by fine geographical unit (i.e., parishes). We then geo-reference these addresses to precise geographical coordinates. Panel 1.A.1a reports the district-level share of successfully geocoded addresses. In Panel 1.A.1b, we weigh each address by the number of people reported to live in that address. The performance of the geo-referencing algorithm is relatively poor in Wales because addresses there are often reported in Welsh.

Figure 1.A.2: Distribution of the Share of Geo-coded Addresses by Census

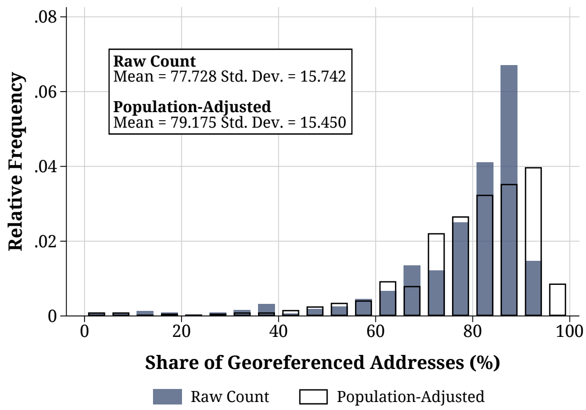
(A) 1851 Census



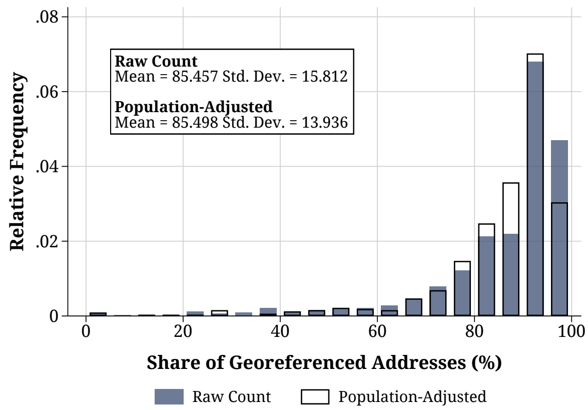
(B) 1861 Census



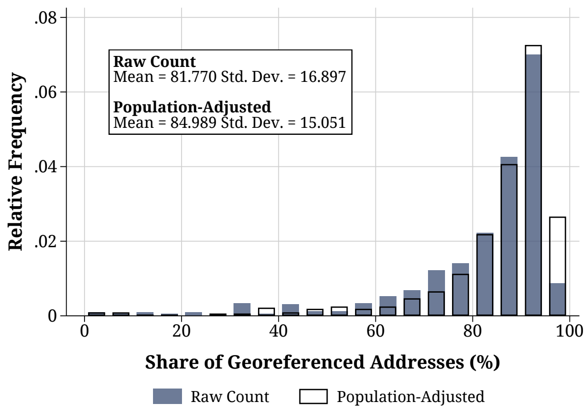
(C) 1881 Census



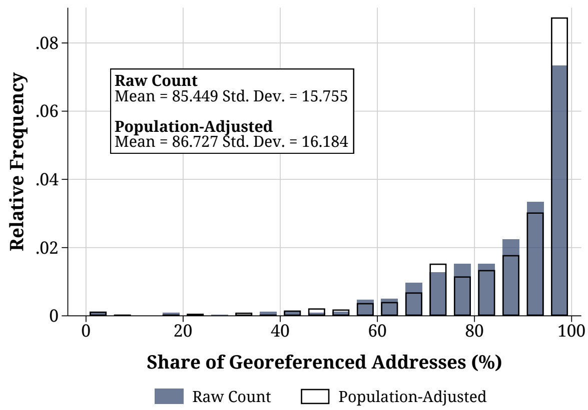
(D) 1891 Census



(E) 1901 Census

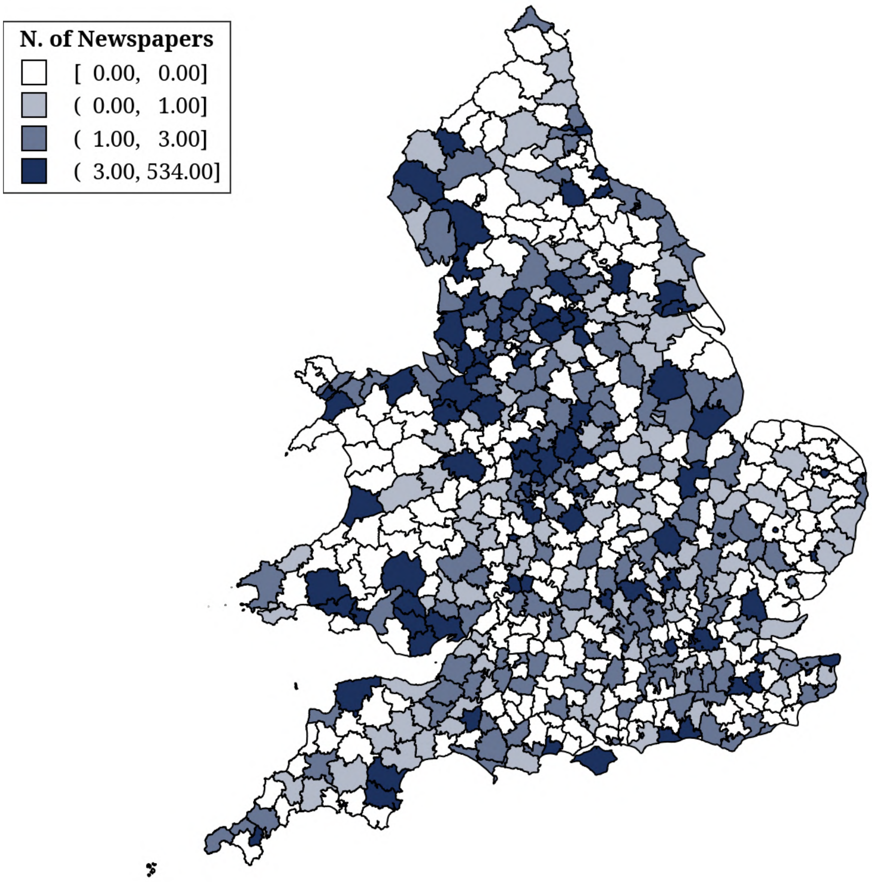


(F) 1911 Census



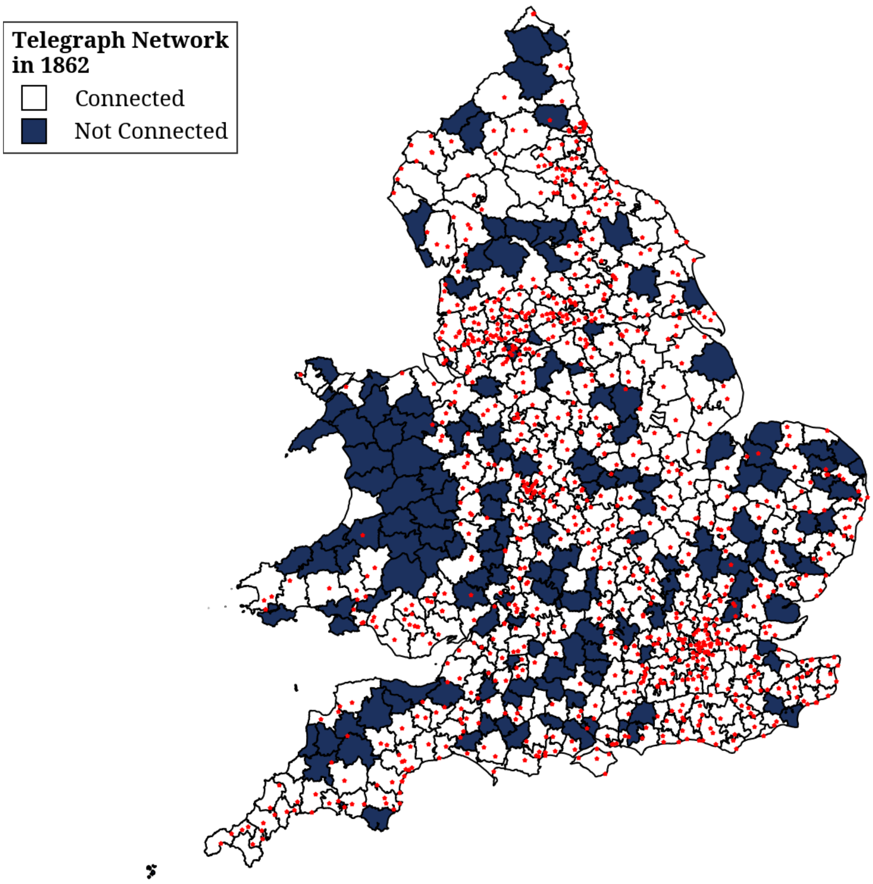
Notes. These figures display the district-level distribution of the share of geo-coded addresses from the UK censuses (1851–1911) by decade. For each census, we obtain a list of more than five million addresses by fine geographical unit (i.e., parishes). We then geo-reference these addresses to precise geographical coordinates. The black-contoured bars report the crude geo-coding rate; the blue bars report the population-adjusted geo-coding rate. Each figure reports the average and standard deviation of the two distributions.

FIGURE 1.A.3: Number of Active Newspapers Over the Period 1880–1940, by District



Notes. This figure reports the spatial distribution of the number of active newspapers across districts over the period 1880–1940. To be included in the data, a publication must be active for at least one year between 1880 and 1940. To retrieve the location of each journal, we geo-reference its publishing address and overlay historical district boundaries to assign it to consistent 1900 districts. The publishing address only lists the city. Hence we cannot distinguish across the eleven London urban districts. We consequently dissolve these districts into a single "London" unit.

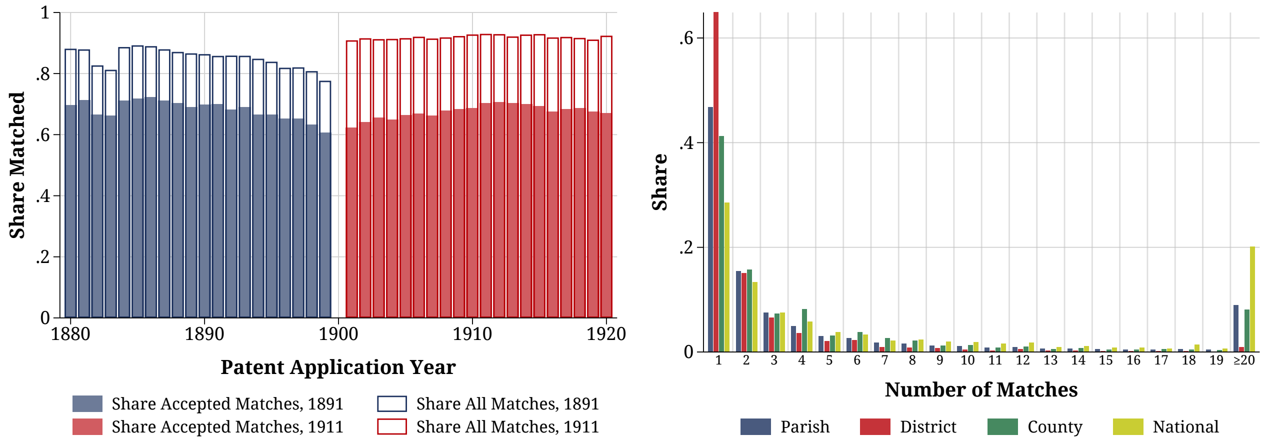
FIGURE 1.A.4: Distribution of Districts Connected to the UK Telegraph Network in 1862



Notes. This figure reports the spatial distribution of telegraph stations across districts in 1862. Red markers display the location of telegraph stations. Districts without any telegraph station are displayed in dark blue. To retrieve the coordinates of each telegraph station, we geo-reference the city where it is located. The list of telegraph stations is taken from the *Zeitschrift des Deutsch-Österreichischen Telegraphen-Vereins, Jahrgang*, volume IX, 1862. This source does not list telegraph stations in London. We thus dissolve urban districts in the London area into a single "London" unit and assume that this unit is connected to the domestic telegraph network.

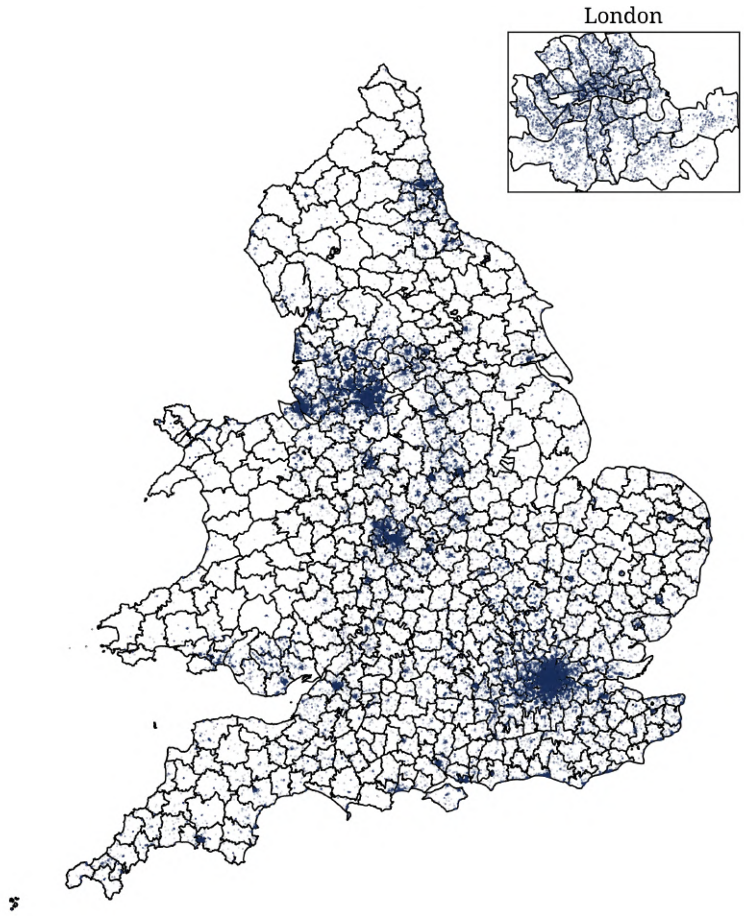
FIGURE 1.A.5: Matching Rate of the Linked Inventors-Census Sample, 1881–1911

(A) Matching Rate Over Time (B) Matching Rate by Geographic Layer



Notes. These figures report the matching rate for the linked inventor-census sample. Panel 1.A.5a reports the matching rate over time for the 1881–1900 sample (blue bars) and the 1901–1920 sample (red bars). Color-contoured bars report the share of records with at least one match; color-filled bars report the share of acceptable linked matches. A record match is acceptable if it has no more than five multiple matches. Panel 1.A.5b reports the share of matches by the number of matches, broken down by geographical layers. In Panel 1.A.5a, we do not show the few matches with quality below .95. In Panel 1.A.5b, the sample is restricted to records with at least one match.

FIGURE 1.A.6: Distribution of Inventors Across UK Districts, 1881–1911



Notes. This figure displays the spatial distribution of inventors across districts between 1881 and 1911. 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, geo-referenced. The background map displays districts at historical borders in 1900.

1.B Novel Patent Data

1.B.1 Sources and Digitization

This section presents the motivation for developing a new patent dataset for England and Wales that spans the second half of the XIX century. Then, we describe the sources we use and how we structure the textual data they contain into a machine-usable dataset. Finally, we describe two data-augmentation routines that we perform to geocode the patents and assign them a modern technology class.

1.B.1.1 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 *et al.*, 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.

1.B.1.2 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 1.B.1: in panel 1.B.1a we show the patent granted to Henry Bessemer for the eponymous process to produce steel, and in panel 1.B.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 professions, (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.

1.B.1.3 Digitization

We perform optical character recognition (OCR) on each patent individually 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, which is 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 implement a script that leverages regular expressions to identify the variables (i–ix). Fortunately, the text of each patent is fairly standardized; hence this routine yields detailed and high-quality results for all variables except (v), which is not used in this paper.

1.B.1.4 Geo-Coding

This exercise results in a database of approximately 800,000 patents granted between 1853 and 1899. 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 1.B.4, we report the resulting distribution of patents (panel 1.B.4a) and patents per capita (panel 1.B.4b). Reassuringly, these are consistent with underlying population and economic development indicators.

1.B.1.5 Technology Class Assignment

Naturally, historical patent documents do not list CPC classes. Yet, technological classification is a key variable in our empirical exercise. To reconstruct the class, we adopt a supervised machine-learning approach. We conjecture, following Xu (2018), that titles are informative of technological classes. We split the PATSTAT data, which covers the years 1900–1939 and for which we observe both titles and classes, in a train and a test set, with a proportion of 4:1. We apply a term frequency-inverse document frequency vectorization algorithm to the titles of both datasets. Then, we estimate a linear support vector machine (LSVC) on the train set. An LSVC is a non-probabilistic classifier that assigns class labels to maximize the width of the gap between classes. Formally, consider a set of points $(\mathbf{x}_i, y_i)_{i=1}^{N}$ where $\mathbf{x} \in \Re^N$ represent the features—in our case, words—and y is the class. For simplicity, assume $y \in \mathcal{Y} = \{-1, 1\}$. An LSVC solves for the hyperplane $\mathcal{W} = \{\mathbf{w} \in \Re^N \text{ s. t. } \mathbf{w}^\top \mathbf{x}_i - \ell = 0\}$ that maximizes the distance between the group i such that $y_i = 1$, and the group where $y_i = -1$. The distance that is most commonly used, that allows for non-linearly separable data, is the hinge loss, which is defined as $d_i = \max\{0, 1 - y_i (\mathbf{w}^\top \mathbf{x}_i - \ell)\}$. In our case, however, we allow for multiple classes, that is, the cardinality of $\mathcal{Y}$ be more than two. We employ an LSVC because the literature notes that it yields particularly robust results. However, the classification outcome would remain fairly unchanged using different algorithms.[50]

On the training set, the LSVC yields a 95% accuracy, measured as the share of patents with a correctly imputed class relative to the total number of patents. This decreases to 85% on the test set, which is not used to train the algorithm. Given that state-of-the-art models trained on modern US data achieve approximately a test 90% accuracy, we interpret these results as rather encouraging (Li *et al.*, 2018). We report the confusion matrix on the test set in Figure 1.B.2. For a given cell, the row label is the true technology class, and the column label is the imputed class. A perfect classifier would thus yield a diagonal confusion matrix. Overall, we find that misclassification errors are evenly distributed, in relative terms, across classes. Hence, even though the classifier is not perfect, there does not seem to be any systematic measurement error in class imputation.

1.B.2 External Validation

To validate our data, we consider the only two series that cover—a portion of—the years 1853–1899. 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 that we use as a comparison 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 are also listed. Absent a way to identify granted patents, we do not report figures after 1883 for the COI series.

[50] In particular, we tested the Naïve Bayes classifier, several Boosting algorithms (e.g., AdaBoost, XGBoost), a random forest classifier, and a simple convolutional neural network. All the above yield similar classification results but slightly lower accuracy than the LSVC. Additionally, we explored alternative vectorization algorithms using transformers (e.g., BERT and RoBERTa) with no significant performance gains.

In Table 1.B.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 for this period.

1.B.3 Summary Statistics and Stylized Facts

We conclude this section by presenting some stylized statistics and facts our new data allow us to uncover. First, as noted in Table 1.B.1, the number of patents granted generally grows over time, although at a somewhat stagnating path. There is, however, a sizable discontinuity between 1883 and 1884, when the number of patents jumps from 6074 to 9873. In 1883 the Patents Act reduced application fees by 83%, as noted by Nicholas (2014). It seems plausible to attribute the discontinuity to this reform.

Second, in Figure 1.B.3, we report the composition of patenting activity by technology class. In each year, we compute the share of patents in a given sector with respect to the total number of patents issued that year. We report such shares over time between 1853 and 1939. The composition of innovation exhibits two clear patterns. First, the share of textiles patents, which in the 1850s represented nearly 20% of the total, shrinks considerably, and in 1939 it accounts for less than 5%. This is consistent with the historical preeminence of textiles during the First Industrial revolution and their subsequent loss of importance. Second, electricity-related innovation grows considerably in the later part of the period. In 1939, it represented more than 20% of the total number of patents issued in the UK. Once more, this echoes historical, anecdotal evidence highlighting the centrality of electricity during the later stages of the Second Industrial revolution and beyond (David, 1990; Mokyr, 1998).

Finally, a crucially novel component of our dataset is that it allows studying the geographical dimension of the innovation process. Thus, in Figure 1.B.4, we report the spatial distribution of the number of patents in absolute number (panel 1.B.4a) and normalized by population (panel 1.B.4b). These maps attest to the importance of duly considering the geography of innovation. The patenting activity appears to be widely dispersed across England and Wales. Heavily industrial areas, such as Lancashire, the Midlands, the Tyne, and South Wales, all feature prominently in terms of issues patents. Similarly, the London area is also a major innovation hub. By contrast, Northern Wales, Anglia, Cornwall, and Cumbria perform poorly. In Figure 1.B.5, we repeat this exercise, but we break down the number of patents by selected technology classes: chemistry (panel 1.B.5a), electricity (panel 1.B.5b), engineering (panel 1.B.5c), engines and pumps (panel 1.B.5d), metallurgy (panel 1.B.5e), and textiles (panel (panel 1.B.5f). While innovation centers remain roughly similar across sectors, some differences emerge. For example, the metallurgy industry was particularly deep-rooted in the Midlands, where we note the largest concentration of metallurgy patenting. Similarly, textile innovation centers in the Lancashire area, the historic “cotton districts”. Our database allows studying a novel, thus far largely unexplored dimension of the innovation and patenting activity. Therefore, the analysis carried out in this paper is one of many that may take advantage of this contribution.

1.B.4 Tables

TABLE 1.B.1: Patents Granted in the UK: Comparison Across Three Datasets

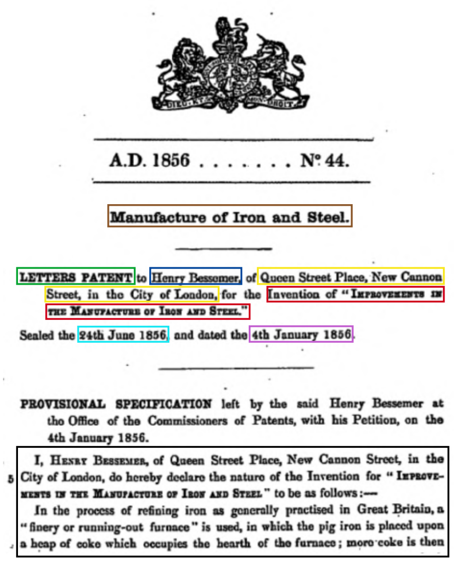
Years 1853-1876				Years 1877-1899			
(1) Year	(2) Our Series	(3) COI	(4) Hanlon	(5) Year	(6) Our Series	(7) COI	(8) Hanlon
1853	3042	3016		1877	4943	4928	4940
1854	2759	2690		1878	5336	5143	5333
1855	2960	2866	2955	1879	5332	5305	5325
1856	3107	2967	3102	1880	5499	5132	5509
1857	3206	3092	3197	1881	5744	5620	5745
1858	3023	2954	2999	1882	6159	6150	6233
1859	3048	2989	2998	1883	6074	6006	5981
1860	3192	3139	3190	1884	9873		
1861	3261	3269	3272	1885	8783		
1862	3482	3459	3486	1886	8999		
1863	3301	3299	3308	1887	9218		
1864	3256	3225	3257	1888	9331		
1865	3378	3364	3378	1889	10325		
1866	3451	3408	3452	1890	10355		
1867	3724	3692	3720	1891	10686		
1868	4008	3908	3984	1892	11429		
1869	3832	3741	3781	1893	11985		
1870	3407	3288	3405	1894	11648		
1871	3525	3479	3525	1895	12198		
1872	3969	3940	3967	1896	13597		
1873	4276	4281	4282	1897	14249		
1874	4494	4516	4491	1898	13100		
1875	4557	4451	4557	1899	13172		
1876	5049	5012	5064				

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.

1.B.5 Figures

FIGURE 1.B.1: Sample Annotated Patent Documents

(A) Henry Bessemer's 1856 Patent



A.D. 1856 . . . . N° 44.

Manufacture of Iron and Steel.

LETTERS PATENT to Henry Bessemer, of Queen Street Place, New Cannon Street, in the City of London, for the Invention of "IMPROVEMENTS IN THE MANUFACTURE OF IRON AND STEEL."

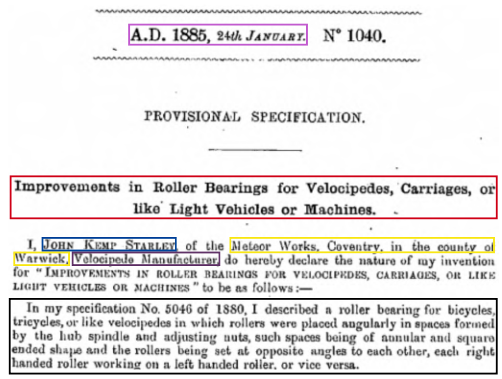
Sealed the 24th June 1856, and dated the 4th January 1856.

PROVISIONAL SPECIFICATION left by the said Henry Bessemer at the Office of the Commissioners of Patents, with his Petition, on the 4th January 1856.

I, HENRY BESSEMER, of Queen Street Place, New Cannon Street, in the City of London, do hereby declare the nature of the Invention for "IMPROVEMENTS IN THE MANUFACTURE OF IRON AND STEEL" to be as follows:—

In the process of refining iron as generally practised in Great Britain, a "finery or running-out furnace" is used, in which the pig iron is placed upon a heap of coke which occupies the hearth of the furnace; more coke is then

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



A.D. 1885, 24th JANUARY. N° 1040.

PROVISIONAL SPECIFICATION.

Improvements in Roller Bearings for Velocipedes, Carriages, or like Light Vehicles or Machines.

I, JOHN KEMP STARLEY, of the Meteor Works, Coventry, in the county of Warwick, Velocipede Manufacturer, do hereby declare the nature of my invention for "IMPROVEMENTS IN ROLLER BEARINGS FOR VELOCIPEDES, CARRIAGES, OR LIKE LIGHT VEHICLES OR MACHINES" to be as follows:—

In my specification No. 5046 of 1880, I described a roller bearing for bicycles, tricycles, or like velocipedes in which rollers were placed angularly in spaces formed by the hub spindle and adjusting nuts, such spaces being of annular and square ended shape and the rollers being set at opposite angles to each other, each right handed roller working on a left handed roller, or vice versa.

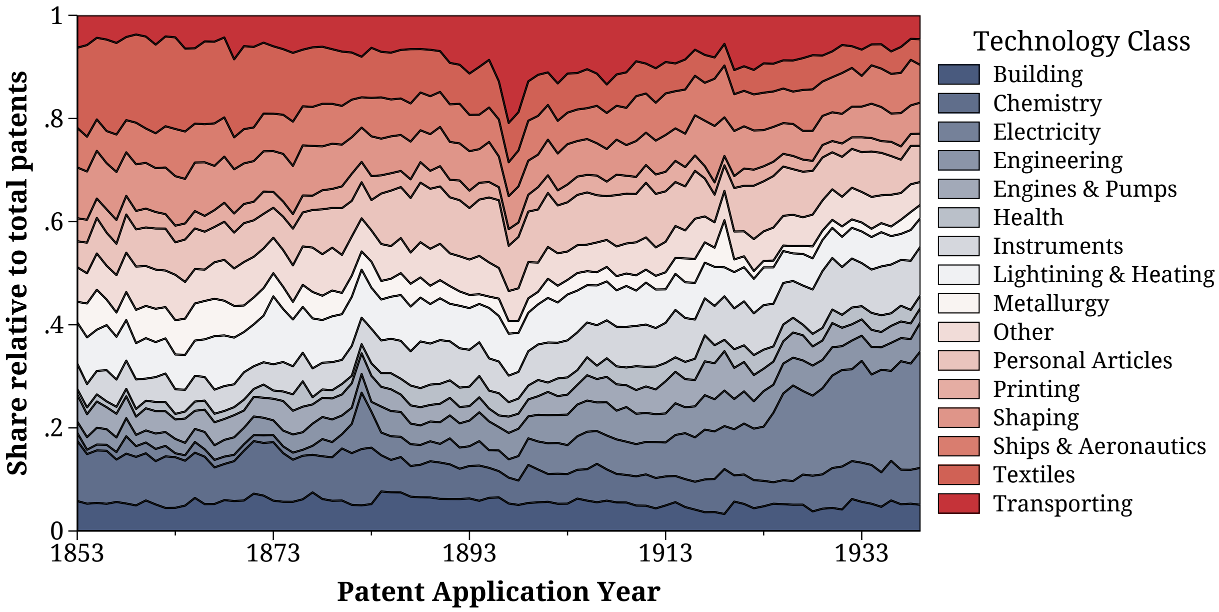
Notes. This figure displays two sample patent documents in our dataset. Panel 1.B.1a was granted to Henry Bessemer in 1856 for the invention of the famous eponymous process for the mass production of steel from the molten pig iron. Panel 1.B.1b was granted to John Starley in 1885 for the invention of the first modern bicycle, which would soon revolutionize mobility in Europe and in 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.

FIGURE 1.B.2: Confusion Matrix of the Technology Sector Classifier

True Technology Class \ Predicted Technology Class	Agriculture	Building	Chemistry	Electricity	Engineering	Engines, Pumps	Food	Health, Amusement	Instruments	Lightning, Heating	Metallurgy	Personal Articles, Furniture	Printing	Separating, Mixing	Shaping	Ships, Aeronautics	Textiles	Transporting
Agriculture	951	10	23	2	2	4	24	8	5	6	0	45	2	21	34	21	6	12
Building	9	2310	42	11	58	6	0	14	25	23	4	65	3	12	63	58	7	79
Chemistry	7	56	4704	34	18	21	21	21	19	78	62	39	31	89	133	31	71	9
Electricity	1	9	25	5146	30	43	1	14	163	53	21	20	12	8	34	40	14	63
Engineering	1	41	15	55	1975	105	0	11	37	36	6	62	4	3	100	62	19	151
Engines, Pumps	1	5	13	43	80	2113	0	2	31	47	2	5	0	10	10	40	1	59
Food	11	2	60	1	5	1	936	9	4	32	0	27	2	22	44	46	3	1
Health, Amusement	6	21	57	29	20	7	5	1525	64	24	7	101	5	15	31	54	17	20
Instruments	3	19	28	304	76	43	11	48	4148	69	23	93	78	14	40	115	17	92
Lightning, Heating	10	22	53	65	35	51	18	7	51	2739	31	50	5	40	24	30	16	40
Metallurgy	2	5	85	18	14	6	0	9	26	39	1565	18	8	17	58	37	5	14
Personal Articles, Furniture	12	67	43	15	26	3	26	47	72	40	3	3847	44	13	90	127	60	26
Printing	1	12	49	16	2	3	4	4	80	13	4	49	1567	11	77	58	45	5
Separating, Mixing	12	8	145	10	19	9	13	16	27	44	33	18	17	1057	54	44	18	1
Shaping	20	61	119	47	111	13	11	19	51	35	52	121	55	26	3737	143	57	70
Ships, Aeronautics	16	51	24	38	73	36	37	25	92	24	21	102	60	22	116	3736	100	97
Textiles	4	10	102	11	18	5	4	5	11	6	4	82	27	12	54	40	2898	4
Transporting	14	59	4	88	149	42	0	5	50	59	3	55	1	1	36	60	4	3863

Notes. This figure displays the confusion matrix of the patent technology classifier. The algorithm assigns to each patent an imputed technology class using information contained in the title. Titles undergo pre-processing and term frequency-inverse document frequency (tf-idf) vectorization. The classifier is trained on an 80% sub-sample of the universe of British patents granted over the period 1900–1940. The figure reports the classifier's performance on the remaining 20% test set, which is not used in training. The y-axis reports the true patent class; the x-axis reports the class imputed by the classifier. A perfect classifier would yield a diagonal confusion matrix. The accuracy in the training (resp. test) set is ≈98% (resp. ≈85%). Lighter to darker blue indicates an increasing number of patents in the cell.

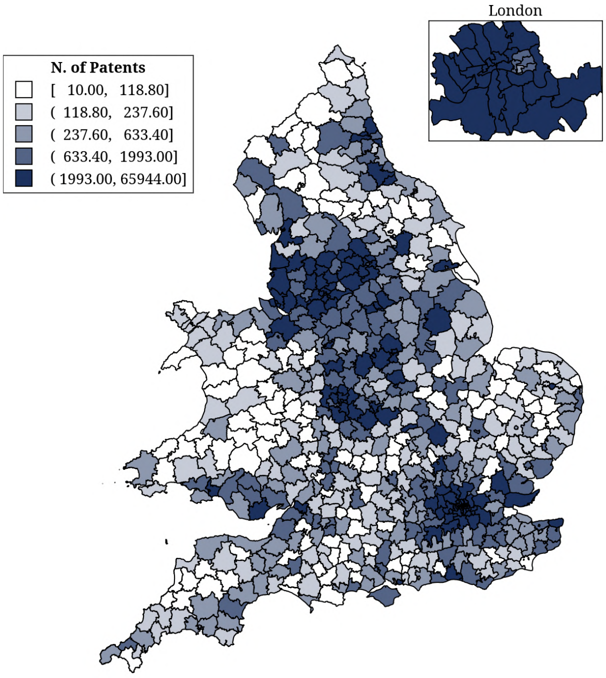
FIGURE 1.B.3: Distribution of Patents Granted in the UK Across Technology Classes



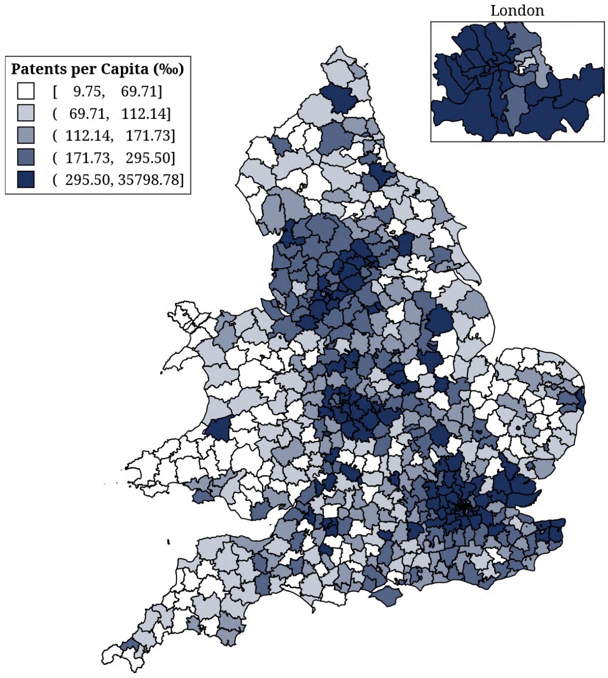
Notes. This figure displays the evolution of innovation in Britain across technology classes from 1853–1939. For each year, we compare the share of patents in each class in our database relative to the total number of patents granted in that year. Data for the period 1853–1899 are from the newly digitized universe of patents; data for the period 1900–1939 are made available by the European Patent Office repository.

FIGURE 1.B.4: Distribution of Patents and Patents per Capita Across Districts, 1880–1939

(A) Number of Patents

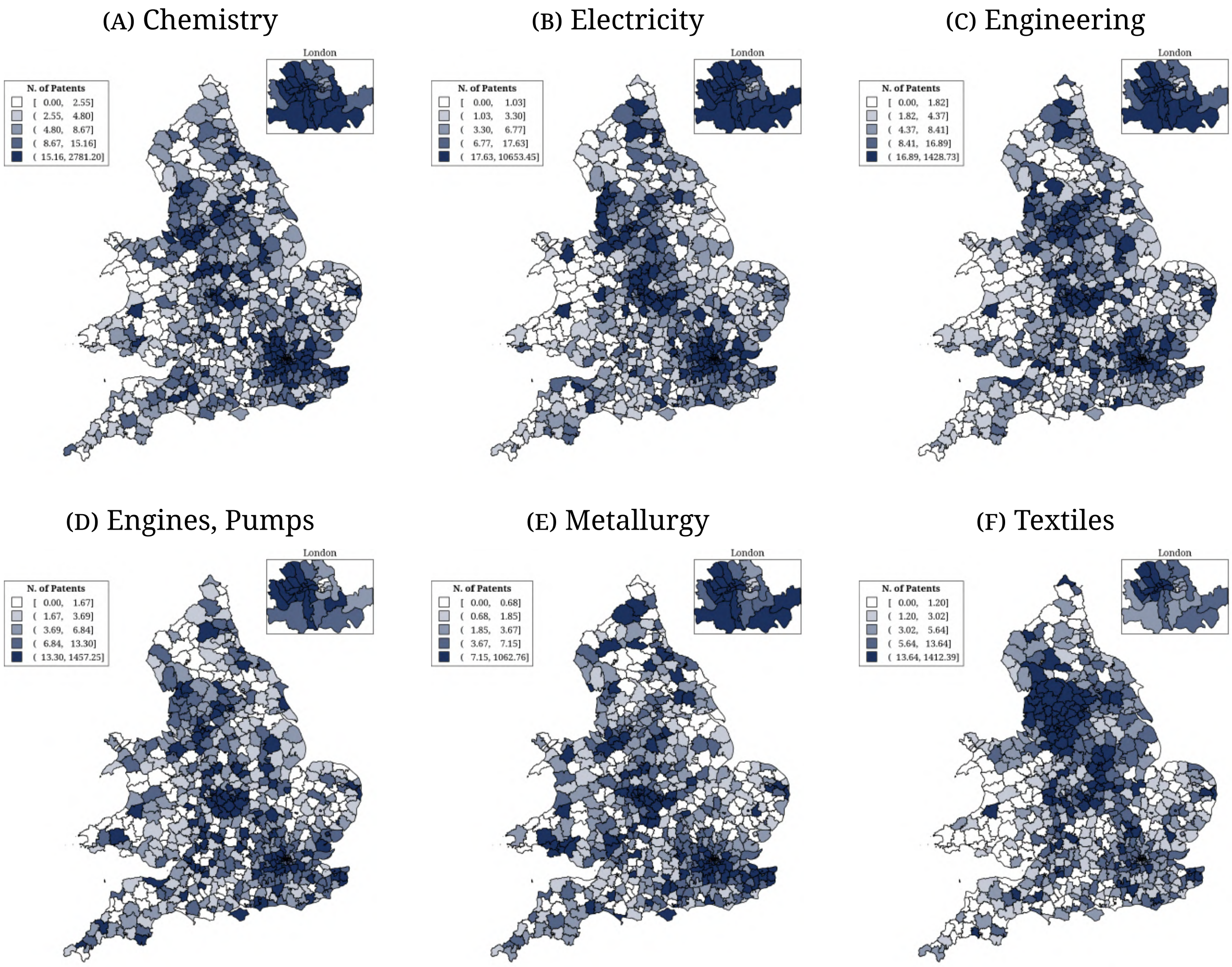


(B) Patents per Person



Notes. These figures report the intensity of patenting activity across districts over the period 1880–1939. Panel 1.B.4a reports the total number of patents granted; Panel 1.B.4b normalizes this by district population in 1900 and expresses the resulting rate in ‰ units. Districts are displayed at 1900 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.

FIGURE 1.B.5: Spatial Distribution of Patents Granted Across Technology Classes, 1880–1939



Notes. These figures report the intensity of patenting activity across districts over 1880–1939 for selected technology classes. Districts are displayed at 1900 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.

1.C Linked International Migrants Sample

This section discusses our methodology to link English and Welsh immigrants to the UK census and present key statistics on the resulting dataset.

1.C.1 Sources and Linking Algorithm

We rely on two sources of externally-compiled data.[51] For the US, we have access to the IPUMS full-count 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.[52] 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 1850-1920. Crucially, we have access to the non-anonymized version of the IPUMS data. Hence, besides publicly available information, we also know each individual's recorded name and surname.

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 decade frequency starting in 1851 and 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.

Our methodology relies on Abramitzky *et al.* (2021). This dataset tackles the problem that neither the US nor the UK—nor other European countries—recorded where British immigrants came from *within* the UK. We thus 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.[53] 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.[54] This implies

[51]We 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.

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

[53]Since women usually change their name upon marriage, we are unable to match them. This is a common problem in linking algorithms (Abramitzky *et al.*, 2021).

[54]Since 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.

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 we link consists of individuals who report, in the US census, either England, Scotland, 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 $\mathcal{M}^A = \{m_1^A, \ldots, m_N^A\}$.

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 $m_i^A \in \mathcal{M}^A$, compute[55]

$$\text{Similarity}_i^A = \alpha \times \text{Name Similarity}_i^A + (1-\alpha) \times \text{Surname Similarity}_i^A \tag{1.13}$$

   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

$$\overline{\mathcal{M}}^A = \left\{ m_i^A \in \mathcal{M}^A \;\middle|\; \text{Similarity}_i^A = \max_{m_{i'}^A \in \mathcal{M}^A} \text{Similarity}_{i'}^A \right\} \tag{1.14}$$

   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:

$$\widetilde{\mathcal{M}}_\tau^A = \left\{ m_i^A \in \overline{\mathcal{M}}^A \;\middle|\; \text{Similarity}_i^A \geq \tau \right\} \tag{1.15}$$

   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

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

of similarities there.

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

$$d_{ij} \equiv \widehat{d_{ij}} + \ell p(1 - \widehat{d_{ij}}) \tag{1.16}$$

where

$$\widehat{d_{ij}} \equiv \begin{cases} 0 & \text{if } m = 0 \\ \frac{1}{3}\left(\frac{m}{|i|} + \frac{m}{|j|} + \frac{m-t}{m}\right) & \text{else} \end{cases} \tag{1.17}$$

where m is the number of matching characters, $|i|$ is the length of string i, and 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.[56]

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. The manual assessment confirmed that the Jaro-Winkler metric outperforms other measures in our setting.

1.C.2 Internal and External Validation

We now present key statistics on the dataset that we assemble. In Figure 1.C.1, we report the matching rate by the number of matches (panel 1.C.1a) and over time (panel 1.C.1b. The matching rate is the ratio between the number of matched individuals and the number of English and Welsh immigrants in the US census. We break down the matching rate by the number of matches every immigrant is associated with. About 40% of the overall immigrant population is matched to one single record in the UK census. Another 10% is matched to two records, and the remaining 50% is matched to three or more records in the UK census. By construction, we can never match someone not appearing in the UK census. This is possible if a child born in, say, 1895 emigrates before 1901, which is the closest subsequent census. In Figure 1.C.1a, we report a corrected matching rate whose denumerator removes these "unmatchable" observations. Overall, 55% of the total number of English and Welsh immigrants is matched to no more

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

than two records in the UK census. This constitutes the baseline sample that we analyze. A 55% matching rate is consistent with standard historical linking algorithms (Abramitzky *et al.*, 2021), although a more precise quantitative assessment is complex because the benchmark statistics refer to intergenerational census linking exercises.

In panel 1.C.1b, we report the matching rate by immigration year. In blue, we report the total number of immigrants; those paired with at least one match are shown in red; the green area reports our baseline sample, which is composed of all those immigrants with no more than two matches. We also impose a quality threshold on names and surnames. Suppose an immigrant is matched to someone born in the same year. In that case, we require both the name and the surname to have a similarity above .85. If an immigrant is matched to someone born either one year before or one year after, we impose a stricter threshold of .9 on both name and surname. We set high thresholds because we are concerned about false positive matches. Following Abramitzky *et al.* (2021), we are thus willing to give up on power to maximize accuracy. In Figure 1.C.2, we report the overall distribution of name (panel 1.C.2a) and surname (panel 1.C.2b) match quality. The solid and dashed red lines superimpose the aforementioned coarse and strict thresholds. The quality distribution is substantially skewed to the left: most matches are of excellent quality. Dropping low-quality ones is, therefore, quantitatively second-order.

Since we match immigrants to the UK census before their migration year, the matching rate decreases over a decade. This is clear from the black line in Figure 1.C.1b, which jumps up at the turn of each decade until 1911. The matching rate before 1881 is relatively low. This is because no census was taken in the UK in 1871. Therefore, we match all those who migrated to the US between 1870 and 1881 to the 1861 census. This mechanically reduces the matching rate, for we cannot match all those born between 1862 and 1881 who migrate during this period. Similarly, the matching rate decreases after 1911. This is because censuses after 1911 are protected by British privacy law. We thus match all those who migrate after 1911 to that census. However, this implies that we cannot match all those who migrated after 1911 and were born after that year. To ensure that our results are not driven by these asymmetries at the edges of the sample, in robustness analyses, we show that restricting the period to the years 1880-1920 does not affect our main findings.

One plausible concern is that instances of migrants with multiple matches in the UK census are not randomly distributed. This may be due to various reasons (Bailey *et al.*, 2020). First, educated individuals are more likely to report their name and surname in full, with consistent spelling over time. This would generate non-classical measurement error because the matching rate would be higher for a selected population sub-sample. This issue does not seem to be relevant in this case, as the matching rate—*i.e.* the share of immigrants that are *eventually* matched, irrespective of the number of matches—approaches the universe of the observations. Second, the number of matches may not be orthogonal to individual characteristics. This may be the case if wealthier individuals give relatively uncommon names, as documented by Olivetti *et al.* (2020). To assess the severity of this concern, we regress the number of matches on a set of individual-level observable variables observed in the US and UK censuses. Under classical measurement error, we would expect no statistically significant correlation between the number of matches and observable characteristics. Table 1.C.1 reports the estimates thus obtained. We find minimal and marginally significant correlations between the number of matches and individual-level characteristics observed in the US census. The number of matches correlates positively with agriculture and low-skilled employment. However, these correlations are very small: one more match is associated with a .01% increase in the probability of being employed in agriculture. This association is marginally larger for low-skilled manufacturing employment (0.03%). These very low magnitudes are unlikely to affect the results we document in this paper quantitatively. Moreover, notice that most correlations are not statistically significant. Most importantly, we do not find any significant association between the number of matches and the location of English immigrants. This is reassuring because our identification assumption crucially hinges on the variation arising from settlement decisions. We believe this is solid evidence of our linking algorithm and the novel database we assemble.

1.C.3 Return Migration Data

Following the logic explained in section 1.C.1, we construct a linked sample of return migrants. This identifies English and Welsh immigrants in the US in decade d and looks for possible matches in the UK census in decade $d + 1$, using a minor variation on the algorithm described previously. Since the last UK census that we have 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.

1.C.4 Summary Statistics and Stylized Facts

The newly developed dataset we develop presents some key novelties compared to available data. It is the first dataset that allows retrieving the origin of US immigrants from England

and Wales at a fine level of geographical aggregation during a period of massive international migrations (1880–1930).[57] The dataset's granular—individual—structure allows us to observe several individual characteristics of immigrants at home and in the US. This section briefly discusses key stylized facts that our new data allow us to document.

In Figure 1.C.3, we explore the origin of English and Welsh emigrants to the US over time. Each figure reports the emigration rate normalized by population in 1900, in thousand units. Two patterns emerge. First, substantial cross-sectional heterogeneity exists in the intensity of out-migration across districts throughout the sample period. Second, we find that the intensity of US emigration flows is initially larger in rural districts, especially in the South West and East regions, but this shifts over time toward industrial and urban areas. By the 1910s, the industrialized Lancashire districts featured as a prominent area of emigration. This finding provides a sound quantitative validation of historical—largely anecdotal—evidence (Erickson, 1972; Baines, 2002).

Additionally, we can study the selection patterns of English and Welsh emigrants along two margins. Specifically, we can compare them to (i) the native US population in the areas where they settled and (ii) the non-migrant population in England and Wales who lived in their origin areas. These exercises extend seminal historical work by Baines (2002), who performed a similar exercise using incomplete information from the population censuses. We defer a discussion of selection patterns to the main text. Here, we only note that our dataset is well-suited to study the selection of British emigrants because it identifies individuals before they migrate, thus conveying a complete picture of selection issues during the period.

[57] Similar data-sets have been produced for Norwegian Abramitzky *et al.* (2014) and Swedish (Andersson *et al.*, 2022) immigrants. Our is the first such effort for a major European country: in 1890, the population in England and Wales stood at more than 27 million inhabitants. This compares to approximately 2 million Norwegians and 4.7 million Swedes.

1.C.5 Tables

TABLE 1.C.1: Correlation Between Number of Matches and Observable Characteristics

	Dep. Var.: Number of Matches		
	(1)	(2)	(3)
Panel A. Occupations			
Agriculture	0.003	0.004*	0.003**
	(0.001)	(0.001)	(0.001)
Low-Skilled Manufacture	0.013**	0.011**	0.008**
	(0.003)	(0.003)	(0.002)
High-Skilled Manufacture	0.006	0.007	0.008*
	(0.003)	(0.003)	(0.003)
Professionals	–0.001	–0.001	–0.001
	(0.001)	(0.001)	(0.001)
Public Administration	–0.004**	–0.003*	–0.002
	(0.001)	(0.001)	(0.001)
Manager	–0.000	–0.000	0.000
	(0.000)	(0.000)	(0.000)
Service Worker	0.001	0.001	0.002
	(0.001)	(0.001)	(0.001)
Panel B. Origin			
Northeast	–0.004		
	(0.006)		
Midwest	0.004		
	(0.004)		
South	–0.002		
	(0.001)		
West	0.001		
	(0.002)		
State FE	No	Yes	No
County FE	No	No	Yes
Year FE	No	Yes	Yes

Notes. This table reports the correlation between observable characteristics of British immigrants in the US census and the number of matches in the linked sample database. In each row, the table displays the correlation between the number of matches and an indicator equal to one if for immigrants that correspond to the row variable and zero otherwise. The sample is restricted to the set of matches we effectively use in the analysis. Column (1) reports unconditional correlations; column (2) includes state and census decade fixed effects; column (3) adds county fixed effects. In Panel A, the characteristics are the occupations; in Panel B, the variables are the Census Bureau region of residence. Standard errors, clustered at the county level, are shown in parentheses.

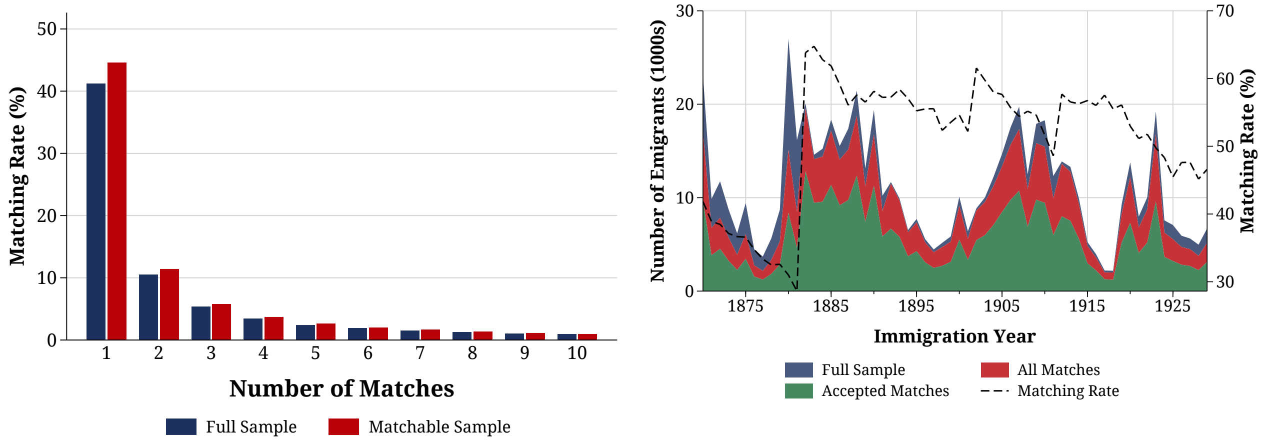
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

1.C.6 Figures

FIGURE 1.C.1: Share of British Immigrants in the US Census Matched to the UK Census

(A) Matching Rate by Number of Matches

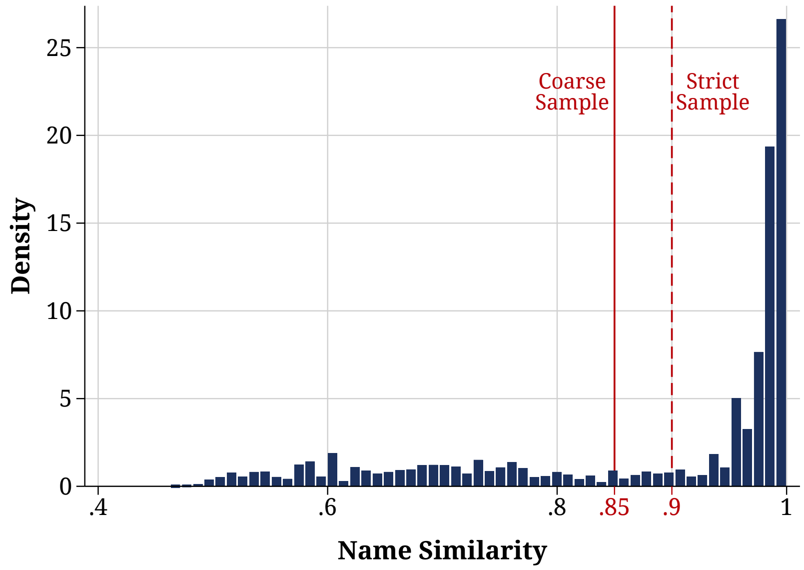
(B) Matching Rate Over Time



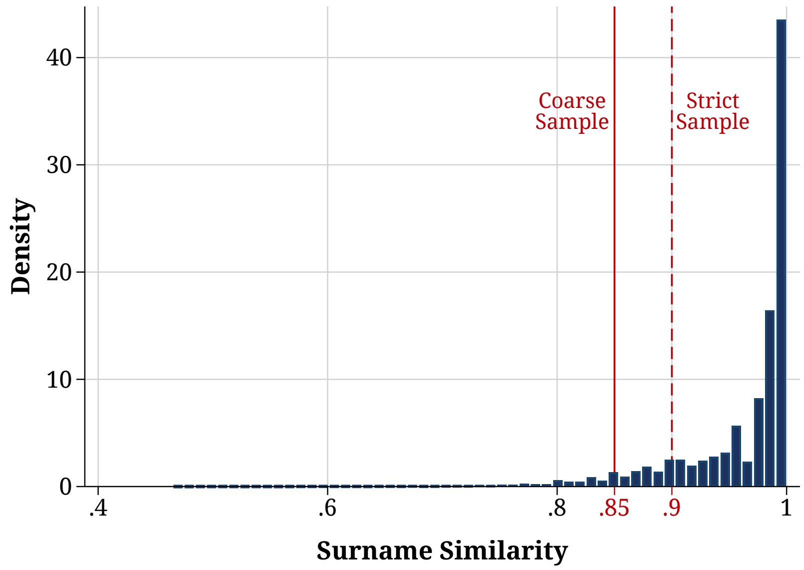
Notes. These figures report the share of English and Welsh immigrants recorded in the US census that we match to the UK census. Panel 1.C.1a plots the share of records that we match to the UK census and whose match quality is such that we retain it in the linked sample, broken down by the number of matches. In the baseline sample, we keep records with no more than two matches. Blue bars report ratios relative to the entire number of immigrants, and red bars restrict the set of immigrants to those we can match. Panel 1.C.1b reports the matching rate over time. The blue area reports the total number of US immigrants, the red area reports the entire number of matches we obtain, and the green area reports the matches that eventually enter our baseline linked sample. The black dashed line on the right y-axis is the ratio between the green and the blue areas.

FIGURE 1.C.2: Quality of Matches in the Complete Linked Sample: Names and Surnames

(A) Name Match Quality

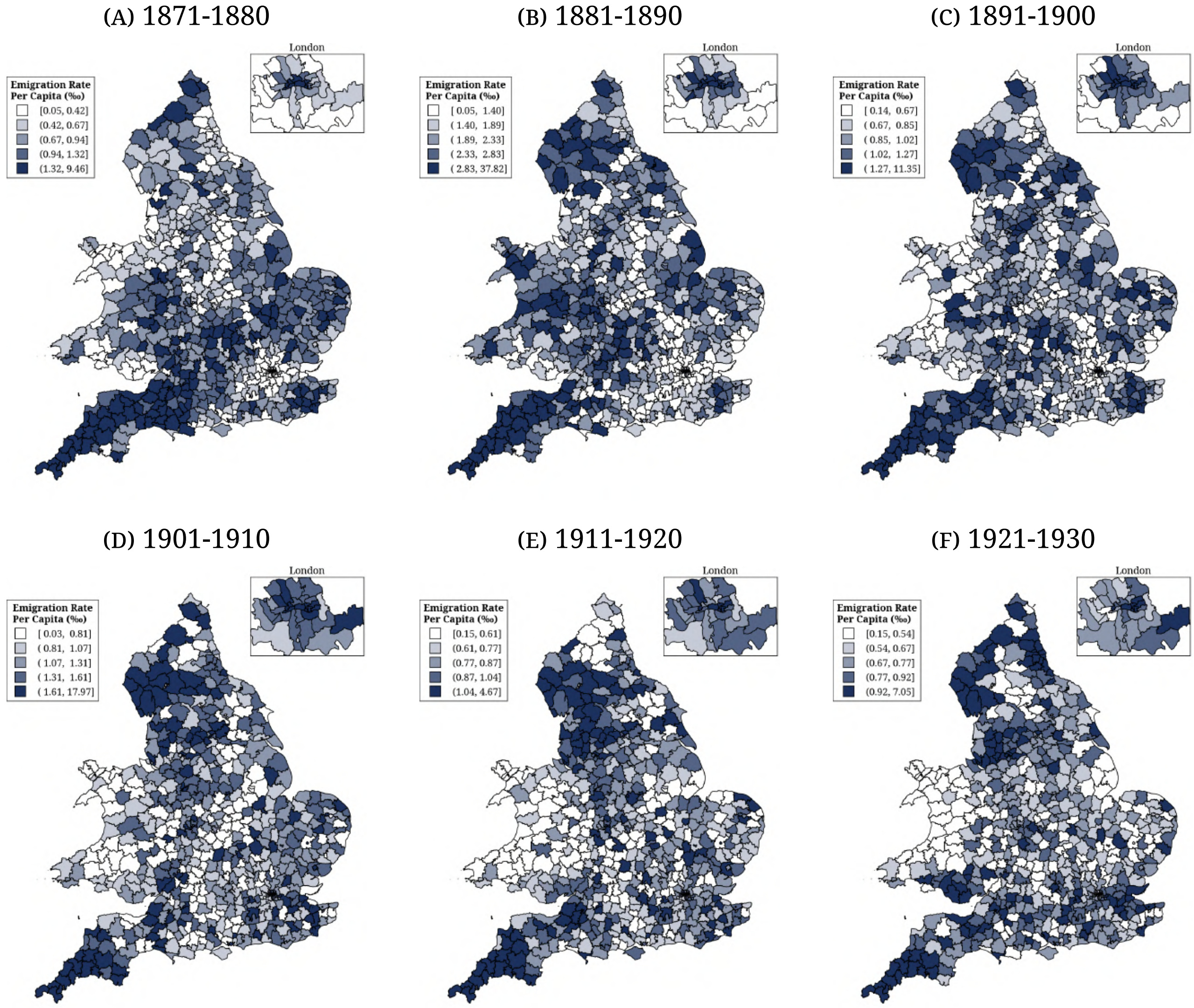


(B) Surname Match Quality



Notes. The figures report 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 1.C.2a reports the distribution of the name similarity; Panel 1.C.2b refers to surnames. The vertical lines mark the quality thresholds we impose for a match to be part of the final linked sample.

FIGURE 1.C.3: Distribution of the Emigration Rate Across Districts, 1871–1930



Notes. These figures report the distribution of US emigrants across districts in England and Wales over the period 1871–1939 by decade. Data are from the matched emigrants' sample. The number of emigrants in each decade is normalized by population in 1900 and is expressed in ‰ units. Districts are displayed at their 1900 borders. Out-migration is also cross-walked to consistent historical borders. Lighter to darker blues indicate higher emigration rates.

1.D Additional Results

This section presents in some detail several additional results that are mentioned in passing in the main text.

1.D.1 Trade-Induced Technology Transfer

Our favored explanation of the return innovation result is that migrants facilitate the flow of knowledge between the areas where they settle and those they originate from. We argue that those flows are fostered by the diffusion of information and by market integration. This section presents one more piece of evidence in this direction. We focus on international trade as a measure of bilateral market integration. Previous research documents that trade fosters innovation, either because of increased import competition (Bloom *et al.*, 2016; Autor *et al.*, 2020), export opportunities (Bustos, 2011; Atkin *et al.*, 2017; Aghion *et al.*, 2018), access to intermediate inputs (Juhász and Steinwender, 2018), and increased market size (Coelli *et al.*, 2022).[58] In our analysis, we interpret trade as a means of facilitating technology transfer between the UK and the US, following Aleksynska and Peri (2014) and Ottaviano *et al.* (2018).

We consider a major shock to trade flows between the US and the UK: the 1931 Smoot-Hawley Act. The Act was a major trade policy reform enacted in response to the Great Depression (Eichengreen, 1986; Crucini, 1994). Importantly for our setting, the Act did not establish a uniform tariff rate. Instead, as we report in Table 1.A.3, tariffs vastly differed across industries before and after the shock. We leverage this variation, interacted with the before-Act knowledge exposure in a difference-in-differences setting.[59] The key idea that underlies this approach is that if migration linkages generate return innovation flows through international trade, then an increase in trade costs is expected to reduce patenting in the UK in the sectors that (i) districts were more exposed to, through migrations, and (ii) were targeted by the tariff increase.

We thus estimate the following double differences model separately for protected and non-protected industries:

$$\text{Patents}_{ik,t} = \alpha_{i\times k} + \alpha_t + \sum_{h=-a}^{b} \beta_h \times \left[\mathbb{1}\left(t = 1931 + h\right) \times \text{Knowledge Exposure}_{ik}\right] + \varepsilon_{ik,t} \qquad (1.18)$$

[58] Shu and Steinwender (2019) provide a critical assessment of the literature studying the effect of international trade on innovation.

[59] We first map sectors defined in the Act to technology classes. We then assign one class to the treatment group if its average *ad valorem* import duty changes by more than fifty percentage points between 1925–1930 and 1931–1936. Yearly tariff rates have been digitized from the *Statistical Abstract of the United States*.

Where i, k, and t respectively denote a district, technology class, and year, and Knowledge Exposure$_{ik}$ is the average sector-level knowledge exposure in the decade before the Act (1920–1930). In the baseline analysis, an industry is protected if its tariff rate increases by more than 50 p.p. between 1925–1930 and 1931–1935. Then, we estimate the triple-differences specification that compares treated and non-treated industries:

$$
\begin{aligned}
\text{Patents}_{ik,t} = \alpha_{i\times k} + \alpha_{i\times t} + \alpha_{k\times t} + \\
+ \sum_{h=-a}^{b} \beta_h \times \left[\text{Tariff}_k \times 1\,(t = 1931 + h) \times \text{Knowledge Exposure}_{ik}\right] + \varepsilon_{ik,t}
\end{aligned}
\tag{1.19}
$$

where Tariff$_k$ is an indicator returning value one for protected industries and zero otherwise.

In columns (1–2) of Table 1.D.3 we report the results of model (1.18). Column (1) presents the estimated coefficient for non-protected industries, while column (2) focuses on protected ones. We find no effect for the former and a negative effect for the latter. This is confirmed when looking at the associated flexible specification, reported in Figure 1.D.1. This also provides evidence supporting the parallel trends assumption for the two groups of technology classes. In columns (3–5), we report the estimates of the triple differences model (1.19). We consider three possible threshold values of the increase in the tariff rate after the Act to define a protected sector (10%, 30%, and 50%). All yield quantitatively similar estimates. Note, however, that the estimated ATE reassuringly increases in absolute magnitudes for larger tariff increases.

The analysis suggests that trade—which we interpret as a proxy for market integration—is a relevant channel through which migration ties generate knowledge flows and technology transfer. However, it is worth noting that the magnitude of the estimated treatment effects of the tariff reform on UK innovation is modest, despite the large increase in tariff duties. We thus interpret trade as one additional, although plausibly not the pivotal, factor driving return innovation.

1.D.2 Selection of British Migrants

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 counties, such as Germany, Sweden, or Italy, US emigration in the UK 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 analysis of incomplete census sources. In this section, we present evidence on the selection of English emigrants to the US and their integration into the US. To construct these statistics, we leverage the novel linked sample that allows us to observe individual-level characteristics before emigrants left—in the UK census—and after they settled—in the US census.

Table 1.D.4 compares US emigrants with the non-migrant population. 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. Migrants are less likely to work in agriculture and as professionals. They are, however, more likely to be employed in industrial sectors, such as textiles and metallurgy. This overall confirms the historical analysis of Baines (2002). Emigrants mainly originated from the North West, including Lancashire, and South West, chiefly, Devon and Cornwall. Similar patterns emerge when looking at return migrants, who are even less likely to be employed as agricultural workers. Return rates in high-emigration areas of the South West appear low compared to the rest of the country, while they are very high in the London area.

In Table 1.D.5, we compare English and Welsh immigrants to the rest of the US population. Column (1) refers to natives, and columns (2) and (5) refer to emigrants and return migrants, respectively. In columns (3) and (6), we compute the difference between natives and emigrants and natives and return migrants, respectively. UK immigrants differ substantially from the rest of the US population: they are less likely to work in agriculture and as civil servants. By comparison, they are more likely to be employed in metallurgy, textiles, and trade. This aligns well with evidence by Erickson (1972), who argues that English immigrants in the US tended to specialize in industries where they had a comparative advantage. Similar patterns emerge for return migrants. Regarding their geographical distribution, UK immigrants settled most commonly in the New England and Mid-Atlantic regions.

1.D.3 Long-Run Effect of Return 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:

$$\text{Patents}_{ik,t} = \alpha_{i\times k} + \alpha_t + \sum_{\tau\in\mathcal{T}} \beta^\tau \left[\text{Knowledge Exposure}_{ik} \times 1\,(t = \tau | t = \tau + 1)\right] + \varepsilon_{ik,t} \tag{1.20}$$

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 1.D.3, 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 1.D.9 we re-estimate model (1.20), sector-by-sector, by decade. Compared to (1.20), 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 almost all sectors and all but a few display significant coefficients after the 1980s.

1.D.4 Further Additional Results

1.D.4.1 Out-Migration and the Volume of Innovation

The main analysis concentrates on the effect of knowledge exposure on the direction of innovation. Knowledge exposure leverages variation in specialization across US counties and bilateral flows between UK districts and US counties. In this section, we briefly comment on the effect of out-migration on innovation's *volume*.

We estimate variations on the following model:

$$\text{Patents}_{i,t} = \alpha_i + \alpha_t + \beta \times \text{US Emigrants}_{i,t} + \varepsilon_{i,t} \tag{1.21}$$

where US Emigrants$_{i,t}$ is the total number of emigrants from district i in decade t. As in the main text, we instrument total out-migration flows with the shift-share instruments constructed using railway-based and leave-out immigration shocks. Compared to the model estimated in the main text, endogeneity concerns in (1.21) are severe. However, if the instruments are valid, then the estimated β coefficients measured the causal effect of out-migration on patenting. A perhaps more crucial concern in regression (1.21) is that we do not have information on emigration to countries other than the US. Suppose emigration rates to, say, Australia or Canada (the second and third most common destinations) were correlated with US out-migration. In that case, we may fail to single out the effect of out-migration.

With these caveats in mind, in Table 1.D.6, we report the estimates of regression (1.21). In panel A, columns (1–3), we report the correlation between measured out-migration and patenting, while columns (4–6) and (7–9) display the reduced form association with, respectively, the railway-based and the leave-out instruments. In panel B we report the 2SLS estimates. We find that the contemporaneous effect of out-migration on innovation is negative. This is reasonable given that out-migration entails a loss of human capital, which, in the light of the selection analysis, was probably relatively skilled and is consistent with the "brain drain" literature. Once we lag emigration by one decade, however, we find a positive effect. This sign reversal is robust across the two instruments in the reduced form and the two-stage least-square estimates. It is tempting to interpret it as evidence of "brain drain", that is, that return innovation increases the volume of innovation (Docquier and Rapoport, 2012). While the results are consistent with this interpretation, they are not conclusive because of the caveats that underlie this exercise.

1.D.4.2 Assortative Matching

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 $\mathbf{P}_{j,t} = \{p_{1j,t}, \ldots, p_{Nj,t}\}$ denote the patent portfolio of county j at 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:

$$\text{Innovation Similarity}_{ij,t} \equiv \frac{\mathbf{P}_{i,t}^{\top}\mathbf{P}_{j,t}}{\|\mathbf{P}_{i,t}\| \cdot \|\mathbf{P}_{j,t}\|} = \frac{\sum_k p_{ki,t}p_{kj,t}}{\sqrt{\sum_k p_{ki,t}^2}\sqrt{\sum_k p_{kj,t}^2}} \leq 1 \tag{1.22}$$

which is a simple cosine similarity. The similarity measure returns value 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:

$$\text{Emigrants}_{i\to j,t} = \alpha_{i\times j} + \alpha_t + \beta \times \text{Innovation Similarity}_{ij,t} + X_{ij,t}^{\top}\Gamma + \varepsilon_{ij,t} \tag{1.23}$$

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. The coefficient β thus yields the correlation between the similarity of innovation activity and migration flows. The dependent variable is measured in logs, and standard errors are two-way clustered by district and county. Under sorting, one would expect $\hat{\beta} > 0$.

We test this prediction in Table 1.D.2. We find no correlation between the similarity of innovation portfolios across county districts and the migration flow between them. This holds irrespective of whether we take the contemporaneous similarity (columns 1–2) or if we lag by one (columns 3–4) or two (columns 5–6) decades. Notably, the standardized beta coefficient of the innovation similarity term is always minimal in magnitude. This suggests that assortative matching based on innovation similarity between origin and destination places is probably not a significant threat to a causal interpretation of our estimates. This notwithstanding, since the similarity of innovation portfolios is measured with error, we do not claim that we can exclude it *tout court*.

1.D.4.3 District-Level Return Knowledge Exposure

To explore the effect of return migration on the return innovation result, we leverage individual-level data and study the possible heterogeneous treatment effect of neighborhood out-migration depending on whether the US migrant return or does not return to the UK. In this section, instead, we construct a measure of "return knowledge exposure" at the district level and show that the association between the baseline knowledge exposure variable and patenting remains significant upon including the return measure. Moreover, we fail to detect any significant association between the return metric and patenting activity. Overall, these results confirm that it is unlikely that return migration is the main mechanism underlying the return innovation result.

The return knowledge exposure metric is analogous to the baseline one, except that it leverages return migration flows:

$$\text{Return Knowledge Exposure}_{ik,t} \equiv \sum_j \left(\frac{\text{Patents}_{jk,t}}{\text{Patents}_{j,t}} \times \text{Return Migrants}_{j \to i,t} \right) \tag{1.24}$$

where Return $\text{Migrants}_{j \to i,t}$ is the number of return emigrants from county j to district i in decade t. We then estimate variants of the following model:

$$\text{Patents}_{ik,t} = \alpha_{i \times t} + \alpha_k + \beta \times \text{Knowledge Exposure}_{ik,t} + \gamma \times \text{Return Knowledge Exposure}_{ik,t} + \varepsilon_{ik,t} \tag{1.25}$$

where we are primarily interested in the relative sizes and significance of the β and γ coefficients.

We report the regression results (1.25) in Table 1.D.10. The baseline regression is displayed in column (1); in columns (2) and (3), we incrementally introduce technology-by-decade and district-by-class fixed effects; in columns (4) and (5) we include the exposure terms with a one-decade and a two-decade lags; finally, in column (6) we include all lagged variables. The Table conveys a very clear message. The estimated β coefficient is always positive, significant, and large in magnitude. By contrast, the estimate of γ is mostly non-significant and even when it is, its magnitude is much smaller than that of the corresponding β. Taken together, we view this exercise as providing additional evidence that return migration is unlikely to be the main driver of the return innovation result.

1.D.5 Tables

TABLE 1.D.1: Zero-Stage Regressions Between Immigrant Shares and Railway Access

	Baseline	Excluding States in...			
	(1)	(2) Northeast	(3) Midwest	(4) South	(5) West
$I^{\text{Rail}}_{t-1} \times$ Immigrant Flow$_{t-1}$	0.372***	0.399***	0.198**	0.461**	0.252**
	(0.102)	(0.111)	(0.097)	(0.198)	(0.101)
I^{Rail}_{t-1}	0.845	4.014	-4.288	-17.024***	3.374
	(2.765)	(2.775)	(2.798)	(5.918)	(2.775)
County FE	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes
N. of Counties	2759	2543	1742	1513	2479
N. of Observations	17308	15803	10919	9222	15980
R^2	0.905	0.903	0.921	0.880	0.915
Mean Dep. Var.	79.842	74.284	55.019	132.174	72.101

Notes. This table reports the results of the zero-stage regressions that we estimate to construct the railway-based county-level immigration shocks. This table largely replicates Sequeira *et al.* (2020). The unit of observation is a county observed at a decade frequency between 1870 and 1930. The dependent variable is the share of the foreign-born population. The main dependent variable is an interaction between the one-decade-lagged national inflow of immigrants and an indicator variable that returns value one if the county was connected to the national railway network in the previous decade and zero otherwise. The regressions also control for the railway indicator, the lagged share of foreign-borns, an interaction between lagged national industrial production and the railway indicator, an interaction between lagged GDP and the railway indicator, population density, the share of the population living in urban centers, and an interaction between the share of the urban population and the national inflow of immigrants. The parameter restriction imposed by the instrument's logic requires that the railway indicator's coefficient be non-positive. In column (1), the sample is the universe of counties; in columns (2), (3), (4), and (5), we drop states in, respectively, the North-East, Midwest, South, and West Census Bureau regions. Each regression includes county and decade fixed effects. Standard errors, clustered at the county level, are displayed in brackets.

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

TABLE 1.D.2: British Immigrants Assortative Matching Across US Counties

	Contemporaneous		10 Years Lag		20 Years Lag	
	(1)	(2)	(3)	(4)	(5)	(6)
Innovation Similarity$_t$	0.083	201.876				
	(2.847)	(155.968)				
Innovation Similarity$_{t-1}$			0.419	333.940		
			(2.800)	(205.476)		
Innovation Similarity$_{t-2}$					1.370	-172.428
					(2.485)	(218.951)
District-County FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
N. of District-Counties	1743283	32176	1665941	31383	1505060	29948
N. of Observations	9029476	88636	7266084	86789	5476833	83220
Sample	All	Non-Zero	All	Non-Zero	All	Non-Zero
R^2	0.473	0.675	0.553	0.675	0.662	0.676
Mean Dep. Var.	0.022	1.617	0.027	1.635	0.034	1.670
Std. Beta Coef.	0.000	0.000	0.000	0.000	0.000	0.000

Notes. This table reports the association between the similarity of innovation activity and migration flows between US counties-UK districts pairs. The unit of observation is a county-district pair, observed at a decade frequency between 1870 and 1920. The dependent variable is the number of emigrants that leave the given district and settle in the given county. The independent variable is the similarity of the innovation portfolios between the county and the district. The innovation similarity is computed as the cosine distance of the respective patent portfolios over the decade. Columns (1), (3), and (5) report results for the universe of county-district pairs; columns (2), (4), and (6) restrict to pairs with non-zero migration flows. Columns (1) and (2) estimate the contemporaneous correlation; in columns (3) and (4), innovation similarity appears with a one-decade lag; in columns (5) and (6), it is included with a two-decade lag. Regressions include district-by-county and decade fixed effects. Standard errors, clustered at the district level, are reported in parentheses.

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

TABLE 1.D.3: Double and Triple Differences Effect of The Smoot-Hawley Act on Innovation

	Double Differences		Triple Differences		
	(1) Not Protected	(2) Protected	(3) +10%	(4) +30%	(5) +50%
Knowledge Exposure × Post	-0.040	-0.463**			
	(0.063)	(0.219)			
Knowledge Exposure × Post × Protected (+10%)			-0.469***		
			(0.078)		
Knowledge Exposure × Post × Protected (+30%)				-0.478*	
				(0.269)	
Knowledge Exposure × Post × Protected (+50%)					-0.685***
					(0.206)
Year FE	Yes	Yes	–	–	–
District-Year FE	No	No	Yes	Yes	Yes
District-Class FE	Yes	Yes	Yes	Yes	Yes
Class-Year FE	No	No	Yes	Yes	Yes
N. of District-Class	632	632	632	632	632
N. of Observations	63200	37920	101120	101120	101120
R^2	0.653	0.563	0.713	0.713	0.713
Mean Dep. Var.	2.125	1.260	1.801	1.801	1.801
Std. Beta Coef.	-0.004	-0.061	-0.054	-0.055	-0.079

Notes. This table reports the estimated effect of an increase in the US tariff rate on innovation in Britain. The unit of observation is a district-technology class pair observed at a yearly frequency between 1920 and 1939. The dependent variable is the number of patents by district technology class. In columns (1–2), the independent variable is the interaction between knowledge exposure over 1910–1920 and a post-reform (1930) indicator variable. The regression in column (1) is estimated over technology classes not targeted by the Act; in column (2), we focus on classes that the Act targets. We define a class as "targeted" if its average tariff rate increases by more than 50% after the Smoot-Hawley Act. In columns (3), (4), and (5), the treatment interacts the previous one with an indicator that returns value one for technology classes whose tariff rates increases by more than, respectively, 10%, 30%, and 50% after 1930. Regressions (1–2) are thus double-difference designs; regressions (3–5) are triple-difference designs. Consequently, in columns (1–2), we include district-by-class and year fixed effects, while in columns (3–5), we add district-by-year and technology-by-year fixed effects. Standard errors, clustered at the district level, are reported in parentheses.

