EMIGRATION RESTRICTIONS AND ECONOMIC DEVELOPMENT

Evidence from the Italian Mass Migration to the United States*

Davide M. COLUCCIA† Lorenzo SPADAVECCHIA‡

This Version: February, 2024 – Click here to download the latest version

Abstract

We study how the restrictive immigration policy enacted by the US in 1921 impacted emigration and technology adoption in Italy. Using a shift-share instrument to predict exposure to the policy across Italian districts in a difference-in-differences setting, we show that 70% of those who would have migrated in the absence of the policy remained in Italy, resulting in a 6.2% population increase. This positive labor supply shock depressed firms' investment in productivity-enhancing technologies by approximately 20%. Consistently with a model of directed technology adoption, we find that firms substituted capital with more abundant labor, as manufacturing employment increased by 15%.

Keywords: Age of Mass Migration, Emigration, Technology Adoption.

JEL Classification: N34, O15, O33.

*This paper supersedes an earlier version circulated as "The Economic Effects of Immigration Restriction Policies: Evidence from the Italian Mass Migration to the US." We are particularly grateful to Mara Squicciarini for her continued guidance and support. We thank Maristella Botticini, Leah Boustan, 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. All errors are our own.

†Northwestern University, 2211 Campus Drive, 60208, Evanston, IL. E-mail: `davide.coluccia@northwestern.edu`.

‡Bocconi University, Via Röntgen 1, 20136, Milan, Italy. E-mail: `lorenzo.spadavecchia@unibocconi.it`.

Introduction

The diffusion of technology across countries is a major engine of global economic growth (Griffith *et al.*, 2006; Comin and Hobijn, 2011).[1] Productivity growth in developing countries, which typically operate far from the technology frontier, especially relies on the adoption of foreign technologies (Suri, 2011; Bryan *et al.*, 2014). This paper examines how out-migration, a common defining feature of the industrialization process, influences the incentive for firms in developing countries to adopt productivity-enhancing technologies, thus impacting long-term prospects of economic prosperity. Specifically, we investigate how restrictions to human mobility imposed by immigration countries influence technology adoption in the emigrants' countries of origin.[2]

The effects of out-migration on technology adoption are *ex-ante* ambiguous and potentially conflicting. On the one hand, emigration entails a loss of human capital—a "brain drain"—that may hamper the ability of countries to adopt new technologies (Kwok and Leland, 1982; Gibson and McKenzie, 2011). On the other, however, higher emigration rates may incentivize the adoption of labor-saving technologies by increasing the relative cost of labor (e.g., see Habakkuk, 1962; Hicks, 1963). From the perspective of directed technical change theory, one can interpret immigration restriction policies (IRPs) as "passive" labor market policies that increase the labor supply in targeted countries. This would, in turn, prompt the substitution of capital with more abundant—hence cheaper—labor, thus depressing investment in capital-intensive technologies. Which effect prevails is, ultimately, an empirical question.

The setting we study is the Italian mass migration, the largest episode of voluntary migration from one single country in recorded history (Choate, 2008). Between 1876 and 1925, approximately thirteen million emigrants left Italy (nearly 70% of the average Italian population in 1900); about half 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, the single most common destination country and the focus of this paper. However, the Italian mass migration to the United States abruptly ended in 1921, when Congress passed the first of a series of restrictive IRPs collectively referred to as the "Quota Acts." The Quota Acts defined numerical quotas on yearly arrivals from European countries, which drastically reduced the inflow of Italians in the US.[3] Between 1920 and 1921, the inflow of Italians in

[1] Economic historians have famously recognized that countries that industrialized relatively late, such as Germany or Italy, relied heavily on innovation produced abroad to catch up with the core industrial countries (Gerschenkron, 1962; Rosenberg, 1982). Recent literature within the tradition of endogenous growth theory embeds technology diffusion dynamics and quantifies its—substantial—contribution to productivity growth (e.g. Eaton and Kortum, 1999; Alvarez *et al.*, 2013; Buera and Oberfield, 2020; Van Patten, 2023).

[2] For brevity, we refer to those policies as immigration restriction policies (IRPs). Data from de Haas *et al.* (2015) suggest that IRPs have become increasingly common since the 1970s and currently account for 40% of the entire corpus of migration laws.

[3] 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

the United States dropped by 85% and never recovered (see Figure 1). We exploit variation arising from this sharp and massive restriction to human mobility.

Throughout this period, a group of countries, including Italy, underwent their first wave of industrialization and structural change as part of a broader transformational phenomenon known as the "Second Industrial Revolution" (Mokyr, 1998). For the first time, between 1895 and 1913, Italy outperformed the leading industrial countries. The postwar years, particularly the Fascist period, however, were marked by economic stagnation and languishing productivity growth. The economic divide between Northern and Southern areas had started to narrow during the economic boom but severely widened during the 1920s and 1930s (Cohen and Federico, 2001). Previous scholarship documents that insufficient investments, especially in the South, hampered the adoption of productivity-enhancing technologies developed in the leading industrial countries. In this paper, we argue that scarce investments partly came as a consequence of the post-1921 restrictive US immigration policy, which, therefore, plausibly contributed to the widened North-South economic gap.