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

TABLE 1.D.4: Selection of US Emigrants Compared to the Rest of the British Population

	Non Migrants	Emigrants			Return Migrants		
	(1) Mean	(2) Mean	(3) Difference	(4) Std. Err.	(5) Mean	(6) Difference	(7) Std. Err.
Panel A. Employment (Dependent variable = 1 if individual employed in:)							
Agriculture	0.281	0.271	-0.009***	(0.001)	0.252	-0.028***	(0.002)
Chemicals	0.008	0.008	-0.001**	(0.000)	0.009	0.001**	(0.000)
Construction	0.141	0.142	0.001	(0.001)	0.145	0.004***	(0.002)
Engineering	0.138	0.138	-0.000	(0.001)	0.148	0.010***	(0.002)
Liberal Profession	0.035	0.032	-0.002***	(0.000)	0.036	0.002*	(0.001)
Metallurgy	0.029	0.034	0.005***	(0.001)	0.032	0.003***	(0.001)
Other Manufacturing	0.074	0.074	-0.000	(0.001)	0.074	-0.001	(0.001)
Public Administration	0.030	0.028	-0.001***	(0.000)	0.032	0.002***	(0.001)
Textiles	0.090	0.099	0.009***	(0.001)	0.082	-0.008***	(0.001)
Trade	0.072	0.079	0.007***	(0.001)	0.078	0.007***	(0.001)
Transport	0.097	0.090	-0.007***	(0.001)	0.103	0.006***	(0.001)
Utilities	0.007	0.006	-0.000	(0.000)	0.009	0.002***	(0.000)
Panel B. Region of Residence (Dependent variable = 1 if individual lives in:)							
East	0.102	0.086	-0.016***	(0.001)	0.089	-0.014***	(0.001)
East Midlands	0.065	0.057	-0.007***	(0.000)	0.058	-0.006***	(0.001)
London	0.132	0.129	-0.003***	(0.001)	0.139	0.006***	(0.001)
North East	0.067	0.070	0.003***	(0.000)	0.070	0.003***	(0.001)
North West	0.179	0.194	0.015***	(0.001)	0.199	0.020***	(0.001)
South East	0.120	0.110	-0.009***	(0.001)	0.117	-0.003***	(0.001)
South West	0.063	0.085	0.022***	(0.001)	0.065	0.002***	(0.001)
Wales	0.070	0.064	-0.006***	(0.000)	0.069	-0.001	(0.001)
West Midlands	0.114	0.110	-0.004***	(0.001)	0.108	-0.006***	(0.001)
Yorkshire	0.088	0.094	0.006***	(0.001)	0.087	-0.001*	(0.001)

Notes. This table compares observable individual characteristics of US emigrants with the rest of the British population. In each row, we define a dummy variable equal to one for individuals in the given employed in the given sector (Panel A) or residing in the given division (Panel B) and compute the average for non-migrants (column 1), emigrants (column 2), and return migrants (column 5). Columns (3) and (6) report the difference between columns (1) and, respectively, columns (2) and (5). Robust standard errors are reported in columns (4) and (7).

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

TABLE 1.D.5: Selection of British Immigrants Compared to the Rest of the US Population

	US Population	Immigrants			Return Migrants		
	(1) Mean	(2) Mean	(3) Difference	(4) Std. Err.	(5) Mean	(6) Difference	(7) Std. Err.
Panel A. Employment (Dependent variable = 1 if individual employed in:)							
Agriculture	0.215	0.121	-0.094***	(0.001)	0.129	-0.086***	(0.001)
Chemicals	0.006	0.010	0.004***	(0.000)	0.006	-0.001***	(0.000)
Construction	0.044	0.096	0.052***	(0.001)	0.087	0.042***	(0.001)
Engineering	0.434	0.199	-0.235***	(0.001)	0.262	-0.172***	(0.002)
Liberal Profession	0.042	0.078	0.036***	(0.001)	0.063	0.021***	(0.001)
Other Manufacturing	0.076	0.159	0.083***	(0.001)	0.148	0.072***	(0.001)
Public Administration	0.014	0.009	-0.005***	(0.000)	0.008	-0.006***	(0.000)
Textiles	0.015	0.076	0.061***	(0.001)	0.080	0.066***	(0.001)
Trade	0.069	0.104	0.035***	(0.001)	0.092	0.023***	(0.001)
Transport	0.056	0.087	0.031***	(0.001)	0.085	0.029***	(0.001)
Utilities	0.028	0.059	0.031***	(0.001)	0.041	0.013***	(0.001)
Panel B. Region of Residence (Dependent variable = 1 if individual lives in:)							
East North Central	0.205	0.210	0.005***	(0.001)	0.192	-0.014***	(0.001)
East South Central	0.087	0.008	-0.079***	(0.000)	0.008	-0.078***	(0.000)
Mid Atlantic	0.208	0.350	0.143***	(0.001)	0.365	0.157***	(0.002)
Mountain	0.030	0.058	0.028***	(0.000)	0.062	0.032***	(0.001)
New England	0.068	0.165	0.097***	(0.001)	0.187	0.120***	(0.001)
Pacific	0.054	0.101	0.047***	(0.001)	0.072	0.018***	(0.001)
South Atlantic	0.130	0.026	-0.104***	(0.000)	0.023	-0.107***	(0.000)
West North Central	0.123	0.067	-0.055***	(0.001)	0.077	-0.045***	(0.001)
West South Central	0.095	0.014	-0.081***	(0.000)	0.013	-0.082***	(0.000)

Notes. This table compares observable individual characteristics of British immigrants with the rest of the US population. In each row, we define a dummy variable equal to one for individuals in the given employed in the given sector (Panel A) or residing in the given division (Panel B) and compute the average for non-migrants (column 1), immigrants (column 2), and return migrants (column 5). Columns (3) and (6) report the difference between columns (1) and, respectively, columns (2) and (5). Robust standard errors are reported in columns (4) and (7).

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

Table 1.D.6: Association Between Out-Migration and the Volume of Innovation

	Dependent Variable: Number of Patents								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. OLS Estimates									
	Measured US Emigration			Railway Instrument			Leave-Out Instrument		
US Emigrants$_t$	-1.453***								
	(0.303)								
US Emigrants$_{t-1}$		0.951***							
		(0.258)							
US Emigrants$_{t-2}$			-0.702						
			(0.646)						
Railway-Predicted Emigrants$_t$				-0.141***					
				(0.038)					
Railway-Predicted Emigrants$_{t-1}$					0.283***				
					(0.074)				
Railway-Predicted Emigrants$_{t-2}$						0.212**			
						(0.107)			
Leaveout-Predicted Emigrants$_t$							-0.307***		
							(0.093)		
Leaveout-Predicted Emigrants$_{t-1}$								0.473***	
								(0.126)	
Leaveout-Predicted Emigrants$_{t-2}$									0.162
									(0.134)
Std. Beta Coef.	-0.228	0.125	-0.086	-0.185	0.284	0.196	-0.095	0.106	0.034
Panel B. Two-Stage Least-Square Estimates									
	Railway Instrument			Leave-Out Instrument			Overidentified 2SLS		
US Emigrants$_t$	-1.479***			-2.351***			-1.433***		
	(0.372)			(0.396)			(0.389)		
US Emigrants$_{t-1}$		1.670***			1.424***			1.662***	
		(0.458)			(0.366)			(0.455)	
US Emigrants$_{t-2}$			-28.128			59.484			-18.895
			(30.883)			(377.778)			(15.412)
Std. Beta Coef.	-0.232	0.219	-3.441	-0.368	0.187	7.276	-0.224	0.218	-2.311
K-P F-stat	71.165	215.018	0.805	10.054	70.998	0.026	43.918	106.788	0.933
Sargan-Hansen J							3.473	2.735	0.397
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of District	620	620	618	618	618	618	618	618	618
N. of Observations	2474	1858	1236	2472	1854	1236	2472	1854	1236
Mean Dep. Var.	179.519	221.154	288.871	179.672	221.617	288.871	179.672	221.617	288.871

Notes. This table reports the association between US out-migration and the number of patents. The unit of observation is a district, at a decade frequency between 1880 and 1939. The dependent variable is the number of patents. In Panel A, we report the association with measures out-migration (columns 1–3), the reduced-form railway instrument (columns 4–6), and the reduced-form leave-out instrument (columns 7–9). In Panel B, we report the two-stage least-square estimates of the railway (columns 1–3), leave-out (columns 4–6), and combined (columns 7–9) instruments. All regressions include district and decade fixed effects; standard errors are clustered at the district level and are displayed in parentheses. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

TABLE 1.D.7: Estimated Effect of the Influenza Pandemic on US Innovation

	Double Differences		Triple Differences			
	(1) Level	(2) Level	(3) Level	(4) Level	(5) Share	(6) Share
Excess Deaths × Post	3.120***					
	(0.751)					
1(Q. of Excess Deaths > 75) × Post		1.474***				
		(0.504)				
Excess Deaths × Post × Pharma			2.641***		0.063**	
			(0.678)		(0.032)	
1(Q. of Excess Deaths > 75) × Post × Pharma				1.311***		0.042***
				(0.451)		(0.016)
County FE	Yes	Yes	–	–	–	–
Year FE	Yes	Yes	–	–	–	–
County-Year FE	–	–	Yes	Yes	Yes	Yes
County-Class FE	–	–	Yes	Yes	Yes	Yes
Class-Year FE	–	–	Yes	Yes	Yes	Yes
N. of County-Class	1272	1272	21624	21624	21624	21624
N. of Observations	50880	50880	864960	864960	864960	864960
Classes in Sample	Pharma	Pharma	All	All	All	All
R^2	0.405	0.405	0.683	0.683	0.114	0.114
Mean Dep. Var.	0.991	0.991	0.534	0.534	0.077	0.077
Std. Beta Coef.	0.296	0.083	0.191	0.041	0.032	0.009

Notes. This table reports the effect of exposure to the Great Influenza Pandemic (1918–1919) on innovation in the US. The units of observation are counties (columns 1–2) and county-class pairs (columns 3–6). Units are observed at a yearly frequency between 1900 and 1939. In columns (1–4), the dependent variable is the number of patents granted; in columns (5–6), the dependent variable is the number of pharmaceutical patents divided by the total number of patents granted. In column (1), a post-influenza indicator is interacted with a measure of excess mortality, namely, the ratio between the average number of deaths during the pandemic (1918–1919) and the three previous years. In column (2), the treatment interacts the post-influenza indicator with a dummy variable equal to one for counties in the top quartile of the excess deaths distribution. In columns (3) and (5), the treatment interacts the excess deaths measure with a post-influenza dummy and an indicator variable for pharmaceutical patents; in columns (4) and (6), the excess deaths variable is coded as binary, and returns value one for counties in the top quartile of the excess mortality distribution. In columns (1–2), regressions include county and year fixed effects; in columns (3–6), regressions include county-by-year, county-by-technology class, and technology class-by-year fixed effects. Standard errors, clustered at the district level, are reported in parentheses.

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

TABLE 1.D.8: Double and Triple Differences Effect of Synthetic Shocks on US Innovation

	Baseline	Excluding States in ...				Innovation Shock Treshold		
	(1)	(2) Northeast	(3) Midwest	(4) South	(5) West	(6) 0.1%	(7) 1%	(8) 10%
Innovation Shock × Post	32.727***	40.194***	26.937***	32.514***	33.336***	94.663***	18.948***	3.958***
	(2.610)	(4.018)	(2.613)	(2.633)	(2.802)	(8.056)	(1.384)	(0.207)
County-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-by-Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of County-Class	51263	47376	32849	28727	46169	51264	51250	51097
Number of Observations	2101783	1942416	1346809	1177807	1892929	2101824	2101250	2093047
Mean Dep. Var.	0.772	0.511	0.730	1.241	0.765	0.772	0.752	0.615

Notes. This table reports the effect of synthetic innovation shock on US innovation. These coefficients are not interpreted as causal but as evidence that synthetic shocks capture relevant variation in county-technology-specific patenting activity. The unit of observation is a county-technology class pair observed at a yearly frequency between 1900 and 1939. The baseline treatment is an interaction between an innovation shock and a post-shock indicator. An innovation shock occurs when the residualized patenting activity in a given county technology is in the top 0.5% of the overall distribution of residualized values. Because the setting is staggered, all regressions are estimated using the methodology of Borusyak *et al.* (2021). Column (1) reports the estimate for the entire panel of counties; in columns (2), (3), (4), and (5), we exclude counties in, respectively, the North-East, Midwest, South, and West Census Bureau regions. In columns (6), (7), and (8), instead, we consider different thresholds for the definition of innovation shocks at the top 0.1%, 1%, and 10% of the overall distribution of residualized patents, respectively. All regressions include county-by-year, county-by-technology class, and technology class-by-year fixed effects. Standard errors, clustered at the county level, are reported in parentheses.

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

TABLE 1.D.9: Long-Run Sector Correlation Between Knowledge Exposure and Innovation

	1940s	1950s	1960s	1970s	1980s	1990s	2000s
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable: Number of Patents in:							
Agriculture	3.183***	2.654***	1.914***	1.909**	1.750**	0.647	0.516
	(0.751)	(0.617)	(0.602)	(0.790)	(0.823)	(0.609)	(0.641)
Building	6.065***	4.748***	5.312***	3.918***	4.540***	3.265***	1.804*
	(1.181)	(1.085)	(1.093)	(0.942)	(1.031)	(1.010)	(1.045)
Chemistry	23.553***	15.352***	23.228***	29.817***	11.953**	5.042	2.337
	(7.640)	(5.371)	(6.691)	(12.031)	(5.354)	(6.310)	(6.625)
Electricity	32.018***	31.842**	22.787***	16.405**	8.087	2.124	2.665
	(8.010)	(13.812)	(7.136)	(8.273)	(6.442)	(7.260)	(7.732)
Engineering	9.870***	7.969***	9.549***	8.374***	4.229***	1.167	0.702
	(1.705)	(1.563)	(1.754)	(2.008)	(1.406)	(1.666)	(1.720)
Engines, Pumps	6.551***	7.387***	5.926***	8.172*	4.480*	0.200	0.784
	(2.090)	(2.559)	(1.960)	(4.461)	(2.299)	(2.228)	(2.263)
Food	10.948***	9.570***	8.089***	10.390***	4.173**	0.495	-0.291
	(1.995)	(2.262)	(2.486)	(3.452)	(1.830)	(2.340)	(2.344)
Health, Amusement	4.430***	4.959***	3.988***	7.074***	6.700***	4.318***	5.210***
	(1.307)	(1.405)	(1.419)	(2.111)	(1.536)	(1.526)	(1.708)
Instruments	14.172***	14.338***	15.127***	14.236***	10.658***	4.819	4.079
	(2.937)	(3.244)	(3.480)	(4.101)	(2.905)	(3.072)	(3.599)
Lightning, Heating	11.553***	8.118***	8.113***	5.359***	3.534***	1.581	0.513
	(2.154)	(1.447)	(1.774)	(1.397)	(1.270)	(1.528)	(1.476)
Metallurgy	18.803***	9.443***	13.905***	10.698***	6.346**	1.834	0.849
	(3.888)	(2.573)	(3.541)	(3.493)	(2.722)	(3.110)	(3.236)
Personal Articles, Furniture	6.810***	6.784***	5.250***	2.813***	1.599**	1.367*	1.674**
	(1.014)	(1.175)	(0.811)	(0.755)	(0.803)	(0.811)	(0.821)
Printing	6.914***	7.830***	8.202***	6.030***	3.245***	1.984*	1.455
	(1.226)	(1.341)	(1.573)	(1.205)	(1.045)	(1.190)	(1.313)
Separating, Mixing	7.892***	7.493***	7.633***	8.032***	5.290***	1.602	1.166
	(1.707)	(1.508)	(1.681)	(1.922)	(1.458)	(1.557)	(1.696)
Shaping	9.833***	7.901***	8.591***	7.795***	3.555***	1.156	0.426
	(1.584)	(1.377)	(1.520)	(1.629)	(1.214)	(1.421)	(1.491)
Ships, Aeronautics	8.433***	8.800***	9.757***	6.624***	3.441***	1.319	0.905
	(1.032)	(1.156)	(1.379)	(1.193)	(0.946)	(1.081)	(1.161)
Textiles	14.865***	12.841***	11.039***	10.100***	3.649***	0.752	0.475
	(2.044)	(1.653)	(1.760)	(2.000)	(1.263)	(1.464)	(1.496)
Transporting	5.102***	4.368***	3.974***	2.988**	1.704*	0.471	0.147
	(1.157)	(1.194)	(1.231)	(1.391)	(0.934)	(1.123)	(1.167)
Number of Districts	632	632	632	632	632	632	632

Notes. This table reports the correlation between knowledge exposure in the years 1900–1930 and subsequent patenting activity by sector. For each class displayed in the rows, we estimate a model that interacts knowledge exposure with decade dummies, and we report the coefficients for each decade in the respective column. The 2010s decade serves as the baseline category. All regressions include district and decade fixed effects. Robust standard errors are displayed in parentheses.

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

TABLE 1.D.10: Association Between Return and Outward Knowledge Exposure

	Dep. Var.: Number of Patents					
	(1)	(2)	(3)	(4)	(5)	(6)
Knowledge Exposure$_t$	2.230***	2.183***	1.368***			2.366**
	(0.565)	(0.572)	(0.392)			(0.862)
Return Knowledge Exposure$_t$	-0.048	-0.019	-0.167			0.359
	(0.031)	(0.034)	(0.114)			(0.247)
Knowledge Exposure$_{t-1}$				2.271***		0.976**
				(0.592)		(0.447)
Return Knowledge Exposure$_{t-1}$				0.124		0.081
				(0.087)		(0.118)
Knowledge Exposure$_{t-2}$					1.152**	0.334
					(0.407)	(0.522)
Return Knowledge Exposure$_{t-2}$					0.278**	-0.268**
					(0.125)	(0.116)
District-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
Technology Class FE	Yes	–	–	Yes	Yes	Yes
Class-Decade FE	No	Yes	Yes	No	No	No
District-Class FE	No	No	Yes	No	No	No
N. of District-Class	11376	11376	11376	11376	11376	11365
N. of Observations	45464	45464	45464	34115	22747	22705
R^2	0.637	0.643	0.844	0.638	0.627	0.647
Mean Dep. Var.	9.869	9.869	9.869	12.175	15.868	15.891
Std. Beta (KE)	1.707	1.672	1.048	1.621	0.782	
Std. Beta (Return KE)	0.340	0.332	0.208	0.248	0.114	

Notes. This table reports the association between innovation and the baseline measure of knowledge exposure, accounting for return knowledge exposure. The unit of observation is a district-technology class pair, observed at a decade frequency between 1880 and 1920. The dependent variable is the number of patents by district-technology decade. Return knowledge exposure is constructed by interacting county-level specialization with district-county return migration flows analogously to the baseline knowledge exposure measure. In columns (1) and (4–6), we include district-by-decade and technology class fixed effects. In column (2), we add technology-by-decade fixed effects; the specification in column (3) is saturated. Standard errors, clustered at the district level, are displayed in parentheses. The Table reports the standardized beta coefficient of both the baseline knowledge exposure term and the return knowledge exposure term.

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

TABLE 1.D.11: Heterogeneous Effects of Neighborhood Out-Migration on Innovation

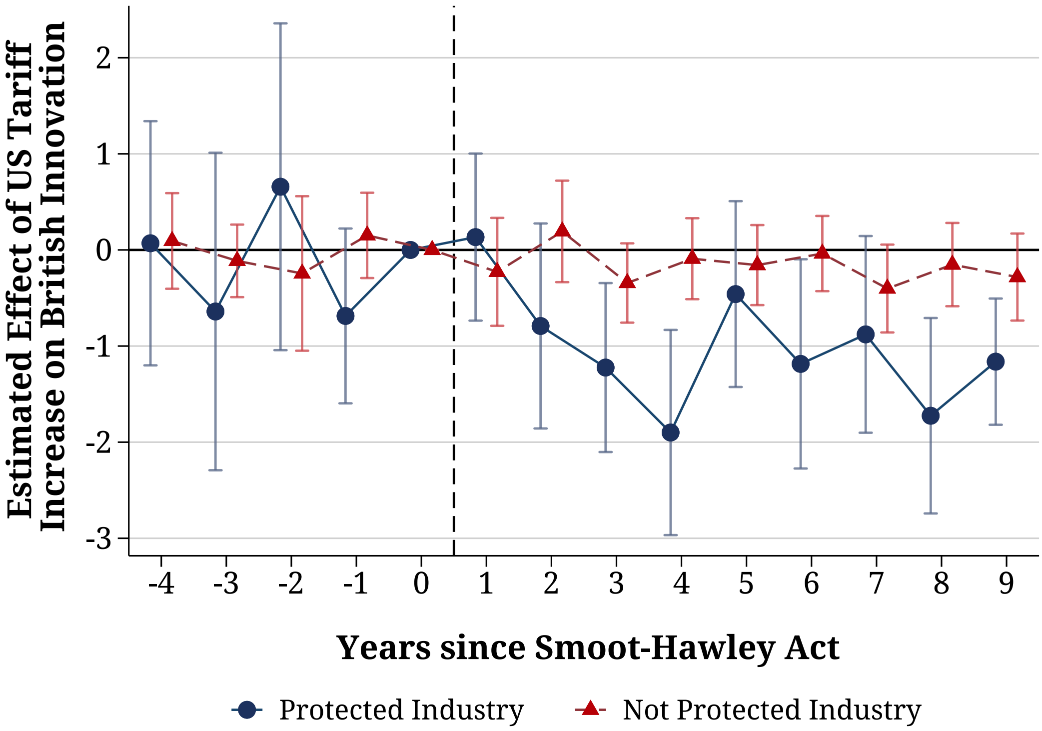
	By Age	By Occupation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Emigrant × Post	0.105*	0.348***	0.595***	0.172*	-0.040	-0.209	-0.030	0.069
	(0.057)	(0.134)	(0.222)	(0.103)	(0.110)	(0.202)	(0.212)	(0.051)
Age ∈ [18, 30) × Emigrant × Post	0.195**							
	(0.092)							
Age ∈ [30, 40) × Emigrant × Post	0.006							
	(0.068)							
Age ∈ [50, 60) × Emigrant × Post	-0.033							
	(0.083)							
Age ≥ 60 × Emigrant × Post	0.040							
	(0.182)							
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	Engineering	Metallurgy	Construction	Textiles	Trade	Pub. Adm.	Agriculture
N. of Individuals	469250	62716	12875	65013	31144	40576	15420	102463
N. of Observations	13608250	1818764	373375	1885377	903176	1176704	447180	2971427
R^2	0.135	0.120	0.097	0.148	0.080	0.097	0.295	0.103
Mean Dep. Var.	0.616	0.871	0.672	0.564	0.548	0.988	0.745	0.253
Std. Beta Coef.	0.002	0.004	0.010	0.003	-0.001	-0.003	-0.000	0.002

Notes. This table reports some heterogeneity analysis on the individual-level effect of neighborhood migration on patenting activity. The units of observation are individuals who are observed yearly between 1900 and 1920. The baseline treatment is an indicator that, for a given individual, returns value one after at least one person that was living in the same neighborhood as the individual migrates to the United States. In column (1), we interact this treatment with age category dummies and normalize the dummy for the age range 40–50 as the baseline category. In columns (2–8), we estimate the baseline double differences model by recorded occupations. Hence, in column (2), we estimate the model only for individuals employed in engineering occupations. All models include individual and year fixed effects. Standard errors are clustered at the district level and are reported in parentheses.

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

1.D.6 Figures

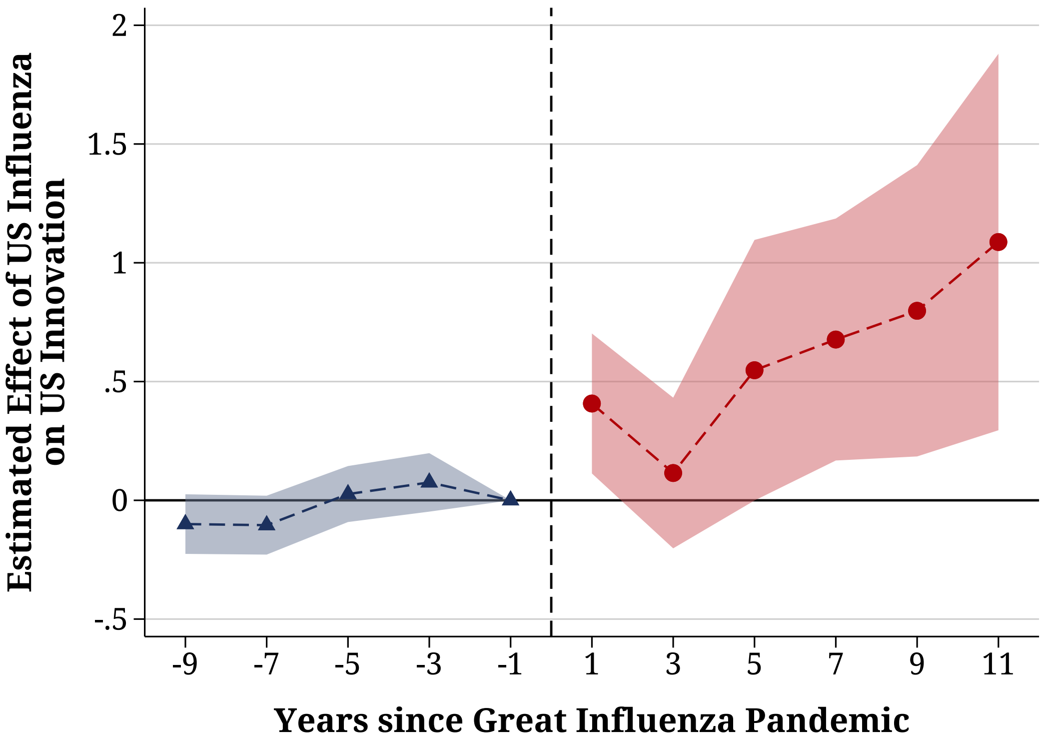
FIGURE 1.D.1: Flexible Double Differences Effect of Tariff Reform on Innovation



Notes. This figure reports the estimated dynamic treatment effects of increased US tariff rates on innovation in Britain. The unit of observation is a district-technology class pair observed at a yearly frequency between 1920 and 1939. The dependent variable is the number of patents. The independent variable is the interaction between knowledge exposure over 1910–1920 and year dummies. The last year before the Reform, 1929, is the baseline category. The blue dots report the estimated treatment effects for technology classes targeted by the Act; the red dots restrict the sample to non-treated technology classes. We define a class as "targeted" if its average tariff rate increases by more than 50% after the Smoot-Hawley Act. Regressions include district-by-class and year fixed effects. Standard errors, clustered at the district level, are reported in parentheses.

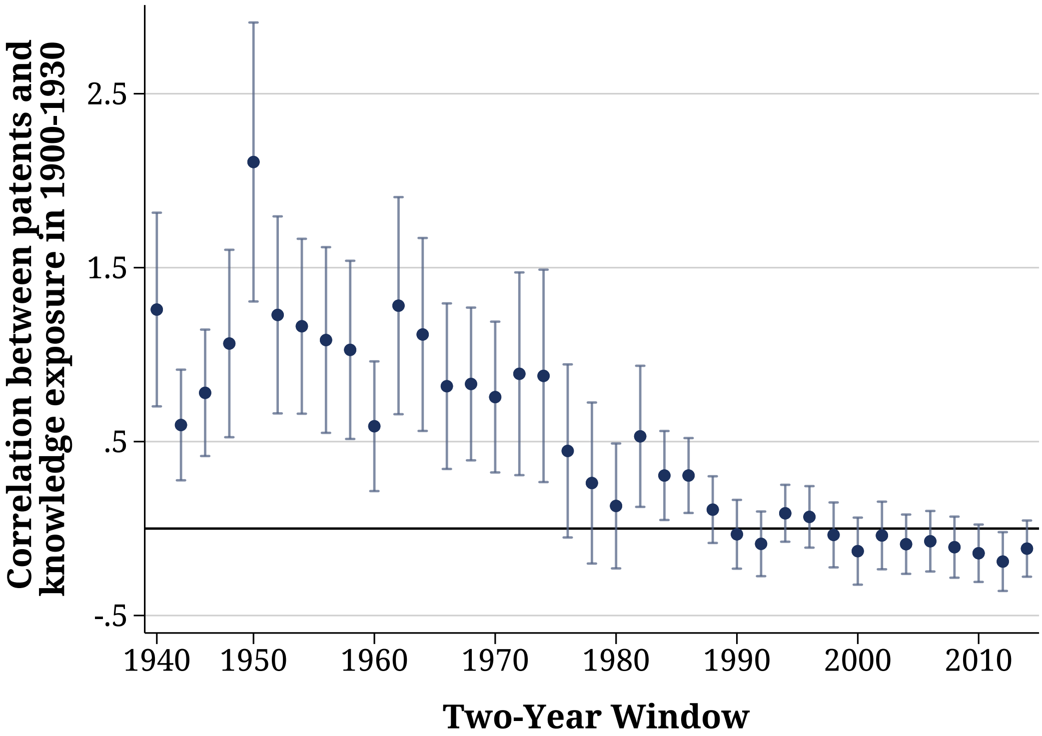
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

FIGURE 1.D.2: Flexible Triple Differences Effect of the Influenza on US Innovation



Notes. These figures report the dynamic treatment effects of exposure to the Great Influenza Pandemic on innovation in the US. The units of observation are county-technology class pairs; units are observed at a yearly frequency between 1900 and 1939. The dependent variable is the number of patents. The treatment is an indicator equal to one for pharmaceutical patents and districts in the top quartile of the excess mortality distribution. The graph displays the interaction coefficients between the treatment and biennial time dummies, where the last dummy before the pandemic—1916–1917—serves as the baseline category. Excess mortality is computed as the average number of deaths during the pandemic over the average number of deaths in the three years before the pandemic. The black dashed line indicates the timing of the treatment. The regression includes county-by-technology class, technology class-by-biennial, and county-by-biennial fixed effects. Standard errors are two-way clustered by district and technology class; bands report 95% confidence intervals.

FIGURE 1.D.3: Long-Run Association Between Knowledge Exposure and Innovation



Notes. This figure reports the correlation between knowledge exposure in the period 1900–1930 and subsequent innovation activity. The unit of observation is a district-technology class pair. Units are observed at a biennial frequency between 1940 and 2015. Each dots report the coefficient of an interaction term between—time-invariant—knowledge exposure and biennial time dummies. The last biennial, 2014–2015, serves as the baseline category. The model includes district-by-technology class and decade fixed effects. Standard errors are clustered at the district level. Bands report 95% confidence intervals.

1.E Robustness Analysis

This section provides details on the technical implementation of the analyses discussed in the main text and briefly describes the exercises we perform to ensure the results' robustness.

1.E.1 Alternative Baseline Specifications

In this section, we list and comment on the alternative specifications of the main equation that we estimate in the main text.

1.E.1.1 Alternative Dependent Variables

In the principal analysis, we use the raw number of patents at varying levels of aggregation as the dependent variable. We thus follow Chen and Roth (2022), who note that under transformations of the dependent variable defined at zero—as would be our case, to avoid dropping zero-patents observations—, the estimates of the average treatment effect are scale-dependent. Since it is common practice in the innovation literature to take the log-transformation, in Table 1.E.1, we show that the results are robust using a battery of alternative transformations.

1.E.1.2 Alternative Definitions of Knowledge Exposure

In Table 1.E.2, 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:

$$\text{Knowledge Exposure}^2_{ik,t} = \sum_j \left(\frac{\text{Patents}_{jk,t}}{\text{Patents}_{j,t}} \times \text{Migrants}_{i \to j,1880} \right) \tag{1.26}$$

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:

$$\text{Knowledge Exposure}^3_{ik,t} = \sum_j \left(\frac{\text{Patents}_{jk,1880}}{\text{Patents}_{j,1880}} \times \text{Migrants}_{i \to j,t} \right) \tag{1.27}$$

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:

$$\text{Knowledge Exposure}^4_{ik,t} = \sum_j \left[\sum_{\tau \le t} \left(\frac{\text{Patents}_{jk,\tau}}{\text{Patents}_{j,\tau}} \right) \times \text{Migrants}_{i \to j,t} \right] \tag{1.28}$$

The idea behind (1.28) is that specialization can be defined in terms of the cumulative number of patents filed before the given period. In Table 1.E.2, we show that all these measures yield quantitatively similar results.

1.E.1.3 Alternative Fixed Effects

In the main text, we report the results for a specification that includes district-by-time and technology class fixed effects. These are intended to capture time-varying unobserved heterogeneity at the district level and time-invariant features of technologies that we do not observe. In Table 1.E.3, we show that the—OLS and 2SLS—results are robust when including a wide array of alternative fixed effects. First, in columns (1) and (6), we report the unconditional correlation between innovation and knowledge exposure. This documents that knowledge exposure alone explains a sizable (30%) share of the variation in patenting activity. Then, in columns (2–5) and (7–10), we incrementally include additional fixed effects and show that the significance and magnitude of the coefficients remain very stable. In particular, in columns (5) and (10), we saturate the model with all couples of fixed effects to non-parametrically control for heterogeneity at the district-time, technology-time, and district-technology levels. The results are confirmed even in this demanding specification.

1.E.2 Instrumental Variable Strategy

This section discusses how we construct the county-level shocks necessary to compute the predicted bilateral flows, as described in section 1.4. We first present the strategy to construct the shocks for the main railway-based instrument. Then, we explain how we compute the shocks for the additional, leave-out instrument.

1.E.2.1 Railway-Based Instrument

The baseline instrument leverages county-level immigration shocks obtained by leveraging variation in the conditional timing when each county was connected to the US railway network. This strategy closely mimics the instrument developed by Sequeira *et al.* (2020) to estimate the long-run effect of immigration in the US.

To construct such shocks, we follow a two-step procedure. We first estimate the following zero-stage equation:

$$
\begin{aligned}
\text{Immigrant Share}_{j,t} = \alpha_j + \alpha_t + \beta \text{Immigrant Share}_{j,t-1} + \gamma I^{\text{Rail}}_{j,t-1} + \\
+ \delta \left(I^{\text{Rail}}_{j,t-1} \times \text{Immigrant Flow}_{t-1} \right) + \zeta \left(\text{Industrialization}_{t-1} \times I^{\text{Rail}}_{j,t-1} \right) + \\
+ \eta \left(\text{GDP Growth}_{t-1} \times I^{\text{Rail}}_{j,t-1} \right) + X^{\intercal}_{j,t-1}\Theta + \varepsilon_{j,t}
\end{aligned}
\tag{1.29}
$$

where (Immigrant Share) is the share of foreign-born individuals, $I^{\text{Rail}}_{j,t}$ is a dummy variable returning value one if county j is connected to the railway network in decade t, and zero otherwise, (Immigrant Flow) is the aggregate immigration inflow computed from Willcox (1928), (Industrialization) is an index of industrial production computed by Davis (2004), and annual average GDP growth is obtained from Maddison (2007) data. The other terms control for confounding factors and non-random connections to the railway network. The term X includes log-population density, lagged urbanization, and an interaction between lagged urbanization and lagged aggregate immigrant flow. The core of the identification strategy that we borrow from Sequeira *et al.* (2020) is to exploit variation generated by the interaction between aggregate immigration inflows and connection to the railway network (δ). The underlying idea is that connection to the railway only induces a larger immigrant inflow if it occurs during a period of high immigration. If this reasoning holds, the estimate of β should be close to zero, and that of δ should be positive. We confirm these predictions in Appendix Table 1.D.1.

We construct a synthetic series of county-level time-varying immigration shocks from equation (1.29) as follows:

$$
\widehat{\text{Immigrant Share}}_{j,t} = \hat{\delta} \left(I^{\text{Rail}}_{j,t-1} \times \text{Immigrant Flow}_{t-1} \right) \tag{1.30}
$$

where $\hat{\delta}$ is simply the OLS estimates from the previous model. We thus generate a set of county-level immigration shocks that are orthogonal to economic development and other characteristics that may induce sorting into the US. Variation, in other words, is solely due to the timing when a county is connected to the railway network.

1.E.2.2 Alternative Instrumental Variable

As further robustness to the railway instrument, we develop a simple leave-out instrument that borrows heavily on the literature that uses shift-share instruments to estimate the effects of immigration (e.g. Card, 2001; Tabellini, 2020). The rationale that underlies this approach is that if assortative matching across counties by British immigrants is the main threat to identification in the baseline regression, then it is possible to leverage the distribution of immigrants from *other* countries to construct county-level immigration shocks that yield consistent estimates because they do not reflect such assortative matching effects.

In practice, let ω_j^M be the share of immigrants from country M that settle in county j in the period 1860-1870, i.e., before the beginning of the analysis years. We then compute the aggregate inflow of immigrants from country M in each subsequent decade and construct the predicted immigrant inflows as

$$\widehat{\text{Immigrant Share}}_{j,t} = \frac{1}{\text{Population}_{j,t}} \sum_{\substack{M \neq \text{UK} \\ M \in \mathcal{M}}} \left(\omega_j^M \times \text{Immigrant Inflow}_t^M \right) \tag{1.31}$$

where $\mathcal{M}$ is a set of origin countries. Both (1.30) and (1.31) yield a set of county-specific immigration shocks that do not conflate immigration patterns of the British. They leverage very different sources of variation, though, which enables us to use the resulting instruments jointly and perform over-identification tests.

We allow multiple sets of origin countries $\mathcal{M}$. The baseline exercise, whose first-stage relevance is shown in Table 1.E.4 and results are displayed in Table 1.E.5, collates all countries except for the UK.[60] To account for possible correlation between British immigrants and those from other nationalities, however, we vary the set of included countries in Table 1.E.6. In particular, we drop all countries in Northern Europe (column 3), which may have been more similar to England and Wales. Moreover, in column (6), we only include non-European countries and show that results hold nonetheless. The coefficients remain relatively stable across all specifications, indicating the possibility that assortative matching may be a quantitatively mild issue.

1.E.2.3 Tests on Instrument Validity

The validity of the shift-share instrument for knowledge exposure that we construct hinges on the exogeneity of the shocks constructed using either (1.30) or (1.31), following Borusyak *et al.* (2022). In practice, they advise conducting two types of falsification tests. First, shocks should be orthogonal to observed county-level characteristics. Second, the instrument should not be systematically correlated with district-level observable variables. The first test provides evidence of the exogeneity of the shocks, while the second should support the exclusion restriction that underlies the instrument.

[60] In Figure 1.E.2 we report binned scatter plots of the association between predicted and actual knowledge exposure using the two instruments.

We perform the first exercise in Figure 1.E.3. Panel (A) displays the correlation of the observed immigration shares with county-level observable characteristics. As expected, immigration is not random as it tends to be concentrated in larger counties, which also display higher patenting activity. In panels (B) and (C), we report the correlation between the predicted immigrant shares using the railway-based and the leave-out approaches, respectively. We fail to detect a statistically significant correlation between the so-constructed immigrant shares and the large majority of county-level observable variables.[61] This provides reassuring evidence in favor of the validity of the instruments.

We report the second exercise in Figure 1.E.4. Each dot displays the correlation between district-level observable variables and actual, railway-based, and leave-out emigration in panels (A), (B), and (C), respectively. Unsurprisingly, districts featuring higher emigration flows are larger, produce more patents, and have a larger share of the population working in agriculture and textiles. On the other hand, synthetic out-migration, whether constructed using the railway or the leave-out shocks, is not correlated with any such variables. Once more, we interpret these results as evidence supporting the validity of the shift-share research design.

1.E.3 Shock Propagation

This section describes the technical definition of the synthetic innovation shocks and exposure to the influenza pandemic, along with two falsification exercises and several sensitivity analyses.

1.E.3.1 Details on the Construction of the Synthetic Shocks

We define a synthetic innovation shock as an unusual deviation from the number of patents granted in a given county, technology class, and year. Formally, we estimate the following fixed-effects regression:

$$\text{Patents}_{jk,t} = \alpha_{j\times k} + \alpha_{k\times t} + \alpha_{j\times t} + \varepsilon_{jk,t} \tag{1.32}$$

where j, k, and t denote a county, technology class, and year respectively, and α is the associated fixed effect. In particular, we include county-by-year fixed effects to remove fluctuations in patenting activity due to, for instance, economic growth. We remove technology-by-year fixed effects to ensure that the shocks do not reflect aggregate changes in the propensity to patent in a given class. Finally, we average out county-by-class fixed effects to remove asymmetries due

[61] Even when the correlation remains significant, the standardized beta coefficient is substantially lower than in the benchmark panel (A).

to initial specialization. We then construct a series of residualized innovation activity from the residuals of (1.32).

In the baseline analysis, we define an innovation shock as an observed residualized patenting activity in the top 0.1% of the overall distribution. Let $\Gamma(\cdot)$ be the cumulative distribution of the residuals of regression (1.32). Then, the set of shocks $\xi(\tau)$, for $\tau = 0.001$, is given by the set $\xi(\tau) = \{\xi \in \text{supp}(\Gamma) \mid \Gamma(\xi) - \Gamma(\tau) \geq 0\}$. In Table 1.E.7, we use two other threshold values of τ (1% and 0.5%). We find that the average treatment effect decreases as τ increases. This is compelling since larger τ flags smaller residualized patenting activity as instances of treatment. In Table 1.D.8, we show the "effect" of synthetic shocks on innovation. This is not a causal effect but rather a measure of the relevance of such shocks. There is a strong and positive association between the number of patents and the period when a shock is observed, and this also holds excluding specific areas (columns 2–5). In columns (6–8), we show that larger levels of τ are associated with a lower increase in patenting.

1.E.3.2 Details on the Construction of the Influenza Shock

To construct exposure to the influenza across counties, we follow Berkes *et al.* (2022). From the mortality statistics collected by the Bureau of Census, we define a metric of excess deaths as the ratio between average deaths during the pandemic (1918–1919) relative to the average in the preceding three years.[62] Formally, we have

$$\text{Excess Deaths}_j = \frac{\frac{1}{2}\sum_{t=1918}^{1919} \text{Deaths}_{j,t}}{\frac{1}{3}\sum_{t'=1915}^{1917} \text{Deaths}_{j,t'}} \tag{1.33}$$

We then code a binary variable equal to one if county j is in the top 25% of the excess deaths distribution to avoid issues of continuous treatment (Callaway and Sant'Anna, 2021).

The baseline estimation equation for US counties is then

$$\text{Patents}_{jk,t} = \alpha_{j\times k} + \alpha_{k\times t} + \alpha_{j\times k} + \delta\,(\text{Excess Deaths}_c \times \text{Pharma}_k \times \text{Post}_t) + \varepsilon_{jk,t} \tag{1.34}$$

where Pharma_k is an indicator variable returning value one if k is pharmaceutical patents, and zero otherwise, and Post_t is an indicator variable returning value one for years after 1918, and zero otherwise. Figure 1.D.2 reports the associated flexible triple differences estimates, which, with no evidence of statistically significant pre-treatment coefficients, suggests that the influenza had a strong, positive, and significant effect on pharmaceutical innovation in the US.

[62] Due to data limitations, this is the pre-pandemic period that maximizes the sample size. Mortality statistics thus allow covering 60% of the US population.

1.E.3.3 Robustness to Synthetic Shock Analysis

We perform two main exercises to ensure that the results using the synthetic shocks are robust. First, in Figure 1.E.7, we ensure that the estimated effect of US innovation shocks on UK innovation remains significant and is quantitatively consistent under different estimators that allow for staggered roll-out of the treatment across units. The estimated ATE remains significant, and its magnitude is preserved under various estimators.

Second, in Table 1.E.7, we vary two margins along which a district is considered to be treated. First, as previously discussed, we consider different thresholds τ (1%, 0.5%, and the baseline 0.1%) above which we flag synthetic innovation shocks. Reassuringly, larger levels τ, which require a lower marginal increase in patenting to flag a synthetic shock, lead to smaller ATEs. This is consistent with the idea that larger innovation shocks in the US should lead to larger innovation shocks in the UK. Second, we vary the threshold of emigration that we impose for a district to be considered exposed to the innovation shock. In our main analysis, we consider a district exposed to the innovation shock in a given county if it is in the top quartile of the distribution of emigration to that county. We consider two additional thresholds (top 50% and top 90%). We find that the baseline result is qualitatively robust to all such thresholds. Moreover, we confirm that larger exposure thresholds lead to larger estimates ATEs. This suggests that the more intense the previous migration tie between a county and a district, the larger the diffusion effect of county-level shocks on district-level innovation.

1.E.3.4 Shock Falsification Checks

The rationale for the analysis discussed in the main text (table 1.3 and Figure 1.5) and thus far is that the influenza only impacted patenting in pharmaceutical patents in the US. If that is the case, then this would ignite an innovation shock that was localized in areas that were more exposed to the influenza, and that could reverberate in the UK to districts whose emigrants had settled in such areas.

We test this assumption in Figure 1.E.6a. Each dot reports an estimated δ coefficient of equation (1.34), except that the treated technology is reported in each row. Thus, the exclusion restriction would require that each coefficient was not statistically different from zero, except for pharmaceuticals. This assumption is confirmed in the data. The ATE for pharmaceuticals is the only one that is positive, significant, and quantitatively large. Figure 1.E.6a thus implies that we expect to observe an increase in pharmaceutical patents only, and only in districts whose emigrants had settled in areas that were more severely exposed to the pandemic.

We test this in Figure 1.E.6b, in which we estimate the baseline triple-difference specification of the main text, except that the treated technology is reported in each row, as before. While estimates are noisier here, we confirm that the estimated ATE for pharmaceuticals is the largest and statistically significant across classes, as expected. Overall, Figure 1.E.6 thus provides convincing evidence that (i) the influenza fostered innovation in pharmaceuticals only in the US, and (ii) that districts whose emigrants had settled in areas that were more severely exposed to the influenza display higher innovation activity in pharmaceuticals.

1.E.4 Tables

TABLE 1.E.1: Knowledge Exposure and Innovation: Alternative Dependent Variables

	Level of Patents					Share of Patents				
	(1) Baseline	(2) $\ln(\cdot)$	(3) $\ln(1+\cdot)$	(4) $\ln(\varepsilon+\cdot)$	(5) $\text{arcsinh}(\cdot)$	(6) Share	(7) $\ln(\cdot)$	(8) $\ln(1+\cdot)$	(9) $\ln(\varepsilon+\cdot)$	(10) $\text{arcsinh}(\cdot)$
Panel A. OLS Estimates										
Knowledge Exposure	1.342***	0.015***	0.067***	0.142***	0.082***	0.005***	0.015***	0.004***	0.169***	0.005***
	(0.143)	(0.002)	(0.007)	(0.016)	(0.009)	(0.001)	(0.002)	(0.000)	(0.020)	(0.001)
R^2	0.772	0.802	0.824	0.766	0.813	0.330	0.625	0.344	0.523	0.334
Std. Beta Coef.	0.299	0.101	0.396	0.495	0.407	0.411	0.139	0.439	0.629	0.418
Panel B. Reduced-Form Estimates										
Knowledge Exposure	0.037***	0.001***	0.001***	0.001***	0.001***	0.000***	0.001***	0.000***	0.001***	0.000***
	(0.007)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.800	0.811	0.816	0.752	0.805	0.347	0.651	0.358	0.510	0.350
Std. Beta Coef.	0.075	0.032	0.043	0.027	0.039	0.046	0.043	0.046	0.022	0.046
Panel C. Two-Stage Least Square Estimates										
Knowledge Exposure	1.224***	0.018***	0.031***	0.034***	0.034***	0.002***	0.018***	0.002***	0.027***	0.002***
	(0.195)	(0.005)	(0.005)	(0.007)	(0.006)	(0.000)	(0.005)	(0.000)	(0.008)	(0.000)
K-P F-stat	109.826	83.266	109.826	109.826	109.826	109.826	83.266	109.826	109.826	109.826
Std. Beta Coef.	0.296	0.116	0.171	0.108	0.153	0.181	0.158	0.181	0.087	0.181
District-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Technology Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of District-Class	11268	8475	11268	11268	11268	11268	8475	11268	11268	11268
N. of Observations	67549	36290	67549	67549	67549	67549	36290	67549	67549	67549
Mean Dep. Var.	10.392	1.795	1.137	-0.005	1.400	0.051	-2.946	0.046	-4.636	0.050

Notes. This table displays the association between innovation and exposure to US knowledge using alternative transformations of the dependent variable. The unit of observation is a district-technology class pair, observed at a decade frequency between 1880 and 1939. In columns (1–5), the dependent variable is the number of patents; in columns (6–10), the dependent variable is the share of patents in a given technology, normalized by the total number of patents. In columns (1) and (6), we do not transform the dependent variable; in columns (2) and (7), we take the log; columns (3) and (8) report the estimates using log(1+), which avoids dropping zeroes; in columns (4) and (9) we take log(0.01+) of the dependent variable; columns (5) and (10) report the estimates using the inverse hyperbolic sine. The main explanatory variable is knowledge exposure. In Panel A, we estimate the correlation through OLS; in Panel B, we report the reduced-form association between the instrument for knowledge exposure and innovation; in Panel C, we display the two-stage least-squares estimates. Each model includes district-by-decade and district-by-technology class fixed effects. Standard errors are reported in parentheses and are clustered at the district level.

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

TABLE 1.E.2: Knowledge Exposure and Innovation: Alternative Independent Variables

	Dependent Variable: N. of Patents				
	(1)	(2)	(3)	(4)	(5)
Knowledge Exposure	1.342***				
	(0.143)				
$\ln(1 + \text{Knowledge Exposure})$		4.175***			
		(0.228)			
Fixed-Emigrants Knowledge Exposure			2.610***		
			(0.300)		
Fixed-Patents Knowledge Exposure				0.063***	
				(0.015)	
Cumulative Knowledge Exposure					0.136**
					(0.067)
District-Decade FE	Yes	Yes	Yes	Yes	Yes
District-Technology Class FE	Yes	Yes	Yes	Yes	Yes
N. of District-Class	11268	11268	11268	11268	11268
N. of Observations	67549	67549	67547	67555	67549
R^2	0.772	0.766	0.770	0.765	0.764
Mean Dep. Var.	10.392	10.392	10.393	10.392	10.392
Std. Beta Coef.	0.299	0.119	0.249	0.104	0.017

Notes. This table displays the association between innovation and exposure to US knowledge, using alternative transformations of knowledge exposure. The unit of observation is a district-technology class pair, observed at a decade frequency between 1880 and 1939. In column (1), we report the baseline estimate. In column (2), we take knowledge exposure in log terms, adding one to avoid dropping zeros since the baseline measure is defined as non-negative. In column (3), we fix bilateral district-county bilateral exposure shares as the number of emigrants from the given district to the given county in the decade 1870-1880. In column (3), instead, we fix county-level specialization as the share of patents in a given field granted in the decade 1870-1880 only. In column (5), for a given decade, we measure specialization as the sum of patents obtained until the end of that decade relative to the total number of patents obtained until the end of that decade. The measure used in column (5) thus considers the cumulative patent stock instead of its decade-on-decade flow. The main explanatory variable is knowledge exposure. Each model includes district-by-decade and district-by-technology class fixed effects. Standard errors are reported in parentheses and are clustered at the district level.

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

TABLE 1.E.3: Knowledge Exposure and Innovation: Alternative Sets of Fixed Effects

	Dependent Variable: N. of Patents									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	10)
Panel A. Correlational Estimates										
	OLS					Poisson				
Knowledge Exposure	2.393***	1.947***	1.936***	1.942***	1.241***	0.038***	0.011***	0.014***	0.014***	0.010***
	(0.212)	(0.194)	(0.184)	(0.191)	(0.149)	(0.004)	(0.002)	(0.003)	(0.003)	(0.002)
R^2	0.284	0.431	0.539	0.547	0.781	0.226	0.718	0.749	0.760	0.864
Std. Beta Coef.	0.533	0.433	0.431	0.432	0.276	0.303	0.084	0.112	0.109	0.080
Panel B. Instrumental Variable Estimates										
	Reduced Form					Two-Stage Least Squares				
$\widehat{\text{Knowlegde Exposure}}$	0.158***	0.080***	0.041***	0.041***	0.033***					
	(0.012)	(0.009)	(0.010)	(0.012)	(0.009)					
Knowledge Exposure						2.053***	1.490***	0.867***	0.779***	1.097***
						(0.157)	(0.167)	(0.192)	(0.208)	(0.271)
R^2	0.104	0.385	0.501	0.509	0.808	0.293	0.111	0.065	0.059	0.028
K-P F-stat						184.700	96.871	169.304	132.033	46.312
Std. Beta Coef.	0.322	0.164	0.083	0.083	0.068	4.196	3.045	1.772	1.591	2.243
District FE	No	Yes	–	–	–	No	Yes	–	–	–
Decade FE	No	Yes	–	–	–	No	Yes	–	–	–
Class FE	No	Yes	Yes	–	–	No	Yes	Yes	–	–
District-Decade FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Class-Decade FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
District-Class FE	No	No	No	No	Yes	No	No	No	No	Yes
N. of District-Class	11268	11268	11268	11268	11268	11268	11250	11250	11250	10081
N. of Observations	67549	67549	67549	67549	67549	67549	67474	65946	65946	59703
Mean Dep. Var.	10.392	10.392	10.392	10.392	10.392	10.392	10.404	10.645	10.645	11.758

Notes. This table displays the association between innovation and exposure to US knowledge. The unit of observation is a district-technology class pair, observed at a decade frequency between 1880 and 1939. The dependent variable is the number of patents. The main explanatory variable is knowledge exposure. In Panel A, in columns (1–5), we estimate the correlation through OLS; in columns (6–10), we estimate the model as a Poisson regression to account for the many zeros in the data; columns (1–5) in Panel B report the reduced-form association between the instrument for knowledge exposure and innovation; columns (6–10) report the two-stages least square estimates. Columns (1) and (6) reports the unconditional regressions; in columns (2) and (7), we include district, decade, and technology class fixed effects; columns (3) and (8) add district-by-decade fixed effects; in columns (4) and (9) we include district-by-decade and class-by-decade fixed effects; models in columns (5) and (10) are saturated with all couples of fixed effects. Standard errors are reported in parentheses and are clustered at the district level.

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

Table 1.E.4: First Stage of the Instrumental Variable Estimation

	Railway-Based (SNQ) Instrument				Leaveout Instrument			
	Baseline	Dropping Districts in...			Baseline	Dropping Districts in...		
	(1)	(2) London	(3) Lancs	(4) S-W	(5)	(6) London	(7) Lancs	(8) S-W
Panel A. Dependent Variable: Bilateral Flows								
SNQ Migrants	0.007***	0.008***	0.007***	0.007***				
	(0.001)	(0.001)	(0.001)	(0.001)				
Leaveout Migrants					0.006***	0.005**	0.005***	0.005***
					(0.002)	(0.002)	(0.002)	(0.002)
District-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of District-Counties	1736040	1653240	1518000	1625640	1786360	1701160	1562000	1672760
N. of Observations	8399666	7999046	7344700	7865506	10403031	9906861	9096450	9741471
Panel B. Dependent Variable: Knowledge Exposure								
SNQ Knowledge Exposure	0.005***	0.005***	0.005***	0.005***				
	(0.000)	(0.000)	(0.000)	(0.000)				
Leaveout Knowledge Exposure					0.169***	0.157***	0.159***	0.145***
					(0.034)	(0.036)	(0.034)	(0.032)
District-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of District-Classes	11322	10782	9900	10602	11304	10764	9882	10584
N. of Observations	56587	53887	49488	52987	67801	64561	59280	63481

Notes. This table reports the first-stage estimates of the two shift-share instruments we propose. In Panel A, the observation units are district-county pairs, observed at a decade frequency between 1870 and 1920 (columns 1–4) and 1930 (columns 5–8). In Panel B, the observation units are district-technology classes, at decade frequency between 1870 and 1920 (columns 1–4) and 1930 (columns 5–8). In columns (1–4), the predicted number of emigrants constructed using the railway-based instrument that leverages shocks *à la* Sequeira *et al.* (2020); in columns (5–8), predicted emigrants are constructed using the leave-out instrument. Columns (1) and (5) report the full-sample estimates; in columns (2) and (6), we exclude districts in the London area; columns (3) and (7) exclude districts in the Lancashire area; in columns (4) and (8) we drop districts in the South-West. In Panel A, all models include district-by-decade and county-by-decade fixed effects; in Panel B, regressions include district-by-decade and technology class-by-decade fixed effects. Standard errors, clustered at the district level, are reported in parentheses.

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

TABLE 1.E.5: Return Innovation Result Using the Leaveout Instrument

	Reduced Form			TSLS			Overidentified TSLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\widehat{\text{Knowledge Exposure}}$	0.007*								
	(0.004)								
$\widehat{\text{Knowledge Exposure}}_{t-1}$		0.018***							
		(0.006)							
$\widehat{\text{Knowledge Exposure}}_{t-2}$			0.029**						
			(0.012)						
Knowledge Exposure				0.093*			0.322***		
				(0.051)			(0.052)		
Knowledge Exposure$_{t-1}$					0.180***			0.082*	
					(0.065)			(0.044)	
Knowledge Exposure$_{t-2}$						0.103**			0.032
						(0.041)			(0.038)
District-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Technology Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of District-Class	11196	11196	11196	11196	11196	11196	11196	11196	11196
N. of Observations	55980	44784	33588	55957	44761	33586	55957	44761	33586
R^2	0.816	0.831	0.850	0.018	-0.006	-0.002	0.054	-0.001	-0.000
Mean Dep. Var.	1.079	1.202	1.312	1.079	1.203	1.312	1.079	1.203	1.312
Std. Beta Coef.	0.007	0.017	0.015	0.051	0.103	0.055	0.175	0.047	0.017
K-P F-stat				27.303	23.391	248.916	62.737	59.305	198.950
Sargan-Hansen J							24.274	6.131	31.636

Notes. This table reports the estimated return innovation effect estimated using the leave-out instrument. The unit of observation is a district-technology class pair, observed at a decade frequency between 1880 and 1939. The dependent variable is the number of patents. The main explanatory variable is knowledge exposure. In columns (1–3), we report the reduced-form association between knowledge exposure constructed using predicted emigration flows using the leave-out instrument and the dependent variable; in columns (4–6), we report the associated two-stage least-squares estimates. In columns (7–9), instead, we exploit the railway and the leave-out instruments to estimate an over-identified instrumental variable regression. This allows us to report the associated Sargan-Hansen J-statistic to test the validity of the over-identifying restrictions. The Sargan-Hansen test does not refute the null that the instruments are valid. Each model includes district-by-decade and district-by-technology class fixed effects. Standard errors are reported in parentheses and are clustered at the district level.

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

TABLE 1.E.6: Return Innovation Result Using the Modified Leaveout Instruments

	Baseline	Excluding Immigrants from...				
	(1)	(2) UK	(3) UK + North Eu.	(4) UK + South Eu.	(5) UK + East Eu.	(6) UK + Europe
Panel A. Second-Stage Estimates						
Knowledge Exposure	1.849***	0.454***	0.589***	0.428***	0.315**	0.487***
	(0.174)	(0.125)	(0.170)	(0.096)	(0.160)	(0.139)
N. of Observations	78876	78876	78876	78876	78876	78876
Mean Dep. Var.	11.768	11.768	11.768	11.768	11.768	11.768
K-P F-statistic		39.267	30.074	346.557	14.672	52.663
Panel B. First-Stage Estimates						
Knowledge Exposure (No Northern Europe + UK)			0.215***			
			(0.039)			
Knowledge Exposure (No Southern Europe + UK)				0.889***		
				(0.048)		
Knowledge Exposure (No Eastern Europe + UK)					0.481***	
					(0.126)	
Knowledge Exposure (No Europe + UK)						3.103***
						(0.428)
N. of Observations		78876	78876	78876	78876	78876
Mean Dep. Var.		3.063	3.063	3.063	3.063	3.063
District-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
Technology Class FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table reports the instrumental variable estimates of the effect of knowledge exposure on innovation using modified versions of the leave-out instrument. The unit of observation is a district-technology class pair, observed at a decade frequency between 1880 and 1939. The dependent variable is the number of patents. The explanatory variable is knowledge exposure. In column (1), we report the OLS correlation. In columns (2–6), we construct predicted bilateral emigrant flows using county-level immigration shocks that exclude immigrants from different parts of the world: in (2), we exclude only immigrants from UK nations; in (3), we exclude the UK immigrants along with those from other Northern Europe countries; in (4), we exclude immigrants from the UK and Southern Europe; in (5), UK and Eastern Europe immigrants are excluded; in (6), we exclude all European immigrants. Panel A reports the second-stage estimates; Panel B reports the associated first-stage estimates. All regressions include district-by-decade and technology class fixed effects. Standard errors, clustered at the district level, are displayed in parentheses.

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

TABLE 1.E.7: Triple Differences Effect of Synthetic Shocks: Varying Thresholds

	Top 1% Synthetic Shocks			Top 0.5% Synthetic Shocks			Top 0.1% Synthetic Shocks		
	(1) Top 50%	(2) Top 75%	(3) Top 90%	(4) Top 50%	(5) Top 75%	(6) Top 90%	(7) Top 50%	(8) Top 75%	(9) Top 90%
Innovation Shock (Above 50%) × Post	0.187**			0.299***			0.617***		
	(0.083)			(0.098)			(0.126)		
Innovation Shock (Above 75%) × Post		0.224***			0.377***			0.617***	
		(0.080)			(0.081)			(0.126)	
Innovation Shock (Above 90%) × Post			0.326			0.825***			0.532***
			(0.269)			(0.229)			(0.175)
District-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-by-Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Counties	189586	362024	426147	217975	381438	434687	431467	431467	445120
Number of Observations	5247	9187	10671	6762	9786	10900	10834	10834	11128
Mean Dep. Var.	1.022	1.106	1.586	1.410	1.3	1.686	2.064	2.064	2.020

Notes. This table displays the effect of US innovation shocks on innovation activity in the UK. The unit of observation is a district-technology class pair observed at a yearly frequency between 1900 and 1939. The dependent variable is the number of patents. The treatment variable is equal to one for district-technology class pairs after a synthetic innovation shock in a given technology class is observed in counties where the district has above k-percentile emigrants. We consider three different thresholds for k: above the median, above the top 25%, and above the top 10%. A synthetic shock is observed whenever the residualized patenting activity in a given county-technology class pair is in the top ℓ-percentile of the residualized patenting activity distribution. We consider three such *ell*: top 1%, in columns (1–3), top 0.5%, in columns (4–6), and top 0.1%, in columns (7–9). Since the treatment timing is staggered, we estimate the models using the imputation estimator developed by Borusyak *et al.* (2021). All models include district-by-year, district-by-technology class, and technology class-by-year fixed effects; standard errors, clustered two-way by district and technology class, are shown in parentheses.

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

Table 1.E.8: Triple Differences Effect of the US Influenza on UK Innovation: Robustness

	Double Differences		Triple Differences				
	(1)	(2)	(3)	(4)	(5) No London	(6) No Lancs	(7) No S/W
Influenza Emigration × Post	0.008* (0.004)						
1(Q. of Influenza Emigration > 75) × Post		0.980** (0.463)					
Influenza Emigration × Post × Pharma			0.004** (0.002)				
1(Q. of Influenza Emigration > 75) × Post × Pharma				0.584*** (0.163)	0.396** (0.139)	0.671*** (0.173)	0.423** (0.156)
District FE	Yes	Yes	–	–	–	–	–
Year FE	Yes	Yes	–	–	–	–	–
District-Year FE	–	–	Yes	Yes	Yes	Yes	Yes
District-Class FE	–	–	Yes	Yes	Yes	Yes	Yes
Class-Year FE	–	–	Yes	Yes	Yes	Yes	Yes
N. of District-Class	631	631	10727	10727	10217	9384	10047
N. of Observations	18930	18930	321810	321810	306510	281520	301410
Classes in Sample	Pharma	Pharma	All	All	All	All	All
R^2	0.544	0.544	0.668	0.668	0.616	0.653	0.679
Mean Dep. Var.	0.927	0.927	0.763	0.763	0.559	0.721	0.706
Std. Beta Coef.	0.082	0.082	0.014	0.016	0.015	0.018	0.010

Notes. This table displays the effect of the Great Influenza Pandemic shock on innovation activity in the UK. In columns (1–2), the observation unit is a district; in columns (3–7), the observation unit is a pair district-technology class; units are observed at a yearly frequency between 1900 and 1939. The dependent variable is the number of patents. In column (1), the treatment variable is an interaction between an influenza exposure term equal to the share of emigrants to counties in the top 25% of the flu-related excess mortality distribution and a post-Influenza indicator; in column (2), we code exposure as a binary variable equal to one for districts in the top 25% of the continuous exposure distribution. In columns (3) and (5–7), the treatment term in column (1) is interacted with an indicator variable for pharmaceutical patents; in column (4), we interact the treatment term in column (2) with the same pharmaceutical indicator. Regressions in (1–4) report full-sample estimates; in columns (5), (6), and (7), instead, we drop districts in the London, Lancashire, and South-West areas, respectively. Regressions in columns (1–2) include district and year fixed effects; regressions in columns (3–7) include district-by-year, technology class-by-year, and district-by-technology class fixed effects. Standard errors, reported in parentheses, are clustered by district in columns (1–2) and two-way by district and technology class in (3–7).

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

TABLE 1.E.9: Double Differences Effect of Neighborhood Emigration: Alternative Threshold

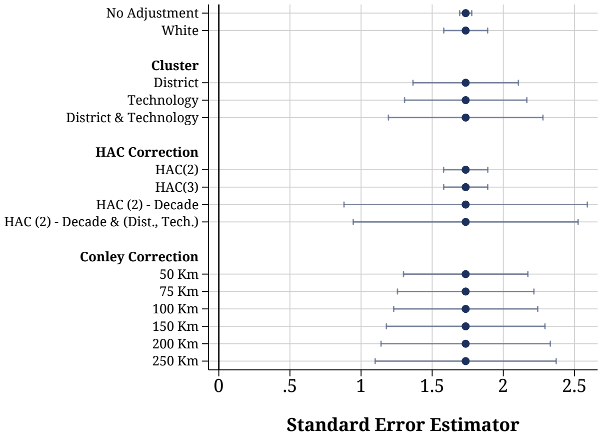
	Baseline Sample				Dropping Individuals in...		
	(1)	(2)	(3)	(4)	(5) London	(6) Lancashire	(7) South-West
Panel A. All Emigrants							
Neighborhood Emigrant × Post	0.120**	0.146**	0.133**	11.846*	0.079	0.142**	0.167***
	(0.059)	(0.068)	(0.059)	(6.144)	(0.065)	(0.062)	(0.061)
Std. Beta Coef.	0.016	0.019	0.018	0.155	0.011	0.020	0.022
Panel B. Only Non-Return Emigrants							
Non-Return Neighborhood Emigrant × Post	0.148***	0.199***	0.160***	14.694**	0.061	0.172***	0.226***
	(0.056)	(0.062)	(0.058)	(6.293)	(0.061)	(0.059)	(0.058)
Std. Beta Coef.	0.019	0.025	0.020	0.186	0.008	0.023	0.028
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	–	Yes	Yes	Yes	Yes	Yes
Parish × Year FE	No	Yes	No	No	No	No	No
Matching	No	No	Yes	No	No	No	No
Sample	Full	Full	Full	Inventors	Full	Full	Full
N. of Individuals	473112	473112	469585	4224	410327	422230	352064
N. of Observations	9462240	9412502	9391700	84480	8206540	8444600	7041280
Mean Dep. Var.	0.890	0.892	0.893	99.716	0.794	0.836	0.893
S.D. Dep. Var.	40.291	40.337	40.351	414.695	37.439	39.126	41.333

Notes. This table reports the effect of neighborhood out-migration on innovation. The units of observation are individuals who are observed yearly between 1900 and 1920. In columns (1–3) and (5–7), the sample consists of the universe of males who did not emigrate over the period and that were at least 18 years old in 1900; in columns (4) and (8), we restrict the sample to inventors. The dependent variable is the number of patents obtained annually. In columns (1–4), the sample consists of individuals residing in all England and Wales divisions; in columns (5–7), we exclude the top tree-patents producing areas: London, Lancashire, and the South-West. In Panel A, the independent variable is an indicator that, for a given individual, returns value one after at least one person that was living in the same neighborhood as the individual migrates to the United States; in Panel B, we restrict to emigrants that never return in the period of observation. In this context, "neighborhood" refers to emigrants within a range of 100 meters from the individual in the sample. Each model includes individual and—at least—year fixed effects; in column (2), we include parish-by-year fixed effects; in column (3), individuals are weighted by their coarsened exact matching weight. The estimates are obtained using the method discussed in Borusyak *et al.* (2021) to account for the staggered roll-out of the treatment across individuals. Standard errors, clustered at the district level, are reported in parentheses.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

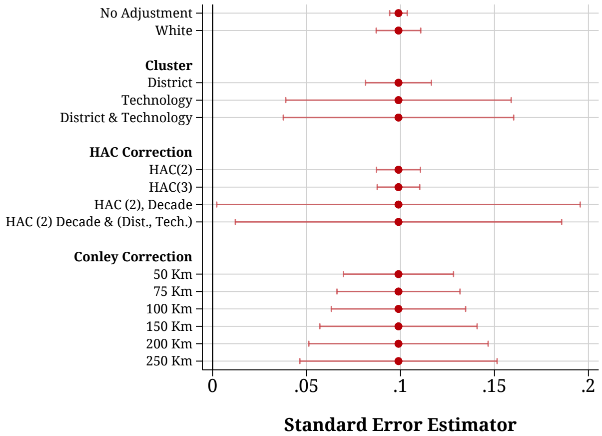
1.E.5 Figures

FIGURE 1.E.1: Alternative Standard Errors Estimators of the Return Innovation Result

(A) Measured Knowledge Exposure



(B) Railway-Based Knowledge Exposure

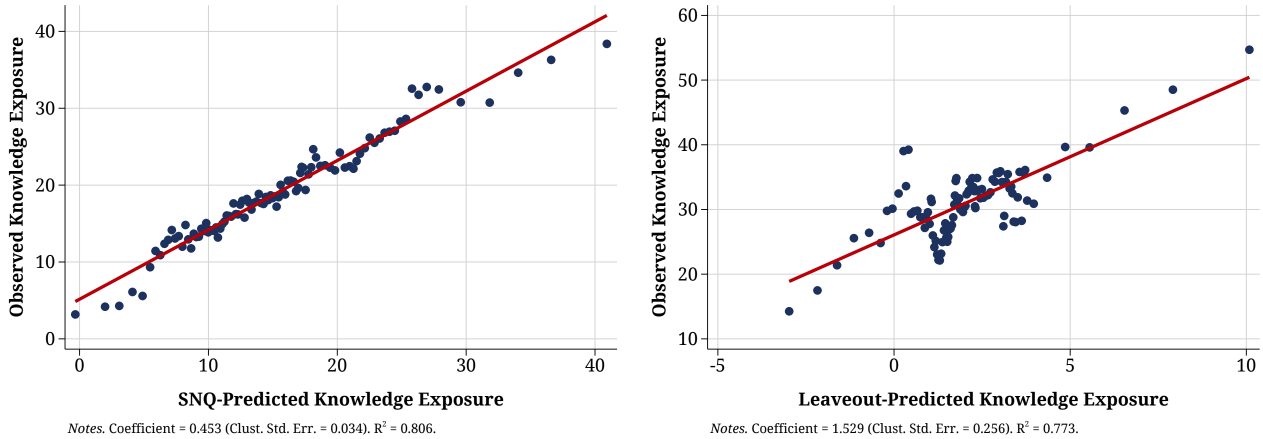


Notes. These figures report alternative estimates for the standard errors (SEs) of the regression between the number of patents and knowledge exposure. The unit of observation is a district-technology pair, observed at a decade frequency between 1880 and 1930. Models include district-by-technology and decade fixed effects. In Panel 1.E.1a, the independent variable is measured knowledge exposure; Panel 1.E.1b reports the estimated reduced-form coefficient between patents and the railway-based instrument. We report unadjusted SEs, robust to heteroskedasticity (White); clustered at the district, technology class, and two-way by district and technology class; robust to heteroskedasticity and autocorrelation of order 2 (HAC (2)), order 3 (HAC (3)); robust to heteroskedasticity and autocorrelation, and clustered by decade (HAC (2) - Decade) and two-way by decade and district-by-technology class (HAC (2) - Decade & (Dist., Tech.). Finally, we also report SEs that account for spatial autocorrelation at various orders (between 50 and 250 kilometers) following Conley (1999). Bands report 95% confidence intervals.

FIGURE 1.E.2: First Stage Binned Scatter Plot

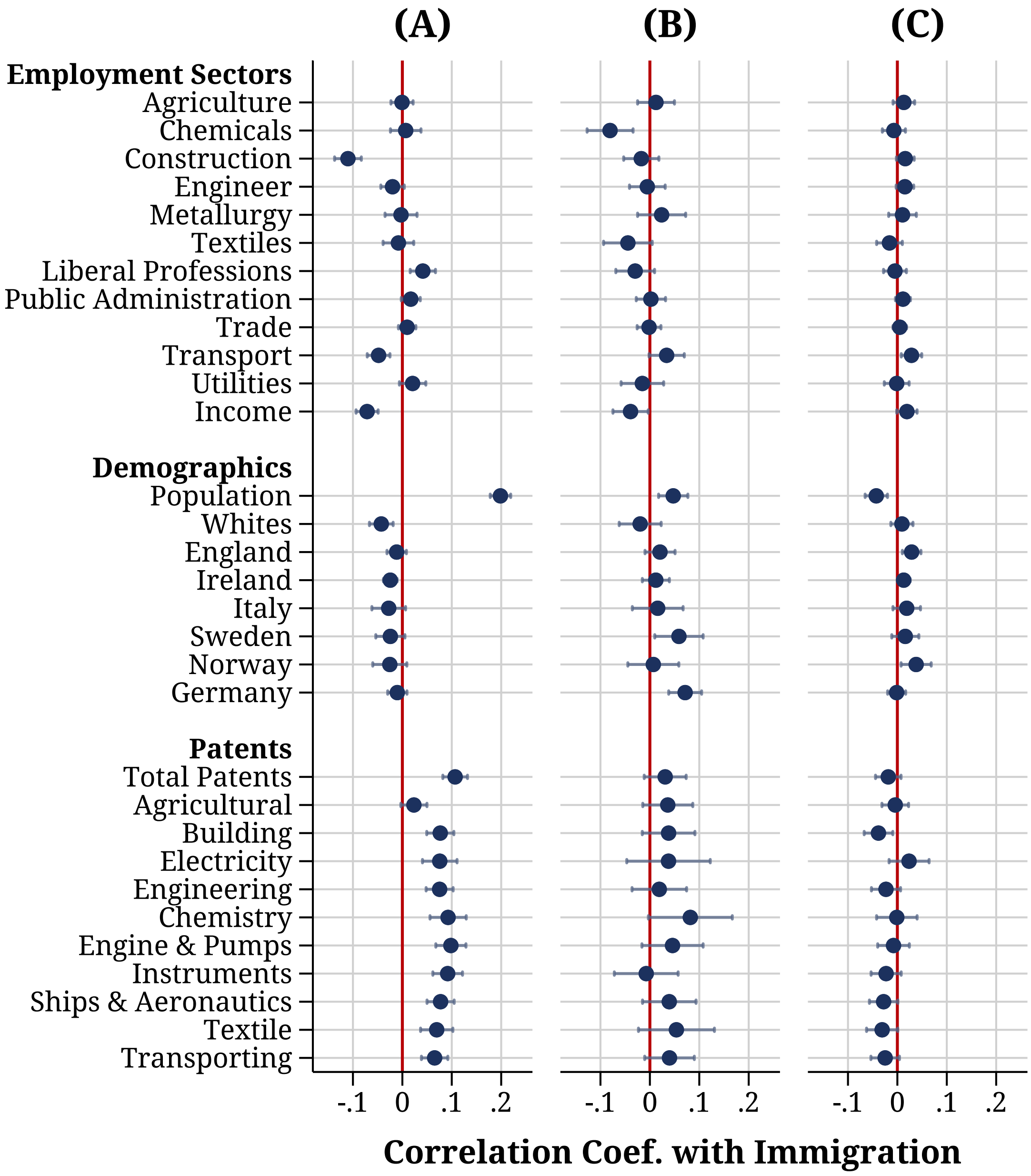
(A) Railway-Based Instrument

(B) UK Innovation



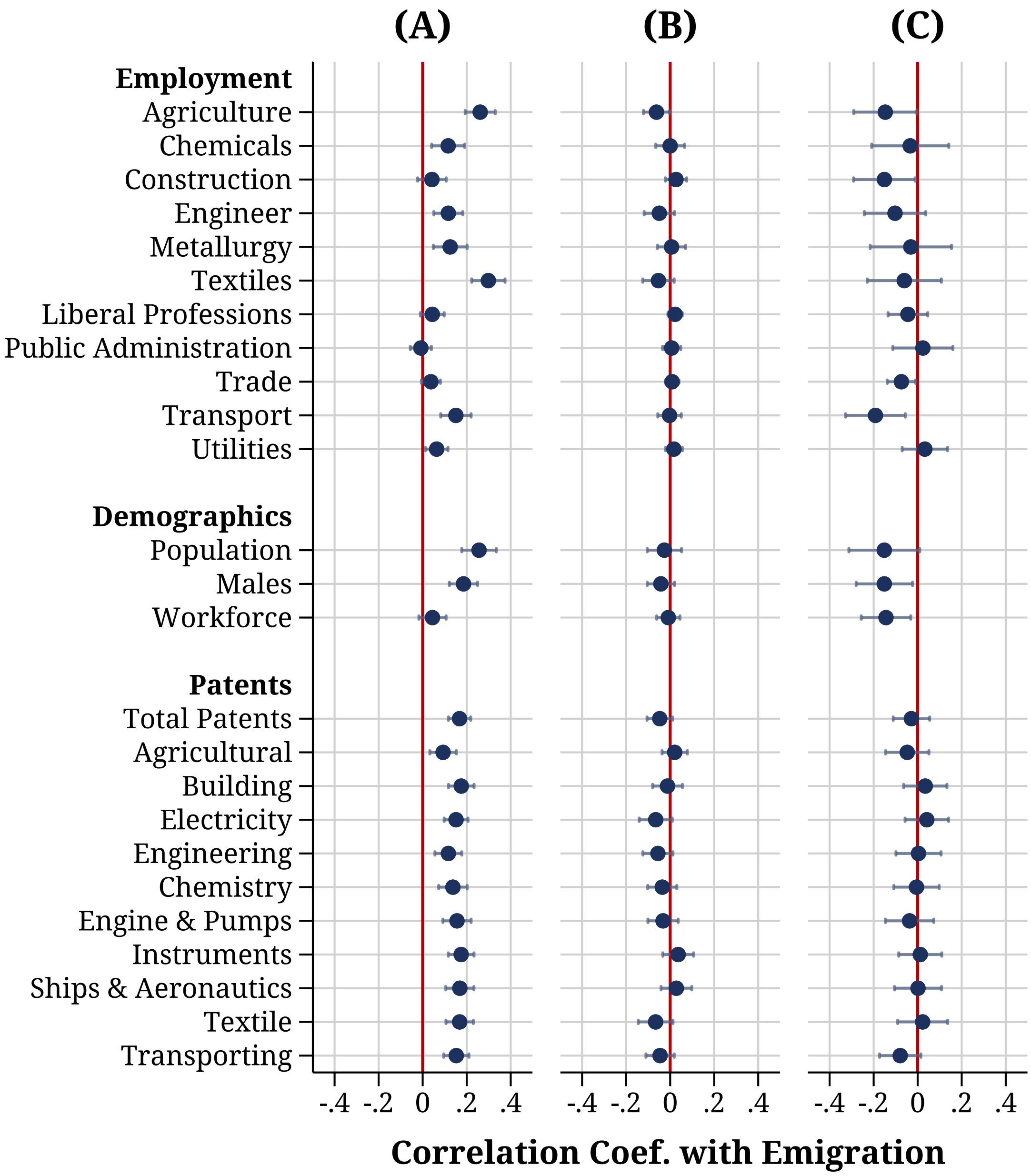
Notes. These figures are binned scatter plots of the association between actual and predicted knowledge exposure obtained using the railway-based instrument (Panel 1.E.2a) and the leave-out instrument (Panel 1.E.2b). The unit of observation is a district-technology class pair, at a decade frequency between 1880 and 1920. Graphs partial out district-by-decade and technology class fixed effects. We report the associated regression coefficients and standard errors, clustered at the district level, below each graph.

FIGURE 1.E.3: Shock-Level Balance Tests for Instrumental Variable Validity



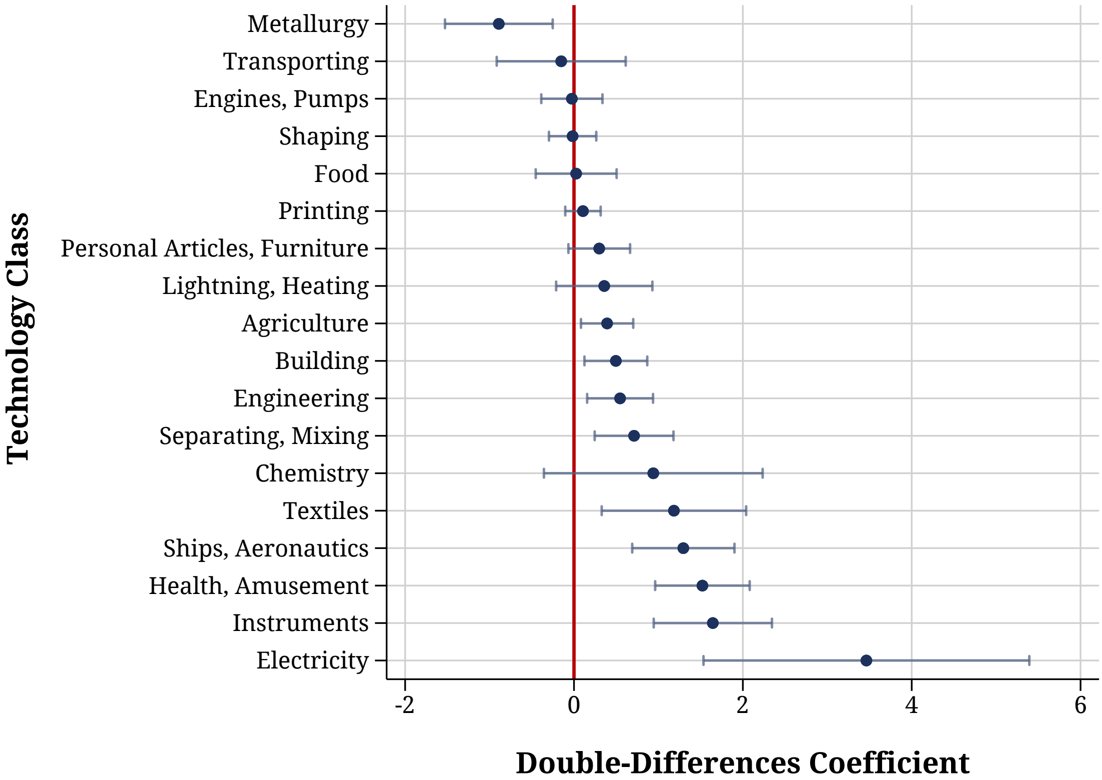
Notes. This figure reports the correlation between county-level observable characteristics and the (predicted) immigrant share. The unit of observation is a county observed at a decade frequency between 1870 and 1920. Panel (A) refers to the observed immigrant share; Panel (B) refers to the immigrant share predicted from the railway-based shock constructed from the zero-stage estimates *à la* Sequeira *et al.* (2020); Panel (C) refers to the leave-out shocks used to construct the alternative leave-out instrument. Each dot reports the correlation between the row variable and the immigrant share, lagged by one decade. Variables are standardized for the sake of readability. Each model includes county and state-by-decade fixed effects. Standard errors are clustered at the county level. Bands report 95% confidence intervals.

FIGURE 1.E.4: District-Level Balance Tests for Instrumental Variable Validity



Notes. This figure reports the correlation between district-level observable characteristics and the (predicted) number of emigrants. The unit of observation is a district observed at a decade frequency between 1870 and 1920. Panel (A) refers to the observed number of emigrants; Panel (B) refers to the predicted emigrant outflow obtained from the railway-based instrument; Panel (C) refers to the leave-out instrument. Each dot reports the correlation between the row variable and out-migration, lagged by one decade. Variables are standardized for the sake of readability. Each model includes district and decade fixed effects. Standard errors are clustered at the county level. Bands report 95% confidence intervals.

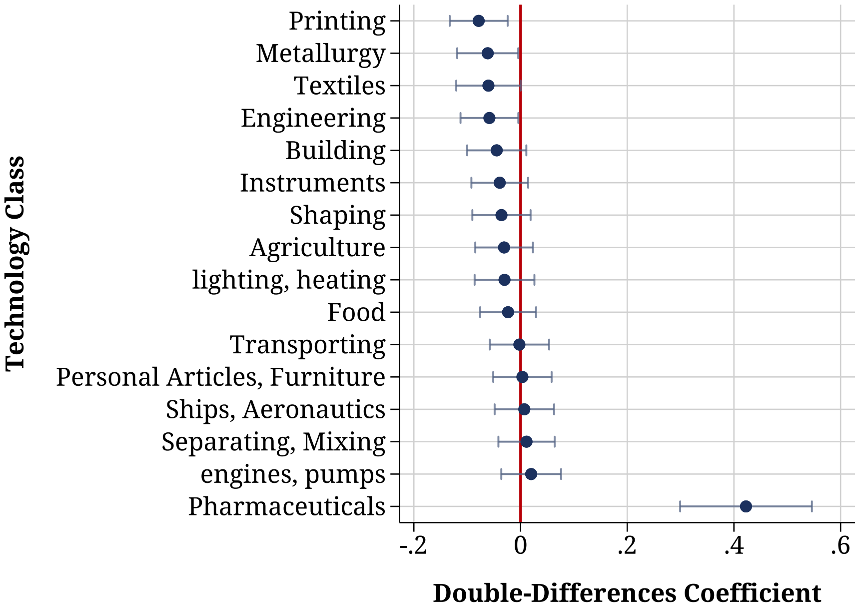
FIGURE 1.E.5: Effect of Synthetic Innovation Shocks Across Technology Classes



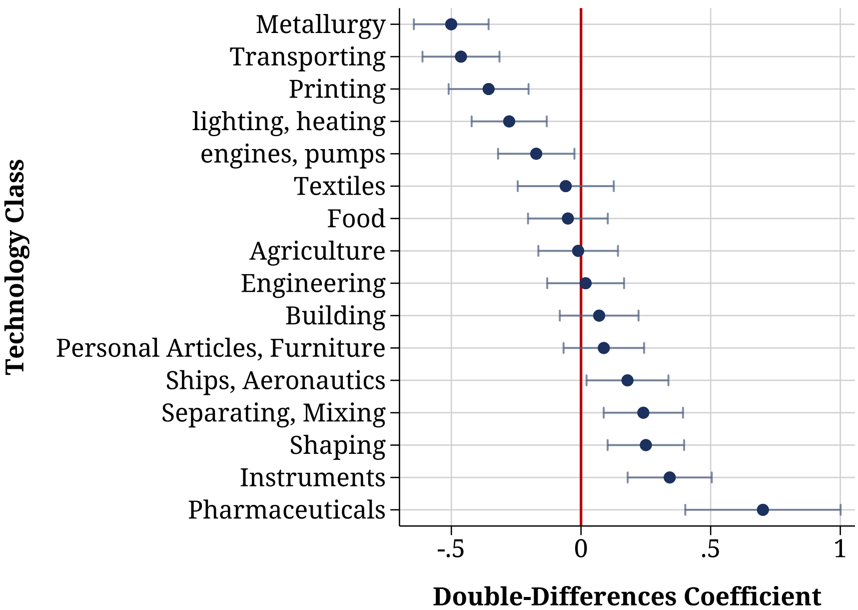
Notes. This figure reports the effect of synthetic innovation shocks on innovation in the UK by technology class. Each dot reports one double-differences estimated effect of the baseline exposure treatment with innovation; in each row, the treatment is activated whenever a district has above-median. The unit of observation is thus a district, observed at a yearly frequency between 1900 and 1993. Regressions include district and year fixed effects, and standard errors are clustered at the district level. Bands report 95% confidence intervals.

FIGURE 1.E.6: Effect of the Influenza Shock Across Technology Classes

(A) US Innovation

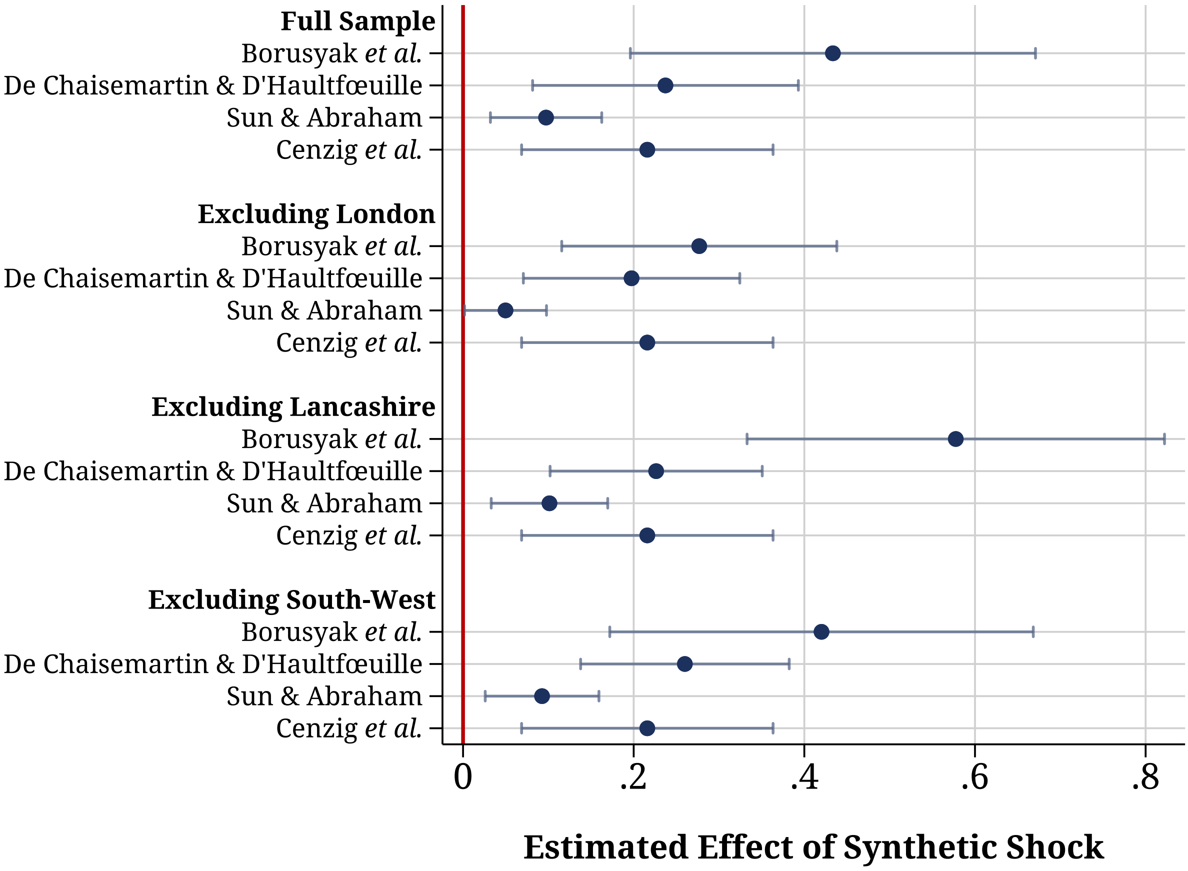


(B) UK Innovation



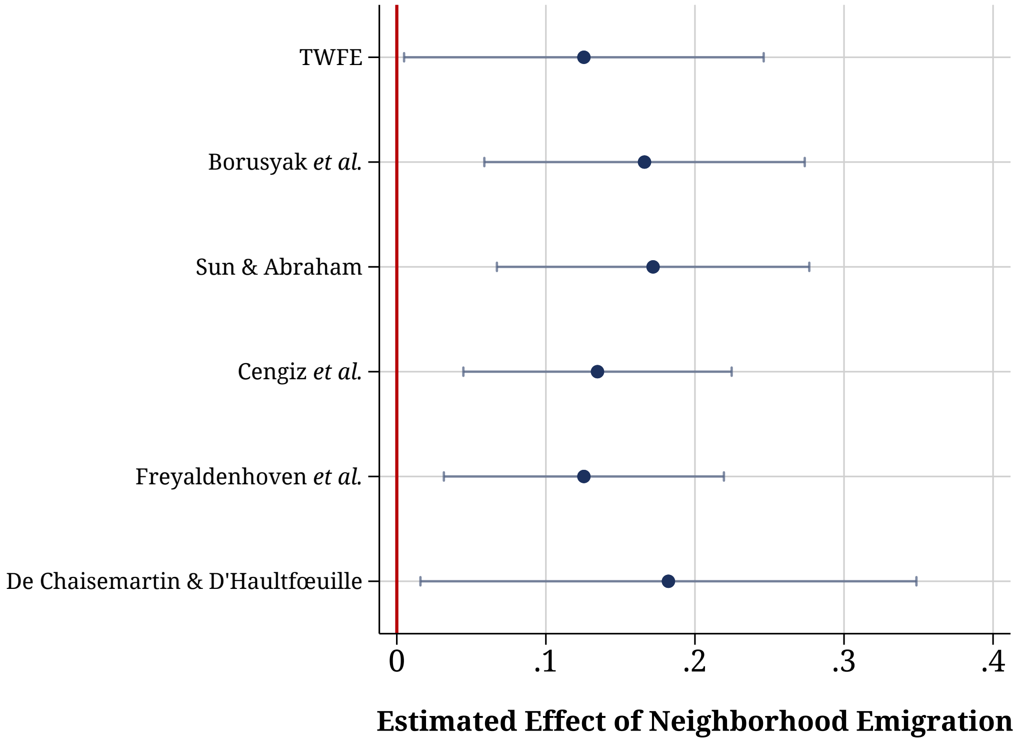
Notes. This figure reports the effect of the Influenza shock on innovation, by technology classes, in the US (Panel 1.E.6a) and in the UK (Panel 1.E.6b). Each dot reports one triple-differences estimated effect of the baseline exposure treatment with innovation; in each row, exposure is interacted with a sector-specific dummy variable. If the shock only impacted innovation in pharmaceuticals, we would expect each coefficient but the pharmaceutical one to be statistically equal to zero. Regressions are saturated with fixed effects; standard errors are two-way clustered at the technology class and county (Panel 1.E.6a) or district (Panel 1.E.6b) level. Bands report 95% confidence intervals.

FIGURE 1.E.7: Alternative Staggered Estimators of Synthetic Shocks



Notes. This figure reports the estimated effect of synthetic innovation shocks in US counties on innovation activity in the UK, using alternative estimators that explicitly allow for the staggered treatment roll-out design. The unit of observation is a district-technology class pair observed at a yearly frequency between 1900 and 1939. The dependent variable is the number of patents. The treatment variable is an indicator that, for a given district-technology, returns value one after a synthetic innovation shock in that technology class is observed in at least one county where the district has above-average out-migration. A synthetic innovation shock is observed whenever the residualized number of patents observed in the country is in the top 0.5% of the overall distribution. We estimate the models on the full sample of districts, as well as excluding the top three areas in terms of patents granted: London, Lancashire, and the South-West. We report the estimates obtained using four estimators that allow for the inclusion of all the triple differences interactions of the fixed effects: Borusyak *et al.* (2021), De Chaisemartin and D'Haultfœuille (2022), Cengiz *et al.* (2022), and Sun and Abraham (2021). Standard errors are clustered at the district and technology class levels. Bands report 95% confidence intervals.

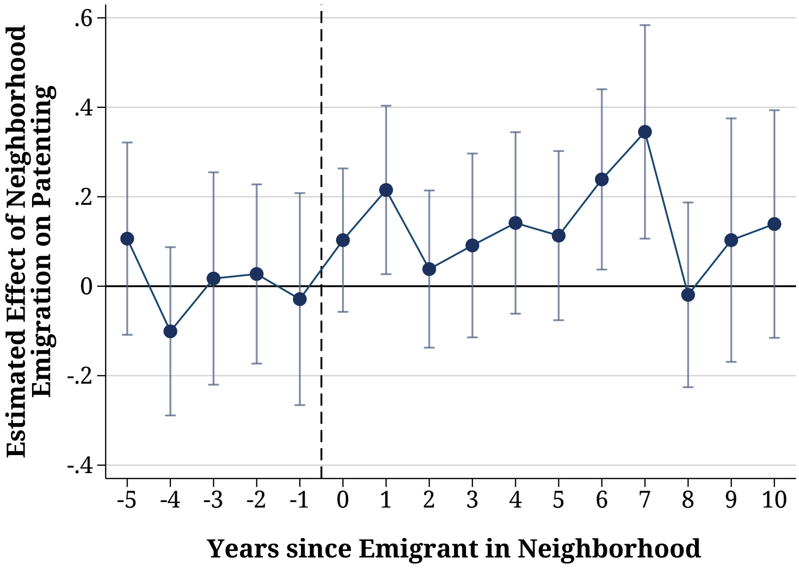
FIGURE 1.E.8: Alternative Staggered Estimators for Neighborhood Out-Migration



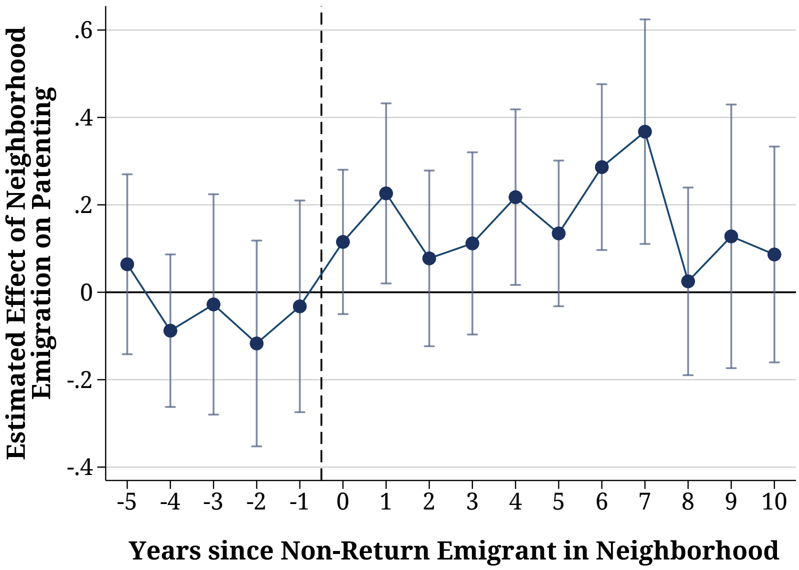
Notes. These figures report the effect of neighborhood out-migration on innovation. The units of observation are individuals observed at a yearly frequency between 1900 and 1920. The sample consists of all males who did not emigrate over the period and aged at least 18 in 1900. The dependent variable is the number of patents obtained every year. The treatment variable is an indicator that returns value one after at least one person living in the same neighborhood as the individual migrates to the United States. We report the estimates obtained using six estimators that allow staggered roll-out of treatment assignment: the baseline two-way fixed effects (TWFE) estimator, Borusyak *et al.* (2021), Sun and Abraham (2021), Cengiz *et al.* (2022), Freyaldenhoven *et al.* (2019), and De Chaisemartin and D'Haultfœuille (2022). Standard errors are clustered at the district level. Bands report 90% confidence intervals.

FIGURE 1.E.9: Flexible Double Differences Effect of Neighborhood Out-Migration

(A) All Emigrants

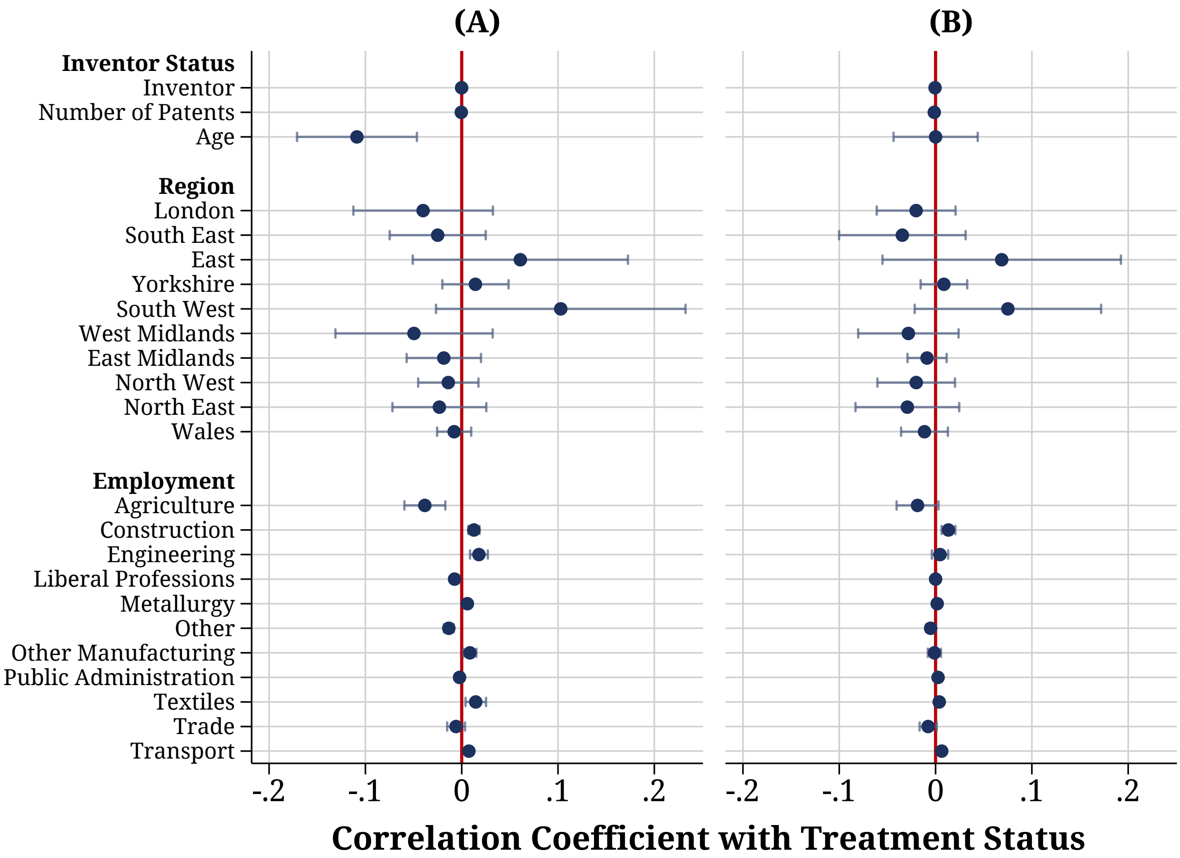


(B) Non-Return Migrants



Notes. These figures report the effect of neighborhood out-migration on innovation. The units of observation are individuals observed at a yearly frequency between 1900 and 1920. The sample consists of all males who did not emigrate over the period and aged at least 18 in 1900. The dependent variable is the number of patents obtained every year. In Panel 1.E.9a, the treatment variable is an indicator that returns value one after at least one person that was living in the same neighborhood as the individual migrates to the United States; in Panel 1.E.9b, we restrict to emigrants that never return in the period of observation. Each model includes individual and parish-by-year fixed effects. Standard errors are clustered at the district level. The estimates are obtained using the estimator discussed in Borusyak *et al.* (2021). Bands report 95% confidence intervals.

FIGURE 1.E.10: Co-variate Balance for Individual-Level Design



Notes. These figures report the correlations between individual-level observable characteristics and treatment status in the individual-level analysis. The units of observation are individuals observed at a yearly frequency between 1900 and 1920. The sample consists of all males who did not emigrate over the period and aged at least 18 in 1900. Variables are observed in the 1911 census. Hence some of them are not pre-determined when the treatment initiates. Each dot reports the correlation between the row variable and a dummy variable equal to one if the individual is treated in the observation period and zero otherwise. Variables are standardized for readability. Panel (A) reports the unweighted correlation; in Panel (B), individuals are weighted by their CEM weights. Standard errors are clustered by division. Bands report 95% confidence intervals.

Bibliography

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.

AGHION, P., A. BERGEAUD, M. LEQUIEN and M. J. MELITZ (2018). "The Impact of Exports on Innovation: Theory and Evidence." *NBER Working Paper*, (No. w24600).

ALEKSYNSKA, M. and G. PERI (2014). "Isolating the Network Effect of Immigrants on Trade." *The World Economy*, 37(3): 434–455.

ANDERSSON, D., M. KARADJA and E. PRAWITZ (2022). "Mass Migration and Technological Change." *Journal of the European Economic Association*, 20(5): 1859–1896.

ATKIN, D., A. K. KHANDELWAL and A. OSMAN (2017). "Exporting and Firm Performance: Evidence from a Randomized Experiment." *The Quarterly Journal of Economics*, 132(2): 551–615.

AUTOR, D., D. DORN, G. H. HANSON, G. PISANO and P. SHU (2020). "Foreign Competition and Domestic Innovation: Evidence from US Patents." *American Economic Review: Insights*, 2(3): 357–74.

BAILEY, M. J., C. COLE, M. HENDERSON and C. MASSEY (2020). "How Well Do Automated Linking Methods Perform? Lessons from US Historical Data." *Journal of Economic Literature*, 58(4): 997–1044.

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 M igration." *Journal of Development Economics*, 102: 23–47.

BEACH, B. and W. W. HANLON (2022). "Historical Newspaper Data: A Researcher's Guide and Toolkit."

BERGEAUD, A. and C. VERLUISE (2022). "A new Dataset to Study a Century of Innovation in Europe and in the US." *Working Paper*.

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 (2022). "Religiosity and Science: An Oxymoron? Evidence from the Great Influenza Pandemic." *Working Paper*.

BERTHOFF, R. (1953). *British Immigrants in Industrial America 1790-1950*. Harvard University Press.

Bloom, N., M. Draca and J. Van Reenen (2016). “Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity.” *The Review of Economic Studies*, 83(1): 87–117.

Borusyak, K., P. Hull and X. Jaravel (2022). “Quasi-Experimental Shift-Share Research Designs.” *The Review of Economic Studies*, 89(1): 181–213.

Borusyak, K., X. Jaravel and J. Spiess (2021). “Revisiting Event Study Designs: Robust and Efficient Estimation.” *Working Paper*.

Bustos, P. (2011). “Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms.” *American Economic Review*, 101(1): 304–40.

Callaway, B. and P. H. Sant’Anna (2021). “Difference-in-Differences with multiple time periods.” *Journal of Econometrics*, 225(2): 200–230.

Card, D. (2001). “Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration.” *Journal of Labor Economics*, 19(1): 22–64.

Cengiz, D., A. Dube, A. Lindner and D. Zentler-Munro (2022). “Seeing Beyond the Trees: Using Machine Learning to Estimate the Impact of Minimum Wages on Labor Market Outcomes.” *Journal of Labor Economics*, 40(S1): S203–S247.

Chen, J. and J. Roth (2022). “Log-like? ATEs Defined with Zero Outcomes are (Arbitrarily) Scale-dependent.” *Working Paper*.

Coelli, F., A. Moxnes and K. H. Ulltveit-Moe (2022). “Better, Faster, Stronger: Global Innovation and Trade Liberalization.” *The Review of Economics and Statistics*, 104(2): 205–216.

Conley, T. G. (1999). “GMM Estimation with Cross Sectional Dependence.” *Journal of Econometrics*, 92(1): 1–45.

Crucini, M. J. (1994). “Sources of Variation in Real Tariff Rates: The United States, 1900–1940.” *The American Economic Review*, 84(3): 732–743.

Curtis, S. J. (1875). *The story of the Marsden mayoralty: with sketch of the mayor’s life*. Leeds (UK): Express Office.

David, P. A. (1990). “The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox.” *American Economic Review*, 80(2): 355–361.

Davis, J. H. (2004). “An Annual Index of US Industrial Production, 1790–1915.” *The Quarterly Journal of Economics*, 119(4): 1177–1215.

De Chaisemartin, C. and X. D’Haultfœuille (2022). “Difference-in-Differences Estimators of Intertemporal Treatment Effects.” *NBER Working Paper*, (No. w29873).

Docquier, F. and H. Rapoport (2012). “Globalization, Brain Drain, and Development.” *Journal of Economic Literature*, 50(3): 681–730.

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).

EICHENGREEN, B. (1986). "The Political Economy of the Smoot-Hawley Tariff." *NBER Working Paper*, (No. w2001).

ENCYCLOPÆDIA BRITANNICA (2018). *David Edward Hughes*.

ERICKSON, C. (1972). *Who were the English and Scots emigrants in the 1880s?*, pp. 87–125. Arnold.

FINISHING PUBLICATIONS, L. (2018). "Early British Patents. A Cradle of Inventions. British Patents 1617 to 1895." MFIS [data collection].

FREYALDENHOVEN, S., C. HANSEN and J. M. SHAPIRO (2019). "Pre-event Trends in the Panel Event-Study Design." *American Economic Review*, 109(9): 3307–38.

HANLON, W. W. (2016). "British Patent Technology Classification Database: 1855-1882." Unpublished [data collection].

JUHÁSZ, R. and C. STEINWENDER (2018). "Spinning the Web: The Impact of ICT on Trade in Intermediates and Technology Diffusion." *NBER Working Paper*, (No. w24590).

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.

LI, S., J. HU, Y. CUI and J. HU (2018). "DeepPatent: patent classification with convolutional neural networks and word embedding." *Scientometrics*, 117(2): 721–744.

MADDISON, A. (2007). *Contours of the World Economy 1-2030 AD: Essays in Macro-Economic History*. Oxford (UK): Oxford University Press.

MOKYR, J. (1998). "The Second Industrial Revolution, 1870-1914." In V. Castronovo, ed., "Storia dell'Economia Mondiale," pp. 219–245. Rome (Italy): Laterza.

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.

OLIVETTI, C., M. D. PASERMAN, L. SALISBURY and E. A. WEBER (2020). "Who Married,(To) Whom, and Where? Trends in Marriage in the United States, 1850-1940." *NBER Working Paper*, (No. 28033).

OTTAVIANO, G. I. P., G. PERI and G. C. WRIGHT (2018). "Immigration, Trade and Productivity in Services: Evidence from UK Firms." *Journal of International Economics*, 112: 88–108.

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*.

SCHURER, K. and E. HIGGS (2020). “Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access.” [data collection] Second Edition, UKDS.

SEQUEIRA, S., N. NUNN and N. QIAN (2020). “Immigrants and the Making of America.” *The Review of Economic Studies*, 87(1): 382–419.

SHU, P. and C. STEINWENDER (2019). “The Impact of Trade Liberalization on Firm Productivity and Innovation.” *Innovation Policy and the Economy*, 19(1): 39–68.

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.

TABELLINI, M. (2020). “Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration.” *The Review of Economic Studies*, 87(1): 454–486.

WILLCOX, W. F. (1928). *International Migrations, Volume I: Statistics*. Cambridge (MA): National Bureau of Economic Research.

XU, S. (2018). “Bayesian Naïve Bayes classifiers to text classification.” *Journal of Information Science*, 44(1): 48–59.

Chapter 2

The Economic Effects of Immigration Restriction Policies*

Evidence from the Italian Mass Migration to the US

2.1 Introduction

In recent years, attitudes towards immigration in developed countries have considerably deteriorated (*e.g.*, Guriev and Papaioannou, 2022). Immigration restriction policies (henceforth, IRPs) are becoming increasingly common, reinforcing an upward trend that has been documented since the 1970s.[2] A large literature in economics studies the potential effects of these pieces of legislation in countries receiving migrants. Evidence on emigration countries remains, however, comparatively scant.[3]

In this paper, we focus on one crucial dimension of economic growth: technology adoption. Since developing and emigration countries typically operate far from the technology frontier, the adoption of new technologies represents a major source of productivity gains.[4] This

*This chapter is jointly authored with Lorenzo Spadavecchia. We are particularly grateful to Mara Squicciarini for her continued guidance and support. We thank Maristella Botticini, Gaia Dossi, Stefano Fiorin, Michela Giorcelli (discussant), Simon Görlach, Thomas Le Barbanchon, Nicola Limodio, Jaime Marques Pereira, Luke Milsom (discussant), Joel Mokyr, Nathan Nunn, Gianmarco Ottaviano, Sebastian Ottinger, Elena Stella, Marco Tabellini, and seminar participants at Alghero, Bari, Bocconi, CESifo, EEA/ESEM, IZA, OECD, and Warwick for insightful comments and discussions. We acknowledge financial support from Bocconi. Nicola Fontana, Marco Manacorda, Gianluca Russo, and Marco Tabellini kindly shared data with us.

[2] Data from (De Haas *et al.*, 2015) show that immigration restriction policies make up for approximately 40% of the entire corpus of migration laws. This share has been steadily increasing since the beginning of the 1970s.

[3] Clemens (2011) notes that in the RePEc archive, papers on emigration account for 25% of the overall migration literature.

[4] Several papers highlight the centrality of technology adoption for economic growth, especially in countries farther from the technology frontier, both theoretically (*e.g.* Parente and Prescott, 1994; Foster and Rosenzweig, 1995; Eaton and Kortum, 1999) as well as empirically (*e.g.* Suri, 2011; Bryan *et al.*, 2014; Juhász *et al.*, 2020). Historically, Gerschenkron (1962) famously discusses the technological catch-up of countries at the periphery of the industrial world in the XIX century.

notwithstanding, the effects of out-migration – and policies attempting to restrict it – on technology adoption are *ex-ante* ambiguous and potentially conflicting. On the one hand, emigration entails a loss of human capital – the so-called brain drain – that may hamper the ability of countries to adopt new technologies (Kwok and Leland, 1982; Gibson *et al.*, 2011; Docquier and Rapoport, 2012).[5] On the other, however, emigration may incentivize the adoption of labor-saving technologies because it increases the relative cost of labor (among others, see Hicks, 1932; Habakkuk, 1962; Acemoglu, 2002). The former interpretation implies that IRPs would bolster technology adoption and prove beneficial for long-run growth. The latter theory, however, predicts that IRP-induced labor supply shocks would dampen the incentive to adopt labor-saving technologies, thus hampering economic development. In this paper, we offer a causal quantification of the effects of restrictive immigration policies on technology adoption in emigration countries.

We investigate this question in the context of the Age of Mass Migration, the largest episode of voluntary migrations in recorded history (Choate, 2008). Specifically, we focus on Italy, the archetypal sending country during this period. From 1876 to 1925, approximately 17 million emigrants left Italy (nearly 70% of the average Italian population in 1900); about half of them never returned. Italy had one of the highest emigration rates and, since the 1890s, it was the leader in sheer emigration numbers (Hatton and Williamson, 1998). On average, 40% emigrants headed toward the United States, and the remaining 60% were split between South America and Europe. The United States was therefore the single most absorbent emigration destination. Italian mass migration to the United States, however, abruptly ended in 1921, when Congress passed the first of a series of restrictive IRPs that we refer to collectively as the "Quota Acts." The Quota Acts defined numerical quotas for European countries that were based on how many citizens from each country were recorded living in the United States at a given point in time.[6]

We leverage the differential exposure to this shock across Italian districts to estimate the economic effects of emigration on industrialization and technology adoption. Comparable empirical exercises face three major limitations. First, emigration seldom flows into only a few destinations; hence, it is difficult to observe large restrictive policy shifts. Second, migration

[5]Emigration has been shown to influence, among others, human-capital accumulation through remittances (Fernandez-Sanchez, 2020), return migration (Dustmann *et al.*, 2011), and increased returns to schooling (Beine *et al.*, 2008). In this paper, however, we focus on technology adoption as one major determinant of long-run growth.

[6]The 1921 Emergency Quota Act restricted the annual number of immigrants admitted into the United States to no more than 3% of the number of residents from that country, as recorded in the 1910 census. The 1924 Johnson-Reed Act reduced the quota to 2%, and pegged the reference date to the 1890 census. These laws explicitly targeted Southern and Eastern European countries, which until the early 1900s hardly took part in the Age of Mass Migration and whose immigrants were perceived by the public as a threat to America's economic welfare and cultural values (Higham, 1955).

dynamics are often affected by co-evolving regulations enacted by both receiving and sending countries which were absent during the period we study (Abramitzky and Boustan, 2017). Third, it is often difficult to retrieve information on emigrants in their home country (Dustmann *et al.*, 2015).[7] Our unique historical setting allows overcoming these difficulties.

Our empirical strategy relies on different exposure to the Quota Acts across Italian districts. Consider, for the sake of argument, two districts A and B, both of which had high emigration rates. However, most migrants from district A went to the United States, whereas none from district B did. Our key observation is that district A will be highly exposed to the Quota Acts, whereas district B will not. This is because emigration flows displayed substantial time and spatial persistence. Local information diffusion and social networks shaped the dynamics of Italian mass migration more than home-destination wage gaps (Gould, 1980b).[8] Formally, our identification assumption thus requires that districts with similar emigration rates but different destinations would not have undergone different development trajectories had the Quota Acts not been enacted, *i.e.* they were on parallel trends in terms of the outcomes we consider. We provide several pieces of evidence supporting this assumption. In Figure 2.2, we plot emigrants as a fraction of the total population, showing that Northern, as well as Southern regions, experienced varying emigration intensities. By contrast, the share of emigrants heading to the United States is prevailing in the *Mezzogiorno* (South of Italy). The figure also shows that exposure to the Quota Acts reflects these heterogeneous patterns once we control for the extensive margin of emigration.[9] It is straightforward to conceive this context in terms of a simple difference-in-differences (DiD) framework with a continuous treatment defined by some measure for quota exposure at the district level, where we control for the share of emigrants relative to the total population.

Existing data from official statistics are not suitable for this exercise because (i) digitized US and Italian censuses and complementary historical statistics do not report the origin of Italian migrants at a granular level of spatial aggregation, and (ii) disaggregated indicators of economic performance for Italy remain scarce. We thus construct a novel dataset linking administrative

[7] Aydemir and Borjas (2007) and Mishra (2007) overcome this issue by studying Mexican emigration to Canada and the United States, exploiting that about 95% of Mexican emigrants go to the United States. Meanwhile, Dustmann *et al.* (2015) study this in the context of Poland. These studies all lack exogenous variation to credibly identify the causal impact of migration policy on economic development in sending countries.

[8] Recently, Spitzer and Zimran (2020) formally validated the original information-diffusion hypothesis formulated by Gould (1980b). Further, Brum (2019) argues that the location choice of pioneers was a key determinant of future emigration outflows across districts. These findings confirm the original result from Hatton and Williamson (1998), who noted that pull factors, rather than push factors, explain the bulk of variation in Italian emigration.

[9] In Figures 2.C.1, 2.C.2, and 2.C.3, we show that more-exposed districts were not on different development trajectories before the Quota Acts, conditional on total emigration. This is key for valid causal inference of our estimates, as we explain later.

records of Italian emigrants who arrived at Ellis Island between 1892 and 1930 to their district of origin, and we complement it with newly digitized detailed data from industrial and population censuses. These data allow us to document three sets of results.[10]

We first show that industrial firms located in districts more exposed to the Quota Acts substantially decreased investment in capital goods. We measure investment in capital-intensive production technologies with the number of installed engines, and we further distinguish between traditional mechanical engines and cutting-edge electrical ones. The electrical engine—a defining technology of the Second Industrial Revolution—could yield sizable productivity gains (David, 1990; Mokyr, 1998). We show that in more-exposed districts, the adoption of engines slowed. This effect is particularly strong in magnitude for electrical engines, either measured in absolute number or weighted by the horsepower they generated. This is relevant for our argument because electrical engines were a decisively labor-saving technology (Gaggl *et al.*, 2021). We also show that the worker-per-engine ratio, a proxy for the labor intensity of production technologies, increased in firms located in more-exposed districts. This result is consistent with findings by Andersson *et al.* (2022), who show that emigration fosters the adoption of labor-saving technologies because it dampens labor supply, hence increasing the relative cost of labor. Since technology adoption is a key driver of long-run growth (e.g., Juhász *et al.*, 2020), our evidence suggest that the Quotas had possibly detrimental effects on Italian economic development.

To rationalize these findings, we advance and validate the hypothesis that IRPs induce a geographically segmented labor supply shock.[11] This is because, following an IRP, all those who would have migrated had the policy not been enacted are—at least partly—forced to join the local employment pool. More abundant (thus cheaper) labor dampens the incentive for firms to adopt capital-intensive technologies, as we observe. Under this interpretation, in Italy, the Quota Acts effectively implied that more-exposed districts experienced a disproportionate increase in labor supply, relative to less-exposed ones. Districts that experienced more emigration until 1924 were more exposed to the quotas because pull factors were disproportionately more effective there.[12] We document that population in these districts grew comparatively more relative to districts that were less exposed to the Quota Acts. We provide supportive evidence

[10]In Section 2.D.1, we develop a simple theoretical framework to explain our results in the context of labor-saving directed technical adoption, in the spirit of Zeira (1998) and San (2022).

[11]This approach mirrors that of Abramitzky *et al.* (2019), who document that the Quota Acts induced a negative labor supply shock in *U.S.* counties whose intensity depended on the prevailing origin of immigrants across European countries. In a similar spirit, Beerli *et al.* (2021) show that a reform that granted free access to the Swiss labor market to European workers increased natives' wages and benefitted Swiss firms.

[12]Several studies have documented that emigration location choices tend to persist over time (e.g. Gould, 1980b; Brum, 2019; Fontana *et al.*, 2021; Spitzer and Zimran, 2020).

of this mechanism, showing that (i) emigrants did not substitute the United States with other arrival destinations—neither internal nor international—and (ii) emigration outflows toward unrestricted countries, i.e., countries that did not promulgate IRPs, did not increase. Hence, districts that had been supplying relatively more U.S.-bound emigrants ended up having more "missing" migrants, i.e., people who would have migrated had the Quota Acts not been enacted. This mechanism generates a spatially segmented positive labor supply shock. If our directed technical adoption interpretation is correct, we would expect to observe increased industrial employment in more-treated districts.

To further assess the soundness of the directed technical adoption hypothesis and validate it against alternative mechanisms, we study how employment across sectors reacted to the IRP-induced labor supply shock. We focus primarily on the two biggest sectors at the time, agriculture and manufacturing.[13] We find that employment in manufacturing grew considerably in districts that were comparatively more exposed to the Quota shock. This finding is consistent with directed technical adoption: firms in manufacturing substituted capital goods with more abundant, therefore cheaper, labor provided by missing migrants. By contrast, in agriculture, we find no sizable increase in employment. A possible explanation for this finding is that agriculture in this period was a largely labor-intensive activity, hence the incentive for manufacturing firms to enlarge their labor stock following the Quota shock was larger than for agriculture firms. Because industrial employment grew and agricultural employment did not, the share of workers engaged in manufacturing increased.

Identification, and therefore a causal interpretation of our estimates, may fail if conditional variation in U.S. emigration rates was still systematically correlated with economic performance. Historical evidence provided by Spitzer and Zimran (2020) suggests that this is unlikely. Information diffusion and local social networks were the decisive factors influencing emigrants' location decisions. While we cannot test the baseline identification assumption, we develop two instrumental variables (IVs) to deal with residual endogeneity concerns. In the first validation exercise, we develop an IV along the lines of Tabellini (2020). This allows us to fix the cross-sectional variation in emigrant origin to a given—early—point in time, and to predict a district's emigration using the time-series variation in aggregate outflows, dropping emigrants from that district. Our second IV exploits variation stemming from the timing of when districts became connected to the railway system, in the spirit of Sequeira *et al.* (2020). Because railways drastically reduced transportation costs, they fostered out-migration. Moreover, U.S. emigration boomed as districts got "closer" to transoceanic emigration ports. We thus leverage time

[13]We repeat the entire analysis at the manufacture-sector level. We find that sectors where technology adoption drops the most, are also the ones where employment increases the most.

variation in the evolution of the railway network to instrument U.S. emigration, and we confirm all the baseline results. Both instruments confirm the main results.

This paper is related to three streams of literature. First, we speak to the several contributions seeking to investigate the impact of emigration on sending countries, as opposed to the much more developed literature studying the economic and social effects of immigration.[14] This literature identifies human-capital accumulation as the key driver of economic growth fostered by emigration; it is fueled either by return migrants or by increased returns to schooling (Beine *et al.*, 2008; Dustmann *et al.*, 2011; Dinkelman and Mariotti, 2016; Akram *et al.*, 2017; Fernandez-Sanchez, 2020). Evidence by Becker *et al.* (2020) in the context of forced migrations echoes these findings. We inform this literature by studying a different mechanism whereby emigration fosters the adoption of labor-saving technologies. We emphasize that this channel operates plausibly independently from human-capital accumulation.

Second, we contribute to the literature that studies the relationship between technology adoption and the supply of production inputs. Beyond the path-breaking contributions by Hicks (1932) and Habakkuk (1962), Hornbeck and Naidu (2014), Clemens *et al.* (2018), and Hanlon (2015) all study historical settings where changes in the availability of labor and other factors of production altered the direction of innovation activity. Lewis (2011) offers similar evidence in a modern setting. Our paper is closest in spirit to Andersson *et al.* (2022), who show that labor-saving innovation emerged in response to migration-induced labor shortages in 19th-century Sweden. Similar to their paper, we emphasize the labor supply-shock mechanism. However, we focus on technology adoption, and leverage exogenous variation in a DiD framework.[15] Several studies document the importance of technology adoption as a key driver of long-run growth, particularly in developing countries (Suri, 2011; Bryan *et al.*, 2014; Juhász and Steinwender, 2018). Gerschenkron (1962) argues that technology adoption was a pivotal factor that enabled countries at the periphery of the industrialized world, such as Italy, to catch up with leading industrial nations. Moreover, while Andersson *et al.* (2022) study the effect of a labor *shortage*, this paper documents how excess labor stemming from immigration restriction policies shapes the adoption of new technologies.

[14] Borjas (1995, 2014) produced two influential reviews of this literature. Dustmann and Görlach (2016) discuss why empirical works studying immigration reach conflicting conclusions. Abramitzky and Boustan (2017) surveyed papers studying historical and contemporary U.S. immigration. Hatton *et al.* (2005) and Ferrie and Hatton (2015) provided two complementary works studying the role of immigration from the standpoint of global economic history. Clemens (2011) instead surveyed the literature studying the effects of emigration on sending countries.

[15] We do not cover innovation, both because Italy performed poorly by standard indicators of innovation and because Italian firms were not on the technological frontier during this period (Vasta, 1999; Nuvolari and Vasta, 2015).

Third, by virtue of its setting, this paper is related to the large, growing literature investigating the exceptionally broad social phenomenon represented by the Age of Mass Migration (for a review, see Abramitzky and Boustan, 2017). We owe our baseline empirical strategy to the approach pioneered by Abramitzky *et al.* (2019), who leverage differential exposure to the Quota Acts to study how labor scarcity affected the United States. Several papers study both the short-run (Abramitzky *et al.*, 2014; Tabellini, 2020) as well as the long-run (Burchardi *et al.*, 2020; Sequeira *et al.*, 2020) effects of Transatlantic migration. Focusing on emigration countries, Karadja and Prawitz (2019) document that the mass migration fostered the demand for political change in Sweden. Circling back to Italy, Hatton and Williamson (1998) study the aggregate determinants of Italian emigration. Spitzer and Zimran (2020) validate the Gould (1980b) theory, whereby social networks exerted substantial influence on Italian emigration dynamics. Pérez (2021) compares the assimilation dynamics of Italian emigrants to the United States with those who moved to Argentina. Our contribution to this literature is twofold. In terms of methodology, we build the first highly comprehensive geographically disaggregated dataset of Italian emigrants during the years when the bulk of Italian mass migration took place (1900–1914). We also present newly digitized district-level data from population and industrial censuses. In terms of new findings, we show that the massive outflow of unskilled labor leaving Europe toward the Americas was unlikely to have hampered the structural shift towards manufacturing, even at the periphery of the (slowly) industrializing Old World. Our results suggest that the opposite impact prevailed: immigration *restriction* was what likely threatened economic modernization in Italy.

We structure the paper as follows. Section 2.2 describes Italian mass migration, the policies that shaped it, and the key economic characteristics of early 20th-century Italy. In Section 2.3, we discuss our data-collection contribution and our sources. In Section 2.4, we detail our empirical strategy, and we present our three sets of results. Section 2.5 presents our key robustness checks and our IV exercises. Section 2.6 concludes.

2.2 Historical Context

2.2.1 The Italian Mass Migration

The Italian mass migration (1870–1925) was the largest episode of voluntary migration in recorded history (Choate, 2008). Between 1880 and 1913, 17 million —corresponding to 65% of the Italian population in 1900—emigrated; most headed toward continental Europe and the Americas. Along with Ireland, Italy had the highest per capita emigration rate (Taylor and Williamson, 1997). Even though Bandiera *et al.* (2013) document that return rates were equally among the highest in Europe, the Italian mass emigration has long been recognized as a focal feature of the country's development process (Hatton and Williamson, 1998, p. 75). Gould (1980a) vividly describes late-19th-century Italy as the archetypal case of mass migration.

Italy was a latecomer to large-scale mass migration. Northern European countries had been experiencing substantial population outflows since the 1840s. By contrast, Italy, along with other Southern and Eastern European countries, didn't start experiencing mass emigration until the 1880s. The country's migration patterns over the 1870–1925 period display substantial time variation. Until the 1880s, its emigration rate remained relatively modest, and the bulk of its migrants hailed from Northern regions. Prohibitively high transportation costs and prevailing poverty in rural Southern areas largely inhibited migration from the *Mezzogiorno*.[16] During the 1880s, Northerners chiefly moved to neighboring countries on a temporary, seasonal basis (Sori, 1979). The widespread adoption of steamships and an agrarian crisis kicked off the Southern mass emigration (Keeling, 1999). A decade later, the script had flipped: most migrants were now coming from Southern regions. Though the share of migrants from Northern regions declined as the share from Southern regions grew, emigration rates from *both areas* rose steadily from 1870 to 1913 (Hatton and Williamson, 1998, p. 100). By the 1890s, Italy had become the global leader both in sheer numbers of emigrants and in emigration rate, which grew from 5‰ in 1880 to a peak of 25‰ in 1913 (Hatton and Williamson, 1998, p. 95). Again, only Ireland had emigration rates comparable to Italy's during the Age of Mass Migration.

Italian emigration collapsed during World War 1 (WW1) but quickly regained momentum in the years immediately following the war. The epoch effectively came to an end by the early 1920s, when the U.S. Congress enacted a series of restrictive immigration policies that effectively halted mass emigration to the United States. Emigration toward other transoceanic and European destinations nonetheless endured until the outbreak of WW2.

In the 1880s, Italy was a young nation rife with regional disparities spanning cultural and economic dimensions (Smith, 1997). The resulting geographically segmented migratory patterns largely reflected this substantial heterogeneity and provide the backbone of our empirical strategy. Until the early 1880s, the vast majority of migrants from Northern regions moved to European countries. Most of the rest steamed across the Atlantic, to Argentina and Brazil. This pattern is completely reversed for Southern migrants, whose primary destination was the United States. The share of U.S.-bound migrants increased substantially over time in every Italian region. By the 1910s, the United States had become the primary transoceanic destination for

[16] This term refers to Southern Italy, corresponding to NUTS-2 areas ITC and ITH. Regions within these areas are Lazio, Abruzzi e Molise, Campania, Puglie, Basilicata, Calabrie, Sicilia and Sardegna.

all of Italy, though Northern migrants still tended to prefer continental European destinations.

To explain why destinations with low relative wage gaps such as Argentina and Brazil received sizeable migration inflows, Gould (1980b) hypothesizes that local emigration dynamics were driven by a process of information diffusion. Information about emigration opportunities required time to spread across the country, and this diffusion accelerated as the volume of emigration increased. This process implied that emigration from different localities followed an S-curve, whereby emigration started slow, then picked up the pace, until eventually leveling off at saturation. Gould (1980b) provides convincing evidence suggesting that declining regional emigration-rate inequality is consistent with this mechanism. An indirect consequence of the Gould hypothesis is that local emigration rates displayed relatively little sensitivity to economic and demographic conditions, instead featuring high persistence (Hatton and Williamson, 1998, p. 99). Gould's hypothesis further strengthens our identification scheme. We leverage differential exposure of Italian districts to the U.S. Quota Acts to estimate the impact of a restrictive migration policy on economic development. Had migration decisions been exclusively driven by local economic conditions in the first place, our exclusion restriction may have turned weaker.[17]

Transportation costs may have also influenced international migration patterns. Systematic data on ticket fares are, to the best of our knowledge, lacking. Anecdotal evidence suggests that the price of a ticket from Naples to New York be around 170-190 *lire* at 1900-prices (Gomellini and O'Grada, 2011). By contrast, a third-class train from Naples to Milan would cost 100 *lire*, and one to Paris or Berlin would make another 100 (dei Deputati, 1907, p. 14873). Gomellini and O'Grada (2011) suggest that a Southern unskilled laborer would make about 500 *lire* if he stayed at home, while in New York the figure would be around 2000-2500 *lire*. Compounding wage differentials between Italy and the US, these figures highlight that for the Southern population transatlantic migration was a far cheaper option than both internal relocations as well as continental out-migration. Differences in transportation costs, however, are unlikely to explain the choice between transoceanic destinations. Pérez (2021) documents that a ticket from Naples to Buenos Aires in 1902 would cost 170 *lire*. For Southern emigrants, social networks rather than transportation costs, therefore, influenced the preferred emigration destination.

[17] Spitzer and Zimran (2020) provide evidence consistent with Gould's diffusion hypothesis. They show that emigration began in a few districts in the 1870s and 1880s, then subsequently spread to nearby districts over time through immigrants' social networks.

In the United States, Italian emigration was part of the "second wave" of immigration, coming mostly from Southern and Eastern Europe. Compared to first-wave countries such as England and Germany, poorer second-wave nations tended to supply less-educated, less-skilled migrants who experienced harder living conditions, assimilated more slowly and played economic catch-up with the natives for longer (Daniels, 1959; Abramitzky and Boustan, 2017; Albert *et al.*, 2021). Italian emigrants, typically unskilled agricultural workers, were no exception. Because we exploit a migration policy shift to assess the impact of emigration on economic development, the potential endogenous selection of migrants may be relevant for our results.[18] Spitzer and Zimran (2018) nonetheless show that migrants from Southern regions, who constituted the bulk of transoceanic migration, were positively selected.

One last, largely overlooked component of labor migration in Italy during the Age of Mass Migration is internal migration. Current data limitations hinder a quantitative study of internal migration from 1870 to 1925. In the rest of this study, we abstract from explicitly accounting for internal migrations for three reasons (beyond data availability). First, Gallo (2012) shows that internal migrants were easily outnumbered by international migration flows, particularly during the Age of Mass Migration. Second, internal mobility was largely temporary and seasonal, inherently different from transoceanic migration (Gallo, 2012, p. 32). Third, internal migrations reflected historically deep-rooted, persistent economic relationships between regions that are unlikely to influence our results on economic modernization in the 1930s (Gallo, 2012, p. 38).

2.2.2 Migration Policy in Italy and the United States

Newly unified Italy had virtually no emigration policy until 1873. Occasional, largely ineffective provisions were enacted between 1873 and 1887 that reflected the perceived need to deal with labor agents and recruiters, the so-called *padroni*, but did not form a corpus of migration law (Gabaccia, 2013, p. 55). The first such attempt at that was the 1888 Crispi-De Zerbi law, which introduced and regulated the contract of emigration between the migrant and the migration agency. The law was manifestly inadequate, however, to deal with the waves of migration that unfolded starting in the 1890s: it regarded emigration as an artificial phenomenon instigated by migration agencies and attempted to centralize its governance. Apart from a small measure to control ticket fares, it effectively failed (Foerster, 1919, p. 477).

Italian policymakers came to realize that emigration was more likely to *make* laws, rather

[18]Consider the case of negative migrant selection. The additional manpower forced to remain in Italy by the restrictive U.S. migration policy shock would be of relatively low quality. This would confound and downward bias our estimated impact of migration on economic development.

than *abide* them (Foerster, 1919, p. 475). The 1901 emigration law was passed under the renewed understanding that emigration was no artificial phenomenon and that it could bear positive effects on Italy. As such, the law sought to protect migrants from exploitation, rather than restricting their movement. The law established a Commissioner-General of Emigration to oversee the protective institutions and collect data on migrants. Only companies licensed by the Commissioner-General could sell tickets, whose rates were reset every three months. Comparatively minor subsequent legislation further protected remittances (1901), strengthened the authority of the Commissioner-General (1910), and regulated citizenship (1913) (Rosoli, 1998, p. 43).

Throughout this period, Italy either failed at or abstained from, enforcing emigration restrictions (Foerster, 1919, p. 501). The open-border policy enacted by the Italian government, coupled with (if not driven by) the overwhelming tide of migration flows, implies that emigration featured as a first-order dimension of Italian economic and social development.

The United States, for its part, maintained an open border between 1775 and the early 1920s, interrupted only by isolated outbreaks of anti-immigration policy interventions. During the Age of Mass Migration, some 30 million migrants entered the United States. By 1910, 22% of the labor force was foreign-born, the highest such share ever since (Abramitzky *et al.*, 2014). The first lasting attempt to limit immigration was the Chinese Exclusion Act, which effectively halted Chinese immigration until its repeal in 1943.[19] In 1895, a bill was introduced by Henry Cabot Lodge requiring that a literacy test be administered to each immigrant upon arrival. Congress voted for the bill, but it was vetoed by President Cleveland in 1897. Two other such proposals were vetoed by Presidents Taft and Wilson in 1912 and 1915, respectively (Koven and Götzke, 2010, p. 130). A literacy-test law was eventually passed in 1917, but it was largely ineffective thanks to rising literacy rates in Europe (Goldin, 1994).

In 1907, the United States Congressional Joint Immigration Commission, also known as the Dillingham Commission after its chairman, was formed to study, among other things, the economic and social conditions of immigrants. The Commission's 41-volume report favored "old" immigration countries such as England and Germany over "new," mainly Southern and Eastern European ones. The commission opined that because immigration from second-wave countries displayed higher return rates, migrants were less likely to assimilate (Higham, 1955). The highly influential report shaped numerous migration policy interventions. When immigration ramped up again after WW1, nativist demands for restrictions surged, and the Emergency Quota Act was passed in 1921. It was modified by the 1924 Immigration Act, which further

[19] The Chinese Exclusion Act was built on the 1875 Page Act, which banned Chinese women from immigrating. To date, these are the only U.S. laws to have explicitly targeted one ethnic group.

tightened immigration restrictions on second-wave countries.

The 1921 Emergency Quota Act envisaged a (temporary) annual quota of 360,000 immigrants from Europe.[20] Importantly for our identification, entry quotas were assigned to each country as 3% of that country's nationals living in the United States in 1910, as recorded in that year's census. The 1924 Immigration Act made the quota system permanent, lowered the inflow from 3% to 2%, and shifted the census baseline year to 1890. The last provision, in particular, disadvantaged countries newer to mass migration, consistent with the recommendations of the Dillingham Commission.

Abramitzky *et al.* (2019) note that the 1924 Immigration Act had a highly heterogeneous impact on immigration across different sending countries. Flows from Southern and Eastern Europe were heavily curtailed because the share of foreign-born individuals from those countries who lived in the United States in 1890 was extremely small. The quotas assigned to Northern and Western European countries were comparatively generous. For our purposes, the 1921 and 1924 laws (henceforth, the Quota Acts) effectively halted Italian mass migration to the United States. Since the 1890s, America had been absorbing 30% to 40% of all Italian emigration, so the Quota Acts represented a major policy shock for Italy.

2.2.3 Technology Adoption and Economic Growth in Italy

Italy entered the Age of Mass Migration in the 1880s. The country was in the midst of an agrarian crisis (Toniolo, 2014, pp. 60-73) that followed two decades of stagnation. The period from 1895 to 1913 was the only time until the 1950s "economic miracle" in which Italy managed to outperform and narrow the income gap with the leading industrial nations. In the 1920s and 1930s, during the Fascist period, Italy was still a mainly agricultural country, featuring low income per capita and stagnating productivity (Cohen and Federico, 2001, p. 23). During the first half of the Fascist *Ventennio*, economic policy was aimed primarily at fiscal and monetary consolidation. Agricultural policy—which formed an integral part of the Fascist propaganda—centered on boosting agricultural productivity, which had been stagnating since WW1, and draining marshlands. However, sheer numbers attest that agricultural policies resulted in neither substantial intervention nor sizeable progress (Zamagni, 1990, p. 262). All in all, growth slowed after 1925 and regional disparities further widened (Cohen and Federico, 2001, p. 15). Historical evidence is thus consistent with our finding that following the 1921–1924 U.S. emigration restrictions, Italy underwent a period of economic distress and rising regional inequality.

[20] U.S. immigration peaked in 1907, at 1,285,349 entrants. The number of entrants during the 1910s averaged around 800,000.

We relate the migration shock to diminished investment in capital goods, especially technologically advanced ones, and to a shift to labor-intensive production routines. Italy was nowhere near the technological frontier throughout the period, and skill premia actually *declined* from the 1890s onward (Vasta, 1999; Federico *et al.*, 2021). Like today's developing countries, Italy lagged behind large industrial nations in research-and-development expenditures, and it imported substantial amounts of foreign technology, both patents, and machinery. Whenever possible, Italian firms bundled different vintages of capital, adding new machines to existing ones instead of renovating the whole stock (Cohen and Federico, 2001, p. 51). The large pool of unskilled workers made it more profitable for Italian entrepreneurs to adopt labor-intensive technologies relative to the highly capital-intensive German and British ones. Consistent with this narrative, we find that the migration policy shock increased the stock of unskilled workers in regions with high emigration. There, firms opted out of investment in capital goods and became more labor-intensive, thus hampering the process of modernization they had been undergoing prior to the Quota Acts.

2.3 Data

Our analysis spans the years 1881 to 1936. We collected data from a number of sources; we stacked the data by census years and analyzed them at the *circondario* (henceforth, "district") level of aggregation.[21] In 1921, there were 216 districts, each consisting of a variable number of municipalities (see Online Appendix section 2.A for a complete description of the data). Because districts were abolished in 1927, all subsequent data are collected at the municipality level and aggregated at the 1921-district boundaries. Table 2.1 reports summary statistics for the variables in our final dataset.

2.3.1 Emigration

Italian official emigration statistics are of limited scope because out-migration flows were recorded at the province-level of aggregation (Hatton and Williamson, 1998). Province-level data are not suited for quantitative analysis, because provinces were relatively large: in 1921, there were only 60 provinces that together contained a population of approximately 20 million. This

[21] Population censuses were taken in 1881, 1901, 1911, 1921, 1931, and 1936. We do not include data prior to 1901 in our baseline analysis, except for population. Districts were instituted in 1859 as the middle administrative unit between municipalities and provinces. They had mainly statistical and judiciary purposes and were granted little administrative autonomy. In Online Appendix section 2.A.2 we discuss more in detail the sources that we digitized and present a visual summary of all the variables we analyze.

naturally limits the use of official statistics for an econometric exercise. We nonetheless digitize province-level emigration outflows and use them to validate the series we derive from the dataset that we assemble (see Online Appendix section 2.A.1.3).

To overcome this issue and study the Italian mass migration to the United States, we collected administrative records of Italians who entered the country between 1892 and 1930 through the Ellis Island immigration station.[22] This was by far the largest, though not the only, immigration gateway during this period.[23] Administrative records report, for the vast majority of migrants, name and surname, year of arrival, age, municipality of origin, and sailing ship. In this study, we concentrate on the migration year and the municipality of origin. Ultimately, we collected approximately 2.7 million individual observations spanning the years 1890 to 1930.

Because all data were recorded by U.S. officials, the municipality variable displays frequent coding errors. We adapted the matching procedure from Abramitzky *et al.* (2014), using a sound-spelling similarity metric to account for orthographic and misspelling errors[24]. We then set a threshold measure below which we accepted the best-matched municipality and above which we dropped the observation; we then ran robustness checks around this threshold. In our preferred specification, we were able to match 1.6 million migrants to their origin municipality. Among those, 800,000 are coded with no error. We mapped each municipality to the district it belonged to in 1921, then we computed district-level yearly flows. To the best of our knowledge, this is the most comprehensive data spanning the whole Age of Mass Migration for Italy, at this level of aggregation.[25] In figure 2.1, we plot the overall country-level yearly inflow of emigrants who landed in Ellis Island from 1890 to 1930. Emigration took off in the mid-1890s and peaked between 1905 and 1913. It collapsed during World War 1 (WW1), quickly regained momentum in 1920, then was definitively shut down by the Quota Acts in 1921 and 1924. Our data are consistent with both comprehensive U.S. immigration data and overall Italian migration patterns (Brum, 2019; Sequeira *et al.*, 2020). In Figure 2.2, we plot the geographical

[22]These records are freely available at `heritage.statueofliberty.org`. We run queries over a comprehensive pool of 20,000 Italian surnames over 1890–1930 period. In Online Appendix section 2.A.1.3 we document that our newly constructed series correlates well with existing—albeit less granular—emigration data from official statistics.

[23]According to official U.S. statistics, between 1892 and 1924, a total of 14,277,144 migrants entered the country through Ellis Island, out of a total immigration inflow of 20,003,041 (Unrau, 1984, p. 185). Thus, Ellis Island alone accounted for 71.4% of the total immigrant inflow. Some 95% of all Italian immigrants passed through Ellis Island.

[24]In section 2.A.1.1 in the Online Appendix we discuss more in detail the methodology we used to correct coding errors. In section 2.A.1.2 we show that immigrants whose origin municipality was not recorded represent, in every year, less than 1% of the overall sample.

[25]The only other geographically disaggregated data available to date for this period are those collected by Brum (2019) and Fontana *et al.* (2021). Both, however, focus on the pre-1900 period. Our dataset is thus the only one covering the years when the bulk of the mass migration took place (1900–1914).

distribution of migrants across districts. The upper panel displays variation in the emigrants-to-population ratio, i.e., the emigration rate. The lower panel reports unconditional variation in the U.S. emigrants-to-population ratio, which is the baseline measure for treatment exposure. Both figures normalize emigration by population in 1880 and report the resulting standardized series.

2.3.2 Population

We digitize information from six population censuses: in 1881, 1901, 1911, 1921, 1931, and 1936. The main outcome variable is the share of workers in industrial sectors. This variable, as well as total employment in several other sectors, is available for each district between 1901 and 1921. We digitized the 1931 and 1936 census data at the municipality level, then aggregated them at the district level. More granular data on employment for major manufacturing sectors are, unfortunately, only available until 1921. For the remaining years, we digitized them from manufacturing censuses, with the caveat that these are at the province level and are imputed to districts, as described in the next paragraph. The population of each municipality was compiled by the Italian statistical office (ISTAT), and we aggregated it by districts. We computed the k-urbanization rate of a given district as the share of people living in municipalities of population k or higher in that district, relative to the district's population. In some robustness checks, we control for the altitude, area, and population density of the districts.

2.3.3 Economic Activity

To measure shifts in the adoption of capital-intensive technology, we digitized province-level data from the 1911, 1927, and 1937 manufacturing censuses. Manufacturing censuses gathered information on the universe of firms operating in each province at the time of census completion; they provide valuable information about the amount and vintage of capital goods employed by firms. We collect data on (i) the number of operating firms, (ii) the number of operating firms employing inanimate horsepower, (iii) the number of mechanical engines, (iv) the number of electrical engines, (v) the amount of horsepower generated by mechanical engines, and (vi) the amount of horsepower generated by electrical engines. We distinguish between electrical and mechanical engines because the former were at the forefront of technological progress in those years (Gaggl *et al.*, 2021). This allows us to disentangle the possibly differential impact of the labor supply shock induced by the migration shock on different technology vintages. Industrial census data are available only at the province level. To impute them to districts, we regressed province-level outcome variables against the number of workers in each sector, controlling for population, province, and year-fixed effects. Then, from the resulting OLS estimates, we predicted the associated district-level variables.[26]

2.3.4 Other Data

Italy participated in WWI between 1915 and 1918. Because the war took place between two census years and ended just three years before the Emergency Quota Act, it can potentially confound our estimates. We, therefore, collect WW1 death records to measure the geographical variation in the cost imposed by the war across districts.[27] The dataset provides rich information on Italian military personnel who died during WW1. Importantly for our analysis, it includes the municipality of origin of each soldier. Because we conducted our analysis at the district level, we collapse the dataset from municipalities to 1921 districts, and we measured the war's severity in a given district as the ratio between deaths and population in 1910. In Tables 2.B.16 and 2.B.7, we report all our results, further controlling for this measure interacted with a posttreatment indicator, and we confirm our baseline estimates.

To implement our railway instrumental variable, we digitized the entire Italian railway network over 1839–1926 period.[28] For each railway section, we know all the stations it is connected to. Stations are generally labeled in terms of the municipality they were located in. Further details are included for stations located in municipalities with more than one station. We also know the exact date when each trunk was built and opened to public use, as well as the distance it covered and the traction system the trains employed. We use these data to construct the Italian railway network. To capture its evolution over time, we took snapshots of the network at decade frequency.

[26]In Online Appendix section 2.A.2.1 we explain how we conduct the imputation of province-level data to districts. We then validate our imputation methodology by comparing imputed and measured variables.

[27]Death records were collected by the Fascist regime for propaganda purposes. They are available at cadutigrandeguerra.it. This dataset is maintained by the *Istituto per la storia della Resistenza e della società contemporanea*. Acemoglu *et al.* (2022) were among the first to use them in the economics literature.

[28]The data come from the volume *Sviluppo delle ferrovie italiane dal 1839 al 31 dicembre 1926*, edited by the Italian Statistical Office (*Ufficio Centrale di Statistica*) in 1927. To the best of our knowledge, this is the first paper to use these data.

2.4 Results

2.4.1 Empirical Strategy

In this section we explain the baseline empirical strategy we apply to estimate the causal impact of the Quota Acts on technology adoption and the dynamics of labor supply. Our identification relies on geographic variation in emigration patterns and intensity across districts in the pre-quota period.[29] Consider for the sake of argument two ideal districts; call them *A* and *B*. From 1890 to 1924, many Italians emigrated from both districts. However, most emigrants from district *A* headed toward the United States, whereas none from district *B* did. District *A* will thus be more exposed to the emigration restriction shock relative to district *B*. This is the case because social networks and information diffusion exerted a powerful pull, influencing potential emigrants through previous generations' emigrants (Spitzer and Zimran, 2020). This induced substantial persistence in emigration patterns by country of destination. Districts that had experienced higher emigration toward the United States before the Quota Acts were therefore comparatively more exposed to the migration restriction shock relative to those districts whose emigrants headed mainly toward European and South American countries.

Reality was more nuanced than our example. Emigrants left from all districts and headed to numerous destinations, hence the intensity of quota exposure varies smoothly with respect to the relative emigrant outflows to the United States. Importantly, the existing dispersion of U.S.-bound emigrants by district of origin shown in Figure 2.2 ensures that emigration location choices were not systematically correlated with economic development. In other words, we allow the decision to emigrate to be correlated with economic performance at home. What we restrict to be conditionally orthogonal to economic performance is the decision of *where* to emigrate.[30] Our identification assumption—in jargon, parallel trends—thus relies on the key assumption that districts with similar relative emigration outflow but with different destinations would not have undergone differential development patterns had the Quota Acts not been enacted. The wide divide between Northern and Southern regions could threaten our identification scheme. In Online Appendix Tables 2.B.16 and 2.B.7, we show that our baseline results are robust if we include a large set of covariates measured before the Acts, interacted with a year time trend, as further controls. In particular, we show that including an interaction between

[29] This identification scheme therefore mirrors that of Abramitzky *et al.* (2019), who exploit different immigration patterns by country of origin across U.S. counties and the Quota Acts shock to estimate the economic effects of *immigration*.

[30] In Section 2.5.2, we present a simple instrumental variable that further addresses the possible residual correlation between intensity of exposure to the Quota Acts and economic performance of districts.

a Southern dummy and a posttreatment indicator does not qualitatively alter the results. This implies that our estimated effects do not critically depend on a Northern-Southern comparison.

We measure quota exposure of district c as

$$\text{QE}_c = \frac{1}{\text{Population}_{c,1880}} \sum_{t=1890}^{1924} \text{US Emigrants}_{c,t} = \frac{\text{US Emigrants}_c}{\text{Population}_{c,1880}} \tag{2.1}$$

where $\text{Population}_{c,1880}$ is the population of district c in 1880, and $\text{US Emigrants}_{c,t}$ is the number of emigrants who headed to the United States over the period. Since mass outmigration started in the 1890s, in equation (2.1) we normalize the total number of U.S. emigrants with district population in 1880 to ensure that the measure for quota exposure does not conflate confounding variation due to aggregate emigration. Quota exposure in equation (2.1) can be further decomposed as

$$\text{QE}_c = \underbrace{\frac{\text{US Emigrants}_c}{\text{Emigrants}_c}}_{\text{Intensive margin} \equiv \text{IM}_c} \times \underbrace{\frac{\text{Emigrants}_c}{\text{Population}_{c,1880}}}_{\text{Extensive margin} \equiv \text{EM}_c} \tag{2.2}$$

where Emigrants_c is the total number of emigrants. The intensive margin (IM) of exposure measures the relative importance of the United States as an emigration destination; the extensive margin (EM) measures the relative importance of emigration overall. For a district to have high quota exposure, we thus require that (i) cumulative emigrants are a non-negligible share of the 1900 population, and (ii) a non-negligible share of those emigrants headed toward the United States. By contrast, districts with both little overall and little U.S.-bound emigration are at the bottom of the distribution of QE. In our preferred specification, we control for the extensive margin to compare districts with similar emigration rates but different destinations, hence exposure. This is because, while the decision to emigrate is likely endogenous to economic development, the destination should be conditionally quasi-random. In Section 2.5, we show that results are robust to two different instrumental variables exploiting a shift-share design, as well as time-varying access to the railway network. We construct a measure for EM using province-level data of total emigration available in the census, and we assume constant emigration rates within each province.[31] Figure 2.2 plots the geographical variation in EM and QE. We view the figure as supportive evidence that variation in QE is quasi-exogenous upon conditioning on the extensive emigration margin.

[31]Since district-level data on overall migration do not exist, we cannot test this assumption. However, using district-level U.S. emigration figures, we find that within-province U.S. emigration rates do not substantially differ across districts.

Quota exposure defined in equation (2.1) serves as our baseline treatment. Our dataset is a panel of districts, observed every census year between 1901 and 1936. Throughout the rest of the paper, we estimate variations on the following DiD model:

$$y_{c,t} = \gamma_c + \gamma_t + \boldsymbol{x}'_{c,t}\boldsymbol{\beta} + \delta_1 \left(\text{EM}_c \times \text{Post}_t\right) + \delta_2 \left(\text{QE}_c \times \text{Post}_t\right) + \varepsilon_{c,t} \tag{2.3}$$

where y is the log-difference of a generic outcome variable, $\boldsymbol{x}$ is a vector of additional controls, and Post_t is an indicator that is equal to one if $t > 1924$.[32] The baseline specification includes district and time fixed-effects, and standard errors are heteroskedasticity-robust and clustered at the district level unless otherwise specified. Baseline controls are labor market slackness and population. The geographic variation in the treatment is shown in the bottom panel of Figure 2.2, where we normalize total U.S. emigration outflows by 1880 population. The term δ_2 then captures the impact of the emigration restriction shock on the outcome variable y. In all regressions, we control for the emigration rate (EM) because our identification scheme relies on the fact that districts with similar emigration rates but different destinations would not have undergone differential development patterns had the Quota Acts not been enacted. In a series of robustness checks (discussed in detail in Section 2.5), we control for variation due to WW1, measurement errors in the years following the Quota Acts due to changes in registration procedures at Ellis Island, and possible correlation between QE and the error term. There is evidence, moreover, that emigration fosters economic ties, chiefly through international trade, between immigration and emigration countries (*e.g.* Dunlevy and Hutchinson, 1999). We account for this by including the interaction between US GDP and Quota exposure as a further control. This captures demand-type shocks which US emigration districts could be exposed to, depending on the state of the US business cycle.

Causal inference on estimates of model (2.3) requires that the treatment and control groups were on the same trend before the treatment (the Quota Acts) occurred. Because no census was taken in 1891, to test the parallel trends assumption we need to interpolate data points between 1881 and 1901. In the Online Appendix—in Figures 2.C.1, 2.C.2, and 2.C.3—we report the results of these event-study regressions and provide convincing evidence in favor of the parallel-trends assumption. All figures report the estimated coefficient of our baseline treatment interacted with decade dummies. Under the parallel-trends assumption, we expect all coefficients before the treatment period not to be statistically significantly different from zero, as we observe at standard confidence levels. In Table 2.2, we instead report correlations between the outcome

[32] Congress passed the first restrictive migration law—the Emergency Quota Act—in May 1921. The Immigration Act of 1924 further restricted the number of Italians allowed in the US every year. The choice between 1921 and 1924 as the treatment year is however immaterial since we do not observe districts within the two Acts.

variables we collect and the measure for quota exposure, conditional on the extensive emigration margin, population, and province fixed effects for 1911 and 1921. This exercise is not ideal in that we cannot clean for year-fixed effects, but it nonetheless strongly suggests that the treatment and control groups are comparable at all standard confidence levels before the treatment period. In fact, we find that none of the outcome variables we examine has a significantly different-from-zero correlation with the treatment before 1921.

2.4.2 Emigration and Technology Adoption

We study how technology adoption and investment in capital goods by manufacturing firms responded to the IRP shock. To do this, we collect several proxies for capital investment from the manufacturing census, and we report estimates of model (2.3) for these various outcomes. Our two baseline measures of investment in capital goods are the number of engines and their installed horsepower capacity. We distinguish between traditional mechanical engines and technologically advanced electrical ones. The electrical engine, in particular, was a defining innovation of the Second Industrial Revolution, yielding substantial productivity gains relative to older mechanical engines (Mokyr, 1998). Importantly, electrical engines were more labor-saving than electrical ones. We, therefore, interpret investment in electrical engines as a proxy for the adoption of advanced, labor-saving technology, a key driver of long-run economic growth (Juhász *et al.*, 2020).

U.S. observers evocatively described the turn of the 20th century as the Age of Electricity. In 1900, horsepower produced by electrical engines accounted for a mere 5% of overall consumption for production purposes. Two decades after, this figure had risen to 50% (David, 1990). Though productivity growth was relatively slow to manifest, it nonetheless became apparent starting in the early 1920s.

Italian firms were latecomers to technology adoption (Cohen and Federico, 2001). Hence, it seems plausible that well into the 1930s, electricity represented a major source of potential productivity growth. Despite the large productivity gains they could yield, Italian firms were slow to adopt electrical engines. Capital stocks in the early phase of adoption were a patchwork of different engine vintages. All these implied that, in the United States, capital-per-worker increased following the introduction of electrical engines (David, 1990). We document a different pattern in Italy in the aftermath of the IRP shock.

Table 2.3 reports the baseline results. We employ six outcome variables to measure investment in capital goods and technology adoption, and we estimate the causal impact of the

Quota Acts in model (2.3), controlling for the extensive emigration margin, population, labor-market slackness, and district and year fixed effects. From left to right, the columns display the total number of firms, the number of firms with at least one engine of any vintage, the sheer number of mechanical and electrical engines, and the horsepower of mechanical and electrical engines. As in all other regression tables, the first row displays the DiD coefficient δ_2.[33] We find that investment in mechanical and electrical engines alike declined substantially in more-exposed districts, whether such exposure is measured as the sheer number of installed engines or in terms of generated horsepower. In terms of magnitude, however, the effect of the IRP shock is stronger for electrical engines. Our results are qualitatively unchanged if we restrict the estimation sample to Southern regions.[34]

To rationalize this finding, we build on Andersson *et al.* (2022), who hypothesized that emigration fosters invention and adoption of labor-saving technology because it makes labor a relatively scarce production input. We take the specular perspective, arguing that the Quota Acts, and IRPs more broadly, induced a geographically segmented positive labor supply shock. Districts that before the Acts had experienced high U.S.-bound emigration rates were more exposed to the policy shock, because they ended up having disproportionately more "missing migrants." If missing migrants at least partly joined the local employment pool, then those districts were subject to a positive labor supply shock. On the other hand, districts whose emigrants headed toward destinations other than the United States did not undergo any such shock, because emigration to those countries remained unrestricted after the Quota Acts. Directed technical change and adoption theory thus suggests that firms in treated districts would be motivated to decrease investment in capital goods and to substitute capital with labor, which had become a more abundant production input following the IRP-induced shock. We devote the rest of the paper to validating this hypothesis.

An obvious corollary of this hypothesis is that production technologies in more heavily treated districts should become more labor-intensive. We assess this in Table 2.4. To measure labor intensity in production, we calculate the ratio of the number of workers employed in manufacturing to all the previous outcome variables. We thus measure how labor-intensive production technologies were across districts. We find that the number of industrial workers per unit of capital increased. This again holds if we measure capital in terms of the number of installed engines, or in terms of horsepower generated. In terms of magnitude, the effect of the IRP is comparable across vintages—a 1% increase in QE leads to a 0.6% increase in the

[33]The negative coefficient associated to the interaction between the extensive emigration margin and the post-treatment indicator could reflect the fact that emigration districts were negatively selected.

[34]Southern regions include all but EU NUTS 2 ITC and ITH regions. In other words, we drop Aosta Valley, Piedmont, Lombardy, Liguria, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, and Emilia-Romagna.

worker-to-capital ratio for both electrical and mechanical engines.

Finally, we ask whether the effects of the IRP shock are distributed evenly across industrial sectors. To answer this, we repeat the exercise of Table 2.3 for each sector recorded by the manufacturing census.[35] We end up with six sectors, whose estimated DiD coefficients for the various outcomes we report in Figure 2.3. We document sizable heterogeneity across sectors. Firms in relatively backward First Industrial Revolution sectors, particularly textiles and construction, reduced investment in capital goods. This effect is stronger for more-advanced electrical engines. On the other hand, we find that capital investment and adoption of electrical engines by firms in modern sectors, such as chemicals and metallurgy, display a less-marked decrease.[36] The sector-level analysis yields sharper predictions for our directed technical adoption hypothesis. Under this interpretation, we would expect employment in First Industrial Revolution sectors to grow more than in modern ones because firms in the former sectors were apparently eager to substitute capital for newly available labor. We evaluate this prediction in Section 2.4.4.

2.4.3 Emigration and Population Growth

Here, we document that districts more exposed to the migration shock experienced subsequent higher population growth. We view this as evidence confirming our narrative, whereby emigration restriction imposes a positive labor supply shock on the emigrants' country of origin. We thus estimate model (2.3), setting population growth as the outcome variable; we report the resulting estimates in Table 2.5. We compare the estimates obtained from the baseline continuous treatment, as well as those with a categorical dummy treatment equal to one for districts whose exposure is above the median, and zero otherwise. In all regressions, we control for the extensive emigration margin, population, labor-market slackness, and district and year fixed effects.

The estimated DiD coefficient (δ_2) confirms that districts that were more exposed to the Quota Acts experienced higher population growth. This effect is always statistically different from zero. Importantly, significance does not vanish if we restrict the sample only to Southern districts, where the exclusion restriction is sharper. We view this result as confirming that our measure of quota exposure is sound. Districts with more outstanding U.S.-bound emigrant

[35] We do not include "other industries" or "public service industries" in the analysis—the former is a residual category with little economic meaning, and data for the latter are not available in later censuses.

[36] We broadly classify manufacturing sectors based on narrative historical evidence presented by Mokyr (1998). Textiles and construction are therefore more closely associated with the First Industrial Revolution, whereas chemicals and steel-working refer to the Second Industrial Revolution (sometimes called the technological revolution).

stocks experienced less emigration, which triggered higher population growth in the years following the Quota Acts. Though studying the precise mechanism driving this result is beyond the purpose of this paper, this finding is consistent with pull factors, such as social networks and information diffusion, exerting better influence in more-exposed districts. Table 2.5 shows that the significance and magnitude of the DiD coefficient δ_2 both increase once we control for the extensive margins of U.S.-bound emigration.

Implicitly, Table 2.5 provides evidence against mechanisms that could threaten our source of identifying variation. The mechanism we emphasize relies on the fact that at least some of the missing migrants join the local workforce. This may not hold if potential U.S.-bound migrants substituted their decision by either (i) emigrating to unrestricted countries or (ii) migrating internally. In Online Appendix Table 2.B.2 and Figure 2.C.4, we provide evidence against both interpretations. However, if either international or internal substitution were in place, we would not observe any positive effect of IRP exposure on population growth, because missing migrants in exposed districts would not be missing altogether.

2.4.4 Emigration and Industrialization

In the previous subsection, we provide evidence that the Quota Acts increased labor supply in exposed districts. We now ask whether this translated into increased employment and, if so, whether there is heterogeneity across sectors. Historical scholarship suggests that emigrants could, potentially, take on industrial jobs. First, Italian emigrants to the United States were largely unskilled workers who took low-qualification jobs in manufacturing (Abramitzky and Boustan, 2017). Second, Italian firms during this period relied mostly on unskilled workers and employed labor-intensive production technologies (Cohen and Federico, 2001, p. 60). Hence, the increased supply of unskilled labor could be compatible with the demand by firms. To test this, we estimate model (2.3), taking as outcome variables changes in the number of workers employed in agriculture and manufacturing, as well as changes in the share of workers employed in both sectors as a fraction of overall employment.[37] As an alternative measure for broader modernization, we use the urbanization rate, calculated as the share of citizens living in municipalities with more than 5,000 inhabitants.[38]

[37] We harmonize the definition of industrial firms across censuses. For instance, transportation firms were not recorded as industrial firms in 1931, though they were in all other censuses.

[38] Urbanization has been widely used as a proxy for economic modernization. Among others, see Boustan *et al.* (2018) and Sequeira *et al.* (2020). We set the urban threshold at 5,000 inhabitants as this was the median city size before the Mass Migration (1881).

In Table 2.6, we show that while agricultural employment did not significantly react to the Quota Acts, industrial employment increased substantially.[39] This effect is consistent with the evidence presented in Table 2.3, which documents that firms in *manufacturing* decreased their investment in capital goods following the IRP shock. Taken together, these results suggest that manufacturing firms in exposed districts took advantage of more abundant labor unleashed by the IRPs and substituted capital investment with (now cheaper) labor. This evidence is therefore consistent with our directed technical adoption narrative.

In Table 2.7, we repeat the exercise but consider changes in the *share* of industrial and agricultural workers as the main outcome variable. We interpret the share of industrial workers as one further indicator of industrialization, whereas the opposite holds with respect to the share of workers employed in agriculture.[40] Because overall employment hardly reacts to the Quota Acts, industrial employment grows and agricultural employment does not, and the share of workers employed in manufacturing increases. Similarly, the share of workers employed in industry surged. Because industrial firms were the driving force behind economic and social progress during this period, Table 2.7 may suggest that the Quota Acts contributed to the modernization of the Italian economy, pushing comparatively more workers into modern industrial sectors. Finally, in the last column, we report that urbanization hardly increased in exposed districts. This might be driven by the fact that manufacturing firms were not located in urban centers, as shown in figure 2.C.7.

In Figure 2.3, we document sizable heterogeneity in capital investment and technology adoption decisions across sectors. We now ask whether the directed technical adoption mechanism allows the reconciliation of these dynamics with changes in sector-level industrial employment. We therefore estimate the baseline DiD model for the six sectors whose employment was collected in the population and manufacturing censuses. The outcome variable in each regression is the growth rate in sector employment, and we control for aggregate manufacturing employment growth. This is because we are interested in understanding which industrial sectors grew more relative to the increase in aggregate industrial employment. We report the results of this exercise in Table 2.8, where the first row displays the estimated impact of quota exposure. Employment dynamics reflect the heterogeneity in capital investment decisions. Employment in agriculture and fishing in more-exposed districts decreased. On the other hand, firms in First

[39]The OLS estimates report a modest decrease in agriculture employment. The estimated coefficient is marginally significant at the 10% level and small in magnitude. Moreover, the IV estimates of the agriculture coefficient are not significant. We conclude that agriculture employment did not react to the Quota shock.

[40]Our theory predicts that the *number* of workers employed in manufacturing in exposed districts should increase, whereas we do not expect any such effect on agriculture. In turn, this implies that the *share* of workers employed in manufacturing should increase and that the share in agriculture should decrease.

Industrial Revolution sectors—chiefly textiles, but construction as well—increased their labor stock. Moreover, we find that employment in the two distinctively Second Industrial Revolution sectors—namely chemicals and metallurgy—reacted less to the IRP shock, although we still find an increase in comparatively more-exposed districts. These results are entirely consistent with evidence reported in Figure 2.3. Our results suggest that faced with more-abundant unskilled labor, firms in textiles and construction substituted capital with labor, increasing employment and cutting investment in capital goods. By contrast, industrial firms in the agriculture sector reduced their overall labor stock and increased investment in capital goods. High value-added sectors did not respond as much to the labor supply shock, displaying smaller changes in their employment stock and investment in physical capital. All these findings are consistent with the baseline directed technical adoption narrative, and therefore provide evidence in favor of our proposed mechanism.

2.4.5 Discussion and Alternative Mechanisms

We have documented that the Quota Acts, arguably one of the most sudden and restrictive immigration restriction policies in modern history, led to decreased investment in capital goods and hampered technology adoption in more-exposed districts. To rationalize these findings, we showed that the IRP induced a larger positive labor supply shock in more-exposed districts. Throughout the paper, we have interpreted this evidence through the lens of directed technical change and adoption theory. In this section, we discuss some alternative mechanisms that could be compatible with our findings, and we touch on how data limitations might preclude some additional and potentially relevant analysis. We then briefly elaborate on the external validity of our results.

Human-capital spillovers ignited by out-migration have traditionally received sizable attention in the literature. Evidence by Spitzer and Zimran (2018) suggests that Italian emigrants to the United States were positively selected within Southern regions, implying that emigration was exerting a “brain drain” effect on Southern Italy. Under this interpretation, our estimated effects of the Quota Acts would be partially confounded by human-capital dynamics triggered by the IRP shock. More specifically, the drop in capital investment and technology adoption that we estimate might be driven by substitutability between capital goods and the upper tail of the skill distribution of workers, rather than by directed technical adoption. Even though this mechanism does not necessarily conflict with the one we propose, we view this as second-order in our setting, for two reasons. First, we find that the bulk of employment gains and capital investment losses materialized in First Industrial Revolution sectors. These occurred in traditionally low-skilled and labor-intensive manufacturing, especially in Southern regions

(A'Hearn, 1998). Hence it is unlikely that high-skilled workers would be comparatively more productive there. Second, we run a battery of robustness checks—see Online Appendix Tables 2.B.16 and 2.B.7. When we include the literacy rate as a proxy for average human capital in our regressions, results hold.

Along with the brain-drain effect, remittances are a traditionally major research topic within the emigration literature. Despite sizable global flows, Clemens (2011) argues that remittances can have at best a second or third-order effect on economic growth in sending countries when compared to the welfare effects of immigration restriction barriers. Building on this insight, we consequently abstracted from including remittances in our analysis, more so given that existing data are of questionable reliability at best. Remittance dynamics nonetheless represent a competing mechanism. More-exposed districts were receiving more remittances before the Quota Acts, hence they suffered the most from the border closure. Inasmuch as within-household cash transfers result in aggregate savings, remittances may accrue to overall investment dynamics (Rapoport and Docquier, 2006). A large literature has nonetheless documented that remittances are largely spent on consumption and invested in human—rather than physical—capital (for a review, see Yang, 2011).

A more sensible interpretation could be that remittances fostered literacy (e.g., Dinkelman and Mariotti, 2016; Fernandez-Sanchez, 2020). Exposed districts would have thus suffered from a relative drop in skilled workers following the Acts, and the labor force would have reshuffled toward unskilled sectors. This pattern would thus move in the opposite direction of the reverse-brain-drain effect. Under this interpretation, this channel does not conflict with the one we propose. If anything, it augments the relevance of exposure to the Quota Acts in generating an excess supply of workers, which triggered the directed technical incentive to abandon investment in physical capital. To quantify this concern, we run several robustness checks where we control for average human capital. The results of these exercises fully confirm our baseline estimates.

A plausible concern for our empirical strategy is that after the Quota Acts, emigrants simply substituted the United States with either internal or international unrestricted destinations.[41] Our main argument against this interpretation is backed by evidence in Table 2.5. If emigrants substituted the United States with other destinations, we would expect no effect of exposure to the Quota Acts on population growth. Given the persistence of demographic dynamics, it is unlikely that alternative explanations can account for such a sharp, sizable increase that is correlated with the conditionally exogenous variation we exploit. Disaggregated emigration

[41] If the Quotas fostered labor mobility within Italy, our estimates may fail to reflect the productivity gains this could induce (Bryan and Morten, 2019).

data toward countries other than the United States does not exist. However, in Figure 2.C.4, we report aggregate outflows toward the four main emigration destinations, before and after the treatment period(s). We show that the United States is the only country where immigration significantly departs from its historical level, except during WWI.[42] Moreover, the sheer numbers of internal migrations cannot account for the drop in U.S.-bound out-migration (Gallo, 2012). In Table 2.B.2, we show that in no Southern region did the gross outflow to Northern regions from 1921 to 1931 exceed 10% of U.S.-bound emigrants from 1910 to 1920. Qualitative and quantitative evidence alike, therefore, call for dismissing the emigration substitution argument.

A second reason precluding a causal interpretation of our estimates would be that—even when conditioning on the decision to emigrate—the choice of *where* to emigrate was systematically correlated with factors inducing an underlying correlation with local economic development. We provide and discuss evidence throughout this paper against this interpretation. Historical scholarship, however, notes that assimilation patterns of Italian immigrants in the United States and Argentina during this period substantially differed (Klein, 1983).[43] If this was caused by pre-migration differences in characteristics, then our identification scheme may fail. Using detailed data from censuses and passenger lists, Pérez (2021) nonetheless documents that the "success" of Italians in Argentina compared to Italians in the United States was unlikely to be caused by pre-migration differences in observable characteristics between the two groups. Emigrants to Argentina and the United States were essentially indistinguishable in terms of occupation and literacy rate, the only difference being that the former chiefly originated from Northern regions, whereas the latter mostly came from Southern areas. Selection patterns across the two groups do not display sizable differences, providing solid evidence in favor of our identification assumption.

Data limitations prevent us from studying two additional, potentially interesting variables, namely wages and output (productivity). Studying wages would be informative because directed technical adoption hinges on the relatively more abundant labor becoming relatively cheaper. An analysis of wages could reveal this pattern, which we currently implicitly assume. Geographically disaggregated data on wages, unfortunately, do not exist. Productivity would, in turn, be key to investigating the welfare effects of the Quota Acts. However, disaggregated data on output were not recorded until 1936; hence, we lack a time series covering the period

[42] These four countries are the United States, France, Argentina, and Brazil. Taken together, emigrants heading toward these destinations accounted for 70% of the total outflow. We predict the number of emigrants after 1924 using historical emigration before 1914. We show that the United States was the only country whose inflow falls relative to the prediction based on historical data after the Quota Acts.

[43] Argentina and the United States were the two leading destinations for Italian emigrants in this period. Klein (1983), among others, noted that Italian immigrants in Argentina had higher home-ownership rates and were more likely to be employed in skilled occupations compared to Italians in the United States.

we study.

It is not obvious that our results lend themselves to further generalization. Some similarities with contemporary settings nonetheless emerge. In terms of emigrant selection, the average Italian emigrant to the United States was slightly positively selected, left a rural area, and took on unskilled industrial jobs once in the United States (Sequeira *et al.*, 2020). This description is remarkably similar to contemporary emigration from poor countries, whereas it is starkly different from emigration from rich countries (e.g., Grogger and Hanson, 2011). While we do not claim that all our findings generalize to contemporary migration relationships, the evidence presented in this paper indicates that IRPs should be evaluated in terms of their joint effects on sending and receiving countries, beyond remittances and human-capital deprivation, as is standard in the existing literature.

2.5 Robustness Checks

In this section, we summarize our main robustness checks.[44] We essentially address two empirical problems. First, we provide evidence that our results so far are robust to alternative measures of treatment exposure across districts. Second, we propose two simple instrumental variables to deal with potential endogeneity issues relating to our estimates.

2.5.1 Alternative Measures of Treatment Exposure

There are two margins along which measured quota exposure may be subject to mismeasurement. First, while most Ellis Island records after 1900 report the district of origin, this is not true for the years 1890 to 1900. Records for these years most often only report "Italy" as the origin of a migrant.[45] Similarly, after the 1924 Emergency Act was enacted, Ellis Island authorities largely stopped recording immigrants' municipalities of origin. If there were systematic patterns underlying whether migrants were recorded with their district of origin or were simply recorded as Italian, then our measure would suffer from bias. Second, as discussed in Section 2.2, though emigration collapsed during WW1, it did not completely dry out. During the war,

[44] See the Online Appendix, Sections 2.B and 2.C, for detailed tables reporting the results which we discuss here.

[45] Online Appendix section 2.A.1.2 reports the number of migrants whose origin we label as missing. The share of Ellis Island immigrants with missing origin never exceeds 1% of the overall number of immigrants in any given year over the period 1892-1924.

districts closer to emigration ports are in fact disproportionately represented relative to previous shares.[46] This induces spurious variation in measured quota exposure, as we would impute higher exposure to some districts by sole virtue of their geographic position.

The first robustness check we thus consider restricts the sample years over which quota exposure is computed. In our baseline specification of equation (2.1), we measure the exposure of a given district as the share of people who migrated from that district from 1890 to 1924, relative to that district's population in 1880. To make sure that emigration registration procedures and WW1 do not induce systematic measurement error in our estimates, we introduce two other treatment variables. As a first alternative, we consider only emigrants who left no later than 1921. Then, we further restrict the sub-sample to the years before the outbreak of WWI. The first alternative measure seeks to control for the fact that the Ellis Island database lacks information about the municipality of origin for a high number of Italian migrants after 1921. We thus aim to clean for possible measurement error due to the nonrandom selection of registered district locations. The second exposure measure drops emigrants who left after WWI started, as emigration opportunities were possibly affected by proximity to departure ports. In particular, emigrants from districts nearer to ports could be over-represented.

Our baseline results are robust to these different measures of quota exposure, as shown in Online Appendix tables 2.B.13-2.B.14-2.B.15. Most likely, this is because the bulk of emigration took place before 1914, hence restricting the sample to the years before WW1 does not substantially affect our estimated treatment exposure. In particular, though districts closer to ports are over-represented in emigration statistics during WW1, the absolute number of emigrants was negligible relative to previous years, as WWI induced a marked collapse in those districts as well. Finally, emigrants lacking a recorded district of origin constitute the majority for the post-1924 period. Yet, we find no noticeable pattern inducing nonrandom recording across districts. Hence, measured quota exposure should not be mismeasured whether we include those years or not, as confirmed by the estimated coefficients. One further concern is that our results might be driven by remote migration patterns. According to the Gould (1980b) hypothesis, in fact, out-migration from any given region would eventually saturate over time. Hence, it might be that our estimated effects are driven by districts whose out-migration stretches back to years before the Quota shock becomes salient. Similarly, one may wonder whether it is instead more recent emigration waves that drive the results. In Online Appendix tables 2.B.13-2.B.14-2.B.15 we address these concerns by constructing two measures of Quota Exposure which assign increasing

[46] Throughout this period, emigrants could sail overseas only from Naples, Palermo, or Genoa. In Online Appendix section 2.A.1.3 we show that the correlation between our newly constructed emigration series and official statistics is lowest during the WW1 years. We thus report robustness regressions excluding those years from our measured Quota Exposure, and confirm all our baseline estimates.

or decreasing weights on more recent out-migration flows. We find that all our baseline results hold.

2.5.2 Shift-Share Instrumental Variable

A possible concern for our identification strategy is that geographical variation in exposure to the U.S. immigration quotas was not conditionally random across districts. While we provided historical and quantitative evidence against this argument, ultimately the exclusion restriction cannot be formally tested. We, therefore, develop an instrument close in spirit to that presented in Card (2001) and Tabellini (2020) to address a similar—although specular—issue.

Let $\omega_{cr}^{T} \equiv \sum_{\tau=0}^{T} \text{US Emigrants}_{c,\tau} / \text{US Emigrants}_{T}$ be the share of emigrants from district c in region—or province—r until time T (US Emigrants$_{c,T}$) relative to total emigration (US Emigrants$_{T}$). We predict total emigrant outflow from district c from the following "zero-stage" equation:

$$\widehat{\text{US Emigrants}}_{cr}^{T} = \omega_{cr}^{T} \times \sum_{\tau=1890}^{1924} \sum_{c' \notin r} \text{US Emigrants}_{c',\tau} = \omega_{cr}^{T} \times \text{US Emigrants}_{-r} \tag{2.4}$$

In the first stage, we instrument QE_{cr} using $\widehat{\text{US Emigrants}}_{cr}^{T}$, then we plug the resulting predicted $\widehat{\text{QE}}_{cr}^{T}$ into the second-stage regression to estimate the baseline model (2.3). To strengthen the validity of our OLS estimates, we pick T to be before the bulk of the Mass Migration period. Thus, predicted district-level U.S. emigration outflows wash out spurious variation in U.S. emigration due to emigration—endogenously—affecting economic development in emigration districts, conditional on district and year-fixed effects.

The instrumental variable (2.4) exploits two sources of variation. Cross-sectional variation is embedded in the (ω_{cr}^{T}) term. It captures heterogeneity in the origin districts of migrants at a given point in time (t). We can modulate the choice of T so that the distribution of emigrants across districts is more plausibly driven by exogenous information diffusion, and less so by economic outcomes (Spitzer and Zimran, 2020). Time series variation, captured by (US Emigrants$_{-r}$), is driven by changes in the aggregate emigration outflow, excluding the instrumenting district c, and possibly all other districts in the same region (or province). This "leave-out" strategy ensures that our instrument is not correlated with the economic performance of districts in region r, hence mitigating the concern that quota exposure could be correlated with district-level economic performance hence inducing endogeneity and bias our estimated coefficients. By changing T, we address the possible concern that WWI altered the composition of Italian emigrants to the United States in a spatially nonrandom fashion.

In Table 2.B.8, we summarize the results of the first-stage regressions, where we vary measured quota exposure as discussed in Section 2.5.1. We also control for different baseline years T in the construction of the Shift-Share Instrument to make sure emigration patterns reflect district-level variation, which is not correlated with economic performance. The first stage is statistically significant because the instrument has high explanatory power, as we would expect for emigration—and immigration—patterns exhibiting substantial persistence. Minor changes arise in the first stage when comparing results for the two different baseline years considered, 1895 and 1900. An advantage of picking T less than 1906 is that we wash out variation induced by the Messina-Reggio Calabria earthquake (Spitzer *et al.*, 2020). Tables 2.9 and 2.10 compare results from the OLS estimation and from the second stage of our IV regression for different outcome variables, specifically, measures for capital investment, industrialization, urbanization, and population growth. No major differences arise between the two estimations. However, IV regression on population growth yields slightly higher estimates: downward bias in the OLS could arise if the conditional identifying variation was regionally clustered within the South. This however affects neither the sign nor the significance of the results.

2.5.3 Railway-Access Instrumental Variable

Several recent papers call for caution on the use of Bartik instrumental variables (Jaeger *et al.*, 2018; Goldsmith-Pinkham *et al.*, 2020). In our context, the proposed shift-share IV suffers from endogeneity issues if the initial spatial variation of migration patterns was correlated with economic development at baseline. To address this concern, we develop an IV based on the timing when Italian districts became connected to the railway network, similarly to Sequeira *et al.* (2020). In general, gaining access to the railway system in this period drastically reduced transportation costs for potential emigrants, hence increasing the total migration outflow. On top of this, the rationale behind our instrument is that transoceanic migration required a district to be connected to an emigration port.[47] Specifically, because U.S.-bound emigrants could leave only from Genoa, Naples, or Palermo, we leverage variation in the timing when districts became connected to one of these ports to instrument actual U.S.-bound migration outflows.

[47] As Calabrese (2017, pp.52, 90) puts it:

> "The lack of railroads contributed to the isolation. [...] It was only between 1880 and 1900 that over 1,250 miles of railroad were constructed in region [Basilicata], making it more accessible for travel and facilitating emigration. [...] From Potenza and towns in the western part of Basilicata, migrants could travel to Naples by railroad. The building up of infrastructure in Basilicata aided emigrants in traveling to their port of departure."

Let $\text{RA}_{cr,t}$ denote an indicator variable that returns the value one if district c in region r is connected to the railway system in decade t, and zero otherwise. We define railway access to emigration ports $\text{RAP}_{cr,t}$ as follows:

$$\text{RAP}_{cr,t} \equiv \text{RA}_{cr,t} \times \min\left\{d_t(c,\text{Naples}), d_t(c,\text{Palermo}), d_t(c,\text{Genoa})\right\}^{-1} \tag{2.5}$$

where $d_t(c,i)$ is the geodesic distance over the railway network in decade t between district c and emigration port i.[48] Because the network evolves over time, we allow the geodesic distance between each district and the closest emigration port to reflect this time variation. A natural test of the hypothesized role of the railway system in shaping the direction of emigration would be to observe a positive correlation between our measured access $\text{RAP}_{cr,t}$ and the relative share of emigrants headed toward the United States.[49] Evidence presented in the next paragraph confirms this.

Following Sequeira *et al.* (2020), we estimate the following "zero-stage" model:

$$\begin{aligned}\text{US Emigrant Share}_{cr,t} = {} & \alpha_c + \alpha_{r,t} + \beta \text{US Emigrant Share}_{cr,t-1} + \gamma \text{RAP}_{cr,t-1} \\ & + \delta\left(\text{Industrialization}_{r,t-1} \times \text{RAP}_{cr,t-1}\right) + \zeta\left(\text{RAP}_{cr,t-1} \times \text{ Emigrants}_{r,t-1}\right) \\ & + \boldsymbol{x}'_{cr,t}\boldsymbol{\eta} + \varepsilon_{cs,t}\end{aligned} \tag{2.6}$$

where t denotes decades spanning the 1890-1920 period; α_c and $\alpha_{r,t}$ denote district and region-by-year fixed effects; $\text{Emigrants}_{r,t-1}$ is the total number of emigrants leaving region r, where $c \in r$, during decade t, normalized by the total population in that region in 1881; $\text{Industrialization}_{r,t-1}$ is the share of workers employed in manufacturing in region r,[50] and $\boldsymbol{x}_{cr,t}$ is a set of controls consisting of lagged population, a South dummy interacted with lagged railway access, and labor-market slackness. The outcome of interest, $\text{US Emigrant Share}_{cr,t}$, is the share of U.S. emigrants from district c in region r in decade t over district c's population in 1881, and $\text{US Emigrant Share}_{cr,t-1}$ is its lagged value. Our main coefficient of interest is ζ. This captures

[48] In graph theory, the geodesic distance is defined as the shortest path between two nodes. More formally, let the railway system in decade t—call it $\mathcal{N}_t$—be defined as the pair (V,E), where V is the set of nodes, and $E = \{(u,v)|(u,v) \in V^2,\ u \neq v\}$ is the set of edges. Let $\mathbf{A}$ denote the adjacency matrix associated to E, where for every couple of vertices $v,u \in V$, $A_{uv} = 1$ if there is an edge between u and v, and zero otherwise. The (geodesic) distance $d(u,v)$ between the two vertices is the minimum r such that $[\mathbf{A}^r]_{uv} = 1$ (Newman, 2018).

[49] Clearly, emigration toward South America would have equally benefited from railway connection to emigration ports. However, U.S.-bound emigrants easily outnumbered emigrants bound for South America in this period.

[50] Controlling for the share of workers employed in manufacturing serves a twofold purpose. On one hand, it washes out variation in U.S. emigration due to more affluent districts being granted access to the railway system relatively sooner than backward ones (Sequeira *et al.*, 2020). Second, the timing of connection to the railway may itself affect economic development, for instance through increased specialization and industrialization (*i.a.* Donaldson and Hornbeck, 2016; Donaldson, 2018). This would generate endogenous variation, which we wash out when constructing the instrument.

how changes in railway closeness to emigration ports influenced U.S.-bound emigration during periods of high *vis-à-vis* low overall aggregate emigration, accounting for the district population in 1881, i.e., before the mass emigration began. We thus expect the estimate of ζ to be positive. In turn, we expect the estimate of γ to be close to zero, because it reflects how railway access affected U.S. emigration in decades with little overall emigration. The estimated coefficients of regression (2.6) confirm these predictions (for the sake of brevity, we do not report them). One may suspect that the construction of the railway was not random across districts, because more-affluent areas were connected before poorer ones, so we include the interaction between the share of industrial workers and railway access as one further control.

The estimation equation (2.6) yields a set of estimated coefficients that allow us to construct a predicted aggregate series of the share of U.S. emigrants, which we then aggregate up across decades as follows:

$$\widehat{\mathrm{QE}}_{cr} \equiv \sum_{t=1890}^{1920} \hat{\zeta} \left(\mathrm{RAP}_{cr,t-1} \times \mathrm{Emigrants}_{r,t-1}\right) \tag{2.7}$$

We instrument quota exposure with $\widehat{\mathrm{QE}}_{cr}$, then we estimate the resulting instrumented DiD model in a standard two-stage-least-squares setting.

Table 2.B.8 reports the results of the first-stage regressions. The “RA region” column reports the results of the baseline instrument, whereas the “RA total” column uses a variation on equation (2.6) where, instead of the aggregate number of emigrants in the region, we plug in the overall nationwide number of emigrants. We find that there is a strong and positive association between the synthetic and the actual series of U.S.-bound emigrants. Although the F statistics using the railway instrument are not as high as those of the Bartik IVs, these nonetheless provide evidence suggesting that the instrument is not weak. Tables 2.9 and 2.10 compare the second-stage results with the OLS estimates for, respectively, technology adoption and population and employment variables. The railway IV always confirms the baseline estimates in sign and magnitude and, in most cases, preserves their significance.

2.6 Conclusion

In recent years, immigration has become an increasingly focal and polarizing theme in the public debate. Policymakers exhibit widely divergent opinions about the effects of increased immigration on the economic, social, and cultural security of native populations. Yet, a common perspective can be disentangled. Both proponents and opponents of harsher immigration-restriction policies judge them in terms of their effects on their own country, that is the country *subject to* immigration. Few mention, possibly due to relatively scarce evidence, that immigration policies may entail important, even determinant, effects on sending countries. This asymmetric attention in favor of receiving countries is worrisome, given that sending nations often experience greater economic hardship and social distress.

In this study, we explore how restrictive immigration policies shape economic development in sending countries. This poses two empirical challenges. First, emigration is seldom directed toward one—or very few—countries, hence it is difficult to identify the effect of a single immigration policy shift in one such receiving country. Second, migration dynamics are likely affected by preexisting regulations enacted by both receiving and sending countries. To tackle both issues, we study the Italian emigration to the United States during the Age of Mass Migration (1850–1914). Through the 1921 and 1924 Quota Acts, the United States adopted a harshly restrictive immigration policy, which starkly contrasted with the open-border approach that it had maintained almost uninterruptedly since the 1810s. Comparing districts with similar emigration rates but different destinations, we leverage identifying variation in exposure to the Quota Acts to estimate the impact of immigration restriction laws in a difference-in-differences framework.

We find that industrial firms in more-exposed districts underwent sizable reductions in capital investment and a slowdown in technology adoption. These effects are larger for more advanced capital vintages and in relatively backward manufacturing sectors. To rationalize these findings, we advance and validate the hypothesis that IRPs induce a positive labor supply shock on countries sending migrants. Through the lenses of directed technical change and adoption theory, more-abundant labor dampens the incentives for firms to invest in labor-saving, possibly productivity-enhancing, production technologies (e.g., Zeira, 1998; Acemoglu, 2007). We document that population growth increased in comparatively more-treated districts, consistent with the idea that the Quota Acts prevented people who would have migrated from doing so. Our empirical results endorse the directed technical adoption mechanism—we observe that in highly exposed districts, industrial employment increased while agricultural employment did not. Shifting our analysis to manufacturing sectors, we find that sectors where capital investment decreased the most were also the ones that absorbed the bulk of the labor supply shock induced by the Quota Acts. This is consistent with the idea that firms in relatively backward industrial sectors substituted capital-intensive production technologies with labor, which the IRP shock made more abundant (and cheaper).

Taken together our results indicate that immigration restriction policies exert substantial effects on the economic development of sending countries. An immigration restriction shock impresses upward pressure on the labor supply driven by foregone migrants in the sending country. In our setting, this dampened the incentive for manufacturing firms to adopt productivity-enhancing technology. Faced with more abundant labor, firms substituted capital with more labor-intensive production technologies. Because technology adoption is a well-known driver of long-run growth (Juhász *et al.*, 2020), evidence in this paper suggests that immigration restriction policies have potentially long-lasting effects on the economic development of sending countries. The external validity of these findings is not obvious. However, we argue that neither the Italian economy nor emigrants' characteristics during the 1920s were fundamentally different from many of today's developing countries. Hence, we believe history can inform the contemporary debate on this crucial issue.

Tables

TABLE 2.1: Summary Statistics

	(1) N. of Obs.	(2) Mean	(3) Std. Dev.	(4) 10 pct.	(5) 50 pct.	(6) 90 pct.
Panel A: Demographics and Geography						
Area	1070	121.08	77.12	45.93	98.31	240.66
Altitude	1070	0.33	0.22	0.07	0.31	0.63
Population	1066	165.25	156.88	53.37	122.36	319.56
5-Urbanization	1066	0.60	0.26	0.25	0.59	0.95
10-Urbanization	1066	0.37	0.27	0.00	0.31	0.80
15-Urbanization	1066	0.28	0.26	0.00	0.24	0.63
Panel B: Emigration						
Emigration (1890-1930)	1080	284.82	266.57	57.82	238.57	496.41
Emigration (1890-1921)	1080	259.69	241.90	52.57	212.73	453.95
Emigration (1890-1914)	1080	230.64	226.87	42.55	185.68	389.12
US Emigration (1890-1930)	1080	73.20	81.88	7.41	43.51	164.73
US Emigration (1890-1921)	1080	67.26	74.89	6.88	40.40	152.31
US Emigration (1890-1914)	1080	57.71	64.79	5.66	36.08	130.94
Panel C: Employment						
Agriculture Workers	1062	42.70	26.99	16.23	37.45	75.12
Manufacture Workers	1069	21.54	32.80	3.97	11.74	45.64
Trade Workers	1070	5.78	9.93	1.09	2.95	10.88
Liberal Professions	1062	2.48	4.46	0.38	1.28	4.66
Public Administration	1062	3.88	7.86	0.59	1.84	7.34
Panel D: Capital and Technology						
Firms	1061	8278.04	9725.53	587.70	5262.13	19054.43
Firms with Engine	1061	1336.61	2032.29	137.06	679.12	3038.19
Mechanical Engines	1061	816.69	672.69	250.94	554.21	1782.77
Electrical Engines	1061	6051.59	21620.63	84.29	1055.33	12809.84
Mechanical Horsepower	1061	96168.86	163951.19	6021.24	26237.85	310569.50
Electrical Horsepower	1061	53083.77	142887.45	660.49	9552.07	134462.30

Notes. This table reports symmary statistics for the variables in our dataset, except sector-specific capital and employment. All variables are in levels. Area, altitude, population, employment, and emigration are expressed in thousands. Section 2.3 explains how we impute province-level data to districts, and provides details on the sources employed.

TABLE 2.2: Balance Table

	Level		Growth Rate	
	(1)	(2)	(3)	(4)
	1911	1921	1911	1921
All Firms	0.017	0.007	0.029	-0.022
	(0.019)	(0.027)	(0.032)	(0.032)
Firms with Engine	0.033	0.027	0.048	-0.012
	(0.036)	(0.066)	(0.087)	(0.110)
Mechanical Engines	0.005	-0.016	0.089	-0.168
	(0.072)	(0.088)	(0.177)	(0.202)
Electrical Engines	0.005	0.004	0.005	-0.001
	(0.010)	(0.009)	(0.020)	(0.022)
Mechanical Horsepower	-0.038	-0.021	-0.095	0.056
	(0.029)	(0.051)	(0.078)	(0.098)
Electrical Horsepower	-0.007	-0.012	-0.004	-0.026
	(0.026)	(0.039)	(0.053)	(0.070)
Population	-0.000	0.000^{**}	-0.037	0.106
	(0.000)	(0.000)	(0.166)	(0.193)
Manufacture Workers	0.005	0.007	-0.028	0.017
	(0.103)	(0.094)	(0.101)	(0.075)
Agriculture Workers	0.031	0.006	-0.144	-0.048
	(0.096)	(0.125)	(0.153)	(0.127)
Trade Workers	-0.050	-0.032	-0.151	0.099
	(0.092)	(0.094)	(0.133)	(0.075)
Liberal Professions	-0.017	0.006	0.005	0.120
	(0.114)	(0.070)	(0.113)	(0.230)
Public Administration	0.065	0.088	0.027	0.036
	(0.128)	(0.204)	(0.105)	(0.129)

Notes. This table reports the correlation between the treatment measure (QE) and the covariates we use as outcome variables, before the Quota Acts were enacted. Quota exposure is defined as the ratio between US emigrants 1890-1924 and 1880-population. All regressions control for the emigration rate, defined as the ratio between emigrants 1890-1914 and 1880-population, and province fixed effects. Standard errors are clustered at the district level. In the first two columns, the outcome variable is in level; in the last two columns, it is defined in growth rate. Dependent variables are standardized to compare coefficients across models. Under validity of the parallel trends assumption, we require all coefficients not to be statistically different from zero.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.3: Investment in capital goods and emigration

	Firm		Engine		Horsepower	
	(1) All	(2) Engine	(3) Mechanic	(4) Electric	(5) Mechanic	(6) Electric
Quota Exposure × Post	0.128	0.299	-1.015***	-1.098***	-0.613**	-1.268***
	(0.235)	(0.401)	(0.162)	(0.327)	(0.308)	(0.294)
Extensive Margin × Post	-0.093	0.017	0.184*	0.138	-0.271**	0.010
	(0.108)	(0.186)	(0.101)	(0.130)	(0.114)	(0.148)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	207	208	208	207	209	208
Observations	783	785	785	783	787	785
R^2	0.457	0.737	0.844	0.473	0.663	0.841
F-stat	0.772	0.250	12.756	5.578	3.095	7.101
Mean Dep. Var.	0.139	0.128	0.107	0.248	0.020	0.187
Std. Beta Coef.	0.027	0.030	-0.181	-0.187	-0.059	-0.128

Notes. This table displays the effect of being exposed to the Quota Acts on various measures for capital investment and technology adoption. The first and second columns report the effect on, respectively, the number of all firms, and firms with engines. The third and fourth columns show the effect on the number of mechanical and electrical engines; the fifth and sixth display the effect on mechanical and electrical horsepower. All regressions include district and year fixed effects. Additional controls are the log-population and labor market slackness at baseline interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are always clustered at the district level. Standardized betas refer to the baseline Q.E. coefficient.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.4: Labor intensity and emigration

	Worker/Firm		Worker/Engine		Worker/Horsepower	
	(1) All	(2) Engine	(3) Mechanic	(4) Electric	(5) Mechanic	(6) Electric
Quota Exposure × Post	0.208	0.184	1.135***	1.050***	0.248	1.212***
	(0.239)	(0.396)	(0.174)	(0.339)	(0.353)	(0.300)
Extensive Margin × Post	0.051	-0.072	-0.235**	-0.114	0.421***	0.005
	(0.142)	(0.162)	(0.103)	(0.132)	(0.125)	(0.150)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	209	209	208	207	209	208
Observations	785	787	785	783	786	785
R^2	0.522	0.725	0.837	0.456	0.642	0.836
F-stat	6.364	7.630	23.588	3.584	7.179	5.482
Mean Dep. Var.	-0.082	-0.054	-0.077	-0.258	-0.078	-0.195
Std. Beta Coef.	0.036	0.017	0.188	0.180	0.023	0.123

Notes. This table displays the effect of being exposed to the Quota Acts on various measures for labor intensity in production. The first and second columns report the effect on, respectively, the worker-per-firm and the worker-per-firm with engine ratios. The third and fourth columns show the effect on the ratio between worker and mechanical and electrical engines; the fifth and sixth display the effect the ratio between workers and mechanical and electrical horsepower. All regressions include district and year fixed effects. Additional controls are the log-population and labor market slackness at baseline interacted with a post-treatment dummy. Outcome variables are defined in growth rate. District fixed effects refer to 1921-*circondari*. Standard errors are always robust and clustered at the district level. Standardized betas refer to the baseline Q.E. coefficient.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.5: Population Growth and Emigration

	Continuous QE		Categorical QE	
	(1)	(2)	(3)	(4)
Quota Exposure × Post	0.409***	0.449***		
	(0.113)	(0.124)		
Quota Exposure Dummy × Post			0.021***	0.023***
			(0.006)	(0.007)
Extensive Margin × Post		-0.068		-0.051
		(0.055)		(0.053)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Number of Districts	204	204	204	204
Observations	751	751	751	751
R^2	0.452	0.452	0.445	0.445
F-stat	13.337	9.932	13.298	10.086
Mean Dep. Var.	1.042	1.042	1.042	1.042
Std. Beta Coef.	0.219	0.240	0.194	0.210

Notes. This table displays the effect of exposure to the Quota Acts on population growth. Population growth is defined as the decade-on-decade percentage change in population. Continuous QE is the baseline measure defined in (2.1); Categorical QE equals one if the continuous measure is above 1, and 0 otherwise. All regressions control for log-population and labor market slackness in 1901, interacted with a post-treatment measure. Models in columns (2) and (4) include the emigration rate defined as the number of emigrants 1890-1914 over the 1880-population, interacted with a post-treatment dummy. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are always clustered at the district level. Standardized betas refer to the baseline Q.E. coefficient.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.6: Employment in Industry and Agriculture

	Industry Growth		Agriculture Growth	
	(1)	(2)	(3)	(4)
Quota Exposure × Post	1.827***	1.510***	-0.416*	-0.483*
	(0.427)	(0.475)	(0.159)	(0.176)
Extensive Margin × Post		0.637		0.154
		(0.400)		(0.149)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Number of Districts	205	205	206	206
Observations	742	742	750	750
R^2	0.540	0.542	0.461	0.465
F-stat	6.805	7.004	3.556	3.250
Mean Dep. Var.	0.060	0.060	-0.041	-0.041
Std. Beta Coef.	0.149	0.123	-0.116	-0.135

Notes. This table reports the effect of exposure to the Quota Acts on industrial and agricultural employment growth. Sector employment growth are defines as the decade-on-decade changes in employment. All regressions include district and year fixed effects. Further controls include log-population and labor market slackness in 1901 interacted with a post-treatment dummy. Columns (3) and (4) control for the emigration rate, defined as the number of emigrants 1890-1914 relative to 1880-population, interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are robust and clustered at the district level. Standardized betas refer to the baseline Q.E. coefficient.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.7: Urbanization and Share of Workers Employed in Industry and Agriculture

	Industrialization		Agriculture		Urbanization	
	(1)	(2)	(3)	(4)	(5)	(6)
Quota Exposure × Post	1.457***	1.152***	-0.580***	-0.605***	0.218	0.252*
	(0.356)	(0.410)	(0.145)	(0.156)	(0.145)	(0.148)
Extensive Margin × Post		0.598*		0.066		-0.086
		(0.350)		(0.085)		(0.099)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	205	205	204	204	201	201
Observations	729	729	743	743	742	742
R^2	0.476	0.478	0.510	0.510	0.174	0.174
F-stat	6.085	6.494	5.470	4.049	1.125	1.025
Mean Dep. Var.	0.051	0.051	-0.022	-0.022	1.039	1.039
Std. Beta Coef.	0.153	0.121	-0.172	-0.180	0.094	0.109

Notes. This table reports the effect of exposure to the Quota Acts on urbanization and changes in the share of industrial and agricultural workers relative to overall employment. Urbanization is defined as the share of the population living in cities no smaller than 5,000 inhabitants. The share of sector employment is defined as the ratio between sector and aggregate employment. All regressions include district and year fixed effects. Further controls are log-population and labor market slackness in 1901 interacted with a post-treatment dummy. Columns (2), (4) and (6) control for the emigration rate, defined as the number of emigrants 1890-1914 relative to 1880-population, interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are robust and clustered at the district level. Standardized betas refer to the baseline Q.E. coefficient.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.8: Changes in Industry Employment by Sector

	(1) Mining	(2) Agriculture	(3) Steel	(4) Construction	(5) Textile	(6) Chemical
Quota Exposure × Post	0.442	-2.459*	1.379	6.103***	5.651***	0.017
	(0.388)	(1.261)	(1.573)	(1.626)	(1.398)	(0.308)
Extensive Margin × Post	-0.000	1.029	-1.124	-2.693**	-0.715	0.181
	(0.287)	(1.257)	(1.576)	(1.293)	(0.991)	(0.277)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	194	200	198	200	200	195
Observations	685	776	775	778	774	681
R^2	0.071	0.424	0.106	0.317	0.449	0.450
F-stat	8.152	5.645	5.030	16.662	4.555	1.849
Mean Dep. Var.	0.724	0.422	0.250	0.553	0.291	0.751
Std. Beta Coef.	0.008	-0.134	0.096	0.180	0.124	0.000

Notes. This table displays the effect of exposure to the Quota Acts on changes in employment by manufacture sector. Hence, column “Agriculture” reports the impact of QE on employment in manufacture firms working in agriculture, not that on agriculture. We do not show the “public utility” sector due to data availability, and a residual sector of unassigned firms. All regressions include district and year fixed effects. Further controls are log-population, changes in industrial employment, the emigration rate and 1901 labor market slackness interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are clustered at the district level. Standardized betas refer to the baseline Q.E. coefficient.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.9: Investment in capital goods and emigration - 2sls

	Firm		Engine		Horsepower	
	(1) All	(2) Engine	(3) Mechanic	(4) Electric	(5) Mechanic	(6) Electric
Panel A: OLS						
Quota Exposure × Post	0.128	0.299	-1.015***	-1.098***	-0.613**	-1.268***
	(0.235)	(0.401)	(0.162)	(0.327)	(0.308)	(0.294)
Panel B: 2SLS Shift Share						
Quota Exposure × Post	0.429*	0.661*	-0.850***	-0.857***	-0.568*	-1.098***
	(0.233)	(0.398)	(0.164)	(0.318)	(0.329)	(0.287)
Panel C: 2SLS Railway Regional						
Quota Exposure × Post	0.603	-0.822	-0.895**	-1.472*	-0.178	-1.097**
	(0.454)	(1.272)	(0.374)	(0.868)	(0.955)	(0.552)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	207	208	208	207	209	208
Observations	783	785	785	783	787	785
Mean Dep. Var.	0.139	0.128	0.107	0.248	0.020	0.187
Std. Beta OLS	0.027	0.030	-0.181	-0.187	-0.059	-0.128
Std. Beta SS	0.090	0.067	-0.152	-0.145	-0.055	-0.111
Std. Beta RA	0.060	0.077	-0.170	-0.140	-0.073	-0.138

Notes. This table reports the effect of Quota exposure on various measures of capital investment and technology adoption. Panel A presents reduced form estimates. Panel B reports 2SLS estimates based on the instrument defined in (2.4). Panel C reports 2SLS estimates based on te instrument defined in (2.6). The first and second columns report the effect on, respectively,the number of all firms, and firms with engines. The third and fourth columns show the effect on the number of mechanical and electrical engines; the fifth and sixth display the effect on mechanical and electrical horsepower. All regressions include district and year fixed effects. Additional controls are log-population and labor market slackness at baseline interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are always clustered at the district level. Standardized betas refer to the baseline Q.E. coefficient.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

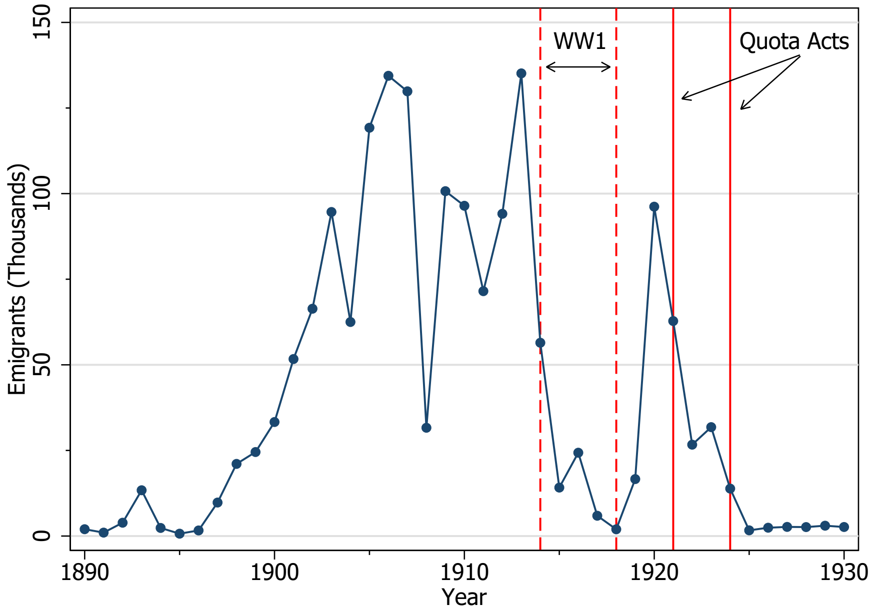
TABLE 2.10: Population Growth, Employment in Industry and Agriculture

	(1) Population Growth	(2) Industry Growth	(3) Agriculture Growth
Panel A: OLS			
Quota Exposure × Post	0.449***	1.510***	-0.483*
	(0.124)	(0.475)	(0.176)
Panel B: 2SLS Shift Share			
Quota Exposure × Post	0.668***	1.673***	-0.138
	(0.138)	(0.544)	(0.222)
Panel C: 2SLS Railway Regional			
Quota Exposure × Post	0.933***	3.385**	-0.733
	(0.248)	(1.347)	(0.479)
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Number of Districts	207	205	209
Observations	754	742	753
Mean Dep. Var.	0.042	0.060	-0.041
Std. Beta Coef. OLS	0.240	0.123	-0.135
Std. Beta Coef. Shift-Share	0.360	0.137	-0.038
Std. Beta Coef. Railway	0.503	0.276	-0.203

Notes. This table reports the effect of exposure to the Quota Acts on industrial and agricultural employment growth. Sector employment growth are defines as the decade-on-decade changes in employment. Panel A presents reduced form estimates. Panel B reports 2SLS estimates based on the instrument defined in (2.4). Panel C reports 2SLS estimates based on te instrument defined in (2.6). All regressions include district and year fixed effects. Additional controls are log-population and labor market slackness at baseline interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are always clustered at the district level. Standardized betas refer to the baseline Q.E. coefficient.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Figures

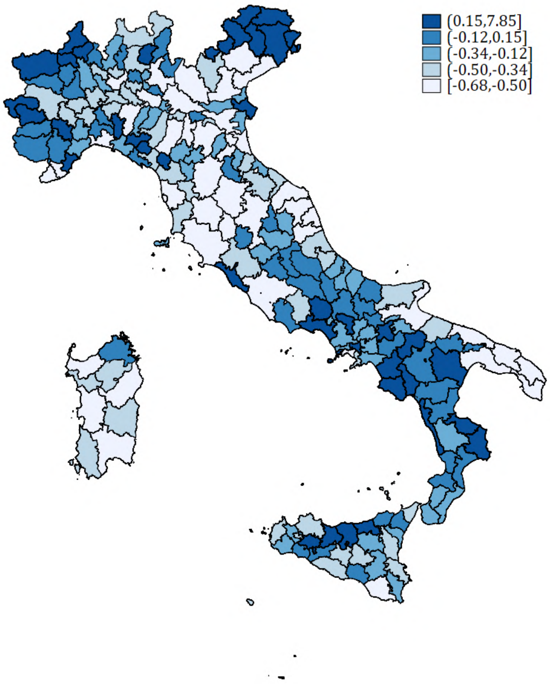
FIGURE 2.1: Total Inflow of Italian Immigrants at Ellis Island



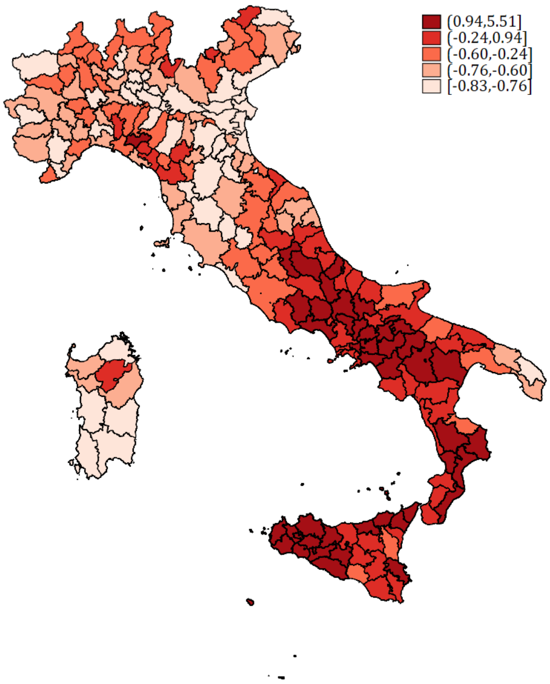
Notes. This figure displays the aggregate number of Italians who registered at the Ellis Island immigration station between 1890-1930. Dashed red lines indicate the period of WW1; solid red lines indicate the 1921 Emergency Quota Act and the 1924 (Johnson-Reed) Immigration Act. Only migrants whose origin we are able to trace are counted in the sum. Refer to the Online Appendix for details on the linking procedure.

Figure 2.2: District-Level Migration Flows, 1890-1924

(a) Emigration Rate

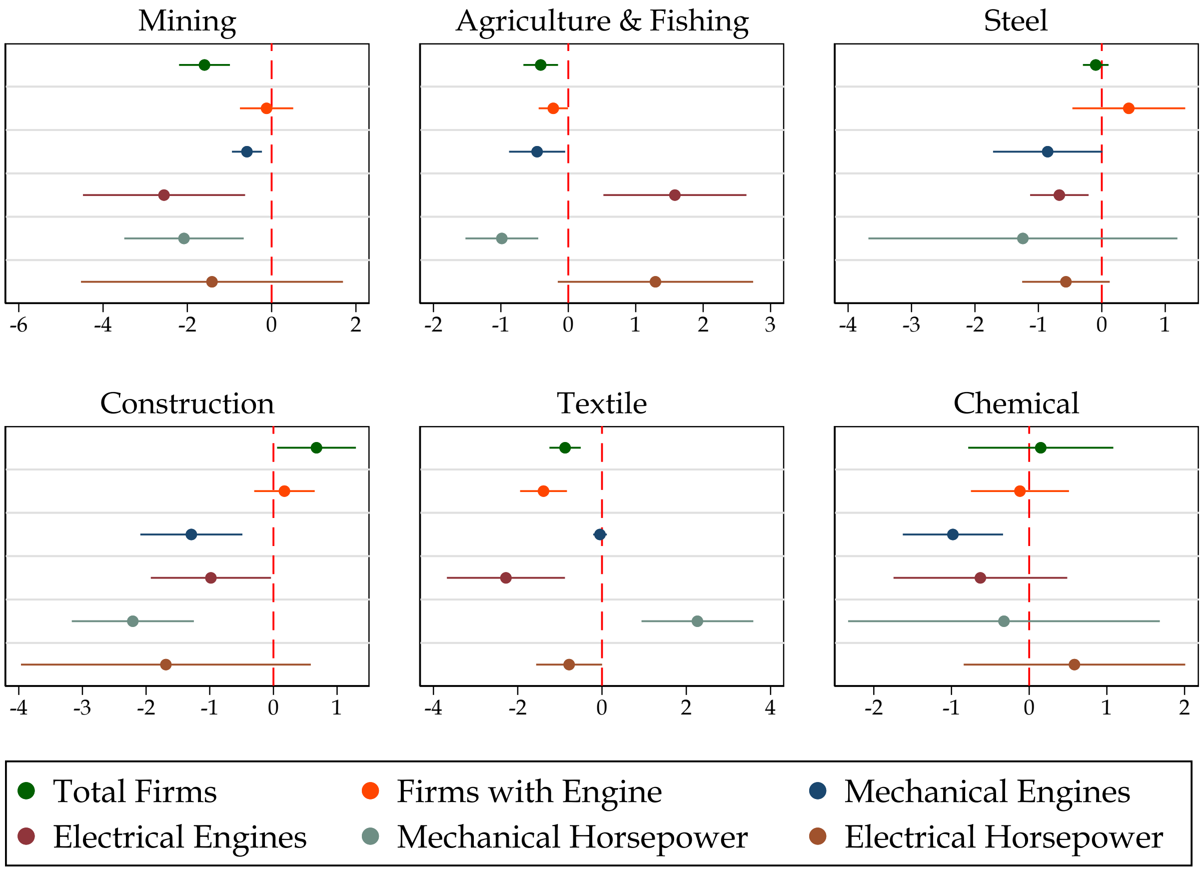


(b) Quota Exposure



Notes. Panel (a) displays variation in the emigrants-to-population ratio (emigration rate). Panel (b) plots the unconditional variation in the US emigrants-to-population ratio (quota exposure). Both figures normalize the number of emigrants by population in 1880, and report standardized variables in log. All figures plot the flows obtained setting $\alpha = .01$ in the matching process. Refer to the Online Appendix for more details and plots for different values of α.

FIGURE 2.3: Capital Investment and Emigration by Industry Sectors



Notes. This figure displays the effect effect exposure to the Quota Acts on capital investment and technology adoption by manufacture sectors. Each marker reports the estimated coefficient in model (2.3) where the outcome is the row-variable. Outcomes are the raw count of firms and firms with engines; the number of electrical and electrical engines; mechanical and electrical horsepower. All regressions include district and year fixed effects. Further controls are log-population, average industrial employment growth, the emigration rate and 1901 labor market slackness interacted with a post-treatment dummy. Standard Errors are clustered at the district level. Bands report the 95% confidence intervals.

Bibliography

Abramitzky, R., P. Ager, L. P. Boustan, E. Cohen and C. W. Hansen (2019). "The Effects of Immigration on the Economy: Lessons from the 1920s Border Closure." *NBER Working Paper*, (No. w26536).

Abramitzky, R. and L. Boustan (2017). "Immigration in American Economic History." *Journal of Economic Literature*, 55(4): 1311–45.

Abramitzky, R., L. P. Boustan and K. Eriksson (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy*, 122(3): 467–506.

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

——— (2007). "Equilibrium Bias of Technology." *Econometrica*, 75(5): 1371–1409.

Acemoglu, D., G. De Feo, G. De Luca and G. Russo (2022). "War, Socialism, and the Rise of Fascism: an Empirical Exploration." *The Quarterly Journal of Economics*, 137(2): 1233–1296.

A'Hearn, B. (1998). "Institutions, Externalities, and Economic Growth in Southern Italy: Evidence from the Cotton Textile Industry, 1861-1914." *Economic History Review*, pp. 734–762.

Akram, A. A., S. Chowdhury and A. M. Mobarak (2017). "Effects of Emigration on Rural Labor Markets." *NBER Working Paper*, (No. w23929).

Albert, C., A. Glitz and J. Llull (2021). "Labor Market Competition and the Assimilation of Immigrants." *CEPR Discussion Paper*, (No. DP16432).

Andersson, D., M. Karadja and E. Prawitz (2022). "Mass Migration and Technological Change." *Journal of the European Economic Association*, 20(5): 1859–1896.

Aydemir, A. and G. J. Borjas (2007). "Cross-Country Variation in the Impact of International Migration: Canada, Mexico, and the United States." *Journal of the European Economic Association*, 5(4): 663–708.

Bandiera, O., I. Rasul and M. Viarengo (2013). "The Making of Modern America: Migratory Flows in the Age of Mass M igration." *Journal of Development Economics*, 102: 23–47.

Becker, S. O., I. Grosfeld, P. Grosjean, N. Voigtländer and E. Zhuravskaya (2020). "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers." *American Economic Review*, 110(5): 1430–63.

Beerli, A., J. Ruffner, M. Siegenthaler and G. Peri (2021). "The Abolition of Immigration Restrictions and the Performance of Firms and Workers: Evidence from Switzerland." *American Economic Review*, 111(3): 976–1012.

Beine, M., F. Docquier and H. Rapoport (2008). "Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers." *The Economic Journal*, 118(528): 631–652.

Borjas, G. J. (1995). "The Economic Benefits from Immigration." *Journal of economic perspectives*, 9(2): 3–22.

——— (2014). *Immigration Economics*. Harvard University Press.

Boustan, L., D. Bunten and O. Hearey (2018). "Urbanization in American Economic History, 1800–2000." In L. P. Cain, P. V. Fishback and P. W. Rhode, eds., "The Oxford Handbook of American Economic History," volume 2, p. 75. Oxford University Press.

Brum, M. (2019). "Italian Migration to the United States: The Role of Pioneers' Locations." *Working Paper*.

Bryan, G., S. Chowdhury and A. M. Mobarak (2014). "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica*, 82(5): 1671–1748.

Bryan, G. and M. Morten (2019). "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia." *Journal of Political Economy*, 127(5): 2229–2268.

Burchardi, K. B., T. Chaney, T. A. Hassan, L. Tarquinio and S. J. Terry (2020). "Immigration, Innovation, and Growth." *NBER Working Paper*, (No. w27075).

Calabrese, V. (2017). *Land of Women: Basilicata, Emigration, and the Women Who Remained Behind, 1880-1914*. CUNY, Unpublished Dissertation.

Card, D. (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64.

Choate, M. I. (2008). *Emigrant Nation: The Making of Italy Abroad*. Harvard University Press.

Clemens, M. A. (2011). "Economics and Emigration: Trillion-dollar Bills on the Sidewalk?" *Journal of Economic perspectives*, 25(3): 83–106.

Clemens, M. A., E. G. Lewis and H. M. Postel (2018). "Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion." *American Economic Review*, 108(6): 1468–87.

Cohen, J. and G. Federico (2001). *The Growth of the Italian Economy, 1820-1960*. Cambridge University Press.

Daniels, R. (1959). *Coming to America: A History of Immigration and Ethnicity in American Life*. Harper Collins Publishers.

David, P. A. (1990). "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox." *American Economic Review*, 80(2): 355–361.

De Haas, H., K. Natter and S. Vezzoli (2015). "Conceptualizing and Measuring Migration Policy Change." *Comparative Migration Studies*, 3(1): 1–21.

dei Deputati, C. (1907). *Atti Parlamentari*, volume Legislatura XII, Prima Sessione, Discussioni, Seconda Tornata.

DINKELMAN, T. and M. MARIOTTI (2016). "The Long-Run Effects of Labor Migration on Human Capital Formation in Communities of Origin." *American Economic Journal: Applied Economics*, 8(4): 1–35.

DOCQUIER, F. and H. RAPOPORT (2012). "Globalization, Brain Drain, and Development." *Journal of Economic Literature*, 50(3): 681–730.

DONALDSON, D. (2018). "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*, 108(4-5): 899–934.

DONALDSON, D. and R. HORNBECK (2016). "Railroads and American Economic Growth: A "Market Access" Approach." *The Quarterly Journal of Economics*, 131(2): 799–858.

DUNLEVY, J. A. and W. K. HUTCHINSON (1999). "The Impact of Immigration on American Import Trade in the Late Nineteenth and Early Twentieth Centuries." *The Journal of Economic History*, 59(4): 1043–1062.

DUSTMANN, C., I. FADLON and Y. WEISS (2011). "Return Migration, Human Capital Accumulation and the Brain Drain." *Journal of Development Economics*, 95(1): 58–67.

DUSTMANN, C., T. FRATTINI and A. ROSSO (2015). "The Effect of Emigration from Poland on Polish wages." *The Scandinavian Journal of Economics*, 117(2): 522–564.

DUSTMANN, C. and J.-S. GÖRLACH (2016). "The Economics of Temporary Migrations." *Journal of Economic Literature*, 54(1): 98–136.

EATON, J. and S. KORTUM (1999). "International Technology Diffusion: Theory and Measurement." *International Economic Review*, 40(3): 537–570.

FEDERICO, G., A. NUVOLARI, L. RIDOLFI and M. VASTA (2021). "The Race Between the Snail and the Tortoise: Skill Premium and Early Industrialization in Italy (1861–1913)." *Cliometrica*, 15(1): 1–42.

FERNANDEZ-SANCHEZ, M. (2020). "Mass Emigration and Human Capital Over a Century: Evidence from the Galician Diaspora." *Working Paper*.

FERRIE, J. P. and T. J. HATTON (2015). "Two centuries of international migration." In B. R. Chiswick and P. W. Miller, eds., "Handbook of the economics of international migration," volume 1, pp. 53–88. Elsevier.

FOERSTER, R. F. (1919). *The Italian Emigration of Our Times*, volume 20. Harvard University Press and Oxford University Press.

FONTANA, N., M. MANACORDA, G. RUSSO and M. TABELLINI (2021). "Emigration and Long-Run Economic Development: The Effects of the Italian Mass Migration."

FOSTER, A. D. and M. R. ROSENZWEIG (1995). "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of political Economy*, 103(6): 1176–1209.

GABACCIA, D. (2013). *Italy's Many Diasporas: Elites, Exiles and the Workers of the World.* Routledge.

GAGGL, P., R. GRAY, I. MARINESCU and M. MORIN (2021). "Does Electricity Drive Structural Transformation? Evidence from the United States." *Labour Economics*, 68: 101944.

GALLO, S. (2012). *Senza attraversare le frontiere. Le migrazioni interne dall'Unità a oggi, publisher=Laterza.*

GERSCHENKRON, A. (1962). "Economic Backwardness in Historical Perspective." *Cambridge University Press*.

GIBSON, J., D. MCKENZIE and S. STILLMAN (2011). "The Impacts of International Migration on Remaining Household Members: Omnibus Results from a Migration Lottery Program." *Review of Economics and Statistics*, 93(4): 1297–1318.

GOLDIN, C. (1994). "The Political Economy of Immigration Restriction in the United States, 1890 to 1921." In "The Regulated Economy: A Historical Approach to Political Economy," pp. 223–258. University of Chicago Press.

GOLDSMITH-PINKHAM, P., I. SORKIN and H. SWIFT (2020). "Bartik Instruments: What, When, Why, and How." *American Economic Review*, 110(8): 2586–2624.

GOMELLINI, M. and C. O'GRADA (2011). "Outward and Inward Migrations in Italy: A Historical Perspective." *Bank of Italy Economic History Working Paper*, (8).

GOULD, J. D. (1980a). "European Inter-Continental Emigration. The Road Home: Return Migration from the USA." *Journal of European Economic History*, 9(1): 41.

——— (1980b). "European International Emigration: The Role of Diffusion and Feedback." *Journal of European Economic History*, 9(2): 267.

GROGGER, J. and G. H. HANSON (2011). "Income Maximization and the Selection and Sorting of International Migrants." *Journal of Development Economics*, 95(1): 42–57.

GURIEV, S. and E. PAPAIOANNOU (2022). "The Political Economy of Populism." *Journal of Economic Literature*, 60(3): 753–832.

HABAKKUK, H. J. (1962). *American and British Technology in the Nineteenth Century: the Search for Labour Saving Inventions*. Cambridge (MA): Cambridge University Press.

HANLON, W. W. (2015). "Necessity is the Mother of Invention: Input Supplies and Directed Technical Change." *Econometrica*, 83(1): 67–100.

HATTON, T. J. and J. G. WILLIAMSON (1998). *The Age of Mass Migration: Causes and Economic Impact*. Oxford University Press.

HATTON, T. J., J. G. WILLIAMSON *et al.* (2005). *Global Migration and the World Economy: Two Centuries of Policy and Performance*. MIT University Press.

HICKS, J. (1932). *The Theory of Wages*. Macmillan.

Higham, J. (1955). *Strangers in the land: Patterns of American nativism, 1860-1925*. Rutgers University Press.

Hornbeck, R. and S. Naidu (2014). "When the Levee Breaks: Black Migration and Economic Development in the American South." *American Economic Review*, 104(3): 963–90.

Jaeger, D. A., J. Ruist and J. Stuhler (2018). "Shift-Share Instruments and the Impact of Immigration." *NBER Working Paper*, (No. w24285).

Juhász, R., M. P. Squicciarini and N. Voigtländer (2020). "Technology adoption and productivity growth: Evidence from industrialization in france." *NBER Working Paper*, (No. w27503).

Juhász, R. and C. Steinwender (2018). "Spinning the Web: The Impact of ICT on Trade in Intermediates and Technology Diffusion." *NBER Working Paper*, (No. w24590).

Karadja, M. and E. Prawitz (2019). "Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States." *Journal of Political Economy*, 127(4): 1864–1925.

Keeling, D. (1999). "The Transportation Revolution and Transatlantic Migration, 1850-1914." *Research in Economic History*, 19: 39–74.

Klein, H. S. (1983). "The Integration of Italian Immigrants into the United States and Argentina: A Comparative Analysis." *The American Historical Review*, 88(2): 306–329.

Koven, S. G. and F. Götzke (2010). *American Immigration Policy: Confronting the Nation's Challenges*, volume 1. Springer.

Kwok, V. and H. Leland (1982). "An Economic Model of the Brain Drain." *The American Economic Review*, 72(1): 91–100.

Lewis, E. (2011). "Immigration, Skill Mix, and Capital Skill Complementarity." *The Quarterly Journal of Economics*, 126(2): 1029–1069.

Mishra, P. (2007). "Emigration and Wages in Source Countries: Evidence from Mexico." *Journal of development economics*, 82(1): 180–199.

Mokyr, J. (1998). "The Second Industrial Revolution, 1870-1914." In V. Castronovo, ed., "Storia dell'Economia Mondiale," pp. 219–245. Rome (Italy): Laterza.

Newman, M. (2018). *Networks*. Oxford university press.

Nuvolari, A. and M. Vasta (2015). "The Ghost in the Attic? The Italian National Innovation System in Historical Perspective, 1861–2011." *Enterprise & society*, 16(2): 270–290.

Parente, S. L. and E. C. Prescott (1994). "Barriers to Technology Adoption and Development." *Journal of political Economy*, 102(2): 298–321.

Pérez, S. (2021). "Southern (American) Hospitality: Italians in Argentina and the United States During the Age of Mass Migration." *The Economic Journal*, 131(638): 2613–2628.

Rapoport, H. and F. Docquier (2006). "The economics of migrants' remittances." volume 2, pp. 1135–1198. Elsevier.

Rosoli, G. (1998). "La political migratoria dell'italia dall'unità al fascismo." *Annali della Fondazione Luigi Einaudi*, 32.

San, S. (2022). "Labor Supply and Directed Technical Change: Evidence from the Termination of the Bracero Program in 1964." *American Economic Journal: Applied Economics*.

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

Smith, D. M. (1997). *Modern Italy: A political history*. University of Michigan Press.

Sori, E. (1979). *L'emigrazione italiana dall'Unità d'Italia alla Seconda Guerra Mondiale*. Il Mulino.

Spitzer, Y., G. Tortorici and A. Zimran (2020). "International Migration Responses to Natural Disasters: Evidence from Modern Europe's Deadliest Earthquake." *CEPR Discussion Paper*, (No. DP15008).

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.

——— (2020). "Like an Ink Blot on Paper: Testing the Diffusion Hypothesis of Mass Migration, Italy 1876–1920." *Working Paper*.

Suri, T. (2011). "Selection and Comparative Advantage in Technology Adoption." *Econometrica*, 79(1): 159–209.

Tabellini, M. (2020). "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration." *The Review of Economic Studies*, 87(1): 454–486.

Taylor, A. M. and J. G. Williamson (1997). "Convergence in the Age of Mass Migration." *European review of economic history*, 1(1): 27–63.

Toniolo, G. (2014). *An Economic History of Liberal Italy (Routledge Revivals): 1850-1918*. Routledge.

Unrau, H. D. (1984). "Ellis Island, Statue of Liberty National Monument, New York-New Jersey, journal=US Department of Interior, National Park Service, volume=1,."

Vasta, M. (1999). *Innovazione tecnologica e capitale umano in Italia (1880-1914): le traiettorie della Seconda Rivoluzione Industriale*. Il Mulino.

Yang, D. (2011). "Migrant Remittances." *Journal of Economic perspectives*, 25(3): 129–52.

Zamagni, V. (1990). *The Economic History of Italy 1860-1990: Recovery after Decline*. Oxford University Press.

Zeira, J. (1998). "Workers, Machines, and Economic Growth." *The Quarterly Journal of Economics*, 113(4): 1091–1117.

Appendix

2.A Data Appendix

2.A.1 Emigration Data

In this section we document in detail the novel emigration data that we collect. The raw data can be found at `https://heritage.statueofliberty.org/`. First, we describe how we deal with spelling mistakes occurring the recorded municipality of origin of immigrants. Second, we report the share of emigrants with missing origin municipality. Last, we provide evidence suggesting that our data squares well with less granular data from official statistics records.

2.A.1.1 Emigration Matching Procedure

This section describes the procedure we follow to match municipalities recorded by Ellis Island US officials to actual Italian *comuni*. Since municipalities changed over time, we first assembled a list of all municipalities that existed between 1890 and 1930 from listed census names. Then along the lines of Abramitzky *et al.* (2014), we run the following matching procedure:

1. Perform manual name cleaning, *e.g.* correcting systematic mistakes and recording shortcuts.
2. Standardize each recorded and actual municipality name using the NYSIIS algorithm trained on Italian phonetics (Atack *et al.*, 1992). This procedure ensures that phonetically identical municipality names have an exact match.
3. For each standardized recorded name which does not have a perfect match in the list of all municipality names, compute the dissimilarity matrix with all those names, according to some metric. Then, pick as a match the *comune* with the lowest dissimilarity.
4. If the distance between a recorded municipality and its best match is lower than some threshold value $\alpha \in [0, 1]$, accept the match. Otherwise, drop the observation.

We evaluate the distance between a recorded municipality name i an actual name j in terms of their Jaro-Winkler similarity d_{ij}:

$$d_{ij} \equiv \widehat{d_{ij}} + \ell p(1 - \widehat{d_{ij}}) \tag{2.8}$$

where

$$\widehat{d_{ij}} \equiv \begin{cases} 0 & \text{if } m = 0 \\ \frac{1}{3}\left(\frac{m}{|i|} + \frac{m}{|j|} + \frac{m-t}{m}\right) & \text{else} \end{cases} \tag{2.9}$$

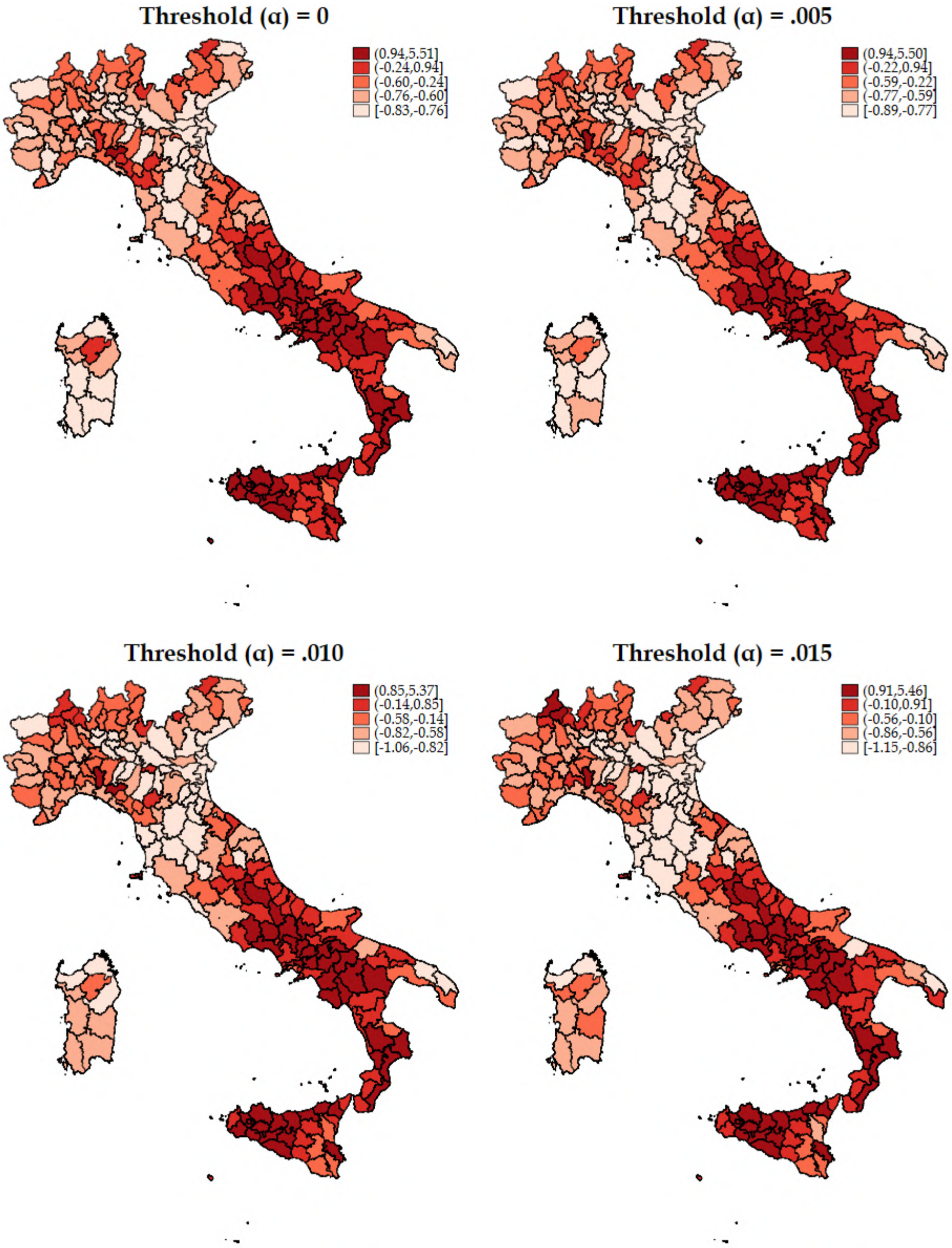
where m is the number of matching characters, $|i|$ is the length of string i, and t is half the number of transpositions, ℓ is the length of common an eventual common prefix no longer than four characters between i and j, and $p = 0.1$ is a constant scaling factor. Two characters are matching only if 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.[51]

The Jaro-Winker distance has been shown to perform relatively well in linking routines (Abramitzky *et al.*, 2021). In our particular case however, this metric outperforms more standard string dissimilarity metrics like the cosine or the Levenshtein because the Jaro-Winkler assigns a “bonus” score to strings starting with closer initial substrings. We noted that coding errors in municipality names are more frequent at the end of names, hence the comparative advantage of the Jaro-Winkler distance.

The matching procedure assigns to each recorded municipality name its best match among the actual names along with their distance d^*_{ij}. We set a threshold $\alpha \in [0, 1]$, pick all matches j with $d^*_{ij} \leq \alpha$, and drop the others.

[51]The Jaro-Winkler distance has been recently employed in the economic history literature for intergenerational linking purposes by, among others, Feigenbaum (2018) and Abramitzky *et al.* (2021).

FIGURE 2.A.1: District-Level Migration Flows varying α



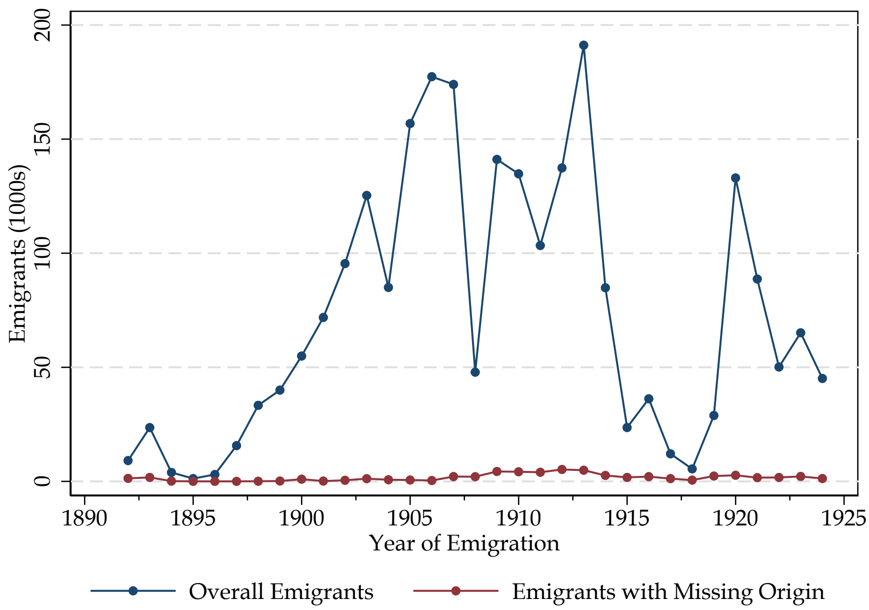
Notes. Each panel plots the number of emigrants across districts over the years 1890-1930. See Appendix 2.A.1.1 for a complete description of the procedure and the meaning of α.

2.A.1.2 Missing Origin

The Ellis Island records report the origin municipality of Italian immigrants starting in 1892 when the immigration station opened. In picture 2.A.2 we report the total number of recorded Italian immigrants at Ellis Island, along with those whose origin municipality is missing. We consider an origin entry as missing if it is either a proper missing or if the record reports coarse geographical aggregates. These include, among others, "Italy", Italian regions, and similar information which make it unfeasible to back out the district of origin of the immigrant. Since our analysis is run at the district level and the non-missing Ellis Island records report origin at the municipality level, to conduct our analysis we aggregate our individual-level data at the district level.

Figure 2.A.2 suggests that missing origins are a minor concern in our dataset. There is no single year when the share of immigrants with missing origin exceeds 1% of the overall immigrants. Throughout our analysis, we therefore drop immigrants with missing origins from our dataset.

FIGURE 2.A.2: Ellis Island Immigration Records: Assessment of Missing Origin



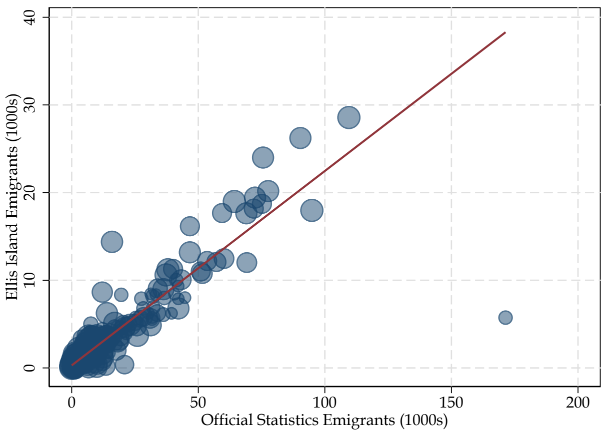
Notes. The blue series reports the total number of Italian immigrants in our sample, over the period 1892-1924. The red series reports the total number of Italian immigrants whose origin is missing in the Ellis Island dataset. We label as "missing" every entry whose origin is either missing, or reports coarse geographical aggregates, such as Italy, and Italian regions or provinces.

2.A.1.3 Validation of the Ellis Island Data

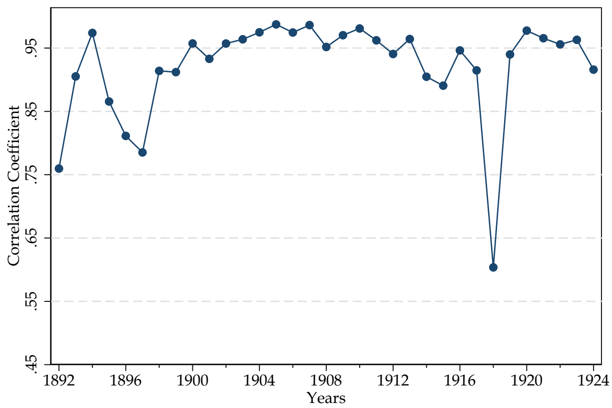
To validate our dataset, we compare it with official statistics data that we digitized from the *Annuario statistico dell'emigrazione italiana dal 1876 al 1925: con notizie sull'emigrazione negli anni 1869-1875*. The data was collected by the Commissioner-General for Emigration, and published by the Italian Statistical Office (ISTAT) in 1926. Data report yearly emigration outflows, broken down by major destination countries, and by region of origin of emigrants. There were 19 regions in Italy before WW1, and 20 thereafter. This implies that official statistics data cannot be used in our analysis, since regions are too few and large. Instead, we use these data as a meaningful validation tool for our dataset.

FIGURE 2.A.3: Validation of the Ellis Island Emigration Dataset

(A) Cross-Sectional Correlation



(B) Time-Series Correlation



Notes. The left panel displays the cross-sectional correlation between region-level US emigration outflows as recorded in the official data—on the x-axis—and in our dataset—on the y-axis. A dot is a region-year, and the size of each dot is proportional to that region's population in the given decade. The right panel reports the R^2 of a regression where the dependent variable is US emigration as recorded in official statistics, and the explanatory variable is US emigration in our dataset. Each regression only considers observations for one given year. The note also reports the overall R^2 of the associated regression for the entire dataset, controlling for year and region-fixed effects.

In the left-hand panel of figure 2.A.3 we report the cross-sectional correlation between US emigration outflows in our dataset—on the y-axis—and in official statistics—on the x-axis. Each observation is a region year, and the size of each dot is proportional to the population of the region in that year's decade, as registered in the population census. The red line reports the fitted values of the associated linear regression. The figure depicts a strong and positive association between the US emigration series in the two data sets. A similar picture would obtain if we bin-scattered observations. A possible caveat is that our dataset consists of fewer emigrants than reported in the official statistics because we searched 30,000 surnames in the Ellis Island Foundation dataset. Although comprehensive, this is not the universe of Italian surnames. This notwithstanding, figure 2.A.3 attests that our dataset is geographically comparable to the official data.

Figure 2.A.3 (B) reports the time-series correlation between the US emigration outflows series in the two data-sets. More specifically, each dot reports the R^2 of the following regression:

$$\text{Ellis Island US Emigration}_{r,t=T} = \alpha + \beta \text{OS US Emigration}_{r,t=T} + \varepsilon_{r,t=T} \qquad (2.10)$$

where Ellis Island US Emigration is the US emigration series measured with data from the Ellis Island database; OS US Emigration is the US emigration series measured with official statistics data; r denotes a region, and $T \in [1892, 1924]$ is a given year. In other words, a dot in a given year T reports the cross-sectional R^2 of a regression including all observations—*i.e.*, one per region—in that year. In the footnote, we also report the R^2 associated with regression (2.10) where we pool observations across years and include year and region-fixed effects. The R^2 is a measure of linear fit between the two series. Hence, we would ideally observe $R^2 = 1$ under perfect collinearity. Results indicate that the correlation between the official series and ours is extremely high over time. Except for the WW1 years, the R^2 is above 75% throughout the sample period and, starting in 1896, always exceeds 90%. In two robustness checks, we confirm that our results remain unchanged even if we drop the periods with relatively low correlation, *i.e.* 1892-1896 and WW1.

The comparison between official statistics data and our series confirms that our measure is a valid proxy for actual US emigration. The main advantage of our dataset is its granularity, which we exploit in our analysis.

2.A.2 Data Sources

We here describe the sources from which we gathered the data needed for our analysis. Analyses are mainly conducted at the district level—aggregation areas comparable to US counties—which were named "Circondario" and are composed of municipalities (whose number ranges from 7900 to 9000 in our sample period). We collected and digitize district- or municipality-level data from multiple historical sources provided by the Italian Institute of Statistics. The main sources are the Population Censuses and Industrial Censuses. As explained in the previous Section, migration flows by municipality were taken from the Ellis Island database.

We here provide a detailed summary of the sources of our variables of interest for each year of our sample, specifying the geographical level at which data were collected. The historical volumes we digitized can be found at this link. Censuses were held on a 10-year basis.

Population Censuses were comprehensive of all information on population, including occupation and alphabetization for the whole period 1901-1921. In 1931 the Census was smaller and did not include information on occupations. The next comprehensive Population Census was held in 1936. In order to fill the gap between the years 1921-1936, we had to take the information on occupation from the 1927 Industrial Census. This resulted in our sample of years for the population's occupations to be: 1901, 1911, 1921, 1927, 1936. As far as it concerns data on the number of firms, engines, and horsepower, they are available in the Industrial Censuses: information was available for the years 1911, 1927, and 1937.

Data on migration flows are gathered at the municipality level from the Ellis Island database, starting from the year 1881. Population at the municipality level was instead collected for all Population Censuses starting from 1861. For the years 1901, 1911, and 1921 data on population by occupation were available at the district level (about 200 units) on the Population Census. For the year 1927, it was instead available in the Industrial Census. In that same year, districts, or "Circondari", were suppressed as administrative units. This means that data on occupations for 1936 had to be collected at the municipality level, for a total of about 8000 municipalities.

Industrial data are from Industrial censuses. The Industrial census was conducted for the first time in 1911, and then again in 1927 and 1937. We digitized these censuses and collected all relevant variables at the province level, *i.e.* the most granular available level of aggregation. Since our analysis is conducted at the district level, we impute these from provinces to districts. In the next section, we explain the details of the imputation procedure.

2.A.2.1 Imputation of Industrial Census Data

The variables we use to proxy capital investment—namely, the number of firms, number of firms with engines, number of mechanical and/or electrical engines, and mechanical and/or electrical horsepower—are digitized from industrial censuses. The most granular level of aggregation available there are provinces. Provinces were composed of several districts, ranging from one to four. In our analysis, we impute these province-level data to districts. In this section, we describe the details of this imputation procedure.

Let subscript p denote a province-level variable, whereas the same variable with subscript d is at the district level. For every variable y_p we need to impute, we run the following simple OLS regression:

$$y_{p,t} = \alpha_p + \alpha_t + \boldsymbol{x}'_{p,t}\boldsymbol{\beta} + \varepsilon_{p,t} \tag{2.11}$$

where $t \in \{1911, 1927, 1937\}$, and α_t and α_p respectively denote year and province fixed effects. Term $\boldsymbol{x}_{p,t}$ includes a set of province-level regressors. These are total employment as well as the number of employed in agriculture, manufacturing, liberal professions, and public administration. Both y and the variables in $\boldsymbol{x}$ are in logs.

We estimate equation 2.11 and retrieve a set of coefficients $\hat{\boldsymbol{\beta}}$. To perform the imputation, we exploit variation of the $\boldsymbol{x}$'s at the district level:

$$y_{d,t} = \boldsymbol{x}'_{d,t}\hat{\boldsymbol{\beta}} \tag{2.12}$$

Notice that, because all regressions include district and year fixed effects, these capture variation which in regressions (2.11) is absorbed by year and province fixed effects.

In table 2.A.1 we compare province-level data from the industrial censuses and imputed variables computed through (2.12), aggregated at province level. The table suggests that there is a strong positive correlation between actual and imputed variables. This is confirmed by a formal test of the statistical significance of the correlation coefficients. These are statistically different from zero—and positive—for all imputed variables, thus suggesting that capital variables computed exploiting district-level variation in the $\boldsymbol{x}$'s correlate with actual province-level variables. We interpret this as evidence supporting our imputation procedure.

TABLE 2.A.1: Comparison Between Actual and Imputed Capital Variables

	(1) ρ	(2) p-value	(3) β	(4) se(β)	R^2
Firms	0.439^{***}	(0.000)	0.160^{***}	(0.023)	0.193
Firms with Engine	0.470^{***}	(0.000)	0.222^{***}	(0.033)	0.221
Mechanical Engines	0.410^{***}	(0.000)	0.105^{***}	(0.022)	0.168
Electrical Engines	0.492^{***}	(0.000)	0.247^{***}	(0.036)	0.242
Mechanical Horsepower	0.469^{***}	(0.000)	0.197^{***}	(0.036)	0.220
Electrical Horsepower	0.468^{***}	(0.000)	0.217^{***}	(0.035)	0.219

Notes. This table compares measured and imputed capital variables. The imputation procedure is fully pinned down by equations (2.11)-(2.12). Each row compares the imputed and the measured row variable. The imputed row variable is predicted at the district level and then aggregated up to provinces. Column ρ reports Pearson's correlation coefficient between imputed and measured variables, along with its Bonferroni-adjusted p-value. Columns β and se(β) respectively display the coefficient and the standard error, clustered at the province level, of a regression where the dependent variable is imputed and the independent variable is measured. Column R^2 reports the coefficient of linear determination of this regression.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

2.A.2.2 Railway data

Data on a district's historical connectivity to the railway network were constructed using information taken from the *Sviluppo delle ferrovie italiane dal 1839 al 31 dicembre 1926* edited by the Italian Statistical Office (*Ufficio Centrale di Statistica*) in 1927. To the best of our knowledge, this is the first paper to use these data. The Italian Statistical Office recorded the year of construction of each railway line connecting two municipalities, providing information on each intermediate station. Hence, we are able to construct the railway network for each year from 1839 to 1926.

As our analysis is carried out at the district level, we obtain a measure of railway access for each district c by aggregating municipality-level data. We build a time-varying dummy—$\text{RA}_{cr,t}$—taking value one if at least one municipality in a given district was connected through the railway to another municipality in a different district, and zero otherwise. We also construct a measure of the capillarity of the presence of the railway in a given district using the number of train stations in that district for each year.

We build the network of districts connected through the railway in order to obtain the distance between each district c and any of the three departure ports: the districts of Genoa, Naples, and Palermo. Each district constitutes a node of the network. An edge is created between two nodes if at least one municipality of the first district is connected to one municipality of the second district. *De facto*, edges connect adjacent districts, as for each year there is no railway line directly connecting two municipalities in nonadjacent districts without stopping in a train station belonging to the intermediate district.

The distance between two adjacent districts is calculated as the geodesic distance between the centroids. The distance $d_t(c,i)$ between any two districts c and i in the network is hence the shortest path, or geodesic path, between the two nodes. We adopt this measure because we interpret the railway network as a weighted graph where edges are weighted by the distance between two nodes. In this context, the shortest path is the minimum sum of edge weights.

TABLE 2.A.2: Visual Summary of Data Sources

(1) **Variable**	(2) **Measurement**	(3) **Observation Unit**	(4) **Source**	(5) **Observed Years**
Demographics				
Population	Measured	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Area	Measured	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Urbanization	Measured	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Literacy	Measured	Municipality	Population Censuses	1881-1936, excl.1891
Employment, by Sector				
Manufacture	Measured	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Agriculture	Measured	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Trade	Measured	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Liberal Professions	Measured	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Public Administration	Measured	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Capital & Industry				
Firms	Imputed	Province, imputed to Districts	Manufacture Censuses	1911, 1927, 1937
Firms with Engine	Imputed	Province, imputed to Districts	Manufacture Censuses	1911, 1927, 1937
Mechanical Engines	Imputed	Province, imputed to Districts	Manufacture Censuses	1911, 1927, 1937
Electrical Engines	Imputed	Province, imputed to Districts	Manufacture Censuses	1911, 1927, 1937
Mechanical Horsepower	Imputed	Province, imputed to Districts	Manufacture Censuses	1911, 1927, 1937
Electrical Horsepower	Imputed	Province, imputed to Districts	Manufacture Censuses	1911, 1927, 1937
Emigration				
US Emigration	Measured	Municipality	Ellis Island Data	1892-1924
Overall Emigration	Imputed	Province, imputed to Districts	Official Statistics of the Commissioner General	1877-1925
Other				
WW1 deaths	Measured	Municipality	Istituto per la storia della Resistenza e della società contemporanea.	1915-1918
Railways	Measured	Municipality	ISTAT – *Sviluppo delle ferrovie italiane dal 1839 al 31 dicembre 1926*	1839-1926

Notes. This table reports all variables used in the paper. The “Measurement” column reports “Measured” if the variable is used in the analysis as it is measured in the source data; instead, it reports “Imputed” if it is measured at a coarser level of aggregation, and is then imputed to districts. The imputation procedure is described in the Data Appendix. The “Observation Unit” returns the level of aggregation at which the variable is measured. The “Source” column displays the type of source the raw data are extracted from. Further references to original sources can be found in the text main body. The “Observed Years” column reports the years when the raw data is available. Literacy data are from Fontana *et al.* (2021).

2.B Additional Tables & Results

TABLE 2.B.1: Regional Emigration

Region	Emigrants to US					Emigrants to all destinations					Share
	76-87	88-99	00-12	13-25	Total	76-87	88-99	00-12	13-25	Total	
Piemonte	5.2	12.3	109.8	43.4	170.8	353.3	332.5	697.2	527.9	1910.8	8.9
Liguria	8.2	10.8	27.2	10.6	56.8	63.0	51.1	89.0	92.9	296.1	19.2
Lombardia	4.4	11.0	56.7	28.6	100.8	237.9	259.7	675.8	441.6	1615.2	6.2
Veneto	1.0	6.0	52.7	48.4	108.1	486.3	1197.6	1298.2	651.0	3633.1	3.0
Emilia-Romagna	1.3	8.4	62.0	24.0	95.8	60.5	137.7	422.4	178.7	799.2	12.0
Toscana	3.3	12.9	89.6	42.0	147.8	110.7	157.5	412.4	230.6	911.2	16.2
Marche	0.2	2.0	62.0	30.6	94.8	12.7	48.0	280.6	131.1	472.3	20.1
Umbria	0.1	0.5	24.1	11.8	36.6	0.5	6.0	129.9	59.4	195.7	18.7
Lazio	0.02	2.3	109.4	50.1	161.9	0.4	14.0	151.4	72.9	238.6	67.8
Abruzzi e Molise	26.9	68.0	371.0	161.6	627.4	58.3	164.1	585.7	241.6	1049.7	59.8
Campania	44.3	157.5	637.8	241.5	1081.2	131.3	339.6	871.0	360.7	1702485	63.5
Puglie	1.3	12.9	164.7	107.9	286.9	8.1	37.2	283.4	172.4	501.2	57.2
Basilicata	28.4	53.3	108.1	38.5	228.3	74.1	106.5	179.8	70.5	431.0	53.0
Calabrie	15.0	58.5	457.7	125.1	656.3	74.1	178.5	539.8	253.6	1046.1	62.7
Sicilia	12.6	117.2	687.7	356.1	1173.6	26.8	170.9	946.5	516.4	1660.6	70.7
Sardegna	0.01	0.03	8.5	5.7	14.2	1.3	6.2	72.8	43.9	124.1	11.5
Total	152.1	533.9	3029.1	1326.0	5041.3	1699.3	3206.9	7635.8	4045.4	16587.4	30.4

Notes. Regional emigration towards US and total emigration during the period 1876-1925. Figures are in thousands. Column “Share” indicates the percentage of total emigrants towards US relatively to all emigrants from that region in the whole period 1876-1925.
Source: our elaboration on data from the *Annuario statistico dell’emigrazione italiana dal 1876 al 1925: con notizie sull’emigrazione negli anni 1869-1875*, Italian Statistical Office (ISTAT), Roma, 1926.

Table 2.B.2: Internal and International Migrations, 1921-1931

Region	Absolute numbers			Share over Population	
	Population	Internal Migrants	Emigrants	Internal Migrants	Emigrants
Abruzzo	1317.2	19.3	170.3	1.5	12.9
Basilicata	524.5	5.6	52.4	1.1	10.0
Calabria	1257.9	8.2	219.4	0.7	17.4
Campania	2896.6	1.2	248.4	0.0	8.6
Emilia Romagna	2183.4	78.7	165.3	3.6	7.6
Lazio	903.5	-133.8	88.2	-14.8	9.8
Liguria	892.4	-60.5	112.7	-6.8	12.6
Lombardia	3680.6	-198.0	460.6	-5.4	12.5
Marche	939.3	25.2	99.2	2.7	10.6
Piemonte	3070.3	-111.9	469.3	-3.6	15.3
Puglia	1589.1	52.9	117.8	3.3	7.4
Sardegna	682.0	2.8	27.7	0.4	4.1
Sicilia	2927.9	31.7	333.4	1.1	11.4
Toscana	2208.9	27.2	198.0	1.2	9.0
Umbria	572.1	-1.0	37.1	-0.2	6.5
Veneto	2814.2	139.8	639.8	5.0	22.7

Notes. This table reports internal migration and out-migration flows over the period 1921-1931. Column "Population" reports population in 1881. Column "Internal migrants" is the net internal migrant flow. To compute net internal migration flows, we take the difference in the outflow of people leaving a given region and the inflow of people arriving in that region during the decade 1921-1931. Since Census data only report the stock of people born in a given region living in another region in 1921 and 1931, to compute the outflow of people leaving a region during that decade, we take the difference across years of the total number of people born in that region and living in any other Italian region. Similarly, to compute the inflow of people arriving in a region during that decade we take the difference across years of the total number living in that region who were born in any other Italian region. Positive (negative) figures imply a net population loss (gain) due to internal migrations. Column "Emigrants" reports the number of international emigrants. Figures are in thousands. Columns "Share over Population" report net internal and international migration figures, relative to 1881-population. Figures are in percentage terms.
Source: our elaboration on data from the *Annuario statistico dell'emigrazione italiana dal 1876 al 1925: con notizie sull'emigrazione negli anni 1869-1875*, Italian Statistical Office (ISTAT), Roma, 1926, and from *Censimento della Popolazione Italiana*, Italian Statistical Office (ISTAT), Roma, 1921 and 1931.

Table 2.B.3: Population Growth

	Dep. Var.: Population Growth							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Quota Exposure × Post	0.408***	0.446***	0.422***	0.443***	0.515***	0.469***	0.342**	0.284**
	(0.113)	(0.124)	(0.120)	(0.120)	(0.134)	(0.132)	(0.134)	(0.136)
Population	0.146***	0.142***	0.165***	0.166***	0.180***	0.183***	0.178***	0.179***
	(0.030)	(0.030)	(0.031)	(0.030)	(0.032)	(0.033)	(0.034)	(0.033)
Extensive Margin × Post		-0.065	-0.091	-0.109*	-0.101*	-0.094*	-0.058	-0.058
		(0.055)	(0.057)	(0.059)	(0.055)	(0.053)	(0.051)	(0.052)
Agriculture × Post			0.095***	0.072***	0.090***	0.078**	0.089***	0.088***
			(0.024)	(0.026)	(0.031)	(0.031)	(0.030)	(0.030)
Urbanization × Post				-0.026**	-0.020	-0.017	-0.017	-0.019
				(0.013)	(0.014)	(0.014)	(0.014)	(0.014)
Literacy × Post					0.024	0.019	0.059***	0.059***
					(0.017)	(0.016)	(0.019)	(0.019)
WW1 × Post						-0.030*	-0.021	-0.020
						(0.017)	(0.015)	(0.015)
South × Post							0.029***	0.029***
							(0.008)	(0.008)
US GDP Growth × QE								0.018**
								(0.008)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	204	204	204	204	204	204	204	204
Observations	751	751	751	751	751	751	751	751
R2	0.453	0.454	0.475	0.479	0.480	0.482	0.495	0.501
F-stat	13.726	10.139	10.400	12.096	14.920	14.928	16.897	15.768
Mean Dep. Var.	1.042	1.042	1.042	1.042	1.042	1.042	1.042	1.042

Notes. This table displays the effect of exposure to the Quota Acts on population growth. Population growth is defined as the decade-on-decade percentage change in population. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are always clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Literacy is the number of people who could read and write as a share of the overall population in 1901. South is a dummy equal to zero if the district is in the EU NUTS 2 ITC or ITH region, and one otherwise. WW1 is the number of deaths due to the First World War, divided by 10,000.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.4: Changes in Industrial Employment

	Dep. Var.: Industry Workers Growth							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Quota Exposure × Post	1.825***	1.497***	1.471***	1.469***	1.457***	1.413**	1.173*	0.996*
	(0.427)	(0.476)	(0.477)	(0.488)	(0.552)	(0.591)	(0.604)	(0.581)
Population	0.206*	0.243**	0.262**	0.261**	0.259*	0.266*	0.255*	0.213
	(0.123)	(0.123)	(0.126)	(0.127)	(0.137)	(0.142)	(0.143)	(0.142)
Extensive Margin × Post		0.652	0.619	0.621	0.616	0.631	0.709*	0.701*
		(0.403)	(0.404)	(0.409)	(0.420)	(0.427)	(0.422)	(0.419)
Agriculture × Post			0.077	0.079	0.075	0.064	0.081	0.068
			(0.082)	(0.094)	(0.108)	(0.111)	(0.112)	(0.110)
Urbanization × Post				0.001	0.000	0.003	0.002	0.000
				(0.058)	(0.061)	(0.062)	(0.062)	(0.061)
Literacy × Post					-0.004	-0.008	0.053	0.052
					(0.072)	(0.073)	(0.085)	(0.084)
WW1 × Post						-0.026	-0.014	-0.009
						(0.065)	(0.065)	(0.065)
South × Post							0.047	0.046
							(0.037)	(0.037)
US GDP Growth × QE								0.136***
								(0.042)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	205	205	205	205	205	205	205	205
Observations	742	742	742	742	742	742	742	742
R2	0.541	0.543	0.543	0.542	0.541	0.540	0.540	0.548
F-stat	6.777	6.951	6.664	5.616	5.194	4.603	4.602	4.748
Mean Dep. Var.	0.060	0.060	0.060	0.060	0.060	0.060	0.060	0.060

Notes. This table displays the effect of exposure to the Quota Acts on changes in industrial employment. Industrial employment growth is defined as the decade-on-decade percentage change in industrial employment. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are always clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Literacy is the number of people who could read and write as a share of the overall population in 1901. South is a dummy equal to zero if the district is in the EU NUTS 2 ITC or ITH region, and one otherwise. WW1 is the number of deaths due to the First World War, divided by 10,000.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.5: Changes in the Share of Industrial Workers

	Dep. Var.: Changes in Share of Industrial Workers							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Quota Exposure × Post	1.455***	1.139***	1.118***	1.237***	1.204**	1.168**	1.154**	0.888
	(0.356)	(0.411)	(0.412)	(0.425)	(0.465)	(0.473)	(0.520)	(0.538)
Population	0.074	0.105	0.124	0.134	0.129	0.134	0.134	0.078
	(0.090)	(0.088)	(0.092)	(0.093)	(0.096)	(0.099)	(0.101)	(0.097)
Extensive Margin × Post		0.613*	0.579	0.509	0.497	0.509	0.513	0.488
		(0.353)	(0.351)	(0.360)	(0.372)	(0.382)	(0.390)	(0.384)
Agriculture × Post			0.072	0.004	-0.005	-0.014	-0.013	-0.028
			(0.059)	(0.075)	(0.096)	(0.101)	(0.104)	(0.101)
Urbanization × Post				-0.077	-0.081	-0.078	-0.078	-0.074
				(0.053)	(0.061)	(0.062)	(0.062)	(0.061)
Literacy × Post					-0.012	-0.014	-0.011	-0.022
					(0.064)	(0.063)	(0.080)	(0.079)
WW1 × Post						-0.020	-0.020	-0.019
						(0.071)	(0.071)	(0.069)
South × Post							0.003	-0.001
							(0.036)	(0.035)
US GDP Growth × QE								0.173***
								(0.035)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	205	205	205	205	205	205	205	205
Observations	729	729	729	729	729	729	729	729
R2	0.477	0.479	0.479	0.480	0.479	0.478	0.477	0.500
F-stat	6.068	6.487	5.568	5.131	4.430	3.894	3.522	7.913
Mean Dep. Var.	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051

Notes. This table displays the effect of exposure to the Quota Acts on changes in the share of industrial workers relative to total employment. The share of industrial workers is defined as the ratio between industrial workers and total employment. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are always clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Literacy is the number of people who could read and write as a share of the overall population in 1901. South is a dummy equal to zero if the district is in the EU NUTS 2 ITC or ITH region, and one otherwise. WW1 is the number of deaths due to the First World War, divided by 10,000.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.6: Technology Adoption in Selected Manufacture Sectors

	Dep. Var.: Mechanical Engines in Construction Firms							Dep. Var.: Electrical Engines in Textile Firms						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Quota Exposure × Post	-1.267***	-1.184***	-1.207***	-1.187***	-1.185***	-1.012**	-0.996**	-2.323***	-2.346***	-2.216***	-2.329***	-2.131***	-2.081***	-2.082***
	(0.385)	(0.396)	(0.393)	(0.383)	(0.411)	(0.423)	(0.413)	(0.692)	(0.720)	(0.705)	(0.715)	(0.758)	(0.760)	(0.760)
Population	0.316**	0.297**	0.316**	0.317**	0.316**	0.337**	0.303**	0.316	0.322	0.199	0.200	0.154	0.160	0.161
	(0.133)	(0.136)	(0.138)	(0.138)	(0.141)	(0.139)	(0.143)	(0.207)	(0.208)	(0.220)	(0.221)	(0.226)	(0.227)	(0.227)
Extensive Margin × Post		-0.181	-0.210	-0.225	-0.225	-0.216	-0.227		0.050	0.205	0.287	0.242	0.245	0.245
		(0.159)	(0.161)	(0.167)	(0.165)	(0.164)	(0.166)		(0.439)	(0.427)	(0.458)	(0.456)	(0.456)	(0.456)
Agriculture × Post			0.094	0.076	0.076	0.078	0.211			-0.530***	-0.421**	-0.362	-0.361	-0.372
			(0.097)	(0.126)	(0.134)	(0.134)	(0.163)			(0.157)	(0.212)	(0.227)	(0.227)	(0.287)
Urbanization × Post				-0.020	-0.020	-0.020	-0.026				0.115	0.090	0.090	0.088
				(0.071)	(0.070)	(0.070)	(0.069)				(0.131)	(0.133)	(0.133)	(0.142)
WW1 × Post					0.002	-0.004	-0.175*					0.171	0.169	0.179
					(0.080)	(0.080)	(0.102)					(0.143)	(0.143)	(0.162)
US GDP Growth × QE						-0.123***	-0.124***						-0.033	-0.033
						(0.033)	(0.033)						(0.025)	(0.025)
Construction Employment × Post							0.001*							
							(0.000)							
Textile Employment × Post														-0.000
														(0.000)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	208	208	208	208	208	208	208	209	209	209	209	209	209	209
Observations	786	786	786	786	786	786	786	791	791	791	791	791	791	791
R2	0.808	0.808	0.808	0.807	0.807	0.811	0.811	0.874	0.873	0.876	0.876	0.876	0.876	0.876
F-stat	5.352	4.724	4.134	3.747	3.590	4.663	4.524	21.263	17.038	16.080	13.346	12.327	10.956	9.865
Mean Dep. Var.	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.472	0.472	0.472	0.472	0.472	0.472	0.472

Notes. This table displays the effect of exposure to the Quota acts on the number of mechanical and electrical engines in construction and textile manufacture firms. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Sector Employment is the 1901-number of manufacture workers. WW1 is the number of deaths due to the First World War, divided by 10,000.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.7: Employment Growth in Selected Manufacture Sectors

	Dep. Var.: Changes in Employment in Construction Firms							Dep. Var.: Changes in Employment in Textile Firms						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Quota Exposure × Post	4.611***	6.103***	6.103***	6.359***	5.520***	6.192***	6.203***	5.247***	5.651***	5.651***	5.977***	5.167***	6.795***	6.724***
	(1.411)	(1.626)	(1.627)	(1.635)	(1.669)	(1.896)	(1.895)	(1.266)	(1.398)	(1.399)	(1.336)	(1.318)	(1.439)	(1.465)
Population	0.027	-0.095	-0.091	-0.103	0.055	0.113	0.146	-0.518	-0.550	-0.549	-0.559	-0.404	-0.288	-0.243
	(0.339)	(0.359)	(0.377)	(0.374)	(0.384)	(0.382)	(0.386)	(0.340)	(0.343)	(0.365)	(0.360)	(0.350)	(0.334)	(0.330)
Extensive Margin × Post		-2.693**	-2.703**	-2.887**	-2.342*	-2.288*	-2.290*		-0.715	-0.720	-0.964	-0.432	-0.364	-0.237
		(1.293)	(1.277)	(1.328)	(1.316)	(1.311)	(1.304)		(0.991)	(1.006)	(1.025)	(1.029)	(1.019)	(0.997)
Agriculture × Post			0.016	-0.128	-0.277	-0.281	-0.424			0.007	-0.170	-0.325	-0.291	-0.711*
			(0.257)	(0.274)	(0.294)	(0.295)	(0.377)			(0.302)	(0.355)	(0.342)	(0.344)	(0.398)
Urbanization × Post				-0.157	-0.073	-0.069	-0.060				-0.192	-0.112	-0.105	-0.196
				(0.167)	(0.164)	(0.164)	(0.163)				(0.181)	(0.178)	(0.177)	(0.183)
WW1 × Post					-0.460**	-0.470**	-0.291					-0.458**	-0.472**	-0.149
					(0.204)	(0.205)	(0.280)					(0.181)	(0.182)	(0.226)
US GDP Growth × QE						-0.369*	-0.369*						-0.807***	-0.805***
						(0.196)	(0.196)						(0.122)	(0.123)
Construction Employment × Post							-0.001							
							(0.001)							
Textile Employment × Post														-0.001**
														(0.000)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	200	200	200	200	200	200	200	200	200	200	200	200	200	200
Observations	778	778	778	778	778	778	778	774	774	774	774	774	774	774
R2	0.315	0.318	0.317	0.316	0.318	0.329	0.328	0.451	0.450	0.449	0.449	0.451	0.499	0.501
F-stat	20.249	16.662	14.117	10.407	8.837	8.022	7.963	5.665	4.555	3.925	4.339	5.170	7.683	8.530
Mean Dep. Var.	0.553	0.553	0.553	0.553	0.553	0.553	0.553	0.291	0.291	0.291	0.291	0.291	0.291	0.291

Notes. This table displays the effect of exposure to the Quota acts on the the the growth rate of workers employed in construction and textile manufacture firms. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Sector Employment is the 1901-number of manufacture workers. WW1 is the number of deaths due to the First World War, divided by 10,000.

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

TABLE 2.B.8: First Stage Regressions

	Shift Share			Railway	
	(1) Pre 1924	(2) Pre WW1	(3) Pre Quota	(4) RAP total	(5) RAP region
IV QE	0.778***	0.833***	0.791***	3.398***	8.255***
	(0.038)	(0.038)	(0.039)	(1.169)	(2.317)
Extensive Margin × Post	0.012	-0.001	0.011	0.205***	0.187**
	(0.015)	(0.012)	(0.015)	(0.077)	(0.072)
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Number of Districts	207	207	207	207	207
Observations	754	754	754	754	754
KP Wald rk F	414.366	483.861	422.069	8.456	12.692

Notes. This table reports the result of the first stage instrumental variable estimation. The instrument (IV Quota Exposure) in the first three columns is defined in (2.4). The first column reports the correlation between QE and its instrument over the full sample (1890-1939). Instrument in column (2) restricts the emigrant outflow to the pre-WW1 period (1890-1914). Column (3) reports the results when considering emigrants over the pre-Quota period (1890-1924). In the last two columns, the instrument is defined as in equation (2.6). Results in column "RA total" use aggregate emigration instead of regional emigration. All regressions partial out district and year fixed effects. Further controls are population, the emigration rate and labor market slackness in 1901 interacted with a post-treatment dummy. K-P F-stat refers to the Kleibergen-Paap F-statistic for weak instrument.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.9: Urbanization and Employment Share - 2sls

	(1) Urbanization	(2) Industrialization	(3) Agriculture
Panel A: OLS			
Quota Exposure × Post	-0.410***	1.316***	-0.606***
	(0.109)	(0.414)	(0.153)
Panel B: 2SLS Shift Share			
Quota Exposure × Post	-0.332***	1.382***	-0.603***
	(0.124)	(0.474)	(0.177)
Panel C: 2SLS Railway Regional			
Quota Exposure × Post	-0.866**	2.379	-1.091***
	(0.359)	(1.545)	(0.393)
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Number of Districts	205	207	208
Observations	995	731	746
Mean Dep. Var.	0.279	0.044	-0.031

Notes. This table reports the effect of exposure to the Quota Acts on urbanization and changes in the share of industrial and agricultural workers relative to overall employment. Urbanization is defined as the share of the population living in cities no smaller than 10,000 inhabitants. The share of sector employment is defined as the ratio between sector and aggregate employment. Panel A presents reduced form estimates. Panel B reports 2SLS estimates based on the instrument defined in (2.4). All regressions include district and year fixed effects. Further controls are log-population, labor market slackness in 1901 interacted with a post-treatment dummy and the emigration rate, defined as the number of emigrants 1890-1914 relative to 1880-population, interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are robust and clustered at the district level.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.10: Labor intensity and emigration - 2sls

	Worker/Firm		Worker/Engine		Worker/Horsepower	
	(1) All	(2) Engine	(3) Mechanic	(4) Electric	(5) Mechanic	(6) Electric
Panel A: OLS						
Quota Exposure × Post	0.208	0.184	1.135***	1.050***	0.248	1.212***
	(0.239)	(0.396)	(0.174)	(0.339)	(0.353)	(0.300)
Panel B: 2SLS Shift Share						
Quota Exposure × Post	0.482*	0.563	1.264***	0.699**	-0.251	0.964***
	(0.276)	(0.428)	(0.190)	(0.327)	(0.403)	(0.294)
Panel C: 2SLS Railway Regional						
Quota Exposure × Post	1.432**	1.453	1.588***	1.725*	-0.337	1.531**
	(0.700)	(1.200)	(0.524)	(0.984)	(1.157)	(0.753)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	209	209	208	207	209	208
Observations	785	787	785	783	786	785
Mean Dep. Var.	-0.082	-0.054	-0.077	-0.258	-0.078	-0.195

Notes. This table displays the effect of being exposed to the Quota Acts on various measures for labor intensity in production. The first and second columns report the effect on, respectively, the worker-per-firm and the worker-per-firm with engine ratios. The third and fourth columns show the effect on the ratio between worker and mechanical and electrical engines; the fifth and sixth display the effect the ratio between workers and mechanical and electrical horsepower. Panel A presents reduced form estimates. Panel B reports 2SLS estimates based on the instrument defined in (2.4). All regressions include district and year fixed effects. Further controls are log-population, labor market slackness in 1901 interacted with a post-treatment dummy and the emigration rate, defined as the number of emigrants 1890-1914 relative to 1880-population, interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are robust and clustered at the district level.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.11: Capital Investment and Emigration by Industry Sectors - 2sls

	Mining		Agriculture		Steel		Construction		Textile		Chemical	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS	(9) OLS	(10) 2SLS	(11) OLS	(12) 2SLS
Panel A: Total Firms												
Quota Exposure × Post	-1.593***	-0.879***	-0.409***	-0.209*	-0.095	0.020	0.677**	0.674**	-0.876***	-0.548**	0.149	-0.264
	(0.306)	(0.323)	(0.130)	(0.119)	(0.103)	(0.111)	(0.314)	(0.328)	(0.188)	(0.216)	(0.474)	(0.434)
Panel B: Firms with Engine												
Quota Exposure × Post	-0.119	0.023	-0.222**	-0.036	0.426	0.715	0.173	0.303	-1.388***	-1.015***	-0.119	-0.255
	(0.321)	(0.357)	(0.110)	(0.112)	(0.451)	(0.467)	(0.241)	(0.247)	(0.282)	(0.325)	(0.320)	(0.283)
Panel C: Mechanical Engines												
Quota Exposure × Post	-0.586***	-0.427**	-0.462**	-0.129	-0.853*	-0.817*	-1.289***	-0.759	-0.047	0.042	-0.982***	-0.897**
	(0.180)	(0.206)	(0.211)	(0.215)	(0.437)	(0.437)	(0.407)	(0.460)	(0.082)	(0.081)	(0.327)	(0.355)
Panel D: Electrical Engines												
Quota Exposure × Post	-2.553***	-1.974**	1.581***	1.525**	-0.669***	-0.465**	-0.982**	-0.226	-2.280***	-1.749**	-0.628	-0.059
	(0.976)	(0.946)	(0.538)	(0.605)	(0.234)	(0.227)	(0.479)	(0.421)	(0.711)	(0.803)	(0.567)	(0.545)
Panel E: Mechanical Horsepower												
Quota Exposure × Post	-2.079***	-1.522*	-0.985***	-0.363	-1.244	-1.528	-2.209***	-1.293**	2.264***	1.346*	-0.324	0.172
	(0.719)	(0.802)	(0.274)	(0.290)	(1.235)	(1.189)	(0.486)	(0.596)	(0.673)	(0.706)	(1.018)	(1.002)
Panel F: Electrical Horsepower												
Quota Exposure × Post	-1.415	-1.041	1.293*	1.606*	-0.565	-0.592	-1.689	-0.330	-0.780*	-0.418	0.583	0.810
	(1.577)	(1.667)	(0.735)	(0.823)	(0.350)	(0.360)	(1.155)	(1.022)	(0.397)	(0.413)	(0.723)	(0.818)
Observations	785	785	782	782	787	787	786	786	787	787	785	785
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	209	209	206	206	209	209	209	209	209	209	209	209

Notes. This table displays the effect of QE on employment by manufacture sector. OLS and 2SLS columns respectively report reduced-form and shift-share instrumental variable estimates. All regressions include district and year fixed effects, log-population and 1901-labor marked slackness interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are clustered at the district level.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.12: Changes in Industry Employment by Sector - 2sls

	(1) Mining	(2) Agriculture	(3) Steel	(4) Construction	(5) Textile	(6) Chemical
Panel A: OLS						
Quota Exposure × Post	0.442	-2.459*	1.379	6.103***	5.651***	0.017
	(0.388)	(1.261)	(1.573)	(1.626)	(1.398)	(0.308)
Panel B: 2SLS						
Quota Exposure × Post	0.419	-2.275	2.757*	5.912***	7.077***	0.158
	(0.494)	(1.583)	(1.575)	(2.183)	(1.327)	(0.361)
Observations	685	776	775	778	774	681
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	194	200	198	200	200	195
Observations	685	776	775	778	774	681
F-stat	8.111	5.319	5.982	15.309	8.373	1.828
Mean Dep. Var.	0.724	0.422	0.250	0.553	0.291	0.751

Notes. This table displays the effect of exposure to the Quota Acts on changes in employment by manufacture sector. Hence, column "Agriculture" reports the impact of QE on employment in manufacture firms working in agriculture, not that on agriculture. We do not show the "public utility" sector due to data availability, and a residual sector of unassigned firms. Panel A presents reduced form estimates. Panel B reports 2SLS estimates based on the instrument defined in (2.4). All regressions include district and year fixed effects. Further controls are log-population, changes in industrial employment, the emigration rate and 1901 labor market slackness interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are clustered at the district level.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.13: Population Growth Varying the Measurement of Quota Exposure

	Baseline	Weighted		Alternative periods		
	(1)	(2)	(3)	(4)	(5)	(6)
QE × Post	0.449***					
	(0.124)					
QE × Post: decreasing weight		1.001**				
		(0.386)				
QE × Post: increasing weight			2.328***			
			(0.551)			
QE× Post: 1902-1905				1.664***		
				(0.411)		
QE× Post: 1906-1909					1.241***	
					(0.414)	
QE× Post: 1910-1913						0.953**
						(0.436)
Extensive Margin × Post	-0.068	-0.045	-0.083	-0.087	-0.047	-0.025
	(0.055)	(0.057)	(0.055)	(0.056)	(0.055)	(0.057)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	204	204	204	204	204	204
Observations	751	751	751	751	751	751
R2	0.452	0.441	0.459	0.459	0.444	0.436
F-stat	9.932	8.538	11.094	11.375	8.965	7.966
Mean Dep. Var.	1.042	1.042	1.042	1.042	1.042	1.042
Std. Beta Coef.	0.240	-0.008	0.268	0.266	0.198	0.149

Notes. This table displays the effect of exposure to the Quota Acts on population growth. Different measures of Quota Exposure are used, as further robustness. Hence, column "Baseline" reports the estimate for Quota Exposure as defined and used throughout the paper. Column "Weighted" reports the coefficients for two measures of Quota Exposure constructed using an exponential smoothing with coefficient 0.9: "decreasing weight" assigns lower weight to US emigration further back in time; "increasing weight" assigns lower weight to more recent US emigration. Column "Alternative periods" shows instead the estimates for Quota Exposure constructed using only US emigration from selected sub-periods of time: we use three different periods, respectively 1902-1905, 1906-1910, 1910-1913. All regressions include district and year fixed effects. Further controls are log-population, changes in industrial employment, the emigration rate and 1901 labor market slackness interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are clustered at the district level.

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

TABLE 2.B.14: Employment in Manufacture Varying the Measurement of Quota Exposures

	Baseline	Weighted		Alternative periods		
	(1)	(2)	(3)	(4)	(5)	(6)
Quota Exposure × Post	1.510***					
	(0.475)					
QE × Post: decreasing weight		4.241***				
		(1.389)				
QE × Post: increasing weight			6.848***			
			(2.304)			
QE× Post: 1902-1905				5.356***		
				(1.615)		
QE× Post: 1906-1909					4.647***	
					(1.483)	
QE× Post: 1910-1913						4.702***
						(1.458)
Extensive Margin × Post	0.637	0.675*	0.626	0.578	0.698*	0.724*
	(0.400)	(0.406)	(0.404)	(0.400)	(0.398)	(0.399)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	205	205	205	205	205	205
Observations	742	742	742	742	742	742
R2	0.542	0.542	0.542	0.543	0.542	0.541
F-stat	7.004	6.400	6.579	7.275	6.653	6.870
Mean Dep. Var.	0.060	0.060	0.060	0.060	0.060	0.060
Std. Beta Coef.	0.123	0.119	0.119	0.126	0.116	0.116

Notes. This table displays the effect of exposure to the Quota Acts on industrial employment growth. Different measures of Quota Exposure are used, as further robustness. Hence, column "Baseline" reports the estimate for Quota Exposure as defined and used throughout the paper. Column "Weighted" reports the coefficients for two measures of Quota Exposure constructed using an exponential smoothing with coefficient 0.9: "decreasing weight" assigns lower weight to US emigration further back in time; "increasing weight" assigns lower weight to more recent US emigration. Column "Alternative periods" shows instead the estimates for Quota Exposure constructed using only US emigration from selected sub-periods of time: we use three different periods, respectively 1902-1905, 1906-1910, 1910-1913. All regressions include district and year fixed effects. Further controls are log-population, changes in industrial employment, the emigration rate and 1901 labor market slackness interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are clustered at the district level.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.15: Investment in Capital Goods Using Time-Weighted Quota Exposure

	Firm		Engine		Horsepower	
	(1) All	(2) Engine	(3) Mechanic	(4) Electric	(5) Mechanic	(6) Electric
QE × Post: decreasing	0.641	1.075	-3.000***	-3.331***	-2.008**	-4.014***
	(0.724)	(1.161)	(0.493)	(0.930)	(0.985)	(0.965)
Extensive Margin × Post	-0.086	-0.037	0.215**	0.015	-0.260**	-0.009
	(0.116)	(0.185)	(0.103)	(0.135)	(0.117)	(0.151)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	208	208	207	207	209	208
Observations	785	785	783	784	785	784
R2	0.963	0.834	0.426	0.953	0.834	0.936
F-stat	0.418	0.272	13.464	4.478	4.915	10.468
Mean Dep. Var.	0.766	0.582	0.018	0.793	0.270	0.828
Std. Beta Coef.	0.012	0.030	-0.367	-0.057	-0.048	-0.088

Notes. This table displays the effect of exposure to the Quota Acts on changes on various measures for capital and investment and technology adoption. Quota Exposure is constructed using an exponential smoothing with coefficient 0.9. In this case, we assigns lower weight to US emigration further back in time. All regressions include district and year fixed effects. Further controls are log-population, changes in industrial employment, the emigration rate and 1901 labor market slackness interacted with a post-treatment dummy. Outcome variables are defined in growth rate. Standard errors are clustered at the district level.

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

TABLE 2.B.16: Changes in Mechanical and Electrical Engines

	Dep. Var.: Changes in Number of Mechanical Engines								Dep. Var.: Changes in Number of Electrical Engines							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Quota Exposure × Post	-0.983***	-1.030***	-1.025***	-0.996***	-0.897***	-0.927***	-0.884***	-0.810***	-0.986***	-1.040***	-1.125***	-1.002***	-1.095***	-1.216***	-1.000***	-0.856**
	(0.171)	(0.180)	(0.182)	(0.182)	(0.204)	(0.223)	(0.236)	(0.233)	(0.331)	(0.337)	(0.317)	(0.306)	(0.311)	(0.336)	(0.357)	(0.353)
Population	-0.096**	-0.077	-0.080	-0.081	-0.066	-0.062	-0.060	-0.052	-0.204**	-0.187**	-0.113	-0.114	-0.127	-0.110	-0.101	-0.085
	(0.048)	(0.048)	(0.050)	(0.050)	(0.051)	(0.052)	(0.053)	(0.052)	(0.090)	(0.092)	(0.088)	(0.088)	(0.088)	(0.086)	(0.085)	(0.085)
Extensive Margin × Post		0.122	0.126	0.108	0.119	0.122	0.110	0.112		0.132	0.046	-0.035	-0.049	-0.037	-0.101	-0.098
		(0.112)	(0.112)	(0.114)	(0.112)	(0.111)	(0.115)	(0.115)		(0.134)	(0.128)	(0.138)	(0.140)	(0.139)	(0.150)	(0.148)
Agriculture × Post			-0.017	-0.047	-0.023	-0.032	-0.035	-0.033			0.317***	0.205**	0.182*	0.149	0.131	0.136
			(0.039)	(0.050)	(0.064)	(0.066)	(0.066)	(0.066)			(0.065)	(0.089)	(0.098)	(0.105)	(0.100)	(0.101)
Urbanization × Post				-0.031	-0.021	-0.019	-0.019	-0.019				-0.118**	-0.127**	-0.119**	-0.120**	-0.120**
				(0.027)	(0.031)	(0.031)	(0.031)	(0.031)				(0.049)	(0.052)	(0.052)	(0.051)	(0.052)
Literacy × Post					0.033	0.031	0.018	0.018					-0.031	-0.041	-0.105	-0.103
					(0.040)	(0.040)	(0.045)	(0.045)					(0.056)	(0.056)	(0.065)	(0.065)
WW1 × Post						-0.019	-0.021	-0.023						-0.078	-0.086	-0.090*
						(0.032)	(0.032)	(0.032)						(0.054)	(0.054)	(0.054)
South × Post							-0.009	-0.009							-0.047	-0.046
							(0.018)	(0.018)							(0.029)	(0.029)
US GDP Growth × QE								-0.046***								-0.086***
								(0.008)								(0.013)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
Observations	801	801	801	801	801	801	801	801	801	801	801	801	801	801	801	801
R2	0.787	0.788	0.787	0.788	0.788	0.788	0.787	0.792	0.482	0.482	0.506	0.514	0.514	0.515	0.517	0.531
F-stat	17.011	13.123	10.585	8.495	7.004	6.078	5.477	7.864	5.144	4.115	7.583	8.404	7.717	6.756	6.827	10.443
Mean Dep. Var.	0.109	0.109	0.109	0.109	0.109	0.109	0.109	0.109	0.253	0.253	0.253	0.253	0.253	0.253	0.253	0.253

Notes. This table displays the effect of exposure to the Quota acts on the number of mechanical and electrical engines. All regressions include district and year fixed effects. Standard errors are clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Literacy is the number of people who could read and write as a share of the overall population in 1901. South is a dummy equal to zero if the district is in the EU NUTS 2 ITC or ITH region, and one otherwise. WW1 is the number of deaths due to the First World War, divided by 10,000.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.17: Changes in Mechanical and Electrical Horsepower

	Dep. Var.: Changes in Horsepower by Mechanical Engines								Dep. Var.: Changes in Horsepower by Electrical Engines							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Quota Exposure × Post	-0.741**	-0.520*	-0.481	-0.492*	-0.243	-0.038	-0.151	0.070	-1.183***	-1.089***	-1.081***	-1.048***	-0.939***	-0.857***	-0.871**	-0.803**
	(0.303)	(0.305)	(0.298)	(0.295)	(0.337)	(0.356)	(0.392)	(0.378)	(0.261)	(0.269)	(0.273)	(0.272)	(0.302)	(0.326)	(0.340)	(0.340)
Population	-0.019	-0.088	-0.125	-0.125	-0.090	-0.119	-0.123	-0.094	-0.111*	-0.149**	-0.156**	-0.155**	-0.141**	-0.151**	-0.152**	-0.142**
	(0.094)	(0.096)	(0.101)	(0.101)	(0.103)	(0.105)	(0.105)	(0.104)	(0.063)	(0.065)	(0.066)	(0.065)	(0.066)	(0.067)	(0.069)	(0.069)
Extensive Margin × Post		-0.520***	-0.478***	-0.471***	-0.433***	-0.457***	-0.423***	-0.416***		-0.235**	-0.228**	-0.251**	-0.239**	-0.244**	-0.241**	-0.238**
		(0.120)	(0.117)	(0.119)	(0.117)	(0.119)	(0.120)	(0.120)		(0.116)	(0.115)	(0.117)	(0.116)	(0.118)	(0.120)	(0.120)
Agriculture × Post			-0.157*	-0.147	-0.084	-0.029	-0.020	-0.011			-0.029	-0.067	-0.040	-0.016	-0.015	-0.013
			(0.088)	(0.099)	(0.103)	(0.109)	(0.110)	(0.109)			(0.059)	(0.072)	(0.088)	(0.094)	(0.095)	(0.094)
Urbanization × Post				0.010	0.036	0.021	0.022	0.023				-0.039	-0.029	-0.034	-0.034	-0.034
				(0.046)	(0.049)	(0.048)	(0.048)	(0.047)				(0.038)	(0.043)	(0.043)	(0.043)	(0.043)
Literacy × Post					0.084	0.101*	0.134*	0.137*					0.036	0.043	0.047	0.048
					(0.059)	(0.059)	(0.074)	(0.073)					(0.056)	(0.058)	(0.066)	(0.066)
WW1 × Post						0.135*	0.140*	0.135*						0.054	0.054	0.052
						(0.074)	(0.075)	(0.074)						(0.051)	(0.053)	(0.053)
South × Post							0.024	0.024							0.003	0.003
							(0.032)	(0.032)							(0.026)	(0.026)
US GDP Growth × QE								-0.133***								-0.043***
								(0.025)								(0.011)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	209	209	209	209	209	209	209	209	208	208	208	208	208	208	208	208
Observations	804	804	804	804	804	804	804	804	802	802	802	802	802	802	802	802
R2	0.635	0.639	0.641	0.640	0.640	0.642	0.642	0.654	0.794	0.796	0.796	0.796	0.796	0.796	0.795	0.797
F-stat	5.579	8.255	7.340	6.731	5.278	5.230	4.718	7.525	8.866	8.101	6.917	5.824	4.978	4.571	4.138	7.539
Mean Dep. Var.	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.189	0.189	0.189	0.189	0.189	0.189	0.189	0.189

Notes. This table displays the effect of exposure to the Quota acts on the horsepower generates by mechanical and electrical engines. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Literacy is the number of people who could read and write as a share of the overall population in 1901. South is a dummy equal to zero if the district is in the EU NUTS 2 ITC or ITH region, and one otherwise. WW1 is the number of deaths due to the First World War, divided by 10,000.

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

TABLE 2.B.18: Labor intensity of Technology: Mechanical and Electrical Engines

	Dep. Var.: Changes in Worker per Mechanical Engine								Dep. Var.: Changes in Worker per Electrical Engine							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Quota Exposure × Post	1.094***	1.212***	1.184***	1.235***	1.207***	1.221***	0.952***	0.840***	1.199***	1.261***	1.287***	1.212***	1.532***	1.629***	1.203***	0.992***
	(0.177)	(0.186)	(0.184)	(0.184)	(0.207)	(0.228)	(0.224)	(0.217)	(0.322)	(0.323)	(0.318)	(0.312)	(0.315)	(0.358)	(0.375)	(0.371)
Population	0.390***	0.357***	0.395***	0.395***	0.392***	0.390***	0.377***	0.364***	0.490***	0.472***	0.439***	0.438***	0.479***	0.466***	0.445***	0.428***
	(0.052)	(0.052)	(0.052)	(0.050)	(0.052)	(0.054)	(0.057)	(0.058)	(0.091)	(0.092)	(0.091)	(0.092)	(0.088)	(0.089)	(0.085)	(0.086)
Extensive Margin × Post		-0.269**	-0.336***	-0.377***	-0.380***	-0.382***	-0.315***	-0.317***		-0.141	-0.086	-0.029	0.013	-0.002	0.120	0.128
		(0.105)	(0.097)	(0.098)	(0.099)	(0.099)	(0.107)	(0.107)		(0.129)	(0.131)	(0.135)	(0.137)	(0.133)	(0.151)	(0.151)
Agriculture × Post			0.175***	0.110**	0.104	0.108	0.128*	0.122*			-0.147**	-0.072	0.003	0.029	0.072	0.065
			(0.044)	(0.050)	(0.064)	(0.067)	(0.066)	(0.065)			(0.069)	(0.092)	(0.105)	(0.113)	(0.103)	(0.103)
Urbanization × Post				-0.067***	-0.070**	-0.071**	-0.072**	-0.072**				0.083	0.114**	0.107*	0.110**	0.109*
				(0.025)	(0.030)	(0.030)	(0.029)	(0.029)				(0.054)	(0.056)	(0.057)	(0.056)	(0.056)
Literacy × Post					-0.009	-0.007	0.068	0.066					0.102*	0.110*	0.245***	0.240***
					(0.039)	(0.040)	(0.047)	(0.047)					(0.059)	(0.061)	(0.071)	(0.071)
WW1 × Post						0.009	0.018	0.021						0.059	0.080	0.084
						(0.040)	(0.037)	(0.036)						(0.066)	(0.061)	(0.061)
South × Post							0.056***	0.056***							0.097***	0.097***
							(0.019)	(0.019)							(0.032)	(0.032)
US GDP Growth × QE								0.069***								0.111***
								(0.012)								(0.013)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
Observations	784	784	784	784	784	784	784	784	784	784	784	784	784	784	784	784
R2	0.787	0.790	0.797	0.799	0.799	0.799	0.802	0.811	0.540	0.540	0.543	0.546	0.549	0.549	0.559	0.579
F-stat	30.935	25.223	24.419	22.054	18.849	16.724	15.967	19.526	13.089	10.230	8.316	7.232	8.589	7.440	8.890	17.873
Mean Dep. Var.	-0.069	-0.069	-0.069	-0.069	-0.069	-0.069	-0.069	-0.069	-0.214	-0.214	-0.214	-0.214	-0.214	-0.214	-0.214	-0.214

Notes. This table displays the effect of exposure to the Quota acts on the worker-per-mechanical engine and worker-per-electrical engine ratios. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Literacy is the number of people who could read and write as a share of the overall population in 1901. South is a dummy equal to zero if the district is in the EU NUTS 2 ITC or ITH region, and one otherwise. WW1 is the number of deaths due to the First World War, divided by 10,000.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

TABLE 2.B.19: Labor intensity of Technology: Mechanical and Electrical Horsepower

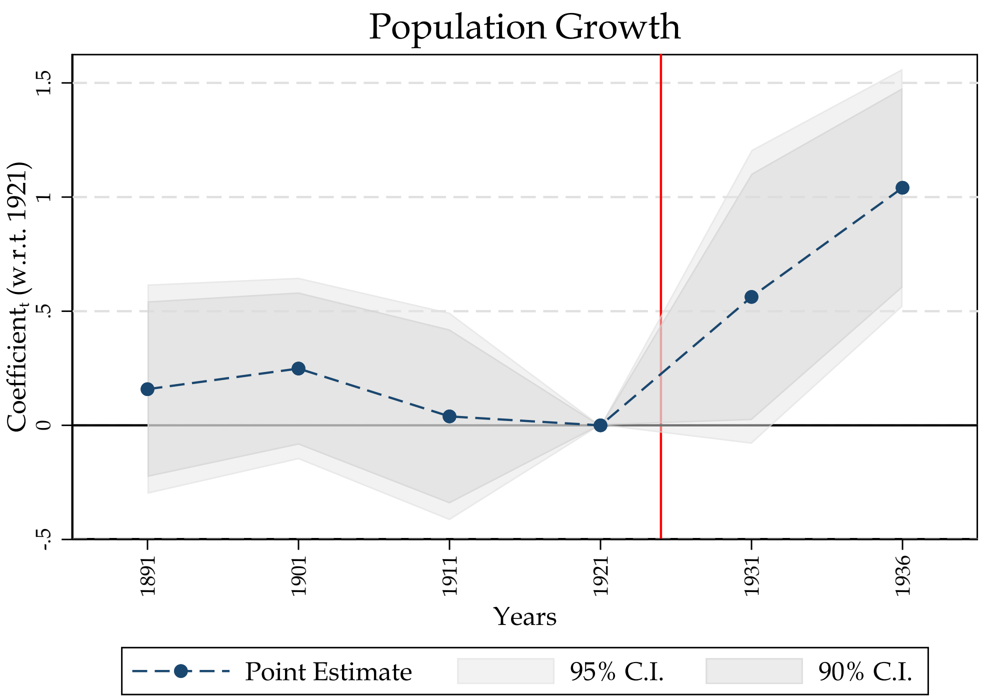
	Dep. Var.: Changes in Worker per Mechanical Horsepower								Dep. Var.: Changes in Worker per Electrical Horsepower							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Quota Exposure × Post	0.960***	0.862**	0.808***	0.897***	0.622*	0.433	0.373	0.066	1.455***	1.495***	1.465***	1.500***	1.323***	1.270***	1.011***	0.868***
	(0.328)	(0.333)	(0.306)	(0.302)	(0.352)	(0.374)	(0.425)	(0.402)	(0.258)	(0.265)	(0.263)	(0.268)	(0.296)	(0.318)	(0.338)	(0.327)
Population	0.275***	0.303***	0.379***	0.379***	0.345***	0.372***	0.369***	0.333***	0.446***	0.434***	0.473***	0.474***	0.452***	0.459***	0.446***	0.432***
	(0.091)	(0.091)	(0.092)	(0.093)	(0.096)	(0.101)	(0.101)	(0.100)	(0.056)	(0.055)	(0.058)	(0.058)	(0.059)	(0.061)	(0.063)	(0.064)
Extensive Margin × Post		0.224*	0.100	0.039	0.005	0.034	0.051	0.047		-0.093	-0.164	-0.190	-0.212	-0.205	-0.135	-0.135
		(0.123)	(0.124)	(0.131)	(0.138)	(0.133)	(0.136)	(0.137)		(0.124)	(0.123)	(0.131)	(0.133)	(0.132)	(0.134)	(0.135)
Agriculture × Post			0.338***	0.251**	0.184	0.131	0.137	0.123			0.184***	0.145**	0.104	0.089	0.113	0.108
			(0.099)	(0.108)	(0.117)	(0.126)	(0.125)	(0.125)			(0.060)	(0.073)	(0.081)	(0.083)	(0.081)	(0.081)
Urbanization × Post				-0.096**	-0.123**	-0.111**	-0.111**	-0.114**				-0.042	-0.059	-0.056	-0.055	-0.056
				(0.045)	(0.049)	(0.049)	(0.049)	(0.048)				(0.037)	(0.040)	(0.041)	(0.040)	(0.040)
Literacy × Post					-0.089	-0.105*	-0.086	-0.095					-0.057	-0.061	0.017	0.015
					(0.057)	(0.058)	(0.073)	(0.072)					(0.047)	(0.048)	(0.054)	(0.053)
WW1 × Post						-0.118	-0.115	-0.109						-0.033	-0.022	-0.018
						(0.082)	(0.082)	(0.080)						(0.052)	(0.048)	(0.047)
South × Post							0.013	0.013							0.057**	0.057**
							(0.035)	(0.034)							(0.023)	(0.023)
US GDP Growth × QE								0.167***								0.079***
								(0.023)								(0.012)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Districts	209	209	209	209	209	209	209	209	208	208	208	208	208	208	208	208
Observations	787	787	787	787	787	787	787	787	785	785	785	785	785	785	785	785
R2	0.658	0.658	0.667	0.669	0.669	0.670	0.670	0.689	0.837	0.836	0.840	0.840	0.840	0.840	0.842	0.847
F-stat	6.525	6.115	7.875	8.833	7.042	6.708	6.242	13.979	31.240	24.003	21.232	18.434	15.649	13.830	14.040	19.302
Mean Dep. Var.	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013	-0.149	-0.149	-0.149	-0.149	-0.149	-0.149	-0.149	-0.149

Notes. This table displays the effect of exposure to the Quota acts on the worker-per-mechanical horsepower and worker-per-electrical horsepower ratios. All regressions include district and year fixed effects. Outcome variables are defined in growth rate. Standard errors are clustered at the district level. Extensive margin is the emigration rate defined as the ratio between 1890-1914 emigration and 1880-population. Agriculture is the number of agriculture workers in 1901. Urbanization is the share of people living in cities no smaller than 10,000 inhabitants in 1901. Literacy is the number of people who could read and write as a share of the overall population in 1901. South is a dummy equal to zero if the district is in the EU NUTS 2 ITC or ITH region, and one otherwise. WW1 is the number of deaths due to the First World War, divided by 10,000.

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

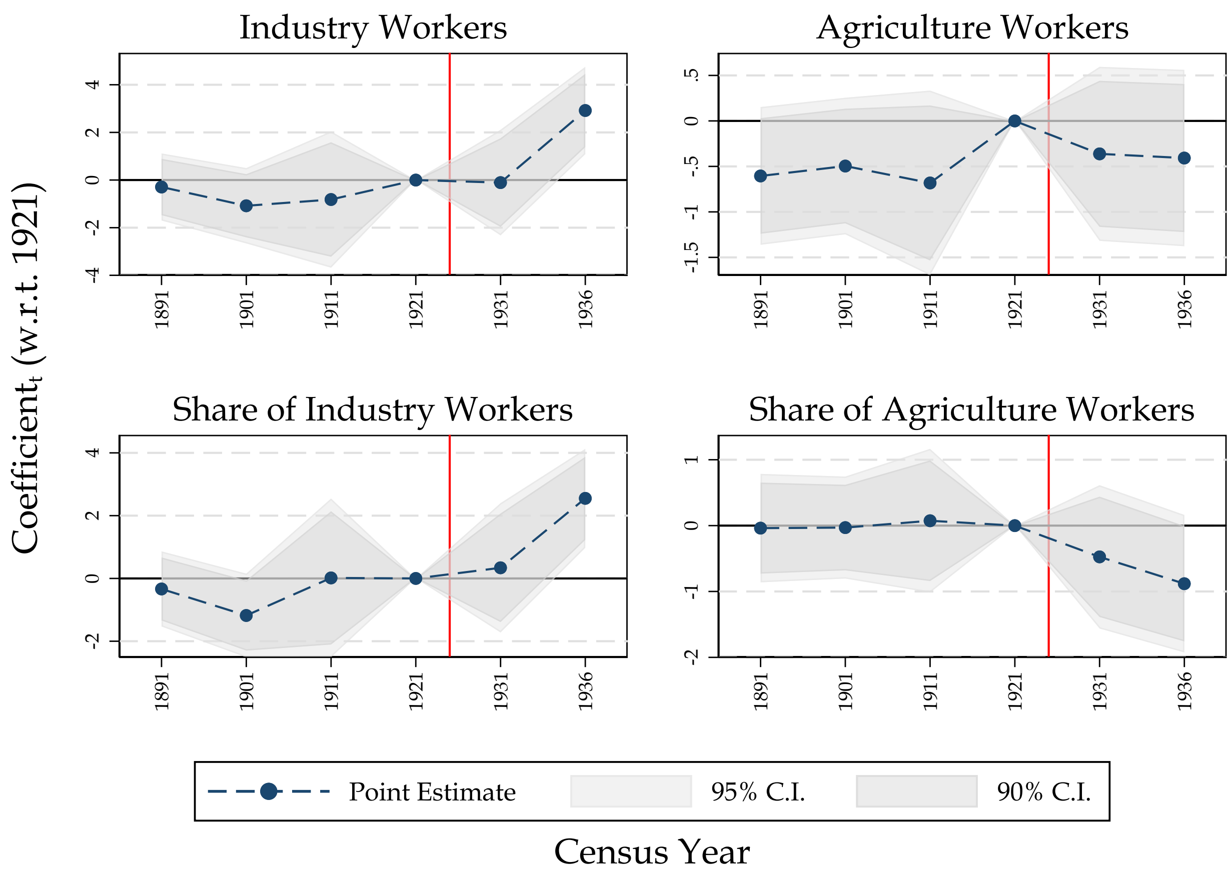
2.C Additional Figures

FIGURE 2.C.1: Event-Study of Population Growth and the Quota Acts



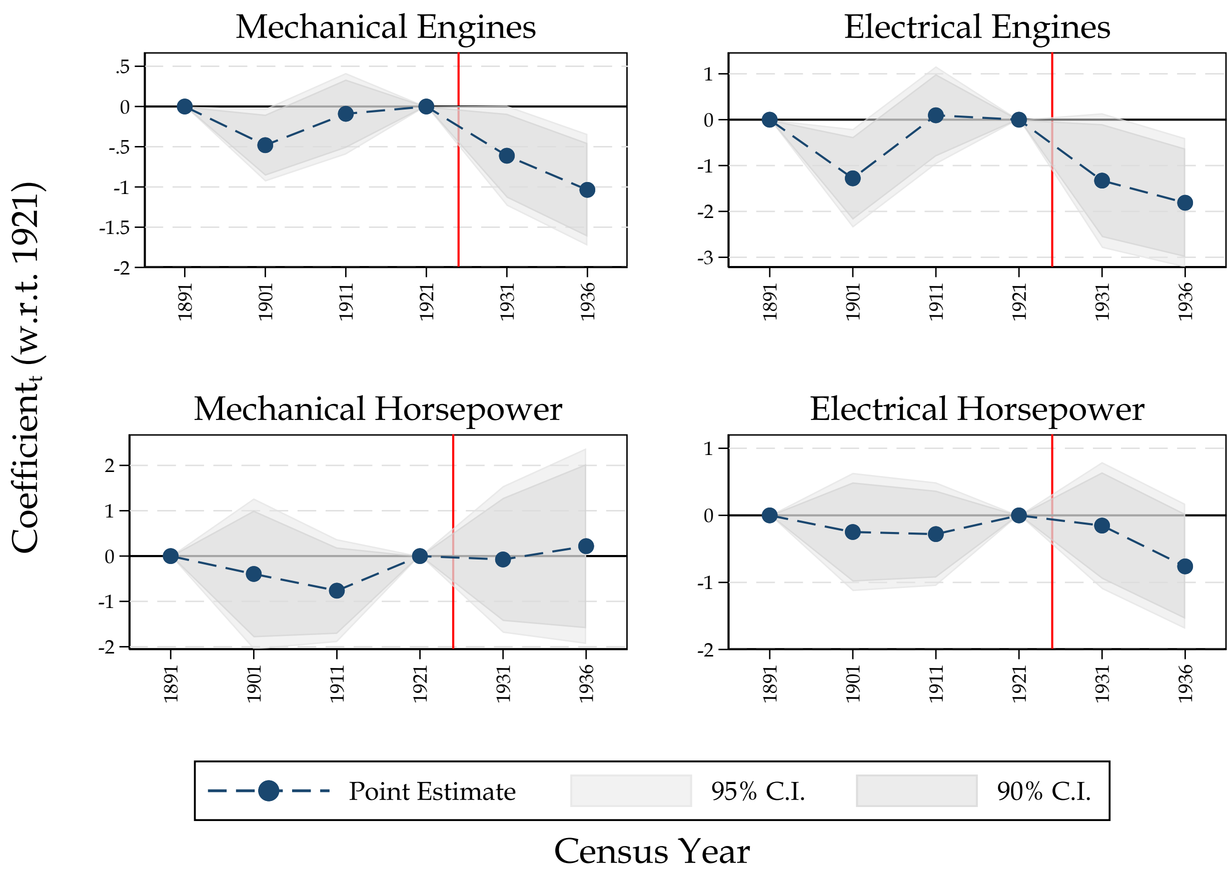
Notes. This figure plots the coefficient of the treatment measure (QE) interacted with census-decade time dummies. Regressions include district and year fixed effects, and region-by-year fixed effects. Further controls are the population in level, and 1901 labor market slackness interacted with census-decade dummies. Standard errors are clustered at the district-by-year level. Bands report 90% and 95% confidence levels. The red line indicates the 1924 (Johnson-Reed) Quota Act.

FIGURE 2.C.2: Event-Study of Industrial and Agriculture Employment



Notes. This figure plots the coefficients of the treatment measure (QE) interacted with census-decade time dummies. Regressions include district and year fixed effects, and region-by-year fixed effects. Further controls are the population in level, and 1901 labor market slackness interacted with census-decade dummies. Standard errors are clustered at the district-by-year level. Bands report 90% and 95% confidence levels. The red line indicates the 1924 (Johnson-Reed) Quota Act.

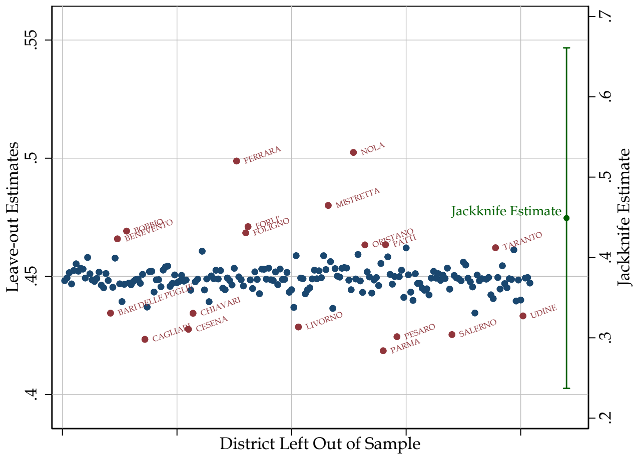
FIGURE 2.C.3: Event-Study of Technology Adoption and Capital Investment



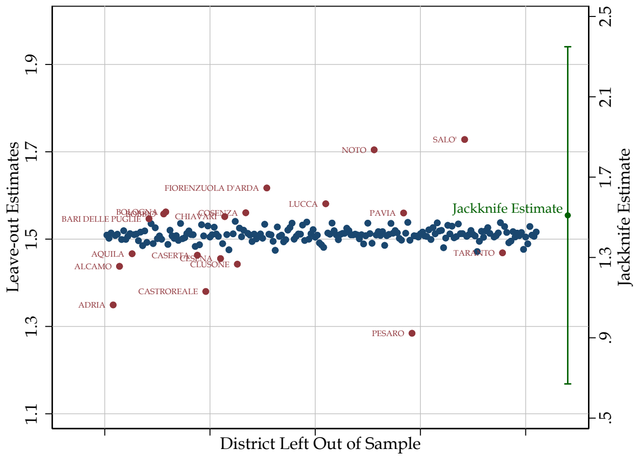
Notes. This figure plots the coefficients of the treatment measure (QE) interacted with census-decade time dummies. Regressions include district and year fixed effects, and region-by-year fixed effects. Further controls are the population in level, and 1901 labor market slackness interacted with census-decade dummies. Standard errors are clustered at the district-by-year level. Bands report 90% and 95% confidence levels. The red line indicates the 1924 (Johnson-Reed) Quota Act. For capital variables, 1931 actually refers to the 1927 Census of Manufacture.

FIGURE 2.C.4: Jackknife Estimation Routine

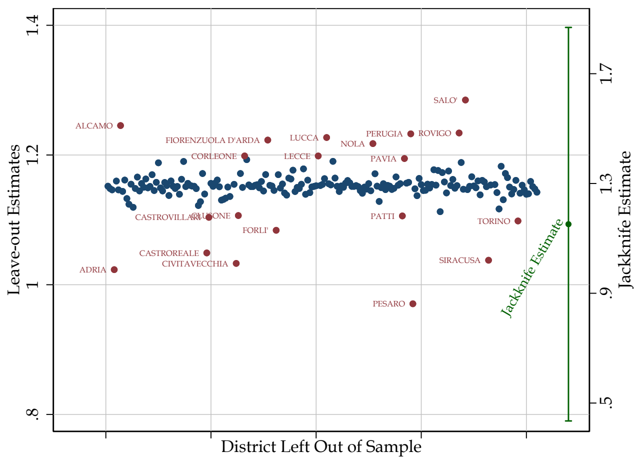
(A) Population Growth



(B) Industrial Employment

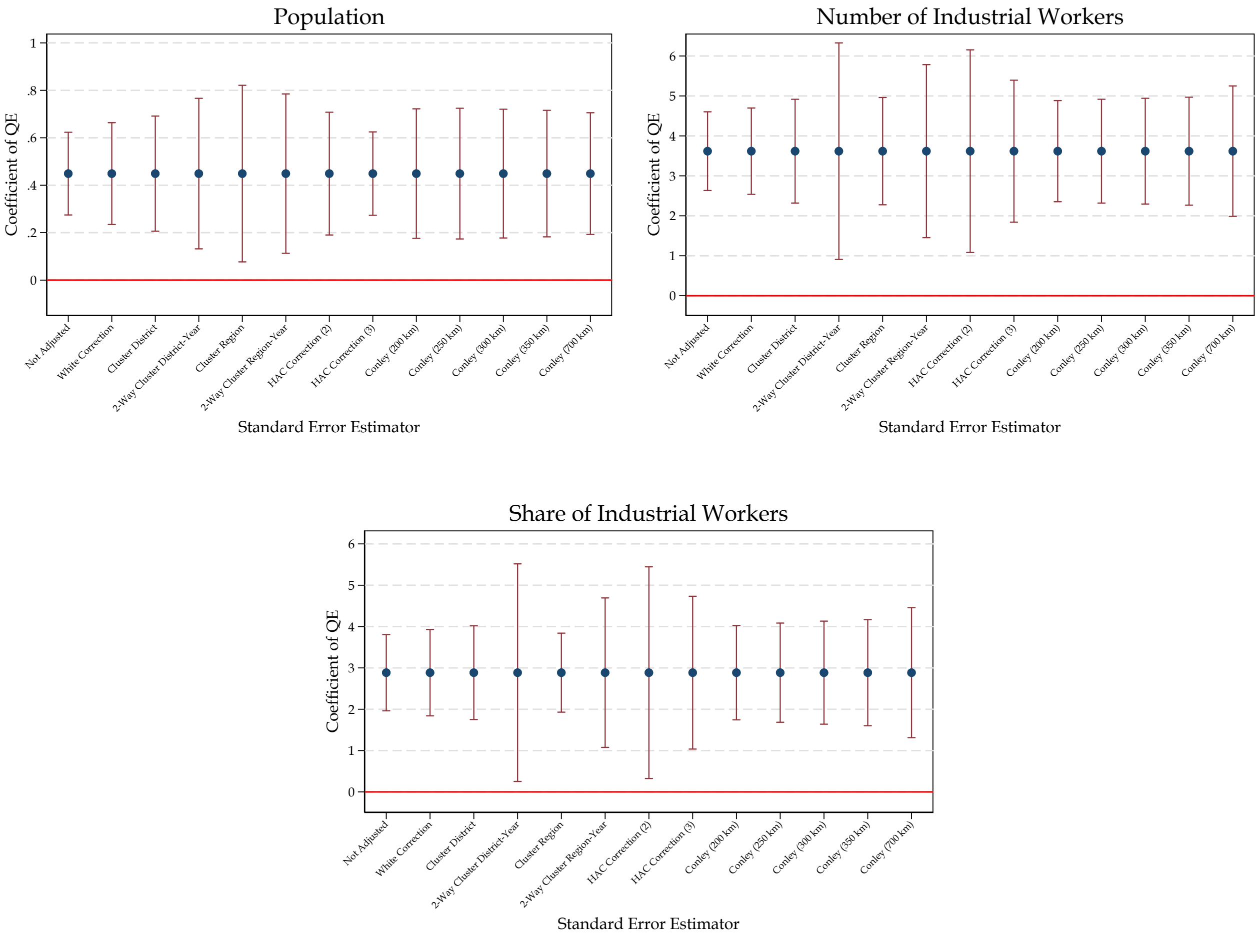


(C) Share of Industrial Workers

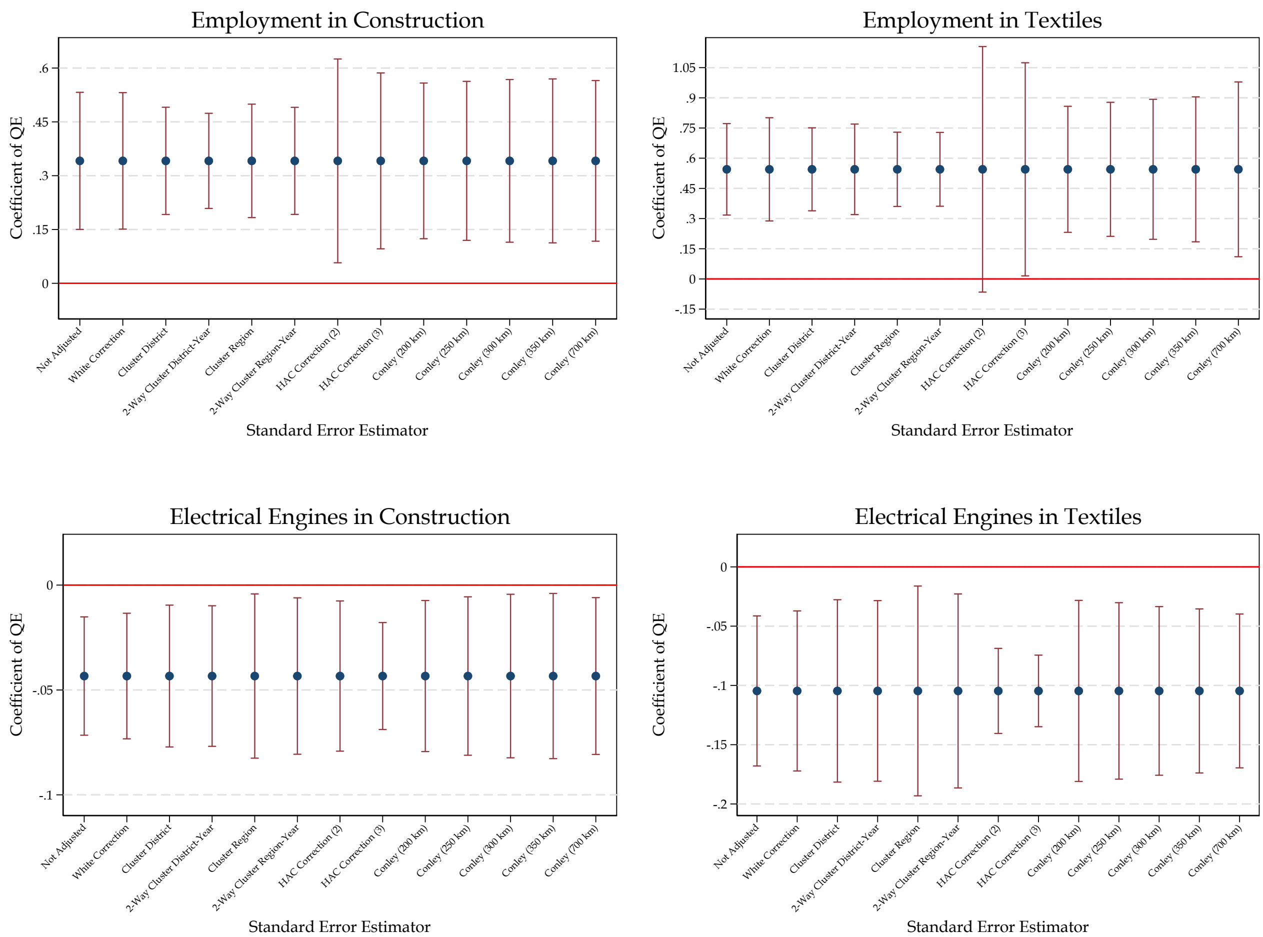


Notes. For each dependent variable shown in the header, each blue dot (on the left y-axis) reports the coefficient of Quota Exposure in the baseline difference-in-differences model dropping one district at a time. Red dots (on the left y-axis) are coefficients above and below respectively the 95th and the 5th percentiles. The green dot (on the right y-axis) reports the Jackknife estimator of the same coefficient, along with its 90% confidence bands.

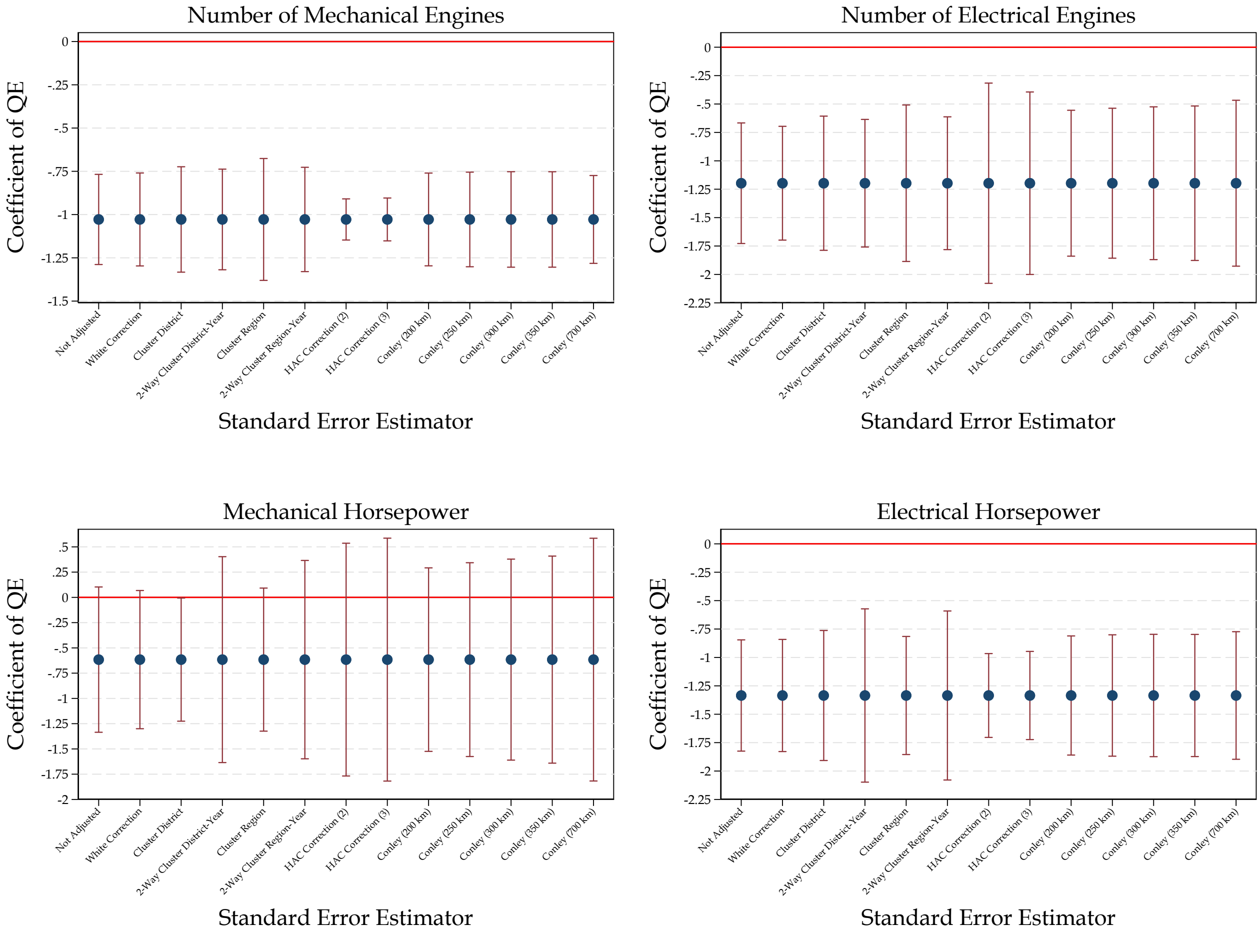
FIGURE 2.C.5: Standard Error Analysis



Notes. For a given outcome variable, the blue dots report the estimate of the coefficient of the treatment (QE) in the baseline difference-in-differences specification. The red bands report the 95% confidence intervals for a set of estimators for the coefficient's standard error. We include White standard errors which allow for heteroskedasticity; several clustered standard errors allowing for within-group autocorrelation; the Driscoll and Kraay (1998) correction for autocorrelation at two different time lags; several Conley (1999) estimates allowing for time and spatial autocorrelation. For the Conley SEs, we set maximal time-autocorrelation at 2 lags, and vary the radius of spatial autocorrelation.

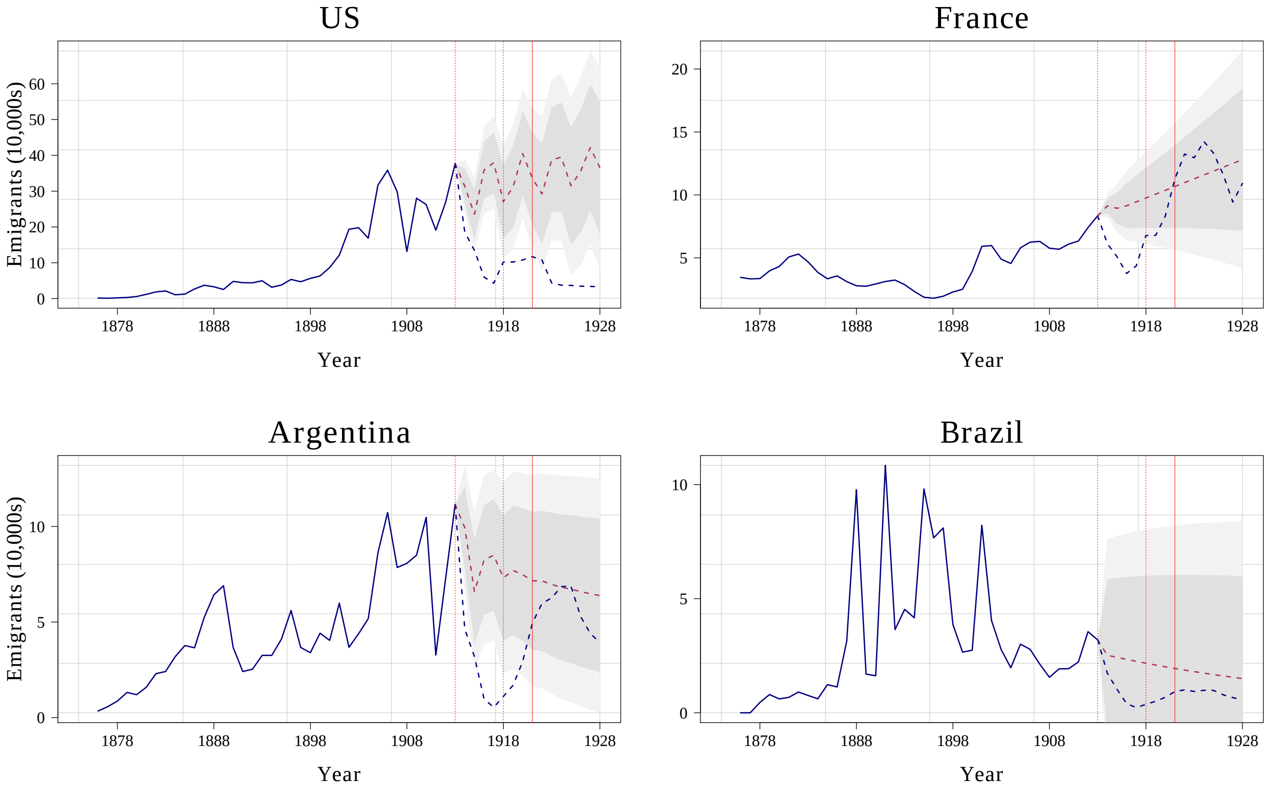


Notes. For a given outcome variable, the blue dots report the estimate of the coefficient of the treatment (QE) in the baseline difference-in-differences specification. The red bands report the 95% confidence intervals for a set of estimators for the coefficient's standard error. We include White standard errors which allow for heteroskedasticity; several clustered standard errors allowing for within-group autocorrelation; the Driscoll and Kraay (1998) correction for autocorrelation at two different time lags; several Conley (1999) estimates allowing for time and spatial autocorrelation. For the Conley SEs, we set maximal time-autocorrelation at 2 lags, and vary the radius of spatial autocorrelation.



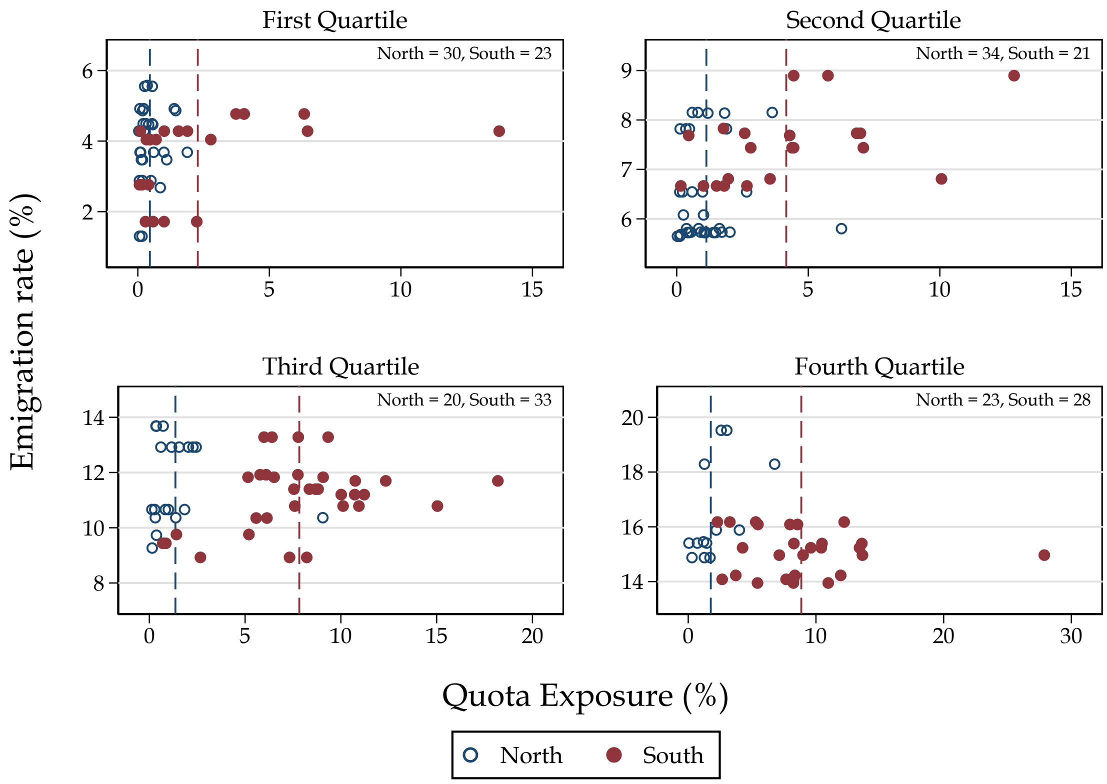
Notes. For a given outcome variable, the blue dots report the estimate of the coefficient of the treatment (QE) in the baseline difference-in-differences specification. The red bands report the 95% confidence intervals for a set of estimators for the coefficient's standard error. We include White standard errors which allow for heteroskedasticity; several clustered standard errors allowing for within-group autocorrelation; the Driscoll and Kraay (1998) correction for autocorrelation at two different time lags; several Conley (1999) estimates allowing for time and spatial autocorrelation. For the Conley SEs, we set maximal time-autocorrelation at 2 lags, and vary the radius of spatial autocorrelation.

FIGURE 2.C.4: Emigration towards main destination countries



Notes. These figures plot the number of Italian emigrants towards the main destination countries over the period 1876-1930. Overall, these countries account for about the 70% of total emigration from Italy during the whole period. The blue line represents the actual number of migrants (and its moving average starting from WWI). The red line reports the predicted number of migrants obtained from an ARIMA model estimated over the historical number of emigrants before WW1. Bands plot 95% and 80% confidence interval for the predicted values. The figures suggest that predictions based on historical emigration patterns reflect variation in the post-Quota period for all destination countries but the US.

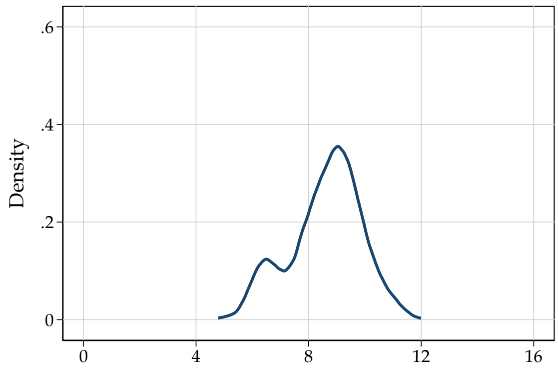
FIGURE 2.C.5: Counties by Quota Exposure and Emigration Rate's Quartile



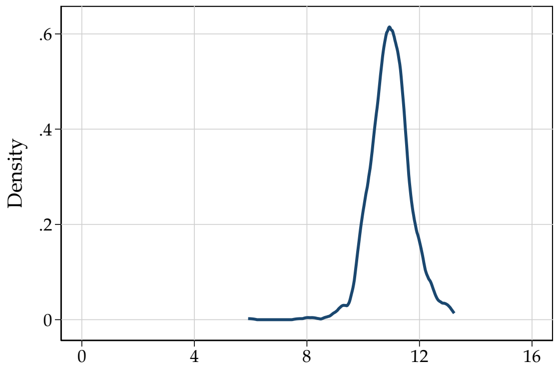
Notes. Each dot represents a district and reports its emigration rate (%, on the y-axis) and its quota exposure (%, on the x-axis). Panels are split by quartiles of the emigration rate. Blue dots are for districts in northern regions; red dots are for districts in southern regions. Red and blue vertical lines display the mean quota exposure for northern and southern regions, respectively. In each panel, on the top-right we report the number of northern and southern districts in the plot. This figure shows that conditional on the emigration rate, northern districts display substantially lower quota exposure despite sizable emigration rate. Hence, our identifying variation conditionally compares northern *vis-à-vis* southern districts, instead of exploiting within-South variation.

Figure 2.C.6: Distribution of Capital and Labor

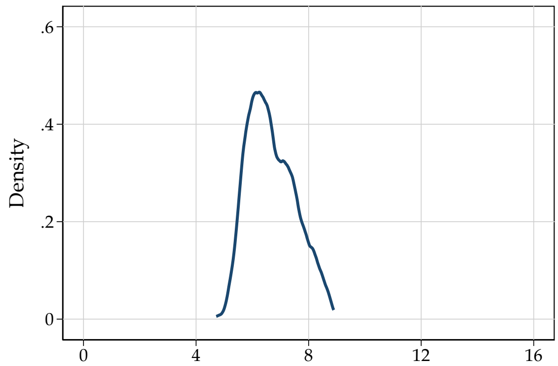
(A) Number of firms



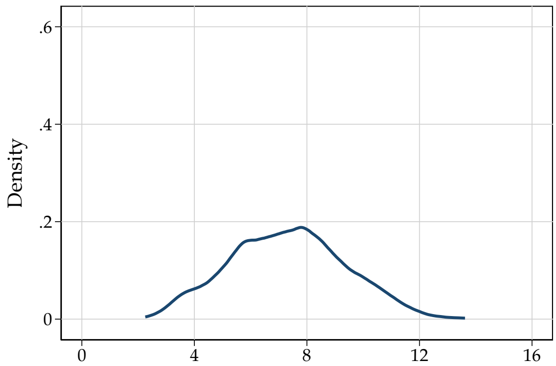
(B) Manufacture Employment



(C) Number of Mechanical Engines



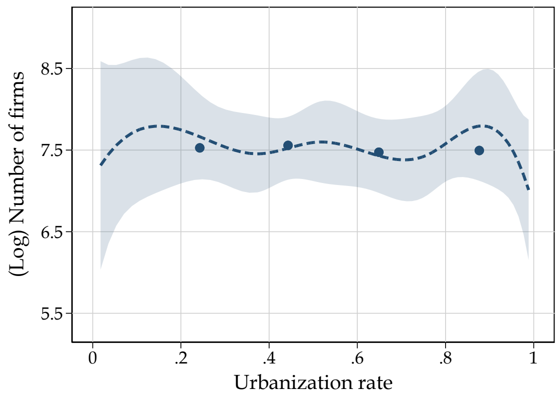
(D) Number of Electrical Engines



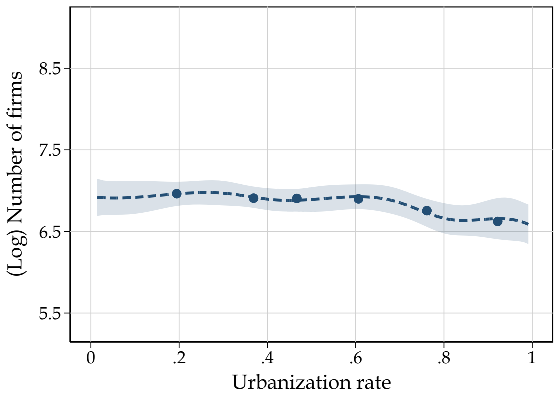
Notes. Each line represents the density plot of capital and labor variables we use throughout our analysis. Variables are expressed in logarithm. We plot the distribution over the whole sample period, 1881-1936. On the top left we show the distribution of the number of firms in each district. On the top right we show the distribution of the total number of workers. On the bottom, we show the distributions for the number of mechanical (left panel) and electrical (right panel) engines.

FIGURE 2.C.7: Number of firms and Urbanization Rate in the Pre Quota Period

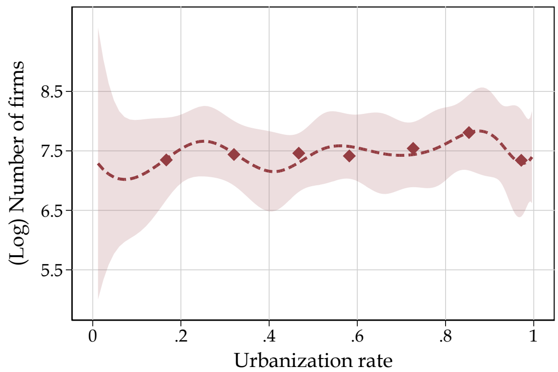
(A) No Control Below Median



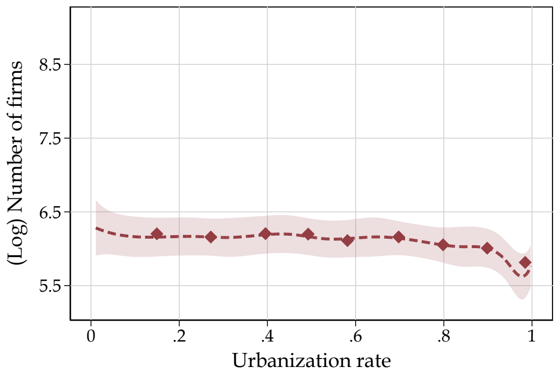
(B) With Controls Below Median



(C) No Control Above Median



(D) With Controls Above Median



Notes. The graphs display binned scatter plots relating the total number of firms (in logarithm) and the urbanization rate at district level in the pre-Quota period (before 1921). The blue lines refer to those district whose Quota Exposure is below the median. Red lines, instead, refer to district with Quota Exposure above the median. The left panels show the results of a binscatter generalized linear regression of the number of firms in a given district to its urbanization rate in the pre-Quota period. For the right panels, we also control for the emigration rate (intensive margin), and for year and province fixed effects. Dashed lines represent the cubic B-spline estimate of the regression function of interest. 95% confidence bands are based on the same spline. The plots show there is no significant difference between the correlation between number of firms and urbanization rate, by exposure to the Quotas.

2.D A Model of Directed Technical Adoption

In this section we develop a simple framework to rationalize our main findings in the context of labor-saving technical change theory. Proofs and further analytical insights on the baseline environment can be found in section 2.D.3.

2.D.1 Theoretical Framework

In this section we develop a simple analytical framework inspired to Zeira (1998) and San (2022) to clarify the empirical implications of directed technical change and adoption theory. The core assumption we make is that capital goods—hereafter, machines—substitute labor as a production input. We thus implicitly restrict technological progress to be labor-saving, differently from *e.g.* Acemoglu (2002, 2007). The decision of the firm to adopt productivity-enhancing machines will depend on their price relative to the cost of labor. In the equilibrium a labor supply shock—such as the one induced by IRPs—dampens the incentive to adopt machines because it pushes down the wage, hence prompting firms to substitute capital with labor.

Consider a closed economy with one consumption good, and a representative household supplying labor. The consumption good is produced by a continuum of tasks $j \in [0,1]$. Each task can be performed with either labor or machines. The amount of machines in task j is denoted by $x(j)$, whereas the amount of labor employed is $e(j)$. Note that each task can be fulfilled with either machines or labor, but not both. This is intended to model in a stylized manner labor-saving machines. To simplify the analysis and following Zeira (1998) we assume that machines fully depreciate at the end of the period, hence the model is essentially static.

The final consumption good is produced by identical perfectly competitive firms with the following production function:

$$Y = A\left[\int_0^{\iota} m x(j)^{\alpha}\, dj + \int_{\iota}^{1} e(j)^{\alpha}\, dj\right] \tag{2.13}$$

where A is a technology parameter, m is the relative productivity of machines and $\alpha \in (0,1)$ is a production parameter. We assume $m \in (0,1)$ following San (2022), and restrict machines to be equally productive across tasks j. The choice variable $\iota \in [0,1]$ denotes *industrialization* defined as the share of automatized tasks, which are those fulfilled by machines. We assume that tasks are ordered by degree of complexity. Because the marginal cost of producing machines—which we define below—is increasing in complexity, the price of machines is non-decreasing in j. It is therefore without loss of generality to assume that the first ι tasks are automatized. This is because the final good producer will first automatize tasks whose machine costs the least, since the relative productivity of machines is constant across tasks. We assume that there is a fixed stock of labor $L > 0$ which is supplied inelastically by the household.

The problem of the representative final good producer is therefore to choose the industrialization level ι, and input quantities $x(j)$ and $e(j)$ for each task, to maximize profits

$$\max_{\iota,\{x(j),e(j)\}_{j\in[0,1]}} Y - \int_0^{\iota} p(j)x(j)\,dj - w\int_{\iota}^{1} e(j)\,dj \tag{2.14}$$

where $p(j)$ is the price of a machine for task j, w is the nominal wage, subject to the technology constraint (2.13). Note that the price of the consumption good is implicitly normalized to one. In section 2.D.3, we formally show that the demand for machines and labor are given by the following demand schedules:

$$x(j) = p(j)^{-\frac{1}{1-\alpha}} (\alpha A m)^{\frac{1}{1-\alpha}} \quad \forall\, j \in [0,\iota] \tag{2.15a}$$

$$e(j) = w^{-\frac{1}{1-\alpha}} (\alpha A)^{\frac{1}{1-\alpha}} \qquad \forall\, j \in [\iota, 1] \tag{2.15b}$$

Combining (2.15a)-(2.15b) with the first order condition for the industrialization rate, it follows that in the equilibrium ι^* is pinned down by the following:

$$m = \left[\frac{p(\iota^*)}{w}\right]^{\alpha} \tag{2.16}$$

The economic intuition behind condition (2.16) is that at the marginal task, *i.e.* the last automatized task, the price of the machine fulfilling the task must be equal to the cost of labor, adjusted by the technology parameter and the relative productivity of machines.

Each machine is produced by a monopolist, following Zeira (1998). The machine producer will seek to set the monopoly price which maximizes its profits subject to demand for machines (2.15a). We assume that the marginal cost of machines $\psi(\cdot)$ is increasing in the complexity of tasks, *i.e.* $\psi'(\cdot) > 0$. Moreover, we assume that the marginal cost function satisfies basic Inada conditions.[52] This is intended to capture the idea that machines substituting low-skill tasks are not as expensive as those replacing tasks on the right side of the skill distribution of workers. The problem of the machine producer is therefore

$$\max_{p(j)} \left[p(j) - \psi(j)\right] x(j) \tag{2.17}$$

[52]In this setting, this simply boils down to $\lim_{j\uparrow 1} \psi(j) = +\infty$ and $\lim_{j\downarrow 0} \psi(j) = 0$. The economic intuition behind these is that it is never profitable for the representative firm to automatize all tasks. Similarly, there is always at least one task that is automatized. Note that while these assumptions are sufficient for the existence of an equilibrium, they are not necessary.

subject to (2.15a). In section 2.D.3, we show that the first-order conditions imply

$$p(j) = \min\left\{mw, \frac{\psi(j)}{\alpha}\right\} \tag{2.18}$$

where the minimum descends from the observation that because each task can be performed by labor as well as by machines, setting a price greater than the productivity-adjusted wage simply pushes the final goods producer not to automatize the task. We now obtain two technical results to ensure existence and uniqueness of the equilibrium. The formal definition of the competitive equilibrium in this economy as well as the proofs of all lemmas and propositions can be found in section 2.D.3.

Lemma 2.D.1. *In the equilibrium, the marginal task ι^* is such that $p(\iota^*) = \psi(\iota^*)/\alpha = wm^{1/\alpha}$.*

Combining this result with the equilibrium conditions of the final goods producer, we derive the following strong existence result.

Proposition 2.D.1. *There exists one and only one $\iota^* \in [0, 1]$ which solves the problem of the final good producer* (2.15a)-(2.15b)-(2.16) *as well as the problem of the machine producers* (2.18) *and verifies labor market clearing. In particular, the equilibrium industrialization ι^* is the solution to the following:*

$$\psi(\iota^*) = L^{\alpha-1}(1-\iota^*)^{1-\alpha}\alpha^2 Am^{1/\alpha}.$$

This concludes our analytical characterization of the environment. We now exploit the model to deliver a number of testable predictions which will guide our empirical analysis.

2.D.2 Empirical Testable Implications

Having established the existence of the equilibrium, we can now derive two key empirical implications of this directed technical adoption setting. First, note that Lemma 2.D.1 conveys the basic intuition of the model. In particular, we have $\psi(\iota^*) = \alpha m^{1/\alpha} w$, hence an increase in the nominal wage induces industrialization to rise because $\psi'(\cdot) > 0$ by assumption. The economic intuition behind this result is that if the cost of labor increases, then the final good producer will seek to automatize more tasks in order to avoid paying the increase in the wage. This is summarized in the following implication statement.

Implication 2.D.1. Following an exogenous increase (resp. decrease) in the nominal wage w, the share of tasks performed by machines ι^* increases (resp. decreases).

A similar comparative static result follows considering an increase in the labor stock. To see it, notice that because the nominal wage is invariant across tasks, from (2.15b) and labor market clearing the total labor stock L is evenly allocated across the $(1 - \iota^*)$ non-automated tasks. Using this insight, we obtain the following empirical prediction.

Implication 2.D.2. Following an exogenous increase (resp. decrease) in the labor supply stock L, the share of tasks performed by machines ι^* decreases (resp. increases).

This is the key implication of the model that we test in the paper. In our setting, we provide evidence that immigration restriction policies induce positive labor supply shocks, hence increasing the labor stock. We show that firms operating in districts which were more exposed to the Quota Acts decreased investment in machinery—section 2.4.2—and increased employment—-section 2.4.4. These findings are fully in line with the empirical predictions 2.D.2 of the model and hence provide evidence in favor of labor-saving directed technical adoption.

Implications 2.D.1-2.D.2 are tested using aggregate data on manufacture employment and investment in physical capital. We provide some results at a more disaggregated sector-level. We refer to relatively backward and modern sectors as respectively "First" and "Second Industrial Revolution" sectors. For concreteness, the former comprise textiles and construction whereas the latter mainly refer to the chemical and metalworking industries. To capture this difference in the model, we assume that machines in the relatively modern sector are more productive than in the relatively backward one. The following result holds.

Implication 2.D.3. Let M and L respectively denote a modern and a backward sector which differ by the productivity of machines $1 > m_M > m_B > 0$. Then, following a positive (resp. negative) labor supply shock, the share of industrialized tasks if $m = m_B$ decreases (resp. increases) more than if $m = m_M$.

We test this prediction using data on employment and technology adoption at the sector level of aggregation. We find that in First Industrial Revolution sectors investment in capital goods and employment respectively decreased and increased considerably more than in Second Industrial Revolution industries. This finding is fully consistent with prediction 2.D.3.

2.D.3 Proofs of Analytical Results

Solution of the problem of the final good producer. Plugging the technology constraint into problem (2.14), the problem of the final good producer reads out as follows:

$$\max_{\iota, \{x(j), e(j)\}_{j \in [0,1]}} A \left[\int_0^\iota m x(j)^\alpha \, dj + \int_\iota^1 e(j)^\alpha \, dj \right] - \int_0^\iota p(j) x(j) \, dj - w \int_\iota^1 e(j) \, dj$$

The—necessary and sufficient—first-order conditions with respect to labor and capital in the generic task j are

$$
\begin{aligned}
x(j) &= p(j)^{-\frac{1}{1-\alpha}}(\alpha A m)^{\frac{1}{\alpha}} \quad \forall\, j \in [0, \iota] \\
e(j) &= w^{-\frac{1}{1-\alpha}}(\alpha A)^{\frac{1}{\alpha}} \qquad \forall\, j \in [\iota, 1]
\end{aligned}
$$

To obtain the first-order condition for the optimal industrialization rate, apply the Leibniz integral rule with respect to ι to get:

$$
x(\iota^*)\left[m x(\iota^*)^{\alpha-1} - p(\iota^*)\right] = e(\iota^*)\left[e(\iota^*)^{\alpha-1} - w\right]
$$

Plugging (2.15a)-(2.15b) into the expression above we get $m = (p(\iota^*)/w)^{\alpha}$. □

Solution of the problem of the monopolist. The solution is trivial upon plugging (2.15a) into the objective function (2.17). □

Proof of Lemma 2.D.1. From (2.18) and (2.16), it is

$$
\begin{aligned}
p(\iota^*) &= \min\left\{\frac{\psi(\iota^*)}{\alpha}, mw\right\} \\
p(\iota^*) &= m^{1/\alpha} w
\end{aligned}
$$

Hence, we have

$$
m = \left[\frac{\min\left\{\frac{\psi(\iota^*)}{\alpha}, mw\right\}}{w}\right]^{\alpha}
$$

We can distinguish two cases. Assume $mw \leq \psi(\iota^*)/\alpha$. This implies that $m = m^{\alpha}$, which is only verified if $m = 1$ or $m = 0$. Since by assumption $m \in (0,1)$, this can never hold. We are left with the case $mw > \psi(\iota^*)/\alpha$. We show that this is consistent with all the parameter restrictions. Note first that since $m \in (0,1)$, it must be $\psi(\iota^*)/\alpha < w$, since otherwise it would be $m \geq 1$. We therefore have $\psi(\iota^*)/\alpha < w$ and $\psi(\iota^*)/\alpha < mw$. Because $m < 1$, the only binding constraint is $\psi(\iota^*)/\alpha < mw$. It is

$$
m = \left[\frac{\psi(\iota^*)}{\alpha} \cdot \frac{1}{w}\right]^{\alpha}
$$

which implies $\psi(\iota^*)/\alpha = w m^{1/\alpha}$. Because $m \in (0,1)$, $m^{1/\alpha} < m$ since $\alpha \in (0,1)$, and therefore $\psi(\iota^*)/\alpha = w m^{1/\alpha} < wm$. This implies that the solution is acceptable. Hence, $p(\iota^*) = \psi(\iota^*)/\alpha$ and this concludes the proof. □

Proof of Proposition 2.D.1. Because $w(j) = w$ for all $j \in [0,1]$, from (2.15b) we get that $e(j)$ does not depend on j and:

$$e(j) = e = w^{-\frac{1}{1-\alpha}}(\alpha A)^{\frac{1}{1-\alpha}} = \frac{L}{1-\iota^*}$$

where the last equality holds by labor market clearing, which requires $(1-\iota^*)e = L$. From Lemma 2.D.1, it is $w = \psi(\iota^*)/(\alpha m^{1/\alpha})$. Plugging this into the previous equation we get

$$\begin{aligned}
\left(\frac{\psi(\iota^*)}{\alpha m^{1/\alpha}}\right)^{-\frac{1}{1-\alpha}}(\alpha\beta)^{\frac{1}{1-\alpha}} &= \frac{L}{1-\iota^*}\\
\frac{\psi(\iota^*)}{\alpha m^{1/\alpha}}(\alpha\beta)^{-1} &= \left(\frac{L}{1-\iota^*}\right)^{-1+\alpha}\\
\psi(\iota^*)L^{1-\beta} &= (1-\iota^*)^{1-\alpha}\alpha^2 A m^{1/\alpha}
\end{aligned}$$

Because $\psi'(\cdot) > 0$, the left hand side is strictly increasing in ι^*. Moreover, because $\alpha \in (0,1)$, the right hand side is strictly decreasing in ι^*. By the Inada conditions, $\lim_{z\uparrow 1}\psi(z) = +\infty$ and $\lim_{z\downarrow 0}\psi(z) = 0$. If $\iota^* = 0$, the right hand side is strictly positive, whereas it is zero if $\iota^* = 1$. Hence, because both are trivially continuous, by the intermediate value theorem there exists at least one ι^* which verifies the equation. Since both are strictly monotone, ι^* is unique. □

Proof of Implication 2.D.1. From Lemma 2.D.1, it is $m^{1/\alpha} = \psi(\iota^*)/(\alpha w)$, or

$$\alpha w m^{1/\alpha} = \psi(\iota^*)$$

Because $\psi'(\cdot) > 0$, an increase in w in the equilibrium implies an increase in $\psi(\iota^*)$, hence in ι^*. □

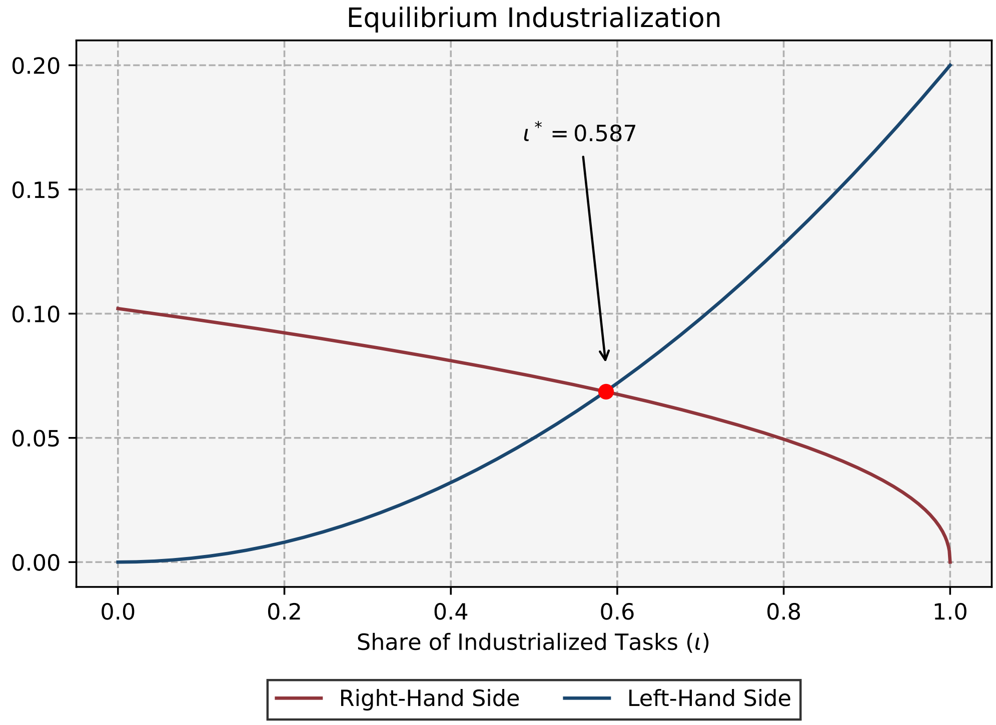
Proof of Implication 2.D.2. First note that because w is invariant across tasks, then by (2.15b) $e(j) = e$ for all j. Moreover, since the productivity of labor is constant across tasks, it is optimal to divide evenly L across the $(1-\iota^*)$ non-automatized tasks. Therefore, by labor market clearing $e = L/(1-\iota^*)$. Plug this in the left-hand side of (2.15b), yielding

$$w^{-\frac{1}{1-\alpha}}(\alpha A)^{\frac{1}{1-\alpha}} = \frac{L}{1-\iota^*}$$

Using Lemma 2.D.1 into the previous equation we get

$$\begin{aligned}
\frac{\psi(\iota^*)}{\alpha m^{1/\alpha}} &= \left(\frac{L}{1-\iota^*}\right)^{\alpha-1}\alpha A\\
L^{1-\alpha} &= \frac{(1-\iota^*)^{1-\alpha}}{\psi(\iota^*)}\alpha^2 A m^{1/\alpha}
\end{aligned}$$

FIGURE 2.D.1: Equilibrium of the Model



This figure plots the equilibrium of the model. The blue and red lines respectively display the left and right-hand side of the final equation of the proof of Proposition 2.D.1. We assume $\psi(j) = \gamma j^2$ even though quadratic costs do not verify the Inada conditions. Parametrization: $\alpha = .55$, $\beta = .45$, $\gamma = .2$, $A = .5$, $L = 1$, $m = .5$.

Because $\alpha \in (0, 1)$ and $\psi'(\cdot) > 0$, the right-hand side is decreasing in ι^*. Therefore, an exogenous increase in L leads to an increase in the right-hand side, hence a decrease in ι^*. Following an increase in the labor supply, the share of automatized tasks decreases. □

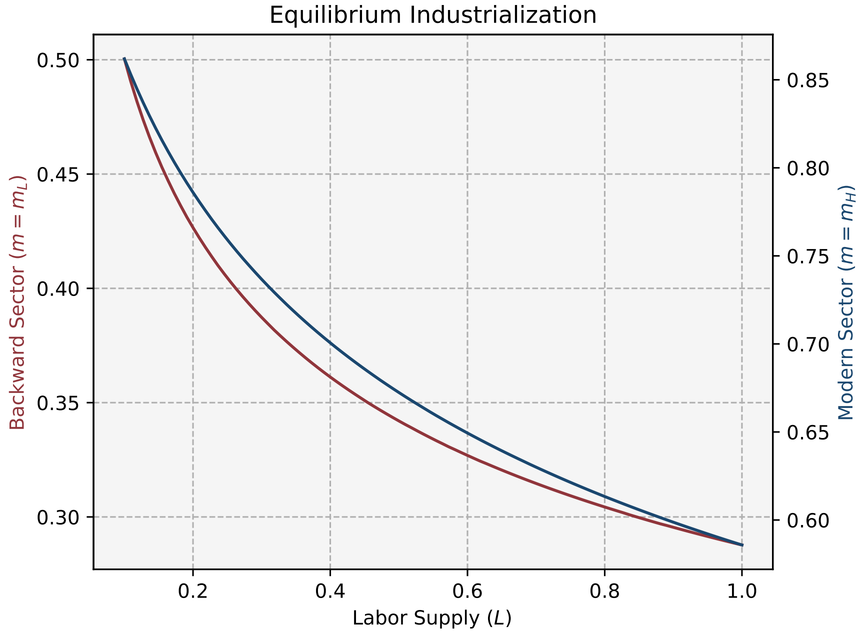
Proof of Implication 2.D.3. Let $m_M > m_B$. From the previous proof, we have

$$\frac{L^{1-\alpha}}{\alpha^2 A m_i^{1/\alpha}} = \frac{(1 - \iota^*)^{1-\alpha}}{\psi(\iota^*)}$$

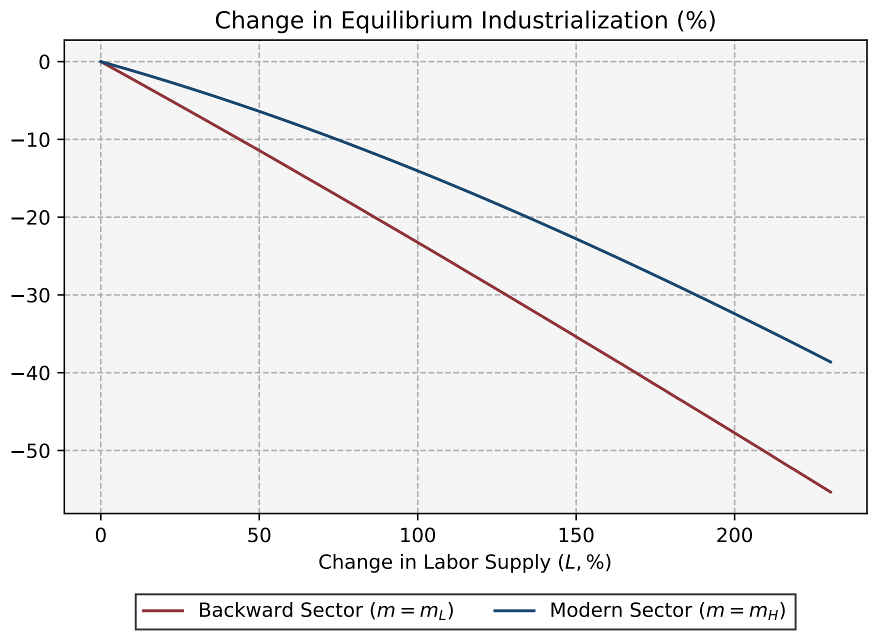
for $i = M, B$. Holding everything else constant, an increase in L translates into an increase in the left-hand side which is smaller if $m = m_M$ than under $m = m_B$ because $m_B, m_M \in (0, 1)$. Therefore, the right-hand side shall increase more under m_B. Hence, the compensating change in ι^* is larger if $m = m_B$, *i.e.* in the relatively backward sector, than if $m = m_M$, *i.e.* in the relatively modern sector. □

FIGURE 2.D.2: Model Comparative Statics

(A) Equilibrium industrialization and the labor supply.



(B) Industrialization response to changes in labor supply.



Figures plot the relationship between industrialization and the labor supply. The red and blue lines respectively display the backward and modern sectors. We assume $\psi(j) = \gamma j^2$ even though quadratic costs do not verify the Inada conditions. Parametrization: $\alpha = .55$, $\beta = .45$, $\gamma = .2$, $A = .5$, $L = 1$, $m_H = .5$, $m_L = .2$.

Bibliography

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.

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

——— (2007). "Equilibrium Bias of Technology." *Econometrica*, 75(5): 1371–1409.

Atack, J., F. Bateman and M. E. Gregson (1992). ""Matchmaker, Matchmaker, Make Me a Match" A General Personal Computer-Based Matching Program for Historical Research." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 25(2): 53–65.

Conley, T. G. (1999). "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45.

Driscoll, J. C. and A. C. Kraay (1998). "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics*, 80(4): 549–560.

Feigenbaum, J. J. (2018). "Multiple Measures of Historical Intergenerational Mobility: Iowa 1915 to 1940." *The Economic Journal*, 128(612): F446–F481.

Fontana, N., M. Manacorda, G. Russo and M. Tabellini (2021). "Emigration and Long-Run Economic Development: The Effects of the Italian Mass Migration."

San, S. (2022). "Labor Supply and Directed Technical Change: Evidence from the Termination of the Bracero Program in 1964." *American Economic Journal: Applied Economics*.

Zeira, J. (1998). "Workers, Machines, and Economic Growth." *The Quarterly Journal of Economics*, 113(4): 1091–1117.

Chapter 3

Dealing with Adversity: Religiosity or Science?*

Evidence from the Great Influenza Pandemic

3.1 Introduction

Throughout history, the occurrence of adverse events—such as natural disasters and pandemics—has posed challenges to societies worldwide and continues to do so today. Understanding how individuals cope with adverse events has key social, economic, and political implications and has been the focus of a vast literature in economics and in other social sciences. Specifically, a strand of research documents that negative shocks lead to an increase in religiosity (Bentzen, 2019). Another strand finds that economies react by boosting innovation efforts (Miao and Popp, 2014; Moscona and Sastry, 2022).[2]

In this paper, we show that these two responses can occur *simultaneously*, making societies both more religious *and* more innovative—a finding at odds with the existing evidence documenting a negative relationship between religiosity and science (e.g., Bénabou *et al.*, 2015, 2022; Lecce *et al.*, 2021). To investigate the possible mechanism behind this pattern, we study how individuals *within* society react to an adverse shock. We uncover heterogeneous responses, with religion and science acting as substitute ways through which different individuals react to adversity. These individual-level findings help reconcile our aggregate results with the existing

*This chapter is jointly authored with Enrico Berkes, Gaia Dossi, and Mara Squicciarini. We are grateful to Ran Abramitzky, Ainoa Aparicio Fenoll (discussant), Sascha Becker, Roland Bénabou, Leah Boustan, Davide Cantoni, Giacomo de Giorgi, Martin Fiszbein, Roberto Galbiati, Sergei Guriev, Reka Juhasz, Nathan Nunn, Daniele Paserman, Michele Pellizzari, Thomas Piketty, Marco Tabellini, John Van Reenen, Nico Voigtländer, Bruce Weinberg, and Katia Zhuravskaya for comments and discussions. We thank seminar audiences at the 2022 ASREC Annual Conference, the ASREC Workshop on the Economics of Religion, Bergen, Bocconi, Bolzano, Boston University, Geneva, Harvard, Louvain la Neuve, LSE, LUISS, the 6th Marco Fanno Alumni Workshop, Northwestern, Parthenope, Pavia, PSE/Sciences Po PEPES, Stanford, UC Berkeley, and the University of British Columbia for feedback and suggestions. We also thank Luca Favero, Lorenzo Pedretti, Luisa Pomarici, and Sara Veronesi for excellent assistance during the construction of the dataset, and Ran Abramitzky, Andy Ferrara, and Marco Tabellini for kindly sharing data.

[2]For example, Bentzen (2019) documents that, across countries and within regions, individuals become more religious when hit by earthquakes. Moscona and Sastry (2022), instead, finds an increase in innovation efforts towards technologies that mitigate environmental distress in U.S. counties more exposed to the Dust Bowl during the 1930s.

literature.

The setting of our study is the Great Influenza Pandemic (1918–1919) in the United States. Historical records document that many people turned to or strengthened their religious faith to cope with the pandemic. At the same time, the period following the pandemic saw an increase in innovation activity and fundamental medical advances.[3] To conduct our empirical analysis, we construct a novel data-driven measure of religiosity at a geographically disaggregated level. This measure is based on naming patterns of babies born between 1900 and 1930 from the historical full-count censuses. Complementing this dataset with information from the Census of Religious Bodies, we empirically identify religious names and construct a measure of "revealed religiosity." The underlying idea is that the first name given to a child conveys information on the religiosity of their parents. Our main metric of scientific progress is the universe of geo-coded patents granted during this period in the U.S. (Berkes, 2018).[4]

Using a difference-in-differences framework, we first show that counties hit harder by the shock experienced an increase in religiosity, an effect stronger for Catholicism. A one-standard-deviation increase in excess deaths—our main measure of intensity of the influenza shock—led to a 0.11 standard deviations increase in overall religiosity. We further document that these same counties also experienced an increase in innovative activities, an effect driven by patents granted in pharmaceuticals. A one-standard-deviation increase in excess deaths led to a 0.21 standard deviations increase in overall patenting activity. In addition, we find that employment in scientific occupations—our alternative indicator of scientific progress—grew in counties hit harder by the pandemic. This effect is mainly due to the occupational choices of young cohorts. Event-study analyses illustrate the absence of pretrends, providing further support for the validity of the research design. As a result of the contemporaneous increase of religiosity and science, their relationship turned from negative before the pandemic to positive afterward. The latter is especially puzzling, because it contrasts with the existing evidence documenting a negative relationship between the two (Bénabou *et al.*, 2015, 2022; Lecce *et al.*, 2021).

What is the mechanism behind the contemporaneous increase in religiosity and science? To answer this question, in the second part of the analysis, we study individual-level responses *within* counties. We obtain three main results.

First, we find that individuals from more religious backgrounds were more likely to turn

[3]An increase in religiosity and innovation activity has also been documented after the COVID-19 outbreak. Bentzen (2021), using Google search data, finds a sharp increase in the intensity of prayers during the early days of the pandemic. Agarwal and Gaule (2022) show that the COVID-19 pandemic catalyzed R&D expenditure on pharmaceuticals and digital technologies.

[4]We refer to science and scientific progress interchangeably, and we use two main proxies: the number of granted patents and the share of individuals in scientific occupations.

to religion in the aftermath of the pandemic, while those from less religious backgrounds were more likely to select a scientific occupation.[5] This suggests that individuals coped with negative shocks in heterogeneous ways: some turned to religion, while others turned to science. Second, we show that science-oriented individuals, who were initially less religious than the rest of the population, became even less religious after the shock. Third, we document that the pandemic widened preexisting differences in religious sentiment. Individuals from more (less) religious backgrounds became even more (less) religious. As a consequence, the distribution of religiosity in counties more exposed to the pandemic became more polarized. Importantly, the individual-level analysis reconciles the county-level findings with the existing literature. In fact, while a county may have become both more religious and more innovative, individuals seemed to react differently to the same shock—based, for instance, on their religious background or on their prepandemic scientific orientation. Religiosity and science appear to have been alternative ways of reacting to the pandemic, with individual becoming even more distant in terms of their religious sentiment than they were before the shock.

We perform several checks to gauge the robustness of our findings. First, we internally validate our measure of religiosity across several dimensions (e.g., by computing our indicator excluding firstborn babies and accounting for potential heterogeneity in fertility patterns). Second, we externally validate our data-driven measure of religiosity by using alternative indicators. In particular, we show that results are robust to using the share of biblical and saints' names, as well as the share of people affiliated with a religious denomination. In addition, to ensure that the increase in religiosity is not driven by internal or external migration, we run a placebo exercise where we test for the impact of the pandemic on the names of adults. The results show no impact of the shock on adults' names, which we interpret as evidence that the observed increase in religiosity was not driven by ex ante more-religious people moving to areas hit harder by the shock. Third, we show that the increase in patenting activity was not driven by low-quality innovations. Patent quality increased after the pandemic in exposed counties, especially in pharmaceuticals. Finally, we address the concern that other factors may be related to the pandemic and may have contemporaneously affected the evolution of religiosity and science, confounding our results. To do so, we start by documenting that initial religiosity and innovation activity are not related to the intensity of the shock. Using an event-study design, we then show that religiosity and innovation were on a similar path across treated and control groups before the shock. Additionally, we rule out that a separate yet overlapping shock—World War I—may partly explain our findings. Taken together, our empirical results,

[5] We measure religious background using individuals' own names (as opposed to their children's), aiming to capture the religious upbringing of a person instead of their current faith.

supported by historical records, provide evidence that the influenza pandemic was conceivably the main driver behind the aggregate increases in both religiosity and scientific progress.

Concerning our within-county results, one key question is why some individuals became more religious while others selected a scientific occupation. Our findings on religiosity are in line with the religious coping hypothesis, which posits that religious faith can represent a coping device to deal with personal distress following a negative shock.[6] What motivated people to turn to science is less obvious. We propose a broad interpretation of "scientific coping," with individuals turning to science either to deal with their psychological distress—as in the case of religious coping—or to try to actively mitigate the negative (e.g., health- and economic-related) effects of the pandemic.[7] While our findings cannot directly uncover the individual-level motivations behind these different behaviors—this would go beyond the scope of this paper—they show that people from different backgrounds may have reacted in different ways to the same shock and that this may have increased the polarization of religiosity within society.

Related Literature This paper is most closely related to the literature studying how societies react to negative shocks. Previous work has shown that, in accordance with the religious coping hypothesis (Pargament, 2001; Ano and Vasconcelles, 2005; Norenzayan and Hansen, 2006), natural disasters are associated with an increase in religiosity, both historically (e.g. Belloc *et al.*, 2016; Bentzen, 2019) and in contemporary scenarios (Sibley and Bulbulia, 2012; Bentzen, 2021).[8] Another set of studies documents that economic crises (Babina *et al.*, 2021), wars (Gross and Sampat, 2021), climate change (Miao and Popp, 2014; Clemens and Rogers, 2020; Moscona, 2021), and pandemics (Gross and Sampat, 2021; Agarwal and Gaule, 2022) all shape innovation activity. To the best of our knowledge, this is the first study to provide evidence that natural disasters may foster a contemporaneous increase in religiosity *and* innovation, and also the first to document the ensuing polarization of religiosity within society.[9]

Additionally, we inform the broad literature on the economics of religion, pioneered by Weber (1905). In particular, we contribute to those studies that analyze the linkage between

[6]An alternative explanation could be that individuals turn to religion as an insurance mechanism against the negative economic effects of the pandemic. While we cannot fully exclude this channel, we believe it is unlikely (as discussed in Section 3.5).

[7]Another possibility is that individuals turned to science because of increased labor demand in STEM occupations. However, the heterogeneity by religious background suggests that, beyond market forces, individual preexisting religiosity played a key role in their decision to turn to science.

[8]The religious coping hypothesis, first developed in the psychology literature, posits that people who are subject to economic and social shocks turn to religious faith as a coping device to deal with personal distress.

[9]Many studies have looked at the impact of natural disasters on, among others, social norms (e.g. Posch, 2022), migration (e.g. Boustan *et al.*, 2012), and economic activity (e.g. Boustan *et al.*, 2020).

religiosity and science.[10] While most papers adopt a historical (Deming, 2010; Mokyr, 2011), theoretical (Bénabou *et al.*, 2022), or cross-sectional perspective (Bénabou *et al.*, 2015, 2022), to our knowledge, we are the first to study the interaction between religion and science in a panel setting and to uncover the individual-level dynamics behind their coevolution.[11]

Finally, we contribute to a growing literature that exploits the informational content of names to capture individuals' characteristics. Names have been used, for example, to measure race and ethnicity (Abramitzky *et al.*, 2016; Fouka, 2020), individualism (Bazzi *et al.*, 2020), socioeconomic background (Biavaschi *et al.*, 2017; Olivetti *et al.*, 2020), and religiosity (Andersen and Bentzen, 2022). While all of these papers assume a preexisting rule to classify names (e.g., whether one has a biblical or saint name), to the best of our knowledge, we are the first to identify the religiosity of names directly from the data.[12]

The rest of the paper is structured as follows. In Section 3.2, we summarize the Great Influenza Pandemic in the United States and discuss the historical evidence concerning its effects on religiosity and innovation. In Section 3.3, we describe the data and our new indicator of religiosity. In Section 3.4, we present the empirical strategy and results. In Section 3.5, we discuss our findings. Section 3.6 concludes.

3.2 Historical Background

In this section, we provide an overview of the Great Influenza Pandemic in the United States and how it influenced religion and innovation.

3.2.1 The Great Influenza Pandemic

Between 1918 and 1919, the Great Influenza Pandemic—also known as the "Spanish Flu"[13]—killed an estimated 40 million people worldwide (approximately 1 in 30 people); it was one of the deadliest natural disasters in modern times (Barro *et al.*, 2020). In the United States, the

[10] Other studies analyze the relationship between religion and accumulation of human capital, more broadly (Becker and Woessmann, 2009; Botticini and Eckstein, 2012; Squicciarini, 2020). For an overview of the literature on the economics of religion, see Iannaccone (1998) and Iyer (2016).

[11] Lecce *et al.* (2021) study how religiosity impacts the birth and migration of scientists in 19th-century French cantons, but they do not analyze how an adverse shock affects society's dual response in terms of religion and science and the underlying individual-level dynamics.

[12] For details on how we construct our religiosity measure, see Section 3.3.

[13] The Great Influenza Pandemic is popularly known as "Spanish Flu" because media in Spain—which was neutral during World War I (WWI)—were free to report news on this disease. Conversely, countries involved in WWI imposed press censorship on the topic. This gave the (incorrect) impression that Spain was either more severely hit by the disease, or that the pandemic had originated in Spain.

pandemic started in the spring of 1918 with sporadic outbreaks. Then a second, more severe wave began in September 1918. The final wave started in January 1919, ending that spring. In total, it killed about 500,000 Americans, corresponding to 0.7% of the U.S. population (Crosby, 1989).[14]

Historical and modern accounts suggest that the pandemic hit the U.S. in a quasi-random fashion. The National Research Council stated that neither demographic characteristics, such as the ethnic composition of the population, nor geographic factors seemed to explain the difference in intensity of the pandemic across the country. Crosby (1989) writes that the states with the highest mortality displayed diverse geographical, climatic, and demographic characteristics. The pandemic hit with varying intensity within states as well. For example, in Minnesota, the death rate in Saint Paul was about 70% higher than in Minneapolis, despite the two cities being just 8 miles apart. In Ohio, Dayton experienced an 80% higher mortality rate than Columbus, even though the two cities had similar demographic characteristics (Huntington, 1923; Almond, 2006).

The infection was caused by strains of the A/H1N1 influenza virus, whose origin is still unknown. Neither antiviral drugs to treat the primary disease nor antibiotics to cure secondary bacterial infections were available. Doctors had to rely on an array of mostly ineffective—sometimes fatal—medicines such as aspirin and quinine (Spinney, 2018). It is debated whether nonpharmaceutical interventions (NPIs)—such as using masks, cancelling public events, closing schools, and implementing isolation measures and quarantines—were effective in limiting the spread of the disease.[15]

3.2.2 The Pandemic and Religion

A large literature documents that individuals become more religious in response to adverse events. One explanation for why comes from the "religious coping hypothesis," which posits that individuals turn to religious beliefs or practices as a way to cope with sudden dramatic circumstances (Pargament, 2001).[16]

[14]By comparison, COVID-19 caused 1.13 million deaths in the United States, approximately 0.3% of the U.S. population, between March 2020 and February 2023 (https://covid.cdc.gov/covid-data-tracker/#datatracker-home; accessed February 12, 2023).

[15]Some authors assert that NPIs were effective in reducing mortality (e.g., Markel *et al.*, 2007; Berkes *et al.*, 2020), while others show that the effect of NPIs on overall deaths was small and statistically insignificant (e.g., Barro, 2022).

[16]For example, Bentzen (2019) documents that individuals become more religious when hit by earthquakes. Religion may also represent an insurance mechanism when negative shocks occur: Ager *et al.* (2016) shows that after the 1927 Great Mississippi Flood, demand for social insurance led to higher churchgoing, while Ager and

The influenza pandemic inflicted substantial emotional and socioeconomic distress and could have acted as a powerful amplifier of religious sentiments (Phillips, 2020). Historical records document that spiritualism gained momentum in the aftermath of the pandemic. Not all confessions reacted in the same way. In the United States, modern evangelism benefited from the pandemic, as evidenced by a sharp rise in the circulation of evangelical magazines (Frost, 2020). Membership in Christian Science also soared during these years, reaching an all-time peak in the 1930s.[17] Catholics and Orthodox Jews identified the influenza as a manifestation of divine anger, the expiation of which called for prayers. On the other hand, some groups of progressive Protestants called for a more scientific interpretation of the pandemic (Phillips, 2020).[18] These heterogeneous responses find empirical support in our analysis discussed in Section 3.4.

3.2.3 The Pandemic and Science

Historical evidence suggests that the period after 1918 was one of sharp intellectual and scientific progress and that the Great Influenza Pandemic was particularly influential in shaping the development of medical sciences (Barry, 2020). Despite being ineffective during the pandemic, medicine evolved enormously in subsequent years. In 1928, Alexander Fleming discovered the medical use of penicillin in treating bacterial infections. By the 1930s, virology had become an established branch of medicine, and the first influenza vaccines were being developed (Spinney, 2018). During this time, medicine became more "scientific" and, hence, effective (Barro *et al.*, 2020).

These advancements in medicine went hand in hand with increased trust in scientific progress. For instance, in her personal journal, Canadian author L. M. Montgomery wrote, "[...] the Spirit of God no longer works through the church for humanity. It did once but it has worn out its instrument and dropped it. Today it is working through Science" (Montgomery, 1992[1924], p. 211). Barry (2020) argues that the pandemic was the key driver behind this paradigm shift because it fostered scientific thinking in the face of such a sudden and dramatic shock.

Ciccone (2018) document that in 19th-century United States, a larger share of the population was organized in religious communities in counties that were exposed to higher common agricultural risk.

[17] Christian Science, founded in 1879, is part of the religious movements belonging to the metaphysical family. It seeks to restore the healing and thaumaturgic virtues of primitive Christianity and has been associated with avoidance of mainstream medicine (Stark, 1998).

[18] There were also conservative Protestant churches, such as those in the Bible Belt—i.e., the region chiefly comprising Alabama, Arkansas, Georgia, Kentucky, Louisiana, Mississippi, Missouri, North Carolina, Oklahoma, South Carolina, Tennessee, and large parts of Florida and Texas—refractory to scientific and medical advancements.

This overview suggests that the 1918–1919 pandemic fostered both scientific progress and religiosity—a result that might seem at odds with theoretical and empirical evidence, which depicts religion and science as opposing forces (e.g., Bénabou *et al.*, 2015, 2022; Lecce *et al.*, 2021). In this paper, we provide causal evidence that the influenza shock led to a simultaneous increase in religiosity and scientific progress, and we reconcile this apparent puzzle by showing that it induced polarization within society, with some people turning to religion and others turning to science.

3.3 Data

To conduct our analysis, we construct a new dataset that combines information on religiosity, on innovation activity, and on the incidence of the Great Influenza Pandemic. This section describes the outcome variables and the main explanatory variables. Appendix 3.A describes the data in detail. In the first part of the analysis, counties are the geographical unit of observation.[19] In the second part of the analysis, we use individual-level data. Table 3.1 provides descriptive statistics of the main variables.

3.3.1 Religiosity Measure

The key challenge when studying religiosity is that it is difficult to measure, both today and in the past. It is especially challenging to find an indicator of religiosity that combines geographical granularity and high-frequency time variation.[20]

In our analysis, we propose a novel measure of revealed religiosity based on naming patterns of newborn babies. The motivating argument is that parents who give comparatively more religious names are more likely to be religious themselves. Therefore, naming patterns provide a measure of "revealed religiosity" of parents, rather than of the children themselves.[21]

[19]To address concerns related to counties changing their boundaries over time, we use 1920 counties as our geography of reference.

[20]This is clear in historical settings—Squicciarini (2020), for instance, uses different measures of religiosity, but these are available for only a few points in time—but it poses substantial limitations to contemporary studies as well. Recent papers leverage information from surveys such as the World Value Survey to measure religiosity (Bénabou *et al.*, 2015, 2022). Yet, because waves are typically years apart and geographically aggregated, survey-based measures are not useful for studying the dynamics of religiosity at high time frequency and fine spatial granularity.

[21]A natural corollary is that names carry informational content on the religiosity of an individual's background: while we cannot infer that an individual called "Paul" is comparatively more religious than one called "Harold," we assume that the parents of "Paul" are likely to be more religious than those of "Harold."

We now describe how we compute the religiosity score associated with first names. The key advantage of this approach is that it allows us to obtain a disaggregated yearly measure of religiosity and to study its changes in the short-to-medium term. The metric we define is conceptually similar to that developed by Andersen and Bentzen (2022) who measure, in premodern and early-modern times, revealed parental religiosity, depending on whether children were named after church-dedicated saints. Our approach differs from theirs: we *empirically* identify our religious names, using data on the entire population of newborns and existing indicators of religiosity.

3.3.1.1 Estimating Religiosity Scores for First Names

We use two main sources to compute religiosity scores. First, we construct naming patterns at the county-cohort level from the full-count U.S. censuses between 1900 and 1930 (Ruggles *et al.*, 2021). More precisely, we take the first name of all babies born between 1896 and 1930 and collapse them at the name-county-cohort level, thus obtaining a panel of name-county pairs at a yearly frequency.[22] Next, we use county-level data from the Census of Religious Bodies. This census—taken once every ten years between 1906 and 1936—allows us to construct, for every county and census-decade, the share of people affiliated with any religious denomination, as well as the share of people affiliated with a Catholic or Protestant one.[23]

To obtain the religiosity scores, we proceed in two steps. First, we compute the relative frequency of names. More precisely, let N_{cd} be the total number of individuals born in county c in decade d. We define the relative frequency of a given name (Name^k) in decade $d \in [t-10, t)$ as the ratio of all babies in that cohort called (Name^k) to the overall size of that cohort N_{cd}:

$$\text{Name Share}^k_{cd} \equiv \frac{1}{N_{cd}} \sum_{i=1}^{N_{cd}} \mathbf{1}(\text{Name}_{icd} = \text{Name}^k) \tag{3.1}$$

where $\mathbf{1}(\text{Name}_{icd} = \text{Name}^k)$ is an indicator function that returns the value one if individual i in county c born in decade d is called (Name^k), and zero otherwise. In the second step, we

[22] A cohort is defined as all babies born in a given year. The first cohort in our sample is composed of all the babies born in 1896. Our reasoning here is that the first Census of Religious Bodies was published in 1906, and we consider the ten cohorts preceding that year.

[23] To gather information on the number of religious members in each county, a report was obtained directly from local churches and congregations. The shares are computed as the number of people affiliated with these groups, normalized by the population of each county. Our analysis focuses on Catholics and Protestants, as they jointly account for more than 90% of the people enumerated by the census.

estimate the following model:

$$y_{csd} = \alpha_c + \alpha_{s\times d} + \sum_{k=1}^{K} \delta^k \times \log\left(1 + \text{Name Share}^k_{csd}\right) + \varepsilon_{csd} \tag{3.2}$$

where y denotes either the share of people affiliated to any denomination, or the share of Catholics, or the share of Protestants; d corresponds to the two prepandemic decades of the religious censuses (1906 and 1916); α_c and $\alpha_{s\times d}$ are, respectively, county and state-by-decade fixed effects.[24] The term K is the total number of names that occur in at least 0.3% of the overall sample.[25] To measure name shares, we include all babies born within ten years before each prepandemic census, hence we restrict the sample to cohorts between 1896 and 1916. Then, we aggregate these shares by decade to estimate equation (3.2).

We label the coefficient (δ^k) as the *religiosity score* associated with name k; we interpret names with larger estimated religiosity scores ($\hat{\delta}^k$) as conveying a more-intense religious sentiment. Because model (3.2) includes county fixed effects, larger religiosity scores are attached to names that become comparatively more frequent in counties that experienced larger increases in religiosity. In Figure 3.1, we report the estimated religiosity scores from model (3.2), where the outcome variable is the share of people affiliated with any religious organization. The figure shows that typically religious-sounding names, such as "Esther," "Paul," and "Grace," all feature positive and large estimated religiosity scores. Because our estimation method seeks to isolate *distinctively* religious names, relatively common ones such as "Mary" or "John" end up not having large scores. A zero-religiosity score does not imply that the name carries no religious content. In the case of "Mary," for instance, its popularity during this period is such that religious and nonreligious people alike used it, thus preventing it from being associated with distinctively religious people. Moreover, we find that names with little connection to saints or biblical episodes are associated with negative religiosity scores. This is the case for Germanic names, such as "Edith", "George," and "Harold". By considering the shares of people affiliated with Catholicism or Protestantism, we can also obtain religiosity scores for both religious denominations separately. Figure B.1 reports the results.

[24] In one of our robustness checks, we compute an alternative measure of religiosity that does not include any fixed effect. The results are robust.

[25] We follow Fouka (2020) and restrict the number of names included in model (3.2) primarily to avoid overfitting. Fouka (2020) uses a threshold of 1,000 for a name to be included in the analysis. In our preferred specification, we instead consider all names whose share in our overall sample is at least 0.3% and run checks around this threshold to assess the robustness of our results.

3.3.1.2 A Yearly County-Level Measure of Religiosity

From model (3.2), we obtain a set of estimated religiosity scores $\{\hat{\delta}^k\}_{k=1}^{K}$, which we use to construct a *yearly* indicator of religiosity at the county level. More specifically, our synthetic measure of religiosity is defined as the predicted values of model (3.2):

$$\hat{y}_{ct} = \sum_{k=1}^{K} \hat{\delta}^k \times \log\left(1 + \text{Name Share}_{ct}^k\right) \tag{3.3}$$

where t denotes a cohort between 1900 and 1930. In addition, by considering religiosity scores associated with different denominations, we can construct synthetic series for Catholic and Protestant religiosity separately.

A concern about our religiosity indicator is how much variation in county-religiosity names explain, net of that captured by fixed effects. In Appendix 3.B, we discuss a number of robustness and validation exercises for our synthetic measure. First, Figure B.2 provides county-binned scatters of synthetic and measured religiosity by denomination. The figure summarizes the results from two distinct exercises. Plots in the left column show in-sample correlations, thus comparing Census-measured and predicted religiosity in 1906 and 1916. Plots in the right column, instead, compare synthetic and measured religiosity in 1926.[26] We refer to this as an "out-of-sample" correlation, as data from the Censuses of Religious Bodies carried out after the pandemic are not used to estimate religiosity scores. All graphs show a positive correlation between actual and predicted religiosity across all denominations. This provides reassuring evidence that naming patterns capture meaningful variation in overall religiosity and further validates our measure.

One caveat of our religiosity measure is that we do not observe the religious affiliation of individuals. If we knew, for every person, their name and religion, we could infer the relative "Catholicism" of a name by measuring how frequent that name occurs within the Catholic population, relative to the overall population.[27] This is not possible using U.S. data, as the census does not contain questions about individuals' religious faith. This information is, however, available in Canadian censuses, which explicitly report the religion of every registered individual (Abramitzky *et al.*, 2020). We therefore construct alternative religiosity scores using the 1881,

[26] Our results do not change if we include data from the 1936 Census of Religious Bodies. However, growing discontent resulted in substantially lower reporting rates in this last Census for some religious groups. Following Stark (1992), we, therefore, consider it less reliable and exclude it from our analysis.

[27] As explained above, in this paper, we compute the intensity of Catholicism or Protestantism conveyed by each name by estimating model (3.2) separately for the (Share of Catholics) or the (Share of Protestants) as reported in the Census of Religious Bodies.

1911, and 1921 Canadian censuses.[28] We focus on Protestantism and Catholicism as the two major denominations in Canada and, for each name, we calculate two separate scores expressing the intensity of Catholicism and Protestantism that each name conveys.[29] In Online Appendix 3.A.6, we elaborate on how we construct this index. Additionally, following Abramitzky *et al.* (2016), we use biblical and saint names as an alternative name-based measure of religiosity.

Finally, we also use as another indicator of religiosity the county-level share of the population with a religious affiliation (for all affiliations, and separately for Catholics and Protestants) recorded by the Census of Religious Bodies for the years 1906, 1916, and 1926.

3.3.2 Measuring Scientific Production

We measure local innovative activities using patent data from the Comprehensive Universe of U.S. Patents (CUSP; Berkes, 2018). The CUSP contains information about the universe of U.S. patents issued between 1836 and 2015. The data for the time period considered in our paper (1900–1930) are extracted from digitized patent documents obtained from the U.S. Patent and Trademark Office.

For the purpose of our analysis, we first assign each patent to a county, based on the residence of its inventor, and a year, based on the year in which the patent was filed. When a patent lists multiple inventors, we give equal weights to the location of each inventor. From the CUSP, we also collect the technology classes associated with each grant (according to the U.S. Patent Classification system) and assign them to technology groupings following the crosswalk provided by the National Bureau of Economic Research (Hall *et al.*, 2001).[30]

In a second step, we build a measure of scientific inclination for a given county by looking at the share of individuals employed in STEM occupations. The underlying idea is that STEM occupations require science-based education. Thus, individuals in STEM occupations are plausibly more science-oriented than non-STEM ones. For each county and census year (1900 to 1930), we compute the share of individuals employed in a STEM occupation relative to (i) the entire

[28] Unfortunately, the 1891 and 1901 individual census records no longer exist. The 1881 census covers the universe of the Canadian population, whereas the 1911 and 1921 censuses cover a 25% sample of the population.

[29] Each score is calculated as the excess frequency a given name appears within that denomination, relative to the overall population.

[30] Whenever a patent is assigned to more than one field, we split it with equal weights across fields. We conflate the “chemical” and “drugs” NBER classes into a single class which we label “pharmaceuticals.” This is because most patents classified as “drugs” would also appear as “chemical,” since each patent is usually assigned multiple US Patent Classification codes. All results for the pharmaceutical class hold also if we consider drug and chemical patents separately. An example of pharmaceutical patent is shown in Figure B.3. For historical consistency, we relabel the “computer and communication” class as simply “communication.”

population; (ii) the number of people employed in high-skilled occupations.[31] We also use these two classifications into STEM and non-STEM occupations when performing the individual-level analysis.

3.3.3 Exposure to the Great Influenza Pandemic

To measure the incidence of the Great Influenza Pandemic across U.S. counties, we use mortality statistics assembled by the U.S. Department of Commerce. These were first collected in 1915 and, throughout the 1915–1918 period, they cover 1274 counties (40% of the total), accounting for more than 60% of the U.S. population. We follow the methodology developed by Beach *et al.* (2020) and measure mortality caused by the flu as average deaths during the flu period (1918–1919) relative to the three years before the pandemic (1915–1917). Formally, excess mortality in county c is defined as

$$\text{Excess Deaths}_c = \frac{\frac{1}{2}\sum_{t=1918}^{1919} \text{Deaths}_{ct}}{\frac{1}{3}\sum_{t'=1915}^{1917} \text{Deaths}_{ct'}} \tag{3.4}$$

This measure represents our baseline treatment. We also report results from a categorical treatment variable equal to one if the baseline treatment (Excess Deaths$_c$) is above its median, and zero otherwise. Figure 3.2 displays the geographical variation in the intensity of the pandemic in terms of excess deaths. We find that the severity of the pandemic varies substantially across counties, even geographically close ones. The rationale behind our excess-mortality measure is that—all else being equal—deaths during the pandemic that exceed those before the pandemic are likely due to the pandemic itself. A possible threat to this argument might be the U.S. involvement in WWI and that WWI deaths are confounding our results. However, this does not seem to be the case. In Figure B.4, we show that there is no significant correlation between deaths from WWI and our measure of excess deaths. In Section 3.4, we show that our results are robust to controlling for a post-1918 time indicator interacted with WWI-related deaths.

3.4 Empirical Results

In this section, we describe two main results. First, we show that exposure to the influenza pandemic led to an increase in both religiosity and innovation activity across counties. Second, we

[31] This second measure increases the comparability of the control group with STEM individuals. Table B.1 lists the set of occupations that we label as STEM (Panel A) and the occupations that we classify as high-skilled (Panel B). By construction, STEM occupations are also high-skilled. Individuals in STEM occupations represent approximately 6% of those employed in skilled professions in the 1930 census.

provide evidence of heterogeneous responses to the pandemic *within* counties. Specifically, we find that individuals from more religious backgrounds further embraced religion, while those from less religious backgrounds were more likely to choose a scientific occupation. In addition, we show that the pandemic widened the distance in religiosity between science-oriented individuals and the rest of the population, and that it led to the polarization of religiosity.

3.4.1 County-Level Evidence

In the first part of the analysis, we study the impact of the pandemic separately on religiosity and innovation at the county level. Our sample consists of a panel of U.S. counties observed over the 1900–1929 period at a yearly frequency. In particular, we leverage quasi-random variation in exposure to the pandemic across U.S. counties in a difference-in-differences (DiD) setting and estimate regression models of the form:

$$y_{ct} = \alpha_c + \alpha_t + \delta \times (\text{Post}_t \times \text{Excess Deaths}_c) + \boldsymbol{x}'_{ct}\boldsymbol{\beta} + \varepsilon_{ct} \tag{3.5}$$

where the subscripts c and t denote county and year, respectively; y_{ct} measures either religiosity or innovation activity; α_c and α_t are county and year fixed effects; Post_t is an indicator variable equal to one if $t \geq 1918$ and zero otherwise; Excess Deaths_c measures the intensity of the pandemic in terms of excess deaths, as explained in Section 3.3.3; and ε_{ct} is the error term. In addition, in all regressions we control for the interaction between 1900-population and the post indicator $\boldsymbol{x}'_{ct}$. Standard errors are clustered at the county level. Our coefficient of interest, δ, captures the impact of the pandemic on religiosity or innovation. To investigate possible heterogeneity of treatment effects over time, we also estimate a more flexible model that, rather than interacting Excess Deaths with the Post indicator, interacts Excess Deaths with biennial time dummies:[32]

$$y_{ct} = \alpha_c + \alpha_t + \sum_{\tau \in \mathcal{T}} \delta^{\tau} \left[\mathbf{1}(\tau \leq t \leq \tau + 1) \times \text{Excess Deaths}_c\right] + \boldsymbol{x}'_{ct}\boldsymbol{\beta} + \varepsilon_{ct} \tag{3.6}$$

where $\mathcal{T} = \{1912, 1914, \ldots, 1928\}$ and $\mathbf{1}(\tau \leq t \leq \tau + 1)$ is an indicator variable that takes value one if t is in the two-year window indexed by τ, and zero otherwise.

Did the influenza spread randomly? We perform three main exercises to test this in the data. First, in Table B.2, we report the correlation between the intensity of the pandemic and

[32]In the dynamic DiD specifications, we code time periods over two-year windows to reduce noisy fluctuations in estimated treatment effects and to improve efficiency by pooling observations.

a set of census-measured county covariates measured in 1910, the last census before the pandemic, accounting for population and state-level fixed effects.[33] Counties more exposed to the pandemic are observationally equivalent with respect to all variables, except for the share of men, and the share of foreigners. This is in line with the pandemic being comparatively more severe in urban areas and for men. Then, to rule out that these differences confound our analysis, we check whether control and treatment counties were on different trends before the shock, and we estimate an event study.

Formally, in Equation (3.6), this implies that the estimates of δ^{τ} would not be statistically different from zero before the pandemic hit,[34] i.e., for all $\tau < 1918$. We find support for the parallel-trends assumption. However, our approach could still be invalid in the presence of shocks correlated with the intensity of the pandemic that positively affected both science and religiosity but that were *not* caused by the pandemic itself. A plausible candidate is the number of soldiers that counties lost in WWI: their deaths might have driven either the religiosity of their families or the ability (or motivation) of a county to produce innovation (or both). To test for this, in Tables B.3 and B.11, we control for the number of deaths in WWI in our regression model and show that the results remain robust.

3.4.1.1 The Effect of the Influenza Pandemic on Religiosity

Table 3.2 displays the DiD estimates obtained using religiosity as dependent variable. In columns (1–3), the dependent variable is the share of individuals affiliated with the religious denominations enumerated in the Census of Religious Bodies, normalized by county population in 1900.[35] In columns (4–6), instead, we use the name-based measure of religiosity described in Section 3.3.1, which allows us to observe counties every year between 1900 and 1929. The estimates reported in columns (1) and (4) show that counties comparatively more exposed to the pandemic experienced an increase in overall religiosity. A one-standard deviation increase in excess deaths led to a 0.11 standard deviations increase in name-based religiosity at the county

[33]State fixed effects control for the fact that the pandemic spread from East to West between August 1918 and November 1918.

[34]Since the setting is not staggered—because the pandemic hit each county in the same period—models (3.5) and (3.6) can be estimated through standard two-way fixed effects (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021). Callaway *et al.* (2021), however, caution against using continuous treatments. We code a binary indicator equal to one for counties with above-median excess deaths. Throughout the paper, we show that the continuous and binary treatments yield qualitatively similar results.

[35]This has the advantage of including the U.S. population across different age groups—not just individuals who had children a decade before and a decade after the pandemic. On the other hand, this measure has two caveats: (i) census-based religiosity is available only at three points in time (1906, 1916, and 1926) and thus does not allow us to study high-frequency variation in religiosity; (ii) the choice to join a religious denomination could be more likely to be affected by social insurance considerations (rather than by religious reasons), thus inducing an upward bias in our results.

level. Similarly, moving from a county at the 25th percentile of the excess mortality distribution to one at the 75th percentile led to an increase in religiosity of 7%. In columns (2) and (5) and (3) and (6), we explore possible heterogeneous effects of the pandemic on, respectively, Catholics and Protestants. We estimate comparable treatment effects for these two denominations.

In Figure 3.3, we report the coefficients of the interactions between the treatment variable and the biennial dummies for overall religiosity. The flexible specification supports the patterns observed in the DiD analysis and confirms the absence of pretrends. In addition, we observe that the increase in religiosity seems to persist over the decade after the pandemic. This is in line with the literature documenting a substantial persistence of religiosity (e.g., Squicciarini, 2020).

In Table B.3, we show that our results hold through a series of robustness checks. First, we code the treatment as a binary variable equal to one if the baseline treatment is above its median, and zero otherwise (column 2). Second, we explicitly control for mortality due to WW1 in column (3). One concern related to our religiosity measure could be that firstborns are often named after a deceased grandparent and thus their names reflect the higher religiosity of previous generations rather than their parents' religious attitudes. If, due to higher mortality, households in areas more affected by the pandemic were also more likely to have recently lost a grandparent, then our results might simply reflect a mechanical effect. Column (4) reports estimates dropping firstborn children in every household. Another concern is that numerous households may display different naming behaviors for later-born children. In column (5) we drop children beyond the fourth. In addition, if religious families displayed higher fertility rates, one may worry that our results are driven by an increase in the number of religious names due to the higher fertility of already-religious households. In column (6) we compute within-household average religiosity to check whether our findings are driven by larger households and differential fertility. All results hold through these alternative specifications. Finally, another concern could be that comparatively more religious people moved into counties where the pandemic had been more severe, perhaps motivated by slacker labor markets. If that were the case, our estimated effect of the pandemic on name-based religiosity would reflect movers' religiosity and their fertility. To deal with this concern, we compute a county-decade measure of religiosity based on the names of the adult population only. The in-migration mechanism would predict a positive impact of the pandemic on this variable. Estimates reported in column (7) show no evidence of any such effect, thus ruling out this potential alternative interpretation. Lastly, one may be worried that the results presented in columns (1–3) of Table 3.2 were driven by small counties, where the variation in both the share of affiliated to religious denominations and in naming patterns may have been more substantial. In table B.9, we weight counties by their population in 1900 and confirm the baseline results.

In a second step, we test whether the results are robust to alternative ways of constructing our religiosity measure. First, in Table B.4 we report the baseline result, but using religiosity scores estimated through equation (3.2) *without* county fixed effects. These scores are thus obtained using the "stock" of religiosity in a given county, instead of its deviations from the mean. The results from this alternative strategy are consistent with our baseline estimates. Second, we test the robustness of our results to the number of names included in the sample. In our baseline analysis, we exclude names appearing in less than 0.3% of the overall population. In Table B.5, we show that our findings are qualitatively unchanged under different frequency thresholds. Finally, a possible concern could be that the results are capturing a "fashion" effect, whereby more-religious names became more fashionable after the pandemic. If this were the case, even though the initial increase in religious names would suggest a positive shift in religiosity, the effect for the following periods would be biased upwards and driven by this fashion effect. In Table B.6, we regress a set of indices reflecting the concentration of the name distribution against our baseline treatment and find no evidence of such mechanism.

In the third set of robustness checks, we perform our analysis using alternative indicators of religiosity. First, we validate the distinction between Catholic and Protestant names by using the Canadian census. The advantage of this census is that, unlike in the United States, individuals were explicitly asked to report their religious affiliation. Columns (1) and (2) of Table B.8 replicate the baseline results using the Canada-based religiosity scores assigned to the names of newborns in the United States—these confirm the increase in the intensity of Catholicism in counties that were more exposed to the pandemic. Since the near-universe of the Canadian population in this period reported being Catholic or Protestant, religiosity scores can measure only the intensity of Catholicism relative to Protestantism, and vice versa.[36] Second, in columns (3)–(5) of Table B.8, we use biblical and saint names as an alternative name-based measure of religiosity, following Abramitzky *et al.* (2016). We find that the pandemic exerted a positive impact on the share of either biblical or saint names. Interestingly, this effect is stronger for saints' names—a result in line with previous findings on Catholicism and Protestantism.[37]

[36]In the Canadian census, fewer than 1% report either a religious affiliation different than Catholic or Protestant or no religious affiliation at all. For details on the construction of the Canadian-census religiosity scores, see Appendix Section 3.A.6.

[37]Perl and Wiggins (2004) argue that historically Catholic parents tended to give newborns the name of a saint, required for the child's baptism. Conversely, Protestants—who stress the centrality of the Bible but do not recognize the cult of saints—tended to give biblical names. In addition, in Figure B.5, we show that the county-level share of Biblical and Saints names, computed using data from Abramitzky *et al.* (2016), is strongly and positively correlated with the religiosity measure constructed using our data-driven approach.

Finally, we explore the effect of the pandemic within urban areas using the city-level sample constructed by Clay *et al.* (2019) and detailed in Appendix 3.A.9. In columns (1–3) of Table B.7 we estimate model (3.6) using a balanced panel of cities observed over 1900–1929. The name-based religiosity measure is computed leveraging variation in naming patterns of children born in each city. The city-level results are consistent with the baseline county-level analysis. Religiosity increased in cities that were more severely affected by the pandemic. We estimate a larger treatment effect for Catholics. This exercise ensures that our results are not driven by individuals residing in rural areas. Moreover, the city-level sample includes several cities in Southern states, which were plausibly more religious.[38] We thus view the city-level exercise as shedding additional internal validity to the county-level analysis.

Throughout different specifications and indicators, we find that the pandemic had a positive effect on religiosity. This finding is consistent with the religious coping hypothesis, which posits that religion may serve as a coping device to deal with mental and psychological distress (e.g., Pargament, 2001; Bentzen, 2019, 2021). In addition, the heterogeneity between Catholics and Protestants is in line with the psychology literature studying the impact of mental distress across confessions (Pargament, 2002), as well as with a recent study on the COVID-19 pandemic, showing that the increase in Google searches for Catholic prayers was substantially higher than for Protestant ones (Bentzen, 2021).

3.4.1.2 The Effect of the Influenza Pandemic on Innovation

We now turn to study how the influenza pandemic impacted innovation. We show that the pandemic had a positive impact on overall innovation (measured by the total number of patents granted during the period), driven mainly by an increase in patents in pharmaceuticals.

In column (1) of Table 3.3, we report the estimated impact of the influenza shock on the volume of innovation—measured as the log(1 + number of patents) in a given county-year. We find that a one standard deviation increase in excess deaths led to an 0.21 standard deviations increase in the number of patents. Similarly, moving from a county at the 25th percentile to one at the 75th percentile of the excess-deaths distribution leads on average to an increase of 19% in the number of patents granted by county-year. Figure 3.4 displays the effect in an event-study framework. Each dot in the plot reports the dynamic treatment effect of the pandemic on innovation in the indicated two-year window, as specified in model (3.6). The coefficients show that the number of patents granted after the pandemic increased significantly more in

[38]Figure B.7 reports the number of cities included in the sample by state and their location.

more-exposed counties, implying that the pandemic induced a sizable increase in innovative activities that persisted for at least one decade after the shock.

We also investigate heterogeneous effects of the pandemic across technology classes. Specifically, we ask whether the influenza shock affected not only the volume but also the *direction* of innovation. To do so, we study the effect of the shock on the number of patents in each sector, controlling for the total number of patents filed in each county-year. Columns (2)–(6) of Table 3.3 show the results of this exercise. For each field, we report the estimated DiD coefficients. We find that the influenza shock has a positive and statistically significant effect only on pharmaceutical patents. Keeping the total number of patents constant, a county at the 75th percentile of the excess-deaths distribution saw an average increase of 11% in pharmaceutical patents, compared to one at the 25th percentile.

In Table B.11, we report a number of robustness checks, separately for the total number of patents irrespective of their field (columns 1–4) and for those in pharmaceuticals (columns 5–9). Columns (1) and (5) report the baseline estimates for comparison. In columns (2) and (7), we restrict the sample to an unbalanced county-year panel that includes only county-years with at least one filed patent. Columns (3) and (8) report results coding the treatment as a binary variable. Columns (4) and (9) control for WWI deaths interacted with the posttreatment indicator and confirm that WWI-related mortality is not driving our result. Column (6) omits the total number of patents as a control, thus reporting the impact of the pandemic on the volume of pharmaceutical patents. The corresponding coefficient should be interpreted as the impact of the pandemic on the total number of pharmaceutical patents. The estimated DiD coefficients remain positive and statistically significant throughout. An additional concern is that our results were driven by small counties, where innovation activity was comparatively low during the pandemic. If this was the case, the actual increase in patenting activity would be more modest than what our estimates would suggest. In columns (4–6) of Table B.10 we weigh counties by their 1900 population and find that the magnitude of the estimated treatment effect is comparable to the baseline, unweighted model.

In the baseline specifications, we take the logarithm of the number of patents, and we add one to avoid dropping zeros. In Tables B.12 and B.13, we show that alternative transformations of the dependent variable yield quantitatively similar results, respectively for all patents and for patents in pharmaceuticals. In particular, while in the baseline regressions we control for the total number of patents—to show that the influenza shock altered the direction of innovation in favor of pharmaceuticals—in columns (7) and (8) of Table B.13, we use the share of patents in pharmaceuticals as our dependent variable. These exercises yield consistent results.

In Table B.15 we show that the positive impact of the influenza shock on innovation was driven both by the higher productivity of existing inventors (intensive margin) and by an increase in the number of inventors (extensive margin).[39] In Table B.15, the dependent variable is the number of patents per inventor (columns 1–3) and the total number of inventors (columns 4–6). We document a large increase in the productivity and number of inventors active in any field (columns 1 and 4), as well as in pharmaceuticals only (columns 2–3, 5–6), even when controlling for average productivity and for the overall number of active inventors.

A plausible concern is that our results may be driven by "low-quality" innovation. Newspapers of the day often advocated remedies for influenza that were not science- or evidence-based, some of which may have been granted a patent in subsequent years. To address this concern, we use the text-based measures of "importance" developed by Kelly *et al.* (2021).[40] Table B.16 shows the results. Column (1) uses the measure of average patent importance in all sectors and shows no significant effect of the pandemic. In column (4), instead, we find that the average importance of pharmaceuticals patents increases. We then focus on "breakthrough" patents. Specifically, we assign to every patent a dummy equal to one if the patent's importance is in the top 20% of the distribution, and zero otherwise. We find that the number and share of breakthrough patents substantially increase in counties hit harder by the pandemic, both in all sectors and in only pharmaceuticals (columns 2–3, 5, and 7). In addition, in column (6) we show that the number of breakthrough patents in pharmaceuticals grows more than the average number in other sectors.

The increase in patenting activity in counties more exposed to the pandemic may reflect demand- or supply-side factors. While we cannot exhaustively distinguish between the two, the data allow exploring one possibly informative margin. Specifically, we distinguish between patents with an assignee from those whose owners are the inventors themselves (approximately 40% of the total). The underlying idea is that patents produced by inventors employed in firms are more likely to reflect demand-driven innovation because firms would be faster to respond to changing market conditions. On the other hand, patents produced by self-employed inventors may reflect their inventors' personal motivation and experience. In Table B.14, we investigate the heterogeneous response to the pandemic of patents with and without an assignee. In columns (1–2), the dependent variable is the share of patents with and without an assignee; in columns (3–4), we divide pharmaceutical patents by the total number of patents; in columns

[39] To disambiguate among homonym inventors, we use the sample of inventors linked to the U.S. full-count census developed by Bazzi *et al.* (2022).

[40] As discussed by Berkes (2018) and Andrews (2021), citation-based quality measures during this period are noisy and mostly uninformative due to the lack of a mandatory reference section until 1947. The measure built by Kelly *et al.* (2021) identifies important patents based on the textual similarity of a given patent to previous and subsequent work. Important patents are those that are distinct from previous work, but are similar to subsequent innovations.

(5–6), we divide pharmaceutical patents by all pharmaceutical patents. We find that the pandemic exerted a statistically significant, positive, and quantitatively large impact on patents without an assignee. By contrast, we find no such effect on patents with an assignee. These results provide suggestive, albeit not exhaustive, evidence that the effect of the pandemic on innovation may have operated primarily by influencing its supply side.

While most innovation activity clusters in urban areas, we perform our baseline analysis at the level of counties. To ensure that the results do not conflate rural-urban disparities, we estimate the effect of the pandemic on innovation at the city level. Columns (4–6) of Table B.7 report the estimates of model (3.2) for the panel of cities described in Section 3.A.9. The results confirm the county-level evidence: despite the smaller sample size, we estimate a positive and statistically significant effect of the pandemic on innovation, especially in pharmaceuticals.

Another concern is that patents may be an imperfect measure for innovation and scientific attitudes, since not all innovation is patented (Moser, 2005). We complement our analysis by using the share of people employed in STEM occupations as an alternative indicator. The rationale for this measure is that a STEM occupation requires that an individual receive a science-oriented education. In turn, receiving a science-based education plausibly correlates with more-favorable attitudes toward, and more trust in, science (Deming and Noray, 2018; Bianchi and Giorcelli, 2020).

We start by running the same specification as in models (3.5) and (3.6) using as dependent variables the share of individuals employed in STEM relative to the overall population. We perform the analysis at the decade level, since this measure is taken from population censuses (1900–1930). Column (1) of Table 3.4 shows an increase in the share of workers in STEM occupations in counties more severely hit by the pandemic. A one standard deviation increase in excess deaths is associated with a 0.86-standard deviations increase in the share of individuals in scientific occupations. Equivalently, moving from the 25th to the 75th percentile of the excess-mortality distribution leads to a 29% increase in the share of individuals in STEM.[41] Column (4) replicates this result, focusing on the skilled sub-sample of the population.

To better understand what drives the change in occupational shares, we use individual-level data on occupations. Specifically, we test whether individuals who were young at the time of the shock, i.e., between 18 and 25 years old, were more likely to be employed in a STEM occupation ten years later compared to older cohorts, in areas that were comparatively more

[41] These coefficients are computed using decade-level data. This explains why the beta coefficients are particularly high, compared to those obtained using yearly-level data, as in Table 3.3.

exposed to the pandemic.[42] We estimate the following linear probability model, where we define as treated individuals aged 18 to 25 in 1918:

$$\text{STEM}_{hct} = \alpha_c + \alpha_t + \delta \times \left(\text{Excess Deaths}_c \times \text{Young}_h\right) + \boldsymbol{x}'_h \boldsymbol{\beta} + \varepsilon_{hct} \tag{3.7}$$

where α_c and α_t respectively denote county and cohort fixed effects, STEM_{ict} is a dummy variable equal to one if head of household h is employed in a STEM occupation, and zero otherwise; $\boldsymbol{x}_h$ includes urban status and race. The categorical variable Young_h is equal to one if individual h is between 18 and 25 in 1918, and zero otherwise. Our coefficient of interest is δ, which measures the causal effect of the pandemic on the probability of being employed in a STEM occupation.

Columns (2)–(3) in Table 3.4 report the results: in counties more exposed to the pandemic, young individuals were significantly more likely to sort into STEM occupations.[43] Why did young cohorts respond disproportionately more to the shock? We have two potential explanations for this finding. The first is mechanical: the pandemic may have affected everyone in similar ways, but young cohorts were the only ones in the process of choosing an occupation. The second is that the pandemic may have specifically affected the attitudes and preferences of individuals in their *impressionable years* (i.e., the young cohorts), and thus the differential occupation choice reflects a change in attitudes occurring only for these cohorts.[44] Next, we replicate the analysis of columns (1)–(3), using non-STEM high-skilled individuals as the comparison group. In particular, in column (4) we use as dependent variable the share of STEM individuals relative to the number of people employed in high-skilled occupations, and in columns (5)-(6), we only include individuals in STEM and other high-skilled in the sample. The results are quantitatively similar.

3.4.1.3 Joint Dynamics of Religiosity and Innovation

After studying the impact of the pandemic separately on religiosity and scientific progress, we now look at their joint evolution. Specifically, we test whether the *same* counties were affected

[42] To construct the sample, we use the cross-section of all individuals in the 1930 full-count census. We drop all individuals born after 1905, as they may have been too young to have already selected an occupation, and we restrict the sample to the working population. We drop individuals who were in prison, retired, or reported no occupation.

[43] In the baseline specification, a young individual is someone between 18 and 25 years old in 1918; in Table B.17, we extend the sample to those aged 18 to 30 in 1918, and the results hold.

[44] According to the "impressionable years" hypothesis—which represents a long-standing argument in psychology—economic, social, and cultural attitudes and beliefs are formed during early adulthood, approximately between the ages of 18 and 25, and change slowly thereafter (Giuliano and Spilimbergo, 2023).

along both dimensions, or whether some counties saw an increase in religiosity while others saw an increase in scientific progress.

We estimate the following model:

$$\begin{aligned} y_{ct} = &\alpha_c + \alpha_t + \delta_1 \times (\text{Excess Deaths}_c \times \text{Post}_t) + \delta_2 \times \text{Religiosity}_{ct} + \\ &+ \delta_3 \times \left(\text{Excess Deaths}_c \times \text{Post}_t \times \text{Religiosity}_{ct}\right) + \boldsymbol{x}'_{ct}\boldsymbol{\beta} + \varepsilon_{ct} \end{aligned} \tag{3.8}$$

where y_{ct} is the log(1+ total patents),[45] and (Religiosity$_{ct}$) is the religiosity measure described in Section 3.3.1. The coefficient δ_1 measures the impact of the pandemic on innovation, δ_2 captures the correlation between the outcome and religiosity before the pandemic, and the term δ_3—alongside δ_2—captures how the correlation between the outcome and religiosity changes after 1918 as a function of exposure to the pandemic. As before, the vector $\boldsymbol{x}_{ct}$ includes an interaction term between county population in 1900 and a posttreatment indicator.

In Table B.18, we report the estimates of model (3.8). The results suggest that counties that were comparatively more affected by the pandemic experienced a joint increase in religiosity and innovation.

Interestingly, as a consequence of this contemporaneous increase in religiosity and science, their relationship shifts from negative to positive—as shown in Figure B.6. In the period before the shock, there was a negative correlation between the intensity of innovation activity (measured as the number of patents per 10,000 individuals) and religiosity at the county level. This is in line with contemporary evidence reported by Bénabou *et al.* (2015). In the lower panel, we show that, in the period after the pandemic, religiosity and science became positively correlated. In Section 3.4.2, we use individual-level data to uncover the possible mechanisms underlying these results.

3.4.2 Mechanisms: Individual-Level Analysis

After observing a contemporaneous increase in religiosity and innovation, two questions naturally arise. Within counties, who turns to religion and who turns to science? Are these the same or different individuals? In this section, we leverage individual-level data to answer these questions. In particular, we focus on individuals who are heads of household in the 1930 census.[46]

[45] Total patents are normalized by county population in 1900, as in Bénabou *et al.* (2022).

[46] The "head of household" variable is provided by the census. During this period, the father and/or husband was usually the head of the household whenever he was present.

First, we show that the pandemic led to an increase in the religiosity of individuals who came from initially more religious backgrounds while individuals from less religious backgrounds were more likely to select STEM occupations. Second, we show that STEM individuals, who were less religious before the pandemic, become even less religious compared to the rest of the population. Third, we document that the pandemic led to the polarization of religiosity.

Taken together, these three results suggest that the pandemic shock led to different reactions within society—based, for instance, on individuals' religious background or initial scientific orientation—with people becoming even more distant in terms of their religiosity than they were before the pandemic. This within-county analysis reveals important heterogeneities in how individuals react to a negative shock, and it helps reconcile our aggregate findings with the existing literature on the negative relationship between religion and scientific progress.

3.4.2.1 Turning to Religion or Turning to Science

We start by studying whether preexisting differences in individuals' religious background could have led to a heterogeneous response to the influenza shock. The full-count census data, on top of covering the universe of the U.S. population, have the advantage of being deanonymized. This allows us to construct two measures of religiosity for each individual: one is their revealed religiosity, based on the names individuals gave to their children; the other is their religious background, based on their own name. Specifically, we interpret an individual's own name as conveying information about the religiosity of their parents and, thus, the religious background of that individual.

Combining these measures, we first study how an individual's religious background shaped their response to the pandemic in terms of religiosity. Next, we explore whether, following the pandemic, the religious background of an individual may have also shaped their propensity to work in a scientific occupation. To measure this, we use an indicator equal to one if they were employed in a STEM occupation, and zero otherwise.[47]

We estimate two triple-differences specifications, one for religiosity and one for the likelihood of selecting a STEM occupation:

$$\begin{aligned}\text{Religiosity}_{jict} &= \alpha_{c\times t} + \alpha_{c\times B} + \alpha_{B\times t} + \\ &+ \delta_1 \times \left(\text{Excess Deaths}_c \times \text{Post}_t \times \text{High Religious Background}_i\right) + \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_{jict}\end{aligned} \tag{3.9}$$

[47]A natural way to construct a measure of scientific background, symmetric to the religiosity one, would be to look at whether individuals had a parent working in a scientific occupation. Unfortunately, due to data limitations, this is not possible, as this exercise would require tracking individuals across several census waves, thus greatly reducing our sample size. The advantage of our measure of religious background is that it can be constructed for every individual without requiring any direct information on, or linking to, their parents.

where j represents a child, i denotes the household head, c and t are respectively county of residence and child birth-year; and

$$\begin{aligned} \text{STEM}_{ict} &= \alpha_{c\times t} + \alpha_{c\times B} + \alpha_{B\times t} + \\ &\delta_2 \times \left(\text{Excess Deaths}_c \times \text{Young}_t \times \text{High Religious Background}_i\right) + \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_{ict} \end{aligned} \tag{3.10}$$

where i denotes a head of household, residing in county c, born in year t.

In both equations, the terms $\alpha_{c\times t}$, $\alpha_{c\times B}$, and $\alpha_{B\times t}$ denote, respectively, county-by-year, religious-background-by-county, and religious-background-by-year fixed effects, and $\boldsymbol{x}_i$ includes urban status and race of the household head. The term "High Religious Background" is a categorical variable returning the value one if the religiosity score of the household head's name is in the top 20% of the religiosity distribution, and zero otherwise. We estimate model (3.9) on the sample of children born between 1900 and 1929. The dependent variable is the religiosity score associated with the name of child j. Children are weighted by the inverse of the number of children in each household. In model (3.10), the sample is composed of heads of households, observed once in the 1930 census. The dependent variable is an indicator variable returning the value one if the head of household i is employed in a STEM occupation in 1930, and zero otherwise. The coefficients δ_1 and δ_2 quantify the effect of county-level exposure to the pandemic, comparing individuals in the top quintile of the background religiosity distribution with the rest of the population on, respectively, religiosity and STEM employment.[48]

Table 3.5 presents the results of the analysis. In columns (1)–(3), the dependent variable is revealed religiosity. Our variable of interest is the interaction between the excess-deaths measure, a dummy "Post" equal to one if a child is born after the pandemic, and the religious background of the household head. In columns (4)–(6), the outcome variable is a dummy equal to one if the household head is employed in a STEM occupation. Our main variable of interest is the interaction between the excess-deaths measure, a dummy "Young" equal to one if a given individual was between 18 and 25 in 1918, and their religious background. [49] All regressions include year-by-county fixed effects, which also absorb the effects of the interaction between excess deaths and the birth year, as well as county-by-background and background-by-year fixed effects.

We find that individuals originating from more religious backgrounds, who were already

---

[48]While in model (3.9) the treatment is at the level of the birth year of the children of the household head (i.e., *Post* refers to a child born after 1918), in model (3.10) the treatment is at the level of the cohort of the household head (i.e., *Young* refers to a household head who turned 25 years old after 1918).

[49]In Table 3.5 columns (1)–(3) we observe multiple realizations—one for each child—of a head of household's religious attitude. In columns (4)–(6), on the other hand, we observe a cross-section of individuals whose scientific attitudes—which we proxy with their occupational choices—are observed only once.

more religious before the influenza shock, became even more religious afterward in more-exposed counties (columns 1–3).[50] By contrast, individuals who were young during the pandemic and came from less religious backgrounds were more likely to choose a scientific occupation (columns 4–6). Evidence in Table 3.5 suggests that an individual's religious background affected their way of dealing with negative shocks. In particular, those who were raised by religious parents were more likely to resort to religion to deal with adversity. On the other hand, growing up in a less religious household made individuals more likely to approach science, possibly as a coping device in the face of the negative shock.

3.4.2.2 Science-Oriented Individuals Became Less Religious

In this part of the analysis, we focus on science-oriented individuals and study whether their religiosity changed after the pandemic, compared to the rest of the population.

In Appendix Table B.19, we show the average religiosity of STEM (column 2) and non-STEM (column 1) individuals before the pandemic, as well as their differences (columns 3–4).[51] STEM individuals are less religious than non-STEM ones. This holds both unconditionally (column 3), and when we condition on county fixed effects, cohort fixed effects, and household-level controls (column 4).

We now turn to study the impact of the pandemic on religiosity for these two types of individuals. We estimate the following triple-differences model:

$$\text{Religiosity}_{jict} = \alpha_{c\times\text{STEM}}+\alpha_{t\times\text{STEM}}+\alpha_{c\times t}+\delta\times(\text{Excess Deaths}_c \times \text{STEM}_i \times \text{Post}_t)+\boldsymbol{x}_i'\boldsymbol{\beta}+\varepsilon_{jict} \quad (3.11)$$

where j denotes a child, i denotes the household head, c and t are respectively county and birth-year of the child; Post_t is a dummy variable taking the value one if child j is born after 1918, and zero otherwise; STEM_i is an indicator variable that takes the value one if the household head is employed in a STEM occupation, and zero otherwise; and $\boldsymbol{x}$ includes urban status and race of the household head. The coefficient δ compares STEM and non-STEM individuals, before and after the pandemic, by county-level exposure to the pandemic. The sample is composed of all children born between 1900 and 1929. Children are weighted by the inverse of the number of children in each household. Table 3.6 shows the results. In columns (1)–(3), the comparison group is the entire population, while in columns (4)–(6), we focus on high-skilled workers. We find that, for both comparison groups, individuals in STEM occupations become less religious

[50]The correlation between revealed religiosity and background religiosity is equal to 0.13 and highly statistically significant ($p < .001$), in line with a large literature on cultural transmission (Bisin and Verdier, 2001).

[51]To construct these variables, we consider only children born before 1918, and we take the within-household average religiosity.

than non-STEM ones in counties more exposed to the influenza shock (columns 1 and 4). This pattern is stronger for Catholics (columns 2 and 5) than for Protestants.

These findings further show that, within society, different groups reacted in different ways to an adverse shock. In particular, STEM individuals appeared to turn further away from religion compared to their non-STEM counterparts.

3.4.2.3 Polarization of Religious Beliefs

In this section, we analyze the impact of the influenza pandemic on the distribution of religiosity within counties. Precisely, we estimate heterogeneous treatment effects of the pandemic across the initial distribution of background religiosity.

To study this question, we discretize the distribution of background religiosity into quintiles, which we label Q^{BR}, and we estimate the following model:

$$\begin{aligned} \text{Religiosity}_{jict} &= \alpha_{c\times t} + \alpha_{c\times Q} + \alpha_{Q\times t} + \\ &+ \sum_{\ell=1}^{5} \delta^{\ell} \times \left[\text{Excess Deaths}_c \times \text{Post}_t \times 1\left(Q_i^{\text{BR}} = \ell\right)\right] + \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_{ict} \end{aligned} \tag{3.12}$$

where j denotes a child, i denotes the household head, c and t are respectively county and child birth-year; Equation (3.12) includes county-by-time, county-by-background, and background-by-time fixed effects, and the term $\boldsymbol{x}_i$ includes urban status and race of the household head. The term $1\left(Q_i^{\text{BR}} = \ell\right)$ is a dummy variable that takes the value one if household head's background religiosity is in the ℓ-th quintile, and zero otherwise. If the pandemic caused an increase in polarization of religiosity, the set of coefficients $\{\delta^{\ell}\}_{\ell=1}^{5}$ in equation (3.12) would be monotonically increasing in ℓ. On the other hand, a decreasing sequence of coefficients would provide evidence that the pandemic led to a convergence of religiosity. In model (3.12), the sample is composed of all children born between 1900 and 1929. Children are weighted by the inverse of the number of children in each household.

In Figure 3.5, we report the set of $\{\delta^{\ell}\}$ coefficients by religious denominations. We normalize the third quintile as the baseline category. The figure provides evidence in favor of an increase in polarization: for individuals with below-median religious backgrounds, the coefficients on exposure to the pandemic are negative, while they are positive for those with above-median religious backgrounds. This suggests that, within the same county, individuals from different religious backgrounds become even more distant in terms of their religiosity, increasing the polarization of religiosity within society. In Table B.20 we report the results of the corresponding regressions.

Taken together, these three individual-level exercises help us understand the contemporaneous increase in both religiosity and science at the county level. They suggest that, within counties, individuals reacted differently to the same shock, based, for instance, on their religious background or on their prepandemic scientific orientation. Thus, while a county may have become both more religious *and* more innovative, individuals seemed to turn either to religion *or* to science, leading to within-county polarization of religiosity.

3.5 Discussion: Interpretation and Limitations of the Results

Our analysis shows two clear patterns: (i) the 1918–1919 influenza pandemic led to an increase in religiosity and production of innovation across U.S. counties and, as a result of the shock, the same counties became both more religious and more innovative; (ii) *within* counties, there was a heterogeneous response to the same shock, with some individuals turning to religion and others turning to science.

One concern is that other factors related to the pandemic may have affected the evolution of religiosity and science, confounding our results. To address this concern, we proceed in three steps. First, we document that neither initial religiosity nor innovation activity are related to the intensity of the shock. Second, our event-study analysis shows the absence of pretrends, suggesting that religiosity and innovation were on a similar path in treated and control groups before the shock. Third, we account for other potentially confounding characteristics, such as differential fertility, WWI deaths, and migration patterns. Our results are robust in all these cases. Taken together, the empirical evidence, supported by the historical records, makes it hard to imagine that the pandemic did not trigger an increase in both religiosity and scientific progress.

A second concern regards our main measures of religiosity and scientific progress. First, does our name-based indicator indeed capture religiosity at the local level? We show that our results are robust to alternative ways of constructing our naming measure and when using alternative classifications of religious names. In addition, we show that in counties hit harder by influenza, the share of people affiliated with a religious denomination increases, providing further evidence that the pandemic led to an increase in local religiosity. Similarly, patents could be an imperfect measure of scientific progress (Moser, 2005). To address this concern, we show that our findings hold when using the share of individuals in scientific occupations as an alternative proxy.

One puzzle emerging when looking at the aggregate patterns is whether these results are driven by individuals becoming both more religious and more innovative or by different individuals reacting differently to the same shock. Our findings suggest that the second mechanism is at play. Individuals from more religious backgrounds further embrace religion, while those from less religious backgrounds are more likely to choose a scientific occupation. This suggests that a group of individuals within society used religion as a coping device, while a separate group turned to science. In addition, we show that the shock widened the distance in religiosity between science-oriented individuals and the rest of the population: people in scientific occupations, already less religious than the rest of the population, moved further away from religion. Finally, the pandemic increased the polarization of religiosity in the population: individuals from more (less) religious backgrounds became even more (less) religious.

One key question regarding our individual-level results is, what explains the increase in religiosity or the choice of a scientific occupation? The findings on religiosity are in line with the religious coping hypothesis, which suggests that religious faith can represent a coping device to deal with personal distress following a negative shock. An alternative explanation for why individuals may turn to religion is for social insurance. While we are not able to fully rule this out (and it goes beyond the scope of our paper), we read our evidence as being in favor of the religious coping hypothesis. First, this is in line with the literature showing that intrinsic religiosity (rather than churchgoing) responds to unexpected events, as noted by Bentzen (2019). Second, as the increase in religiosity persists up to ten years after the shock, it is more likely to be related to a change in beliefs rather than to a temporary increase in the need for social insurance.

What motivates people to turn to science is less obvious. Individuals may turn to science to deal with their psychological distress, similarly to religious coping, or in an attempt to actively mitigate the negative (e.g., health-related or economic) effects of the pandemic. Another possibility could be that individuals turn to science because of increased labor demand in STEM occupations, but our results suggest that, beyond market forces, the individual's religious background plays a key role in the decision to turn to science. While our findings cannot directly speak to the individual-level motivations behind these different behaviors, they provide evidence of a heterogeneous response to the same adverse event.

Finally, one limitation of our individual-level analysis is that, while we can construct the religious background for every individual, we cannot directly measure their scientific one. This is due to our measure of scientific orientation based on occupational choice, which—contrary to our measure of religious background—does not allow us to know an individual's occupation and the parents' occupation from the same census.[52] However, since we know that science-oriented people are less religious than the overall population (Appendix Table B.19), it is plausible to assume that religious background and scientific background are similarly negatively correlated. Taken together, we interpret our results as suggestive evidence that, while individuals from religious backgrounds turned to religion as a coping device in the aftermath of the pandemic, those from a scientific background turned to science.

3.6 Conclusions

In this paper, we provide new evidence on how societies react to adversities, studying an exemplary historical episode: the Great Influenza Pandemic of 1918–1919.

First, we show that society reacted to the pandemic by becoming both more religious and more innovative. Second, using individual-level data from full-count censuses, we suggest that religiosity and science are substitute ways of reacting to the shock. When facing adversity, individuals from more religious backgrounds turned to religion, while those from less religious backgrounds turned to science. Third, we show that the pandemic shock widened the distance in religiosity between scientific-oriented individuals and the rest of the population, and that it increased preexisting differences in religious sentiment. As a consequence, the distribution of religiosity in counties more exposed to the pandemic became more polarized.

Our paper sheds new light on the relationship between religiosity and science. Throughout history, science and religion have often been in conflict, and recent evidence by Bénabou *et al.* (2015, 2022) shows that the two are negatively correlated, both across countries and across U.S. states. We provide novel evidence that—at the individual level—the two are substitute ways of dealing with adversity.

Our analysis helps shed light on modern events such as the reaction of society to the COVID-19 pandemic. Even though the modern context differs in many ways from the one that witnessed the influenza pandemic, including the medical advancements of the past century, the reaction of today's society seems in line with what we document for the 1918–1919 pandemic.[53] In particular, our findings can help explain the opposing views that have emerged since the COVID-19 pandemic on science-based responses to the shock, such as the opposing attitudes toward vaccines.

[52] A natural way to construct a measure of scientific background, symmetric to the religiosity one, would be to look at whether individuals had a parent working in a scientific occupation. Unfortunately, this is not possible due to data limitations; this exercise would require tracking individuals across several census waves.

[53] One key difference between the two pandemics is that no medical remedy or vaccine became available until many years after the earlier pandemic ended.

Finally, our results suggest that, in the aftermath of a negative shock, societies may become more polarized in their religiosity. Because religion has become an increasingly polarizing element in the current political debate, facing adversity may strongly affect not only religious polarization but also the polarization of political views, and more broadly, the polarization of society itself.

Tables

TABLE 3.1: Summary Statistics

	(1) Mean	(2) Std. Dev	(3) Min	(4) Max	(5) Counties
Panel A. Mortality					
Flu excess deaths (%)	1.134	.148	.853	1.779	1220
WW1 deaths	20.864	85.414	.5	2414	1220
Net Flu Excess Deaths (%)	1.099	.159	-.897	1.714	1220
Panel B. County Demographics					
Population	36.905	79.468	.076	1298.405	1217
Area	218.356	283.698	0	5205.831	1217
Share of Whites	.936	.126	.311	1	1217
Share of African Americans	.058	.127	0	.689	1217
Share of Foreign Born	.119	.112	0	.498	1217
Share of Illiterates	.045	.047	.001	.264	1217
Income per Capita	833.501	24.243	746.105	913.328	1217
Panel C. Religious Affiliations					
Total Affiliated	21.635	62.262	.148	982.279	1219
Catholics	9.415	36.48	0	589.856	1219
Protestants	10.386	18.697	.056	309.439	1219
Panel D. Innovation Activity					
All	106.456	409.376	0	5598.142	1220
Pharmaceuticals	13.785	54.099	0	710.225	1220
Communications	3.132	14.761	0	292.194	1220
Electrical	11.224	52.666	0	1039.469	1220
Mechanical	37.472	142.788	0	2026.215	1220
Other	40.843	152.003	0	1872.377	1220

Notes: This table displays the mean, standard deviation, minimum, maximum, and total number of counties of the main variables. Data are measured at the county level. Panel A and B report data from the 1910 census. Data in Panels C and D are at decade level. Hence, for instance, column (1) of Panel C "Total Affiliated" reports the average number of individuals affiliated with any denomination over the period 1900-1930. Column (1) of Panel D "All" reports the average number of patents in any class in each decade between 1900 and 1930. "Excess deaths" is defined as the ratio between total deaths during the pandemic, and total deaths in the three years before. County demographics are measured through the IPUMS full-count census (Ruggles *et al.*, 2021). Income per capita is measured through occupational income scores based on the 1950 Census. Religious affiliation data are from the Census of Religious Bodies. Patent data are from Berkes (2018). All variables are crosswalked to 1920 borders.

TABLE 3.2: The Impact of the Influenza on Religiosity

	Share of Affiliated			Name-Based Religiosity		
	(1) All	(2) Catholics	(3) Protestants	(4) All	(5) Catholics	(6) Protestants
Excess Deaths × Post	0.202***	0.082***	0.082***	0.007**	0.009***	0.006**
	(0.027)	(0.015)	(0.015)	(0.003)	(0.003)	(0.003)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	–	–	–
Year FE	–	–	–	Yes	Yes	Yes
Number of Counties	1219	1219	1219	1201	1201	1201
Observations	3657	3657	3657	36030	36030	36030
R^2	0.861	0.908	0.931	0.450	0.306	0.471
Std. Beta Coef.	0.635	0.337	0.301	0.109	0.184	0.100

Notes: This table displays the impact of exposure to the Great Influenza Pandemic on religiosity. The unit of observation is a county, observed at a decade frequency between 1906 and 1926 (in columns 1–3) and yearly frequency between 1900 and 1929 (in columns 4–6). “Post” is a categorical variable equal to one during and after the pandemic—i.e., over the years 1918 to 1929—or zero otherwise. The baseline treatment “Excess Deaths” is defined in Equation (3.4). In columns (1–3), the dependent variable is the number of individuals affiliated with religious denominations enumerated in the Census of Religious Bodies, normalized by county population in 1900; in columns (4–6), the dependent variable is the name-based religiosity measure described in the main text. Columns (1) and (4) report the effect of the influenza on overall religiosity, whereas columns (2) and (5)—resp. (3) and (6)—display it on the intensity of Catholicism—resp. Protestantism. Regressions include county and time (decades in columns 1–3 and years in columns 4–6) fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors, clustered at the county level, are reported in parentheses.

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

TABLE 3.3: The Impact of the Influenza on the Volume and Direction of Innovation

	Dep. Var.: log(1 + Number of Patents)					
	(1) All Patents	(2) Pharmaceuticals	(3) Communication	(4) Electrical	(5) Mechanical	(6) Other
Excess Deaths × Post	0.503***	0.091***	0.000	0.022	0.014	0.018
	(0.064)	(0.033)	(0.021)	(0.027)	(0.023)	(0.022)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
All Patents	No	Yes	Yes	Yes	Yes	Yes
Number of Counties	1220	1220	1220	1220	1220	1220
Observations	37820	37820	37820	37820	37820	37820
R^2	0.832	0.836	0.717	0.819	0.925	0.935
Std. Beta Coef.	0.211	0.066	0.000	0.018	0.008	0.009

Notes: This table displays the impact of exposure to the Great Influenza Pandemic on the level and direction of innovation. The unit of observation is a county, observed at a yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic—i.e., over the years 1918 to 1929—or zero otherwise. The baseline treatment "Excess Deaths" is defined in Equation (3.4). In column (1), the dependent variable is the (log) total number of patents granted. In the other columns, the dependent variable is the (log) number of patents granted in each column field, controlling for the overall (log) number of patents. In all models, we take $\ln(1 + \text{Patents})$ as the dependent variable to ensure that we do not drop counties without patents. Column (1) estimates the impact of the pandemic on the level of innovation, while columns (2)–(6) display this on the direction of innovation because we control for the total number of patents. Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors, clustered at the county level, are reported in parentheses.

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

TABLE 3.4: Impact of the Influenza on Occupational Choice

	Entire Population			Skilled Population		
	STEM / Population	Dummy = 1 if STEM		STEM / Skilled	Dummy = 1 if STEM	
	(1)	(2) No Controls	(3) Controls	(4)	(5) No Controls	(6) Controls
Excess Deaths × Post	0.004***			0.049***		
	(0.000)			(0.007)		
Excess Deaths × Young		0.003***	0.004***		0.033***	0.031***
		(0.001)	(0.001)		(0.010)	(0.009)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	–	–	Yes	–	–
Cohort FE	–	Yes	Yes	–	Yes	Yes
Household Controls	–	No	Yes	–	No	Yes
Number of Counties	1218	1217	1217	1216	1217	1217
Observations	4868	30679634	30679634	4864	4285423	4285423
R^2	0.656	0.002	0.002	0.584	0.005	0.010
Std. Beta Coef.	0.799	0.021	0.021	0.845	0.076	0.071

Notes: This table displays the effect of the pandemic on occupational choice. In columns (1) and (4), the unit of observation is a county, observed at decade frequency between 1900 and 1930. In columns (2–3) and (5–6), the unit of observation is an individual, observed once in the 1930 population census. In columns (1) and (4), the treatment interacts a "Post" variable equal to one for each census decade after the pandemic, or zero otherwise, with the standard "Excess Deaths" measure defined in Equation (3.4). In column (1), the dependent variable is the (log 1+) share of people employed in STEM occupations, as a share of the population in 1910. In column (4) the share is computed relative to the number of people employed in skilled occupations in 1910. The lists of STEM occupations and of high-skilled occupations are reported in Table B.1. In columns (2–3) and (5–6), the treatment is an interaction between a dummy variable equal to one if the person is employed in a STEM occupation and zero otherwise, and an indicator variable returning value one for all those aged 25 or less at the time of the inception of the pandemic. In columns (2–3) the sample includes the entire population; in columns (5–6) we only include individuals employed in skilled occupations. In columns (3) and (6) we control for race and urban status of the head of household. Regressions in columns (1) and (4) include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator; regressions in columns (2–3) and (5–6) include county and cohort fixed effects. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE 3.5: Religious Background, Religiosity, and STEM Occupations

	Religiosity			STEM Occupation		
	(1) All	(2) Catholics	(3) Protestants	(4) All	(5) Catholics	(6) Protestants
Excess Deaths × Post × High Religious Background	0.066***	0.037***	0.020			
	(0.016)	(0.014)	(0.013)			
Excess Deaths × Young × High Religious Background				-0.003*	-0.004**	-0.003**
				(0.001)	(0.001)	(0.001)
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Background FE	Yes	Yes	Yes	Yes	Yes	Yes
Background-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of Counties	1217	1217	1217	1217	1217	1217
Observations	7641683	7641686	7641678	13569024	13569024	13569024
R^2	0.029	0.023	0.026	0.006	0.007	0.006
Std. Beta Coef.	0.037	0.022	0.013	-0.010	-0.012	-0.013

Notes: This table displays the impact of exposure to the pandemic on religiosity—columns (1)–(3)—and occupational choice—columns (4)–(6)—by individual-level background religiosity. The unit of observation in columns (1)–(3) is a head of household, who is observed once for each child born between 1900 and 1930 in the household. In columns (4)–(6), the unit of observation is an adult. Religiosity is defined as the religiosity score associated with the child's name. "Post" is a categorical variable equal to zero for children born during and after the pandemic—i.e., over the years 1918–1929—or zero for those born before the pandemic—i.e., before 1918. The baseline treatment "Excess Deaths" is defined in Equation (3.4). "STEM" is an indicator variable returning value one if an individual is employed in a STEM occupation—as defined in Table B.1—or zero otherwise. "Young" is an indicator variable equal to one if an individual is younger than 25 years old in 1918, or zero otherwise. "High Background Religiosity" is an indicator variable returning the value one if the religiosity score of the name of the head of the household is in the top 20% of the overall distribution, or zero otherwise. The table displays the coefficient of the interaction between these terms. Each regression includes county-by-cohort, county-by-background, and background-by-cohort fixed effects. We include race and urban status as further household-level controls in each regression. Standard errors are clustered at the county level and are reported in parentheses.

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

Table 3.6: Effect of the Influenza on Individual Religiosity: STEM and Non-STEM

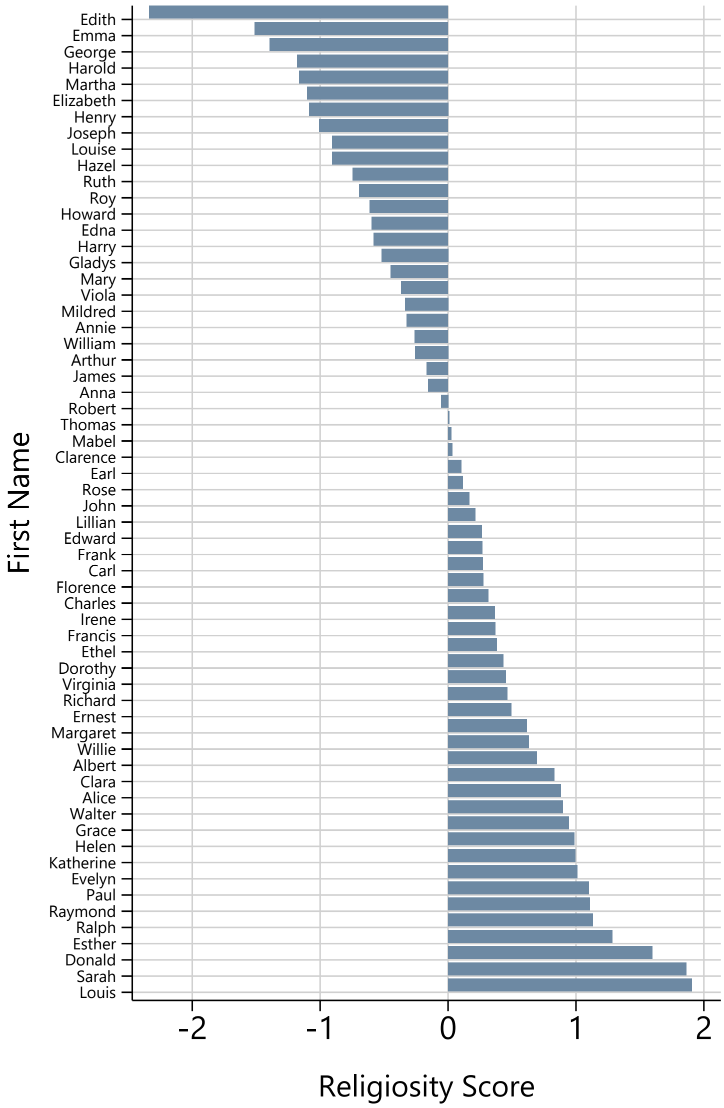
	Entire Population			Skilled Population		
	(1) All	(2) Catholics	(3) Protestants	(4) All	(5) Catholics	(6) Protestants
Excess Deaths × Post × STEM	-0.107**	-0.084***	-0.030	-0.081*	-0.060*	-0.011
	(0.048)	(0.032)	(0.040)	(0.045)	(0.036)	(0.040)
STEM-County FE	Yes	Yes	Yes	Yes	Yes	Yes
STEM-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	Skilled	Skilled	Skilled
Number of Counties	1217	1217	1217	1217	1217	1217
Observations	15096725	15096725	15096725	2275587	2275587	2275587
R^2	0.009	0.006	0.008	0.024	0.022	0.022
Std. Beta Coef.	-0.012	-0.011	-0.004	-0.023	-0.020	-0.004

Notes: This table displays the impact of exposure to the pandemic on STEM and non-STEM individuals' religiosity. The unit of observation is a child, born between 1900 and 1930. Religiosity is defined as the religiosity score associated with the child's name. "Post" is a categorical variable equal to zero for children born before the pandemic—i.e., before 1918—or one for those born after the pandemic—i.e., after 1918. The baseline treatment "Excess Deaths" is defined in Equation (3.4). "STEM" is an indicator variable returning a value of one if one parent of the child is employed in a STEM profession, or zero otherwise. The table displays the coefficient of the interaction between these terms. This estimates the causal effect of the influenza shock on the religiosity of heads of households employed in STEM occupations *vis-à-vis* non-STEM occupations, leveraging variation in county-level exposure to the influenza. All models include STEM-by-county, STEM-by-cohort, and county-by-cohort fixed effects. In columns (1)–(3), the estimation sample includes all individuals; in columns (4)–(6) we include only those employed in skilled occupations, which we enumerate in table B.1. Standard errors are clustered at the county level and are reported in parentheses.

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

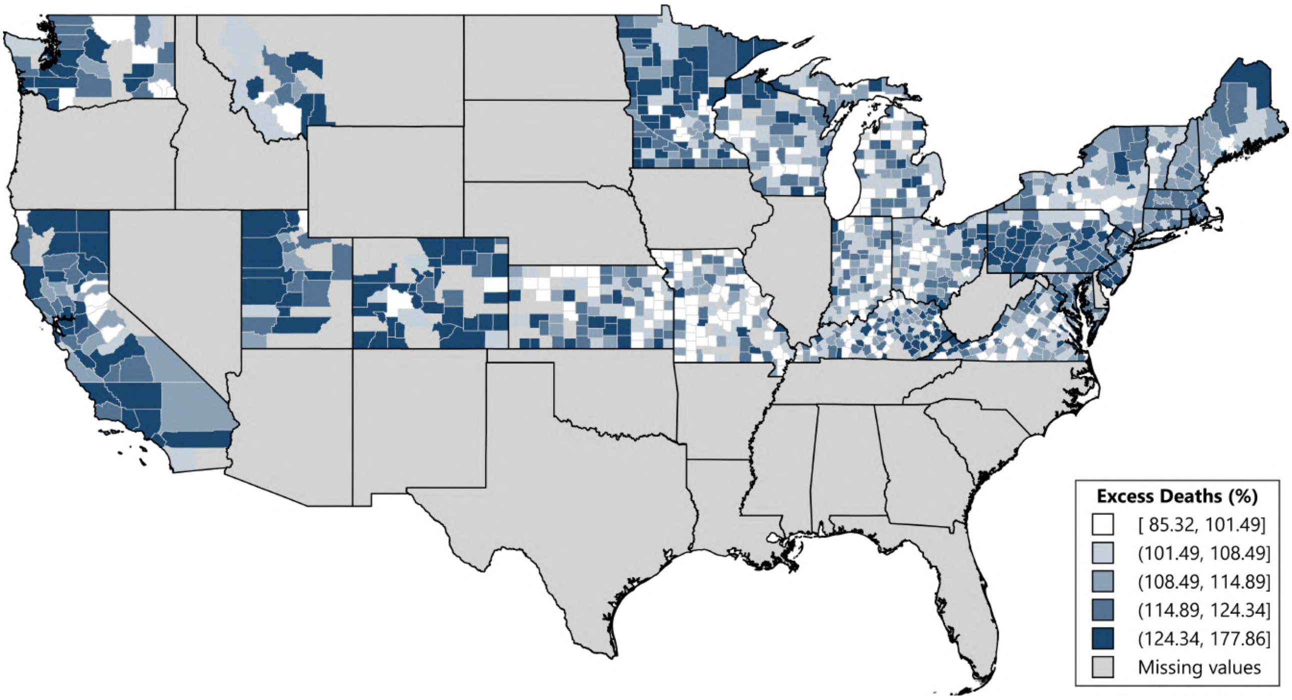
Figures

FIGURE 3.1: Estimated Names Religiosity Scores



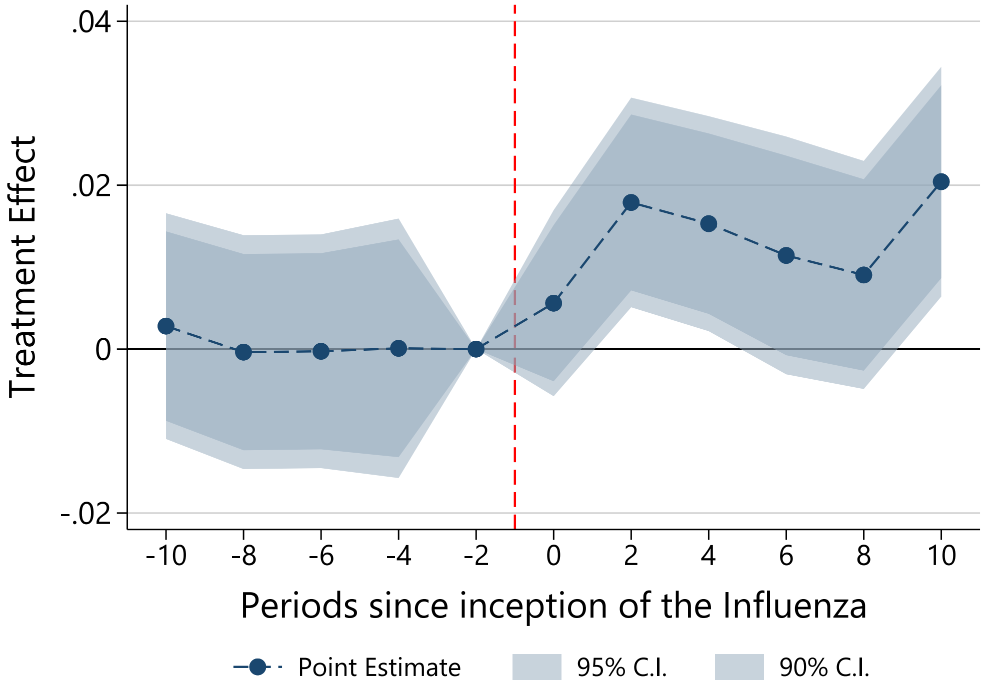
Notes: This figure displays the religiosity scores estimated from model (3.2). Regressions are based on data from the 1906–1916 Censuses of Religious Bodies; they include individuals born between 1896 and 1916. We estimate religiosity scores for names appearing in at least 0.3% of the overall sample. We conflate variations of a single name together—e.g., Anne and Anna—but keep endearments separate—e.g., Anna and Annie. Coefficients are reported in increasing order. In Figure B.1, we report religiosity scores split by confessions.

FIGURE 3.2: Spatial Distribution of Excess Mortality During the Great Influenza Pandemic



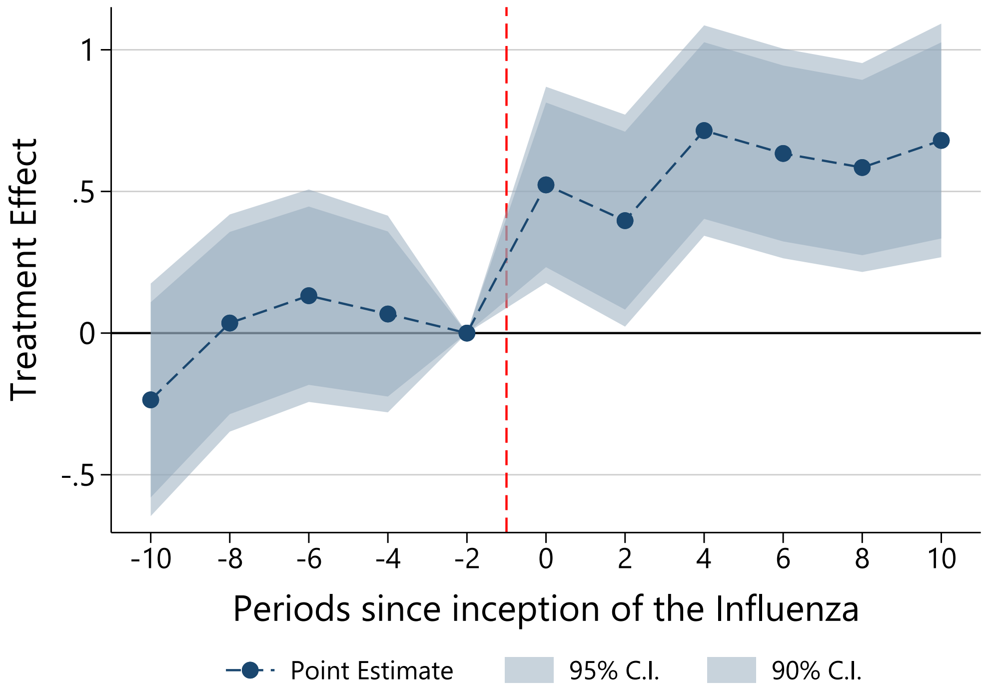
Notes: This figure displays geographic variation in influenza excess deaths, defined in Equation (3.4). Excess mortality is the ratio between the average number of deaths during the pandemic (1918–1919) and the average number of deaths in the three years before the pandemic (1915–1917). Mortality statistics prior to 1915 are not available. Excess mortality is displayed in percentage terms. Lighter to darker blue indicates increasing exposure to the influenza. Counties are displayed at their 1920 borders.

FIGURE 3.3: Impact of the Influenza on Religiosity



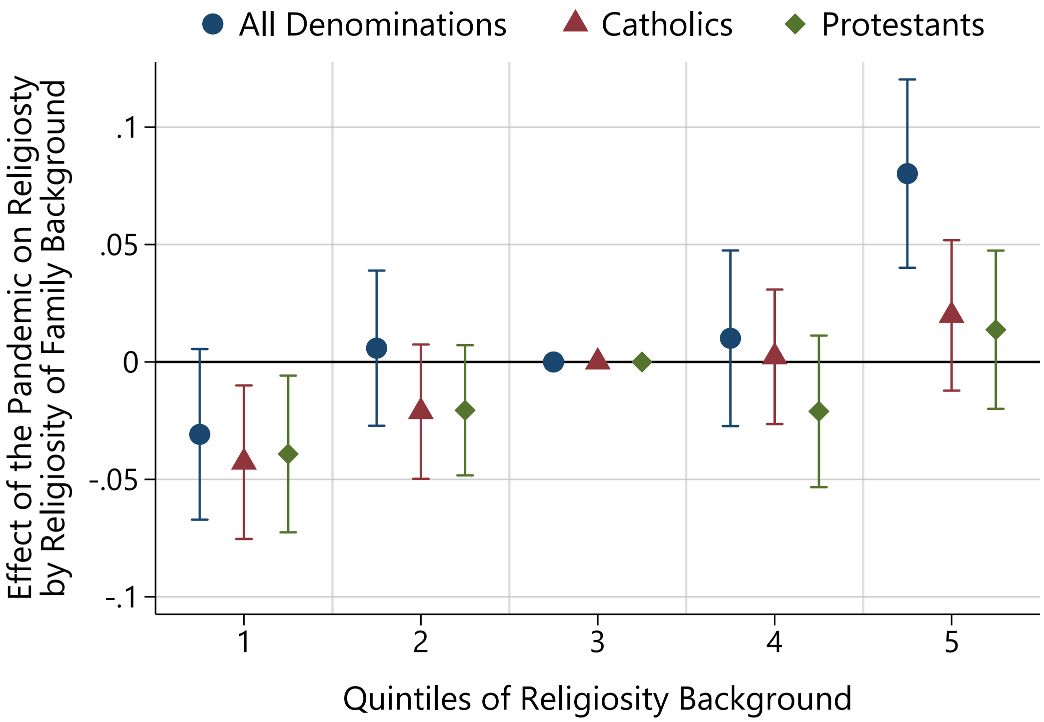
Notes: This figure displays the dynamic treatment effects of the pandemic on overall religiosity. The unit of observation is a county, observed at a biennial frequency. Each dot reports the coefficient of an interaction between the baseline measure of excess deaths, defined in Equation (3.4), and a biennial time dummy. The coefficient for the biennial 1916–1917, i.e., the last two-year window prior to the inception of the Great Influenza Pandemic, serves as the baseline. The model includes county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Bands report 90% and 95% confidence intervals. Standard errors are clustered at the county level. The dashed vertical line indicates the timing of the pandemic.

FIGURE 3.4: Impact of the Influenza on Innovation



Notes. The Figure reports dynamic treatment effects of the pandemic on innovation. The dependent variable is the (log 1+) total number of patents filed in a given year. The unit of observation is a county, observed at a biennial frequency. Each dot reports the coefficient of an interaction between the baseline measure of excess deaths and a biennial time dummy. The coefficient for the biennial 1916–1917, i.e., the last two-year window prior to the inception of the Great Influenza Pandemic, serves as the baseline. The model includes county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Bands report 90% and 95% confidence intervals. Standard errors are clustered at the county level. The dashed vertical line indicates the timing of the pandemic.

FIGURE 3.5: Impact of the Influenza on the Polarization of Religious Beliefs



Notes: This figure reports the estimated impact of the pandemic on the polarization of religious beliefs by religious denomination. Each dot reports the coefficient of an interaction between the baseline measure of excess deaths, a posttreatment indicator, and an indicator for the quintile of background religiosity. The unit of observation is a child, born between 1900 and 1930. Treated children are those born after the influenza, i.e., after 1918. The dependent variable is the religiosity score associated with the child's name. Background religiosity is measured as the religiosity score of the child's head of household. Results are reported by confession, and the third quintile serves as the baseline. Regression models include fixed effects for county by cohort, county by quintile of religious background, and cohort by quintile of religious background. Standard errors are clustered at the county level, and the bands report the 95% confidence interval for each coefficient.

Bibliography

ABRAMITZKY, R., L. P. BOUSTAN and D. CONNOR (2020). "Leaving the Enclave: Historical Evidence on Immigrant Mobility from the Industrial Removal Office." *NBER Working Paper*, (No. w27372).

ABRAMITZKY, R., L. P. BOUSTAN and K. ERIKSSON (2016). "Cultural Assimilation During the Age of Mass Migration." *NBER Working Paper*, (No. w22381).

AGARWAL, R. and P. GAULE (2022). "What Drives Innovation? Lessons from COVID-19 R&D." *Journal of Health Economics*, 82: 102591.

AGER, P. and A. CICCONE (2018). "Agricultural Risk and the Spread of Religious Communities." *Journal of the European Economic Association*, 16(4): 1021–1068.

AGER, P., C. W. HANSEN and L. LØNSTRUP (2016). "Church Membership and Social Insurance: Evidence from the Great Mississippi Flood of 1927." *Southern Denmark University Discussion Papers on Business and Economics*.

ALMOND, D. (2006). "Is the 1918 Influenza Pandemic Over? Long-term Effects of in Utero Influenza Exposure in the Post-1940 US Population." *Journal of Political Economy*, 114(4): 672–712.

ANDERSEN, L. H. and J. S. BENTZEN (2022). "In the Name of God! Religiosity and the Transition to Modern Growth." *CEPR Discussion Paper*, (No. DP16938).

ANDREWS, M. J. (2021). "Historical Patent Data: A Practitioner's Guide." *Journal of Economics & Management Strategy*, 30(2): 368–397.

ANO, G. G. and E. B. VASCONCELLES (2005). "Religious coping and psychological adjustment to stress: A meta-analysis." *Journal of Clinical Psychology*, 61(4): 461–480.

BABINA, T., A. BERNSTEIN and F. MEZZANOTTI (2021). "Crisis Innovation: Evidence from the Great Depression." *NBER Working Paper*, (No. w27851).

BARRO, R. J. (2022). "Non-pharmaceutical interventions and mortality in us cities during the great influenza pandemic, 1918–1919." *Research in Economics*, 76(2): 93–106.

BARRO, R. J., J. F. URSÚA and J. WENG (2020). "The Coronavirus and the Great Influenza Pandemic: Lessons from the "Spanish Flu" for the Coronavirus's Potential Effects on Mortality and Economic Activity." *NBER Working Paper*, (No. w26866).

BARRY, J. M. (2020). *The Great Influenza: The Story of the Deadliest Pandemic in History*. London (UK): Penguin.

BAZZI, S., E. BERKES, M. FISZBEIN and C. FONS-ROSEN (2022). "Individualism and Innovation." *Working Paper*.

BAZZI, S., M. FISZBEIN and M. GEBRESILASSE (2020). "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States." *Econometrica*, 88(6): 2329–2368.

BEACH, B., K. CLAY and M. H. SAAVEDRA (2020). "The 1918 Influenza Pandemic and its Lessons for COVID-19." *NBER Working Paper*, (No. w27673).

BECKER, S. O. and L. WOESSMANN (2009). "Was Weber Wrong? A Human Capital Theory of Protestant Economic History." *The Quarterly Journal of Economics*, 124(2): 531–596.

BELLOC, M., F. DRAGO and R. GALBIATI (2016). "Earthquakes, Religion, and Transition to Self-Government in Italian Cities." *The Quarterly Journal of Economics*, 131(4): 1875–1926.

BÉNABOU, R., D. TICCHI and A. VINDIGNI (2015). "Religion and Innovation." *American Economic Review: Papers & Proceedings*, 105(5): 346–351.

——— (2022). "Forbidden Fruits: The Political Economy of Science, Religion and Growth." *The Review of Economic Studies*, 89(4): 1785–1832.

BENTZEN, J. S. (2019). "Acts of God? Religiosity and Natural Disasters Across Subnational World Districts." *The Economic Journal*, 126(622): 2295–2321.

——— (2021). "In Crisis, We Pray: Religiosity and the COVID-19 Pandemic." *Journal of Economic Behavior & Organization*, 192: 541–583.

BERKES, E. (2018). "Comprehensive Universe of US patents (CUSP): Data and Facts." *Working Paper*.

BERKES, E., O. DESCHENES, R. GAETANI, J. LIN and C. SEVEREN (2020). "Lockdowns and Innovation: Evidence from the 1918 Flu Pandemic."

BIANCHI, N. and M. GIORCELLI (2020). "Scientific Education and Innovation: From Technical Diplomas to University STEM Degrees." *Journal of the European Economic Association*, 18(5): 2608–2646.

BIAVASCHI, C., C. GIULIETTI and Z. SIDDIQUE (2017). "The Economic Payoff of Name Americanization." *Journal of Labor Economics*, 35(4): 1089–1116.

BISIN, A. and T. VERDIER (2001). "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory*, 97(2): 298–319.

BOTTICINI, M. and Z. ECKSTEIN (2012). *The Chosen Few: How Education Shaped Jewish History, 70-1492*. Princeton University Press.

BOUSTAN, L. P., M. E. KAHN and P. W. RHODE (2012). "Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century." *American Economic Review*, 102(3): 238–244.

BOUSTAN, L. P., M. E. KAHN, P. W. RHODE and M. L. YANGUAS (2020). "The Effect of Natural Disasters on Economic Activity in US Counties: A Century of Data." *Journal of Urban Economics*, 118: 103257.

CALLAWAY, B., A. GOODMAN-BACON and P. H. SANT'ANNA (2021). "Difference-in-Differences with a Continuous Treatment." *Working Paper*.

CALLAWAY, B. and P. H. SANT'ANNA (2021). "Difference-in-Differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

CLAY, K., J. LEWIS and E. SEVERNINI (2019). "What Explains Cross-City Variation in Mortality During the 1918 Influenza Pandemic? Evidence from 438 US Cities." *Economics & Human Biology*, 35: 42–50.

CLEMENS, J. and P. ROGERS (2020). "Demand Shocks, Procurement Policies, and the Nature of Medical Innovation: Evidence from Wartime Prosthetic Device Patents." *NBER Working Paper*, (No. w26679).

CROSBY, A. W. (1989). *America's Forgotten Pandemic: The Influenza of 1918*. Cambridge: Cambridge University Press.

DEMING, D. J. (2010). *Science and Technology in World History: Early Christianity, the Rise of Islam and the Middle Ages: 2*. Jefferson, NC: McFarland.

DEMING, D. J. and K. L. NORAY (2018). "STEM Careers and Technological Change." *NBER Working Paper*, (No. w25065).

FOUKA, V. (2020). "Backlash: The Unintended Effects of Language Prohibition in US Schools After World War I." *The Review of Economic Studies*, 87(1): 204–239.

FROST, P. (2020). "An Accelerant of Social Change? The Spanish Flu of 1918-19." *International Political Anthropology Journal*, 13(2): 123–133.

GIULIANO, P. and A. SPILIMBERGO (2023). "Recessions, Lifetime Experiences and the Formation of Political Beliefs." *Working paper*.

GROSS, D. P. and B. N. SAMPAT (2021). "The Economics of Crisis Innovation Policy: A Historical Perspective." *American Economic Association: Papers and Proceedings*, 111: 346–50.

HALL, B. H., A. B. JAFFE and M. TRAJTENBERG (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools."

HUNTINGTON, E. (1923). "Causes of Geographical Variation in the Influenza Epidemic in the Cities of the United States." *Bull. Nat. Res. Council*, 6: 1–36.

IANNACCONE, L. R. (1998). "Introduction to the Economics of Religion." *Journal of Economic Literature*, 36(3): 1465–1496.

IYER, S. (2016). "The New Economics of Religion." *Journal of Economic Literature*, 54(2): 395–441.

KELLY, B., D. PAPANIKOLAOU, A. SERU and M. TADDY (2021). "Measuring Technological Innovation Over the Long Run." *American Economic Review: Insights*, 3(3): 303–20.

LECCE, G., L. OGLIARI and M. P. SQUICCIARINI (2021). "Birth and Migration of Scientists: Does Religiosity Matter? Evidence from 19th-Century France." *Journal of Economic Behavior & Organization*, 187: 274–289.

Markel, H., H. B. Lipman, J. A. Navarro, A. Sloan, J. R. Michalsen, A. M. Stern and M. S. Cetron (2007). "Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic." *Journal of the American Medical Association*, 298(6): 644–654.

Miao, Q. and D. Popp (2014). "Necessity as the Mother of Invention: Innovative Responses to Natural Disasters." *Journal of Environmental Economics and Management*, 68(2): 280–295.

Mokyr, J. (2011). "The Economics of Being Jewish." *Critical Review*, 23(1-2): 195–206.

Montgomery, L. M. (1992[1924]). In M. Rubio and E. Waterston, eds., "The Selected Journals of L.M. Montgomery, Volume III: 1921-1929," Toronto: Oxford University Press.

Moscona, J. (2021). "Environmental Catastrophe and the Direction of Invention: Evidence from the American Dust Bowl." *Working Paper*.

Moscona, J. and K. A. Sastry (2022). "Does Directed Innovation Mitigate Climate Damage? Evidence from US Agriculture." *The Quarterly Journal of Economics*.

Moser, P. (2005). "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs." *American Economic Review*, 95(4): 1214–1236.

Norenzayan, A. and I. G. Hansen (2006). "Belief in Supernatural Agents in the Face of Death." *Personality and Social Psychology Bulletin*, 32(2): 174–187.

Olivetti, C., M. D. Paserman, L. Salisbury and E. A. Weber (2020). "Who Married,(To) Whom, and Where? Trends in Marriage in the United States, 1850-1940." *NBER Working Paper*, (No. 28033).

Pargament, K. I. (2001). *The Psychology of Religion and Coping: Theory, Research, Practice*. New York: Guilford Press.

——— (2002). "The Bitter and the Sweet: An Evaluation of the Costs and Benefits of Religiousness." *Psychological Inquiry*, 13(3): 168–181.

Perl, P. and J. L. Wiggins (2004). "Don't Call Me Ishmael: Religious Naming Among Protestants and Catholics in the United States." *Journal for the Scientific Study of Religion*, 43(2): 209–228.

Phillips, H. (2020). "'17, '18, '19: Religion and Science in three Pandemics, 1817, 1918, and 2019." *Journal of Global History*, 15(3): 434–443.

Posch, M. (2022). "Do Disasters Affect the Tightness of Social Norms?"

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*.

Sibley, C. G. and J. Bulbulia (2012). "Faith After an Earthquake: A Longitudinal Study of Religion and Perceived Health Before and After the 2011 Christchurch New Zealand Earthquake." *PloS One*, 7(12): e49648.

Spinney, L. (2018). *Pale Rider: the Spanish Flu of 1918 and How It Changed the World*. London: Vintage.

Squicciarini, M. P. (2020). “Devotion and Development: Religiosity, Education, and Economic Progress in 19th-Century France.” *American Economic Review*, (110): 3454–3491.

Stark, R. (1992). “The Reliability of Historical United States Census Data on Religion.” *Sociological Analysis*, 53(1): 91–95.

——— (1998). “The Rise and Fall of Christian Science.” *Journal of Contemporary Religion*, 13(2): 189–214.

Sun, L. and S. Abraham (2021). “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects.” *Journal of Econometrics*, 225(2): 175–199.

Weber, M. (1905). *The Protestant Ethic and the Spirit of Capitalism*. London: Routledge.

Appendix

3.A Data: Description and Sources

In this section, we list the sources of the data and describe how we construct the variables used in the analysis.

3.A.1 Patents

3.A.1.1 Patent Data

Patent data are from Berkes (2018), who performed optical character recognition (OCR) on original patent documents issued by the United States Patents and Trademark Office between 1836 and 2010. Information includes the filing and issue year, author name, latitude and longitude of the inventor(s), and inferred USPC technology class. The data contain a set of additional variables, including the complete text of the patent document and the issue year of the patent, not used in our analysis. We geo-code each patent to its 1920 county using boundary shapefiles supplied by NHGIS. When we collapse by county year, we weigh each patent by the inverse of the number of technology classes, as well as by the inverse of the number of authors. Hence, a patent with two authors and two technological classes appears four times in the original patent-level dataset, and each instance is assigned a .25 weight when aggregating at the county level. We code USPC classes to the NBER classification (Hall *et al.*, 2001). We modify the canonical NBER classification and conflate the "Chemical" and "Drugs" categories into a single "Pharmaceuticals" class. Since multiple USPC codes are typically assigned to a single patent, most patents that would fall under "Drugs" would also appear as "Chemical." To avoid this, we simply recast them into one single category. It is worth noting that all the results that we present in terms of pharmaceutical patents also hold if we keep the "Chemical" and "Drugs" classes separate.

3.A.1.2 Quality Data

We measure patent quality using the measure developed by Kelly *et al.* (2021). From their data, we derive two metrics. One is the average quality. The second, which we label "Breakthrough", is an indicator variable returning value one if the patent's quality is in the top 25% of the overall quality distribution, and zero otherwise. Both measures are net of grant-year fixed effects. We take forward and backward similarity within a 5-year window around the issue year of the patent.

3.A.1.3 Linked Inventor-Census Data

Patent data alone do not allow to uniquely identify an inventor. Because in Table B.15 we need to measure inventor productivity as well as the number of unique inventors, we exploit a novel sample of inventors linked to the census by Bazzi *et al.* (2022). This allows us to assign a unique identifier to each inventor in our sample, and compute the related statistics. We defer the interested reader to the accompanying paper describing the data in more detail. An inventor can be matched to multiple census entries. In this paper, we disregard all inventors with more than five matches (about 5% of the overall stock). Then, we weigh the remaining by the inverse of the number of matches, as is standard in the census-linking literature.

3.A.2 Names

We take name data from the individual full-count US 1930 population census (Ruggles *et al.*, 2021). First names require some cleaning. First, we remove non-ASCII characters and drop all those reporting the initial only. Then, we manually identify common diminutives (e.g., "Thos" for "Thomas"). Finally, we agglutinate variations and minor spelling mistakes on the same underlying name. To do so, we code a simple script that collects a set of reference names as those appearing more than 50 times in the entire population census. We then compute the Jaro-Winkler similarity between each name and the reference names, and normalize it to lie between 0 and 1. If, for a given name, there is one reference name with a similarity above .99 we conflate that name to the reference. Otherwise, we just keep the name as is. This simple procedure is not intended to agglutinate either translations (e.g., "Tommaso" and "Thomas) or endearments (e.g., "Willie" and "William"). We thus take a conservative stand as to whether the same name in different languages–or its endearments–may convey different religious attitudes. It is merely an algorithmic approach to correct minor spelling mistakes. Overall, after the manual trimming we are left with 1,366,844 single names, which decrease to 623,792 after the algorithmic trimming procedure. However, weighting these figures by the number of children reveals that less than 20,000 names account for more than 95% of the total number of newborns.

3.A.3 Religious Affiliations

Data on religious affiliations are supplied by NHGIS, and are originally from the Census of Religious Bodies which took place at decade frequency between 1906 and 1936. We discard the 1936 census because previous research shows that the uptake was low and unequal across counties (Stark, 1992). Census enumerators asked churches, congregations, and other local organizations to provide a list of their members. The data was then aggregated at the county level. In our analysis, "Total Religiosity" is computed as the simple sum of religious members across all possible denominations; "Catholics" are enumerated as such. We collectively refer as "Protestants" to a set of denominations which we manually mapped to some branch of Protestantism (including, e.g., the Methodist, Evangelical, and–various–Baptist churches.)

3.A.4 Occupational Structure

Individual-level data on occupations is extracted from the 1930 individual-level population census. More precisely, we use the 1950 harmonized occupation classification. We then manually map occupational codes to STEM occupations as described in Table B.1.

3.A.5 Controls & Mortality Statistics

We extract a battery of individual-level characteristics from the IPUMS full count data. Among those, we use the the race and urban-rural status as additional individual-level controls.

County-level covariates are provided by NHGIS, which in turn aggregates individual-level data from population censuses, and reports data from manufacturing and agricultural censuses. All data come at historical county borders.

Mortality statistics are likewise provided by NHGIS. For the period we are interested in, namely, 1915-1919, they were collected for about 1,200 counties, covering approximately 60% of the US population. We measure Influenza-related mortality as the ratio between deaths during the pandemic, and deaths in the three years which preceded the Influenza.[54]

3.A.6 Canadian Data

Following Abramitzky *et al.* (2020) we use the Canadian Census to construct an alternative measure of religiosity. This has the advantage of reporting information on individuals' religious affiliations as well as their first names. We use three waves of the census: 1881 (full count), 1911, and 1921, and construct religiosity scores by first name for the cohorts born between 1800 and 1916. Using the same procedure outlined in Fouka (2020), we construct the following metric:

[54]The original documents report, for major cities, deaths broken down by (alleged) cause. We do not use this data for two main reasons. First, they are incomplete and are only available for cities. Second, Beach *et al.* (2020) criticize the methodology adopted to impute the cause of deaths.

$$\text{Religiosity score}_{\text{name},r,c} = \frac{\Pr\left(\text{name} \mid I_{r,c}\right)}{\Pr\left(\text{name} \mid I_{r,c}\right) + \Pr\left(\text{name} \mid I_{R\setminus r,c}\right)} \times 100 \tag{3.13}$$

where (name) is first name, r is religion, c is birth cohort, and I is an indicator for individuals of a given religion and birth cohort. $I_{R\setminus r,c}$ indicates individuals of religion other than r. The score ranges from 0 to 100: a score of 0 implies a name is never found among individuals of religion r; a score of 100 implies a name is never found among individuals of a different religion. We define two religious groups in the Canadian Census: Catholics and Protestants.[55] First, we compute the scores in equation 3.13 for each religion and birth cohort. Second, we average the score within-decade (where one decade corresponds to ten birth cohorts), for each religion and name. Finally, we generate a Catholic dummy and a Protestant dummy. Each dummy takes the value 1 if the corresponding religiosity score is larger than its 80, and zero otherwise.

3.A.7 Other Data

In several robustness regressions, we control for WW1 mortality. The underlying data were collected by Ferrara and Fishback (2020).

3.A.8 Boundary Harmonization

County-level data from NHGIS and other sources are typically provided at historical borders. To ensure comparability and consistency, we adopt the method developed by Eckert *et al.* (2020) to compute geographical crosswalks between US counties over time. In a nutshell, their methodology is as follows. Suppose that we know the distribution of a given variable y across counties at decade frequency between 1900 and 1930. To harmonize borders to one single year, Eckert *et al.* (2020) overlay the shapefile of counties in a given year, say, 1900, to that in the reference year, say, 1920. They then compute the percentage of land that a given county shares with itself between the two years, and that which is assigned to other counties. To construct the harmonized variable, one simply multiplies these overlapping area weights by the variable recorded in 1900, and aggregates up by 1920-counties. The underlying assumption is that y is evenly distributed over the county territory. While this may seem untenable in most cases, departures from this assumption are plausibly innocuous in our setting. County borders had in fact undergone major consolidations before 1900 and remained stable thereafter. Moreover, mortality

[55] We only define two groups as, over this period, less than 1% of individuals reported a religious affiliation other than Catholic or Protestant.

data are mostly available for the Northwest and Midwest areas. Boundary changes in these regions were rare and minor after the 1890s. In our application, we map all county-level variables to 1920-borders.

3.A.9 Details on Sample Construction

In this paragraph, we provide additional technical details on the way we construct the estimation samples. The main sample restriction that we impose descends from the fact that we observe mortality for 1265 out of 2917 counties. We then discard 45 counties with implausibly large (above 200%) or low (below 50%) values of excess mortality during the pandemic. Because such figures are due to scarcely-inhabited areas, these 45 counties account for less than 1% of the population in the 1265-counties sample. We are left with a set of 1220 counties. In the rest of the paragraph, we explain why we may not always be able to leverage all 1220 for the estimation.

3.A.9.1 County-Level Religiosity

The county-level religiosity estimation sample is a balanced panel dataset where each county is observed at a yearly frequency between 1900 and 1929. This implies that the number of counties in this balanced panel may not be 1220 as long as at least one county is not observed at least once between 1900 and 1929. This happens because, especially in scarcely-inhabited areas, the name-frequency threshold that we impose may imply that we are not able to match any newborn in a given cohort. If that is the case, the county's religiosity will not be observed every year of the sample, and the county will subsequently be dropped from the estimation sample. This is the case for 19 out of 1220 counties, so the estimation sample, in this case, consists of 1201 counties accounting for 98.5% of the population in the 1220-counties sample.

In one robustness check shown in column (7) of Table B.3 counties are observed at decade frequency instead. In this case, the sample is constructed from adults observed once per census decade between 1900 and 1930, and the post-treatment indicator returns value one for decades 1920 and 1930, and zero otherwise. In table B.4 we employ an alternative measure of religiosity from the 1906 Census of Religious Bodies that does not include fixed effects in the estimation equation of the names religiosity scores. This measure is considerably more volatile than the baseline, so we exclude the top and bottom 5% most extreme observations in the associated synthetic religiosity distribution.

3.A.9.2 County-Level Innovation

The county-level innovation sample is a balanced panel dataset where each county is observed at a yearly frequency between 1900 and 1929. Thus, an observation in the dataset can either be a number above zero (if there are one or more patents observed in that county-year) or zero (if no patents are observed). The estimation sample in this case thus encompasses all 1220 counties for which we observe mortality. In columns (2) and (7) of Table B.11 we do not fill the panel with zeros when no patents are observed. This results in an unbalanced panel dataset where a county may not be observed every year over the estimation time period.

3.A.9.3 Other County-Level Samples

In Table B.6 we use as dependent variables various measure of name concentration. Because these measures display sizable variability, we restrict the sample to exclude counties at the top and bottom 1% of the excess deaths distribution. Results would remain qualitatively unchanged using the full sample, but they would conflate pre-treatment statistically significant–selection–effects that would induce a spurious downward bias in the estimated treatment effects.

3.A.9.4 Individual-Level

We construct two individual-level datasets. In both samples, the unit of observation is the head of the household. In the first, each head of household is observed once. Regressions (3.7) and (3.10) are estimated on this "adult" sample. In the second sample we observe the kids of each head of household. We interpret the kids as realizations of the religiosity of their parent. Regressions (3.9), (3.11), and (3.12) are estimated on this "kid" sample.

3.A.9.5 City-Level

To build the city-level sample, we construct the baseline excess deaths treatment variable from data by Clay *et al.* (2019). The dataset contains mortality information on 976 cities. For 444, however, we observe the number of deaths in one year only, and we do not observe 48 other cities continuously between 1915 and 1919. Moreover, we do not observe population data in 1900 for 41 additional cities. The final sample consists of 443 cities. In Figure B.7 we report the location of each city and the number of cities included in the sample, by state. We geo-code patents to historical city borders and construct the name-based religiosity measure from the individuals that were recorded living in each city in the 1930 census. The city-level sample is used in the regressions displayed in Table B.7.

3.B Additional Tables and Figures

3.B.1 Tables

TABLE B.1: STEM Professions

Occ. Code	Occupation Label	Share (%)	Occ. Code	Occupation Label	Share (%)
(1)	(2)	(3)	(4)	(5)	(6)
Panel A. STEM Occupations					
12	Agricultural sciences	0.00	18	Mathematics	0.00
61	Agricultural scientists	0.03	19	Medical sciences	0.03
13	Biological sciences	0.00	772	Midwives	0.42
62	Biological scientists	0.32	69	Miscellaneous natural scientists	0.08
14	Chemistry	0.04	26	Natural science (n.e.c.)	0.00
7	Chemists	6.07	92	Surveyors	1.05
32	Dentists	12.39	67	Mathematicians	0.01
34	Dietitians and nutritionists	0.64	240	Officers, pilots, pursers and engineers, ship	8.06
16	Engineering	0.01	94	Technicians, medical and dental	1.49
49	Engineers (n.e.c.)	0.88	70	Optometrists	1.33
41	Engineers, aeronautical	0.05	71	Osteopaths	0.73
42	Engineers, chemical	0.61	25	Statistics	0.00
43	Engineers, civil	13.31	75	Physicians and surgeons	30.98
44	Engineers, electrical	9.32	68	Physicists	0.04
45	Engineers, industrial	0.34	23	Physics	0.00
46	Engineers, mechanical	7.13	17	Geology and geophysics	0.00
47	Engineers, metallurgical, metallurgists	0.31	98	Veterinarians	1.77
48	Engineers, mining	1.11	83	Statisticians and actuaries	1.07
63	Geologists and geophysicists	0.33	61	Agricultural Scientists	0.07
Panel B. Other Skilled Occupations					
$1 \leq \cdot \leq 99$	Liberal and Skilled Professions		$200 \leq \cdot \leq 299$	Managers	
$700 \leq \cdot \leq 790$	Service Workers				

Notes: Panel A displays the occupations which we classify as Science, Technology, Engineering, and Mathematics (STEM). Panel B displays the occupations that we classify as skilled: these include all STEM occupations, in addition to the ones listed. Occupation codes and labels are from the IPUMS harmonized 1950 occupation taxonomy. Column "Share" indicates the percentage share of individuals in the given occupation, relative to total employment in STEM occupations in the baseline individual-level sample. STEM occupations account for about 6% of total skilled employment, which in turn accounts for approximately 14% of total employment.

TABLE B.2: Balance Checks Regressions

	(1) Coefficient	(2) Standard Error	(3) 95% C. I.
Panel A. Income and Demographics			
Population Density	−0.098	(0.221)	[−0.531, 0.336]
Income per Capita	0.332	(0.407)	[−0.466, 1.131]
Share of Men	0.520^{***}	(0.163)	[0.200, 0.840]
Share of Illiterates	0.339	(0.362)	[−0.369, 1.048]
Share of Young	0.366	(0.311)	[−0.244, 0.976]
Panel B. Ethnic Composition			
Share of Whites	0.255	(0.246)	[−0.228, 0.738]
Share of African Americans	−0.308	(0.235)	[−0.769, 0.153]
Share of Foreign Population	0.444^{***}	(0.159)	[0.131, 0.757]
Immigrants from:			
Italy	0.287	(0.282)	[−0.265, 0.840]
Ireland	0.088	(0.126)	[−0.159, 0.336]
Austria	0.172	(0.356)	[−0.527, 0.870]
France	0.244	(0.188)	[−0.123, 0.612]
Spain	0.419	(0.365)	[−0.295, 1.134]
Portugal	0.000	(0.291)	[−0.569, 0.570]
Panel C. Religion			
All Denominations	−0.126	(0.246)	[−0.609, 0.356]
Catholics	0.175	(0.226)	[−0.269, 0.619]
Protestants	−0.263	(0.288)	[−0.827, 0.302]
Panel D. Patents			
Total	0.169	(0.103)	[−0.032, 0.370]
Pharmaceutical	0.132	(0.097)	[−0.058, 0.321]
Communication	0.100	(0.101)	[−0.097, 0.297]
Electrical	0.226	(0.172)	[−0.110, 0.562]
Mechanical	0.186^{*}	(0.096)	[−0.003, 0.375]
Other	0.147	(0.091)	[−0.031, 0.324]

Notes: This table displays the correlation between the Excess Death (defined in (3.4)) and a set of covariates in 1910, i.e., the last census year before the pandemic. Column (1) reports the standardized coefficient of a regression between the row variable and our measure of excess deaths; column (2) reports the associated standard error in round brackets; column (3) reports the confidence interval of the point estimate at the 95% confidence level in square brackets. All variables are expressed as shares of total population, except for population density. Regressions control for county population and include state fixed effects.

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

TABLE B.3: Impact of the Influenza on Religiosity: Robustness Analysis

	Baseline Sample			Family Size Cuts		Household	Adults
	(1) Cont. Treat.	(2) Disc. Treat.	(3) WW1	(4) No firstborn	(5) < 5 Kids	(6)	(7)
Excess Deaths × Post	0.007**		0.007**	0.006*	0.007**	0.003**	0.014
	(0.003)		(0.003)	(0.003)	(0.004)	(0.001)	(0.012)
Excess Deaths Dummy × Post		0.003**					
		(0.001)					
WW1 Deaths × Post			0.000				
			(0.000)				
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Baseline	Baseline	Baseline	No Firstborn	< 5 Kids	Household	Adults
Number of Counties	1201	1201	1201	1200	1200	1201	1201
Observations	36030	36030	36030	36000	36000	36030	4804
R^2	0.450	0.450	0.450	0.366	0.370	0.364	0.860
Std. Beta Coef.	0.109	0.021	0.110	0.086	0.099	0.101	0.081

Notes: This table displays the impact of exposure to the Influenza on overall religiosity. The unit of observation is a county, observed at a yearly frequency between 1900 and 1929 in columns (1)-(6), and at a decade frequency in column (7). "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929 in columns (1)-(6) and the decades 1920-1930 in column (7)–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). The dependent variable is the name-based measure of aggregate religiosity described in the main text. Column (1) displays the baseline results. Column (2) reports the results coding the treatment as a binary variable returning value one if the continuous treatment is above its median, and zero otherwise. In column (3) we control for WW1-related deaths. Column (4) drops first-born children in every household. In column (5) we compute religiosity dropping all children beyond the fourth in each household. In column (6) we first compute within-household average religiosity and then aggregate the resulting religiosity series at the county-year level. Column (7) reports results measuring county religiosity using the names stock of adults–which serves as a placebo check. All regressions in columns (1)-(6) include county and year fixed effects; the regression in column (7) includes county and decade fixed effects. Additionally, each regression includes the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.4: Impact of the Influenza on Religiosity: Names Scores without Fixed Effects

	Unweighted			Weighted		
	(1) All	(2) Catholics	(3) Protestants	(4) All	(5) Catholics	(6) Protestants
Excess Deaths × Post	0.037**	0.029*	0.014	0.100***	0.089***	0.021
	(0.018)	(0.016)	(0.011)	(0.034)	(0.030)	(0.025)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Counties	1201	1201	1201	1201	1201	1201
Observations	29612	29612	29612	29612	29612	29612
R^2	0.387	0.540	0.282	0.636	0.804	0.557
Std. Beta Coef.	0.130	0.098	0.082	0.392	0.288	0.128

Notes: This table displays the impact of exposure to the Influenza on religiosity. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). Religiosity is measured using religiosity scores obtained by estimating equation (3.2), except that we do not include the fixed effects in the regression specification. In columns (4)–(6) counties are weighted by their population in 1900. Columns (1) and (4) report the results for total religiosity; columns (2) and (5) refer to Catholics; columns (3) and (6) refer to Protestants. Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.5: Impact of the Influenza on Religiosity: Alternative Thresholds

	All			Catholics			Protestants		
	(τ = 2)	(τ = 3)	(τ = 5)	(τ = 2)	(τ = 3)	(τ = 5)	(τ = 2)	(τ = 3)	(τ = 5)
Excess Deaths × Post	0.012	0.007**	0.004**	0.007*	0.009***	0.004*	0.003	0.006**	0.004
	(0.008)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Counties	1201	1201	1200	1201	1201	1200	1201	1201	1200
Observations	36030	36030	36000	36030	36030	36000	36030	36030	36000
R^2	0.378	0.450	0.356	0.248	0.306	0.483	0.353	0.471	0.416
Std. Beta Coef.	0.115	0.109	0.097	0.117	0.184	0.093	0.047	0.100	0.091

Notes: This table displays the impact of exposure to the Influenza on religiosity. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). Religiosity is measured using religiosity scores obtained estimating equation (3.2). The term τ denotes the frequency threshold a name must exceed to be included in our sample, in ‰ terms. For instance, $\tau = 2$ implies that at least 2‰ children in our sample must be called with a given name, for that name to be included in the sub-sample of names used to compute the religiosity score. We report the baseline results, with $\tau = 3$, as well as those with lower and larger thresholds. As τ decreases, the number of names for which we compute a religiosity score increases. Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.6: Impact of the Influenza on the Concentration of Names

	HHI	CCI	Rosenbluth	C-5	C-7	C-9	C-10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Excess Deaths × Post	-0.090	-0.005**	-0.115*	-0.007	-0.007	-0.007	-0.008
	(0.065)	(0.002)	(0.059)	(0.004)	(0.005)	(0.006)	(0.006)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Counties	1150	1150	1150	1150	1150	1150	1150
Observations	34490	34490	34490	34490	34490	34490	34490
R^2	0.853	0.779	0.875	0.779	0.807	0.824	0.830
Std. Beta Coef.	-0.092	-0.162	-0.109	-0.109	-0.094	-0.085	-0.085

Notes: This table displays the impact of exposure to the Influenza on name concentration. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). The dependent variables measure the concentration of names, and are: in column (1) the Herfindahl-Hirschman (HHI) index; in column (2) the Comprehensive Concentration index (CCI), which relative to the HHI assigns more weight to relatively uncommon names; in column (3) the Rosenbluth index (RI), which further refines the CCI because it is more sensitive to the number of uncommon names. In columns (4)–(7) the dependent variable is the k-concentration ratio, *i.e.* the share of children called with the k most common names. More formally, let s_n denote the share of kids with name n, and let N be the total number of names. Suppose that shares are ranked in increasing order, meaning that $\text{rank}(n) \leq \text{rank}(n')$ if and only if $s_n \geq s_{n'}$, and $\text{rank}(n) < \text{rank}(n')$ if and only if $s_n > s_{n'}$ for all n, n'. Then, $HHI \equiv \sum_{n=1}^{N} s_n^2$; $CCI \equiv s_1 + \sum_{n=2}^{N} s_n^2(2 - s_n)$, $RI \equiv \frac{1}{2\sum_{n=1}^{N} ns_n - 1}$; $C_K \equiv \sum_{n=1}^{K} s_n$. Regressions include county and state-by-year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level, and are reported in parentheses.

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

TABLE B.7: Impact of the Influenza on Religiosity and Innovation: City-Level Analysis

	Religiosity			Innovation	
	(1) All	(2) Catholics	(3) Protestants	(4) All	(5) Pharmaceutical
Excess Deaths × Post	0.007***	0.006***	0.003	0.152*	0.159**
	(0.002)	(0.002)	(0.002)	(0.086)	(0.077)
City FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Number of Cities	443	443	443	478	478
Observations	13290	13290	13290	14340	14340
R^2	0.456	0.375	0.533	0.865	0.782
Std. Beta Coef.	0.125	0.131	0.048	0.062	0.081

Notes: This table displays the city-level effect of exposure to the Influenza on religiosity and innovation. The unit of observation is a city, observed at a yearly frequency between 1900 and 1929. We report the location of each city in the sample in figure B.7. The baseline sample is from Clay *et al.* (2019). We include only cities for which we can construct the baseline excess mortality measure. In columns (1–3), the dependent variable is the name-based religiosity measure constructed on the universe of children born between 1900 and 1929 and residing in each city in the 1930 census; in columns (4–5), the dependent variable is the number of patents and pharmaceutical patents, respectively. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). Each regression includes city and year fixed effects. Standard errors, clustered at the city level, are reported in parentheses.
*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

TABLE B.8: Impact of the Influenza on Canadian and Saint-Biblical Religiosity Scores

	Canada Scores		Biblical and Saints Scores		
	(1) Catholics	(2) Protestants	(3) Biblical/Saints	(4) Saints	(5) Biblical
Excess Deaths × Post	0.013**	0.006	0.055***	0.051***	0.013***
	(0.005)	(0.020)	(0.009)	(0.008)	(0.003)
County FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Number of Counties	1200	1200	1200	1200	1200
Observations	36000	36000	36000	36000	36000
R^2	0.327	0.491	0.959	0.958	0.958
Std. Beta Coef.	0.173	0.016	0.151	0.147	0.095

Notes: This table displays the impact of exposure to the Influenza on religiosity. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). In columns (1)-(3), religiosity is measured using religiosity scores obtained as described in section 3.B from the Canadian census. In column (3), the dependent variable is the share of children by cohort whose name either appears in the bible, or is carried by a saint; in column (4), the dependent variable only includes biblical names; in column (5), it only includes names of saints. Biblical and saints names are from Abramitzky *et al.* (2016). Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.9: Impact of the Influenza on Religiosity: Weighted Regressions

	Share of Affiliated			Name-Based Religiosity		
	(1) All	(2) Catholics	(3) Protestants	(4) All	(5) Catholics	(6) Protestants
Excess Deaths × Post	0.493**	0.212***	0.242***	0.009**	0.009**	0.004
	(0.218)	(0.073)	(0.047)	(0.004)	(0.004)	(0.005)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	–	–	–
Year FE	–	–	–	Yes	Yes	Yes
Number of Counties	1219	1219	1219	1201	1201	1201
Observations	3657	3657	3657	36030	36030	36030
R^2	0.713	0.847	0.779	0.606	0.498	0.640
Std. Beta Coef.	0.198	0.247	0.197	0.191	0.242	0.084

Notes: This table displays the impact of exposure to the Great Influenza Pandemic on religiosity. The unit of observation is a county, observed at a decade frequency between 1906 and 1926 (in columns 1–3) and yearly frequency between 1900 and 1929 (in columns 4–6). "Post" is a categorical variable equal to one during and after the pandemic—i.e., over the years 1918 to 1929—or zero otherwise. The baseline treatment "Excess Deaths" is defined in Equation (3.4). In columns (1–3), the dependent variable is the number of individuals affiliated with religious denominations enumerated in the Census of Religious Bodies, normalized by county population in 1900; in columns (4–6), the dependent variable is the name-based religiosity measure described in the main text. Columns (1) and (4) report the effect of the influenza on overall religiosity, whereas columns (2) and (5)—resp. (3) and (6)—display it on the intensity of Catholicism—resp. Protestantism. Regressions include county and time (decades in columns 1–3 and years in columns 4–6) fixed effects and the interaction between population in 1900 and a post-treatment indicator. Counties are weighted by population in 1900. Standard errors, clustered at the county level, are reported in parentheses.

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

TABLE B.10: Influenza and Innovation: Weighted Regressions

	Dep. Var.: log(1 + Number of Patents)					
	(1) All Patents	(2) Pharmaceuticals	(3) Communication	(4) Electrical	(5) Mechanical	(6) Other
Excess Deaths × Post	0.695***	0.265***	-0.068	0.050	-0.009	-0.014
	(0.166)	(0.094)	(0.198)	(0.106)	(0.059)	(0.055)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
All Patents	No	Yes	Yes	Yes	Yes	Yes
Number of Counties	1220	1220	1220	1220	1220	1220
Observations	37820	37820	37820	37820	37820	37820
R^2	0.950	0.950	0.868	0.939	0.979	0.983
Std. Beta Coef.	0.193	0.098	-0.036	0.019	-0.003	-0.004

Notes: This table displays the impact of exposure to the Great Influenza Pandemic on the level and direction of innovation. The unit of observation is a county, observed at a yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic—i.e., over the years 1918 to 1929—or zero otherwise. The baseline treatment "Excess Deaths" is defined in Equation (3.4). In column (1), the dependent variable is the (log) total number of patents granted. In the other columns, the dependent variable is the (log) number of patents granted in each column field, controlling for the overall (log) number of patents. In all models, we take $\ln(1 + \text{Patents})$ as the dependent variable to ensure that we do not drop counties without patents. Column (1) estimates the impact of the pandemic on the level of innovation, while columns (2)–(6) display this on the direction of innovation because we control for the total number of patents. Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors, clustered at the county level, are reported in parentheses. Counties are weighted by population in 1900.

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

TABLE B.11: Influenza and Innovation: Robustness Regressions

	All Patents				Pharmaceutical Patents				
	(1) Baseline	(2) Unbalanced	(3) Disc. Treat	(4) WW1 Deaths	(5) Baseline	(6) No All Patents	(7) Unbalanced	(8) Dummy	(9) WW1 Deaths
Excess Deaths × Post	0.503***	0.474***		0.503***	0.091***	0.276***	0.134***		0.091***
	(0.064)	(0.087)		(0.064)	(0.033)	(0.047)	(0.051)		(0.033)
Excess Deaths Dummy × Post			0.118***					0.032***	
			(0.020)					(0.010)	
WW1 Deaths × Post				3.739					3.574
				(17.019)					(3.057)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Patents	No	No	No	No	Yes	No	Yes	Yes	Yes
Number of Counties	1220	1184	1220	1220	1220	1220	1184	1220	1220
Observations	37820	23909	37820	37820	37820	37820	23909	37820	37820
R^2	0.832	0.861	0.832	0.832	0.836	0.786	0.824	0.836	0.836
Std. Beta Coef.	0.211	0.223	0.036	0.211	0.066	0.201	0.085	0.017	0.066

Notes: This table displays the impact of exposure to the Influenza on innovation. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. In columns (1)–(4) the dependent variable is the number of patents across all fields; in columns (5)–(9) it is the number of patents in chemical and drugs fields, according to the NBER standard classification. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). Columns (1) and (5) display the baseline results. Columns (2) and (7) report results for the unbalanced panel of counties (*i.e.*, the subsample of county-year observations for which we observe at least one filed patent). Columns (3) and (8) report the results when the treatment is coded as a binary variable equal to one if the continuous variable is above its median, and zero otherwise. Columns (4) and (9) further control for WW1 deaths interacted with the post-treatment indicator. In column (6) we report the estimated effect without controlling for the total number of patents. All regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Columns (5,7-9) further control for the total number of patents. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.12: Influenza and Innovation: Alternative Measures of Overall Innovation

	f (All Patents)		
	(1)	(2)	(3)
	$ln(1+\cdot)$	Count	arcsinh$(\cdot)$
Excess Deaths × Post	0.503***	6.787**	0.617***
	(0.064)	(3.343)	(0.078)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Number of Counties	1220	1220	1220
Observations	37820	37820	37820
R^2	0.832	0.908	0.810
Std. Beta Coef.	0.211	0.069	0.220

Notes: This table displays the effect of the Influenza on overall innovation. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. In column (1), the dependent variable is the log-number of patents, to which we add one to avoid dropping zeros. In column (2) the dependent variable is the raw patent count. In column (3) the dependent variable is the inverse hyperbolic sine of the raw count of patents. Each regression includes county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.13: Influenza and Innovation: Alternative Measures of Pharmaceutical Innovation

	f (Pharmaceutical Patents)							
	(1) $ln(1+\cdot)$	(2) $ln(1+\cdot)$	(3) Count	(4) Count	(5) arcsinh($\cdot$)	(6) arcsinh($\cdot$)	(7) Share	(8) $ln(1+\text{Share})$
Excess Deaths × Post	0.091***	0.246***	0.793***	1.777***	0.117***	0.302***	0.110***	0.071***
	(0.033)	(0.045)	(0.228)	(0.592)	(0.042)	(0.055)	(0.031)	(0.017)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total Patents	Yes	No	Yes	No	Yes	Yes	No	No
Number of Counties	1220	1220	1220	1220	1220	1220	1220	1220
Observations	37820	37820	37820	37820	37820	37820	37820	37820
R^2	0.836	0.788	0.959	0.856	0.820	0.773	0.171	0.228
Std. Beta Coef.	0.224	0.179	0.059	0.132	0.177	0.178	0.190	0.191

Notes: This table displays the effect of the Influenza on innovation in pharmaceuticals. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. In columns (1) and (2), the dependent variable is the log-number of patents, to which we add one to avoid dropping zeros. In columns (3) and (4), the dependent variable is the raw patent count. In columns (5) and (6) the dependent variable is the inverse hyperbolic sine of the raw count of pharmaceutical patents, with and without controlling for the inverse hyperbolic sine of the total number of patents. In column (7) the outcome is the number of pharmaceutical patents, relative to patents in all other fields. In column (8), this is taken in log. Each regression includes county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. In columns (1), (3), and (6) we further control by the total number of patents by county-year, transformed according to the column-specific labeled function. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.14: Influenza and Innovation: Assignee Heterogeneity

	Share of All Patents...		Share of Pharma w.r.t. All Patents...		Share of Pharma w.r.t. Pharma Patents...	
	(1) with Assignee	(2) w/out Assignee	(3) with Assignee	(4) w/out Assignee	(5) with Assignee	(6) w/out Assignee
Excess Deaths × Post	-0.005	0.196***	0.000	0.032***	0.004	0.090***
	(0.008)	(0.023)	(0.002)	(0.008)	(0.003)	(0.019)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Counties	1220	1220	1220	1220	1220	1220
Observations	37820	37820	37820	37820	37820	37820
R^2	0.094	0.596	0.051	0.197	0.067	0.627
Std. Beta Coef.	-0.005	0.196	0.000	0.032	0.004	0.090

Notes: This table displays the impact of exposure to the Influenza on innovation, by assignee status of each patent. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). In column (1) the dependent variable is the share of patents that list at least one assignee (approximately 60% of the sample); in column (2), the dependent variable is the share of patents that do not report an assignee (approximately 40% of the sample). In columns (3–4), the dependent variables are computed as the share of pharmaceutical patents with and without an assignee, relative to the total number of granted patents. In columns (5–6), the share is computed relative to all pharmaceutical patents. Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.15: Influenza and Innovation: Intensive and Extensive Margins

	Patents Per Inventor			N. of Inventors		
	(1) All	(2) Pharma	(3) Pharma	(4) All	(5) Pharma	(6) Pharma
Excess Deaths × Post	0.164***	0.079***	0.038**	0.348***	0.175***	0.070**
	(0.025)	(0.020)	(0.018)	(0.059)	(0.040)	(0.031)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N. of Inventors	No	No	No	No	No	Yes
Patents Per Inventor	No	No	Yes	No	No	No
Number of Counties	1220	1220	1220	1220	1220	1220
Observations	37820	37820	37820	37820	37820	37820
R^2	0.477	0.464	0.509	0.823	0.750	0.791
Std. Beta Coef.	0.249	0.140	0.068	0.164	0.132	0.053

Notes: This table displays the impact of exposure to the Influenza on the (log 1+) number of patents per inventor (intensive margin) and the (log) number of inventors (extensive margin). The unit of observation is a county, observed at yearly frequency between 1900 and 1929. “Post” is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment “Excess Deaths” is defined in equation (3.4). In column (1) the dependent variable is the number of patents per inventor in any field; in columns (2)–(3) we restrict to pharmaceutical patents per inventors; in column (4) the dependent variable is the number of inventors; in columns (5)–(6) we only consider inventors with at least one patent in pharmaceuticals. In column (3) we control for the average productivity, measured as the number of patents per inventor, to capture differential trends in productivity of pharmaceuticals, relative to the aggregate productivity; similarly, in column (6) we control for the number of inventors to disentangle differential patterns for the subgroup of inventors active in pharmaceuticals. Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.16: Impact of the Influenza on the Quality of Innovation

	All Patents			Pharmaceuticals			
	(1) Avg. Quality	(2) Breakthrought	(3) Share Breakthrough	(4) Avg. Quality	(5) Breakthrought	(6) Breakthrought	(7) Share Breakthrough
Excess Deaths × Post	0.099	1.548**	0.021	0.302***	0.991***	0.609***	0.021**
	(0.150)	(0.642)	(0.015)	(0.107)	(0.309)	(0.225)	(0.009)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total Patents	No	No	No	No	No	Yes	No
Number of Counties	1220	1220	1220	1220	1220	1220	1220
Observations	37818	37818	37818	37818	37818	37818	37818
R^2	0.314	0.786	0.131	0.477	0.681	0.793	0.107
Mean Dep. Var.	0.037	1.790	0.080	0.081	0.480	0.480	0.020
Std. Beta Coef.	0.031	0.092	0.062	0.220	0.197	0.011	0.131

Notes: This table displays the impact of the Influenza on the quality of innovation. In the first three columns, the quality indicators refer to the total patent flow; in the last four columns we restrict the sample to patents in pharmaceuticals. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. "Post" is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment "Excess Deaths" is defined in equation (3.4). Quality measures are from Kelly *et al.* (2021). They measure the "innovativeness" of a patent based on textual similarity between that patent and previous and future works, and flag it as important if it is different from previous work, but similar to subsequent ones. In columns (1) and (4), "Avg. Quality" denotes their baseline quality measure (equation (10) in Kelly *et al.* (2021)); in columns (2) and (5)–(6) "Breakthrough" is the raw count of patents in the top quintile of the quality distribution; in columns (3) and (7) "Share Breakthrough" is the share of patents in the top quintile in the quality distribution. Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level, and are displayed in parentheses.

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

TABLE B.17: Impact of the Influenza on Occupational Choice: Alternative Threshold

	Dummy = 1 if in STEM		
	(1) Baseline	(2) No Controls	(3) Controls
Excess Deaths × Younger than 30 in 1918	0.005***	0.006***	0.006***
	(0.002)	(0.001)	(0.002)
County FE	Yes	Yes	Yes
Cohort FE	Yes	–	–
State-Cohort FE	No	Yes	Yes
Household Controls	No	No	Yes
Number of Counties	1217	1217	1217
Observations	13573144	13573098	13573098
R^2	0.003	0.003	0.004
Std. Beta Coef.	0.026	0.030	0.030

Notes: This table displays the effect of the pandemic on the probability of being employed in a STEM occupation. The unit of observation is an individual, observed once in the 1930-population census. For every person, we define a dummy equal to one if the person is employed in a STEM occupation–enumerated in Table B.1–and zero otherwise. We drop individuals born after 1905 because they could still be completing their education spell in 1930, *i.e.* when we observe their occupational choice. An individual is defined to be treated if she is 30 years old or less in 1918, *i.e.* at the beginning of the pandemic. Compared to the baseline estimates, we enlarge the sample of treated individuals to those that were between 25 and 30 at the time of the inception of the pandemic. The baseline treatment "Excess Deaths" is defined in equation (3.4). Column (1) reports the baseline estimates; in column (2) we add state-by-year fixed effects to the baseline model. Column (3) further includes a set of individual-level controls. Individual controls are race and urban status. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.18: Religiosity and the Intensity of Innovation by Exposure to the Influenza

	(log) Patents per Capita		
	(1) All Affiliations	(2) Catholics	(3) Protestants
Excess Deaths × Post	0.032***	0.039***	0.033***
	(0.012)	(0.013)	(0.012)
All Affiliations	-0.042		
	(0.025)		
Excess Deaths × Post × All Affiliations	0.104**		
	(0.041)		
Catholics		0.023	
		(0.026)	
Excess Deaths × Post × Catholics		0.097**	
		(0.047)	
Protestants			-0.079*
			(0.041)
Excess Deaths × Post × Protestants			0.085*
			(0.048)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Number of Counties	1201	1201	1201
Observations	36030	36030	36030
R^2	0.749	0.750	0.749

Notes: This table displays the correlation between innovation and religiosity by exposure to the pandemic. The dependent variable is the log of patents, normalized by county-population in 1900. The unit of observation is a county, observed at yearly frequency between 1900 and 1929. “Post” is a categorical variable equal to one during and after the pandemic–*i.e.* over the years 1918-1929–and zero otherwise. The baseline treatment “Excess Deaths” is defined in equation (3.4). Religiosity by denomination is measured as described in the main text. Counties are weighted by their population in 1900. Regressions include county and year fixed effects and the interaction between population in 1900 and a post-treatment indicator. Standard errors are clustered at the county level and are reported in parentheses.

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

TABLE B.19: Religiosity of Individuals in STEM Compared to the Rest of the Population

	Non-STEM	STEM	Difference	
	(1)	(2)	(3)	(4)
All	0.021	0.015	–0.006*	–0.011***
			(0.086)	(0.001)
Catholics	–0.156	–0.176	–0.020***	–0.017***
			(0.000)	(0.000)
Protestants	0.027	0.010	–0.016***	–0.022***
			(0.000)	(0.000)
County FE	No	No	No	Yes
Birth Year FE	No	No	No	Yes
Controls	No	No	No	Yes

Notes: Columns (1) and (2) report the average religiosity of the non-STEM and the STEM populations; columns (3)–(4) report the difference between the two groups. Denomination varies by row (hence, for instance, average Catholic religiosity for Non-STEM is .181, it is .085 for STEM individuals, and their unconditional difference is -.096). To construct religiosity, we take all children in our baseline sample born before the Influenza, *i.e.* 1917. Observations are weighted by the inverse of the total number of kids in each household. In columns (1), (2), and (3) we report the unconditional statistics. In column (4) we include a set of county and (child) birth year fixed effects and we control for race and urban status. Standard errors are clustered at the county level. In columns (3)–(4), we report in parentheses the p-value associated with the estimates.

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

Table B.20: Impact of the Influenza on the Polarization of Religious Beliefs

	All	Catholics	Protestants
	(1)	(2)	(3)
Overall Religiosity Background=1 × Excess Deaths × Post	-0.031*		
	(0.019)		
Overall Religiosity Background=2 × Excess Deaths × Post	0.006		
	(0.017)		
Overall Religiosity Background=4 × Excess Deaths × Post	0.010		
	(0.019)		
Overall Religiosity Background=5 × Excess Deaths × Post	0.080***		
	(0.020)		
Catholic Religiosity Background=1 × Excess Deaths × Post		-0.043**	
		(0.017)	
Catholic Religiosity Background=2 × Excess Deaths × Post		-0.021	
		(0.015)	
Catholic Religiosity Background=4 × Excess Deaths × Post		0.002	
		(0.015)	
Catholic Religiosity Background=5 × Excess Deaths × Post		0.020	
		(0.016)	
Protestant Religiosity Background=1 × Excess Deaths × Post			-0.039**
			(0.017)
Protestant Religiosity Background=2 × Excess Deaths × Post			-0.021
			(0.014)
Protestant Religiosity Background=4 × Excess Deaths × Post			-0.021
			(0.016)
Protestant Religiosity Background=5 × Excess Deaths × Post			0.014
			(0.017)
County × Background FE	Yes	Yes	Yes
County × Birthyear FE	Yes	Yes	Yes
Background × Birthyear FE	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes
N. of Counties	1217	1217	1217
Observations	7641690	7641690	7641690
R^2	0.026	0.021	0.024

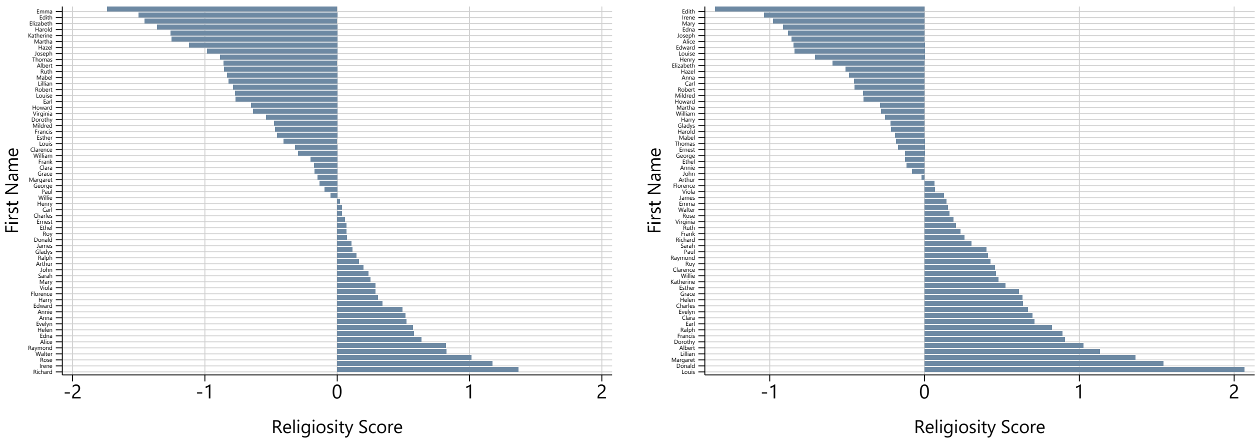
Notes: This table displays the impact of exposure to the pandemic on the polarization of religious beliefs, for all denominations. The unit of observation are children born between 1900 and 1930. "Post" is a categorical variable equal to zero for children born before the pandemic–*i.e.* before 1918–and one for those born after the pandemic–*i.e.* after 1918. The baseline treatment "Excess Deaths" is defined in equation (3.4). Background religiosity is measured as the religiosity score of the name of the head of the household, and it is discretized in quintiles. The third quintile serves as the baseline category and its coefficient is not reported. The dependent variable is overall religiosity (column 1), Catholic religiosity (column 2), and Protestant religiosity (column 3). Each regression includes county-by-background, background-by-year, and county-by-year fixed effects. Children are weighted by the inverse of the number of children within each household. Standard errors are clustered at the county level, and are reported in parentheses.

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

3.B.2 Figures

FIGURE B.1: Estimated Names Religiosity Scores, by Confession

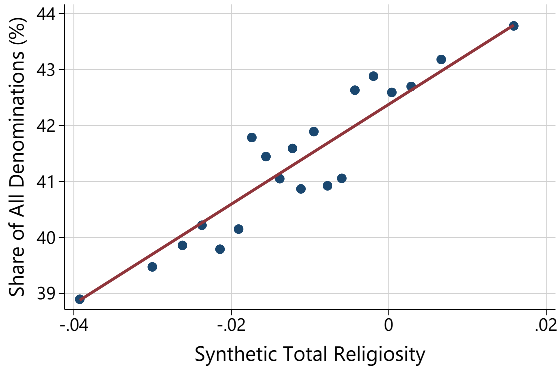
(A) Catholics (B) Protestants



Notes: The Figures display the religiosity scores estimated from model (3.2). Bars report the point estimate of each coefficient. Regressions are based on data from the 1906-1916 Censuses of Religious Bodies, and include individuals born between 1896 and 1916. We estimate religiosity scores for names appearing in at least 0.3% of the overall sample. We conflate variations of a single name together–e.g. Anne and Anna—but keep endearments separate—*e.g.,* Anna and Annie. Coefficients are reported in increasing order. Panel B.1a reports scores for Catholicism; Panel B.1b reports scores for Protestantism.

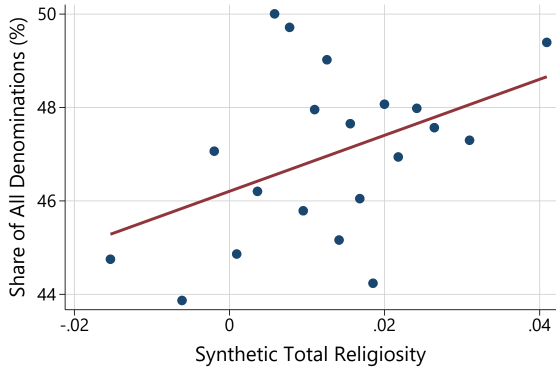
FIGURE B.2: In-sample and Out-of-sample Fit of the Religiosity Measure

(A) In-sample: All Denominations



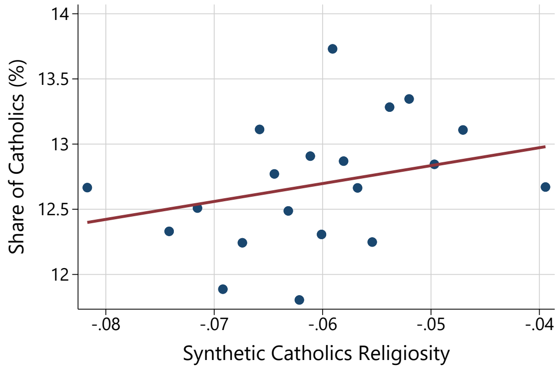
Notes. Regression Coefficient = 89.222 (Std. Err. = 12.079). R^2=0.856.

(B) Out-of-sample: All Denominations



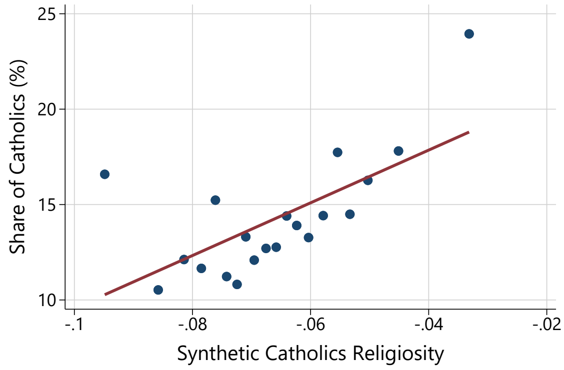
Notes. Regression Coefficient = 59.950 (Std. Err. = 31.614). R^2=0.349.

(C) In-sample: Catholics



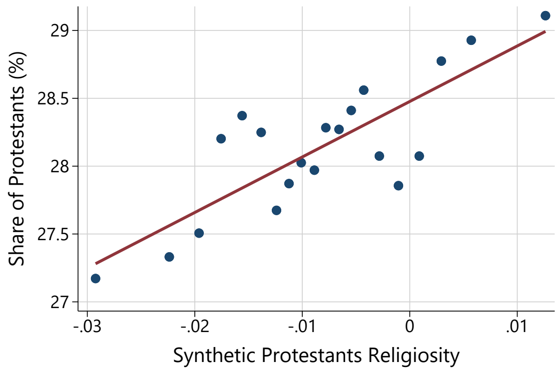
Notes. Regression Coefficient = 13.746 (Std. Err. = 12.634). R^2=0.911.

(D) Out-of-sample: Catholics



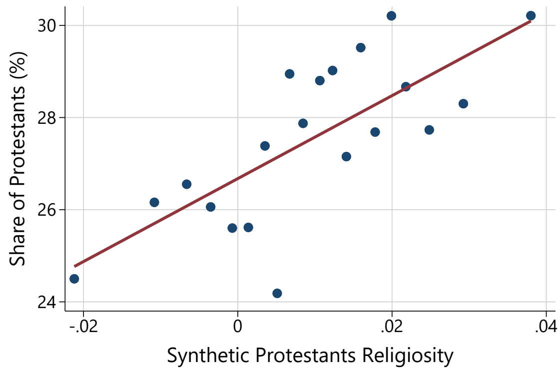
Notes. Regression Coefficient = 138.124 (Std. Err. = 41.434). R^2=0.432.

(E) In-sample: Protestants



Notes. Regression Coefficient = 40.904 (Std. Err. = 7.767). R^2=0.947.

(F) Out-of-sample: Protestants



Notes. Regression Coefficient = 90.009 (Std. Err. = 22.835). R^2=0.606.

Notes: These figures are county-level binned scatter plots reporting the correlation between our religiosity measure and the number of affiliated member to: all denominations (B.2a-B.2b), Catholicism (B.2c-B.2d) and Protestantism (B.2e-B.2f) normalized by population in 1900. In-sample figures report data for 1906 and 1916 censuses of religious affiliations. Out-of-sample figures instead report data for 1926. In-sample regressions control for county fixed effects; out-of-sample regressions include state fixed effects. Counties are weighted by their population in 1900. In the note we report the regression coefficients and the associated R^2.

FIGURE B.3: Example of Pharmaceutical Patent

(A) Text

Patented Mar. 1, 1927. 1,619,005

UNITED STATES PATENT OFFICE.

SAMUEL M. STRONG, OF GARDEN CITY, NEW YORK.

RESPIRATION AND PULSE RECORDER.

Application filed January 11, 1922. Serial No. 528,485.

This invention relates to a device or instrument for recording the character of the actions of the heart and respiratory organs of a person.

The primary object of the invention is to provide an instrument which will produce an accurate graphic representation of the rate, rhythm, and force of respiration and pulse of a human being over a short or a long period of time.

ing plate 13 and a vertical portion 19 adapted to be placed against a side plate of the casing 10. The main bearing plate fits snugly within the casing and one end of the horizontal portion 18 abuts against the cover 11 when the latter is in position. A side bearing plate 20 is located immediately adjacent to the detachable side plate 11 and an intermediate bearing plate 21 is interposed between the bearing plate 20 and the

(B) Figures

March 1, 1927. S. M. STRONG 1,619,005

RESPIRATION AND PULSE RECORDER

Filed Jan. 11, 1922 2 Sheets-Sheet 2

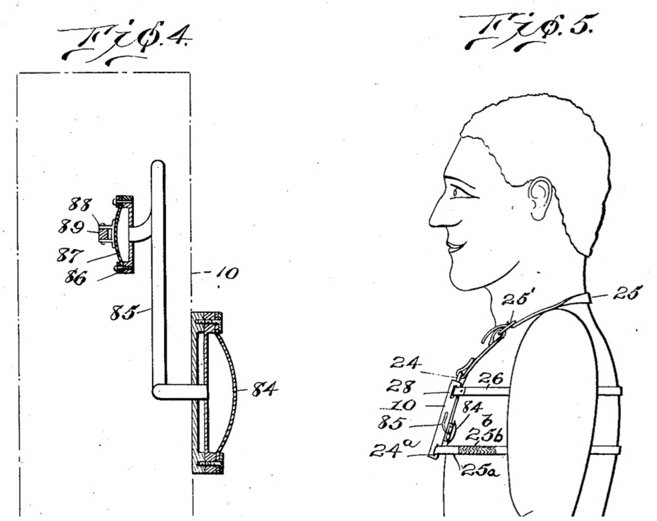
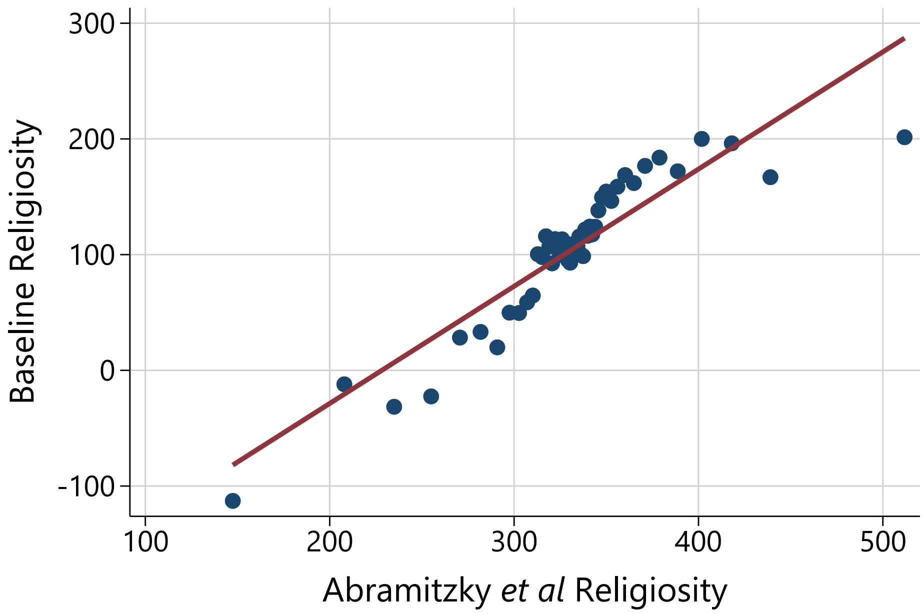


FIGURE B.4: Correlation Between WW1 and Influenza Deaths

Notes. Regression Coefficient = 0.170 (Std. Err. = 0.157). R^2=0.001.

Notes: This figure displays the correlation between WW1 and Influenza-related deaths. Gross Excess Mortality is the baseline treatment. WW1 deaths are taken as logs. In the note, we report the regression coefficient between the two variables, along with the R^2 of the model. Data on WW1 deaths are from Ferrara and Fishback (2020).

FIGURE B.5: Correlation Between Abramitzky et al. (2016) and Our Religiosity

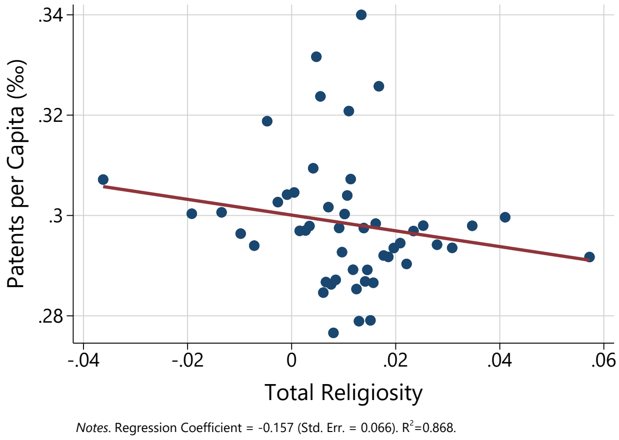


Notes. Regression Coefficient = 1.013 (Std. Err. = 0.022). R^2=0.390.

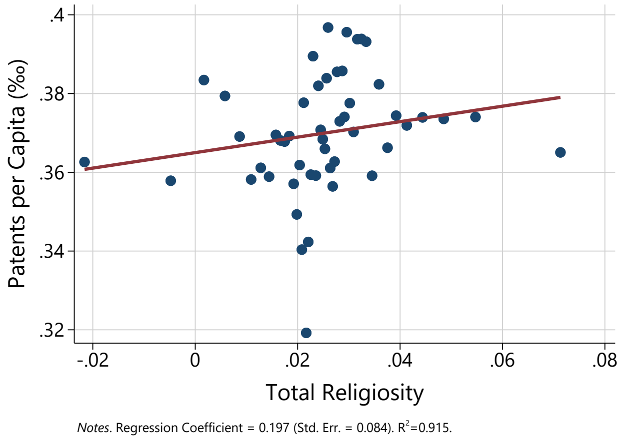
Notes: This figure reports the correlation between our baseline religiosity measure (multiplied by 100) and the share of biblical and saints names, as defined in Abramitzky *et al.* (2016). The unit of observation is a county, observed at a yearly frequency between 1900 and 1930. Counties are weighted by their population in 1900. The graph partials out county fixed effects. We report in note the regression coefficient and the associated standard error, clustered at the county level, and R^2 coefficient.

FIGURE B.6: Correlation Between Religiosity and Science

(A) Before the Great Influenza Pandemic (1910–1917)

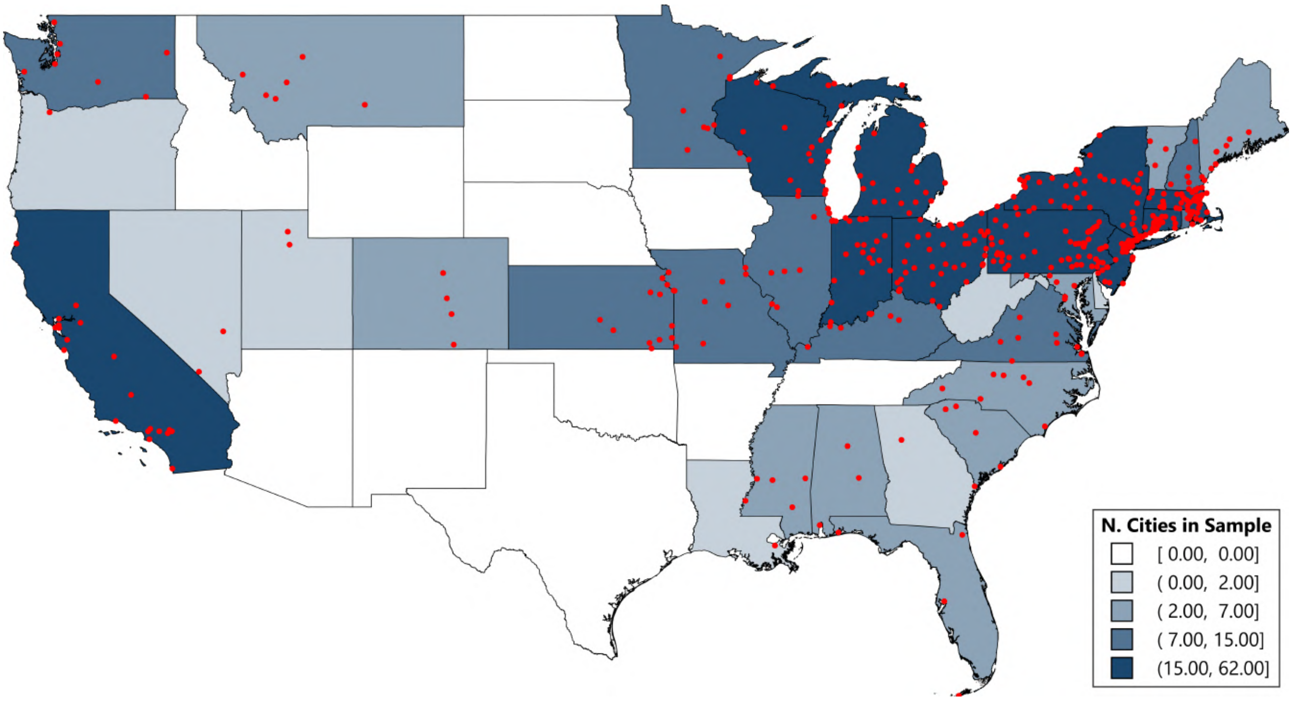


(B) After the Great Influenza Pandemic (1920–1929)



Notes: These figures display county-level binned scatter plots reporting the correlation between science–measured as patenting activity normalized by 1900-county population–and religiosity. The unit of observation is a county, observed at yearly frequency. Counties are weighted by their 1900-population. Religiosity is defined as described in section 3.3.1 and refers to overall religiosity. Graphs absorb for county and year fixed effects. We report the regression coefficients and associated R^2 separately in each graph.

FIGURE B.7: Distribution of Cities in the Alternative Sample



Notes: This figure reports the spatial distribution of the cities in the city-level sample used in Table B.7. We use data from Clay *et al.* (2019), which contains information on 483 large cities. The red dots report the coordinates of the 478 cities for which we can construct the excess mortality treatment measure. Lighter to darker shades of blue indicate the state-level number of cities included in the final sample.

Bibliography

ABRAMITZKY, R., L. P. BOUSTAN and D. CONNOR (2020). "Leaving the Enclave: Historical Evidence on Immigrant Mobility from the Industrial Removal Office." *NBER Working Paper*, (No. w27372).

ABRAMITZKY, R., L. P. BOUSTAN and K. ERIKSSON (2016). "Cultural Assimilation During the Age of Mass Migration." *NBER Working Paper*, (No. w22381).

BAZZI, S., E. BERKES, M. FISZBEIN and C. FONS-ROSEN (2022). "Individualism and Innovation." *Working Paper*.

BEACH, B., K. CLAY and M. H. SAAVEDRA (2020). "The 1918 Influenza Pandemic and its Lessons for COVID-19." *NBER Working Paper*, (No. w27673).

BERKES, E. (2018). "Comprehensive Universe of US patents (CUSP): Data and Facts." *Working Paper*.

CLAY, K., J. LEWIS and E. SEVERNINI (2019). "What Explains Cross-City Variation in Mortality During the 1918 Influenza Pandemic? Evidence from 438 US Cities." *Economics & Human Biology*, 35: 42–50.

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).

FERRARA, A. and P. V. FISHBACK (2020). "Discrimination, migration, and economic outcomes: evidence from World War I." (No. w26936).

FOUKA, V. (2020). "Backlash: The Unintended Effects of Language Prohibition in US Schools After World War I." *The Review of Economic Studies*, 87(1): 204–239.

HALL, B. H., A. B. JAFFE and M. TRAJTENBERG (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools."

KELLY, B., D. PAPANIKOLAOU, A. SERU and M. TADDY (2021). "Measuring Technological Innovation Over the Long Run." *American Economic Review: Insights*, 3(3): 303–20.

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*.

STARK, R. (1992). "The Reliability of Historical United States Census Data on Religion." *Sociological Analysis*, 53(1): 91–95.