To identify the effect of this immigration restriction policy shock on economic development, we define a district in Italy as more exposed to the Quotas if a larger proportion of its emigrants had moved to the United States before 1921, conditional on the overall emigration rate. In a difference-in-differences setting, we thus compare districts with similar emigration rates and leverage variation among destination countries.[4] Formally, our identification assumption thus requires that districts with similar emigration rates but different destinations would not have undergone diverging development trajectories had the Quota Acts not been enacted. We provide a battery of checks to ensure the plausibility of this assumption, but ultimately, we cannot test the conditional exogeneity of treatment intensity. We thus adapt the shift-share instrumental variable design developed by Card (2001) to construct plausibly random variation in the intensity of exposure to the Quotas. The instrument predicts district-level emigration to the US over time by interacting the initial migration flows between Italian districts and US counties with subsequent non-Italian immigration inflows across US counties. The "naïve" and the instrumented difference-in-differences designs yield quantitatively highly consistent results.

The historical setting allows us to overcome several limitations of contemporary scenarios. First, emigration seldom flows into only a few destinations; hence, observing significant restrictive policy shifts is challenging. Second, migration 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).

Existing data from official statistics are not suitable for this exercise because (i) digitized US and Italian

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, 2002).

[4]This intensity-of-treatment design is closely related to the conceptual framework adopted by Abramitzky *et al.* (2023) to study the effect of the Quota Acts on the US labor market.

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 during this period. We thus construct a dataset that links the administrative records of Italian emigrants who arrived at Ellis Island between 1892 and 1930 to their district of origin, we match a subset of these records to the US full-count population census, and we complement it with newly digitized detailed data from industrial and population censuses.

The empirical analysis proceeds in three steps. First, we establish that the Quota Acts had a tangible impact on Italian districts. Theoretically, it is possible that those who would have migrated to the United States in the absence of the policy shift could move to a different country. Looking at aggregate emigration numbers (Figure 1), this seems implausible. Emigration to the United States completely dried up after 1921, but emigration to other countries did not. The total number of emigrants in the 1920s is roughly comparable to that in the 1880s and early 1890s before the US migration gained momentum. We formally test this "imperfect substitution" argument and find that the population in districts that were conditionally more exposed to the Quota Acts increases after 1921. The effect of the Quotas on the Italian population is sizable in magnitude. Districts above median exposure to the Quotas display a 6.2% increase in population between 1921 and 1936 compared to districts below median. In other words, we estimate that approximately 70% of those who would have plausibly migrated to the United States—in the absence of the Quotas—remained in Italy. This finding is consistent with previous studies documenting the spatial persistence of Italian emigration flows (Gould, 1980b; Brum, 2019; Spitzer *et al.*, 2020). More qualitative cross-country evidence confirms that emigration outflows toward countries that did not promulgate IRPs did not increase. Hence, districts 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.

Our second finding is that manufacturing firms in provinces that were more exposed to the Quota Acts substantially decreased investment in capital goods. We consistently estimate adverse effects across several measures of capital goods, ranging from firms with at least one engine to the number of installed engines and the power they generate. For instance, districts above median exposure to the Quota shock display 19% less mechanical engines and 62% less electrical engines relative to districts below median exposure until the mid-1930s. All sectors within manufacturing exhibit similar decreases in capital investment. How can we reconcile the increasing population with the decreasing adoption of productivity-enhancing production technologies? We propose interpreting these findings through the lens of directed technological change theory (Acemoglu, 2002, 2007). As labor becomes a more abundant production input, firms are incentivized to forego investment in capital goods and employ more labor. In Online Appendix D, we develop a simple theoretical framework in the spirit of Zeira (1998) and San (2023) to show how that, in this setting, investment in labor-saving technologies would decrease in response to a positive labor supply shock

To test this interpretation, we study how employment across sectors reacted to the Quota Acts. First, we explore how agriculture and manufacturing employment responded to the policy shock at the district level. We find that manufacturing employment increases in districts more exposed to the IRP shock. Quantitatively, manufacturing employment in districts above-median exposure is 21% higher after the Quotas, while agriculture employment is 8% lower relative to districts below-median exposure. Historical evidence suggests that Italian agriculture in the 1920s was organized as a heavily labor-intensive sector (Cohen and Federico, 2001). It thus seems plausible that manufacturing could be the primary beneficiary of the substantial labor supply shock. Second, we look at how different sectors within manufacturing absorbed the population increase. Analogous to the capital results, we estimate comparable increases in manufacturing employment across industries. Thus, the aggregate and sector-level employment and capital responses to the IRP shock are consistent with the predictions of the directed technical change conceptual framework.

Overall, this paper documents that policies enacted by immigration countries to curtail migrant inflows bear important consequences on countries sending migrants. Foregone emigration, in fact, generates a positive labor supply, which, in turn, dampens technology adoption, thus potentially hampering their long-run prospects of economic growth. Despite the peculiarities of the historical setting, we discuss that the Italian migration to the United States shares similarities with contemporary emigration episodes between developing and developed countries, thus highlighting the policy-relevance of our findings.

Related Literature This paper is related to three streams of literature. First, we speak to the several contributions investigating the impact of emigration on sending countries, as opposed to the larger literature studying the economic and social effects of immigration (Clemens, 2011). 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), the diffusion of novel knowledge (Coluccia and Dossi, 2023), entrepreneurship (Anelli *et al.*, 2023), and social norms (Beine *et al.*, 2013; Bertoli and Marchetta, 2015; Tuccio and Wahba, 2018). We inform this literature by showing that emigration fosters the adoption of labor-saving technologies. We emphasize that this channel operates plausibly independently from human capital accumulation. In addition, we document that this mechanism arises in response to restrictive immigration policies enacted by receiving countries.

Second, we contribute to the literature that studies the relationship between technology adoption and the supply of production inputs. Following the seminal contributions by Hicks (1963) 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 rather than innovation and leverage exogenous variation in an instrumented difference-in-differences framework. Several studies document the importance of technology adoption as a critical driver of long-run growth, particularly in developing countries (Suri, 2011; Bryan *et al.*, 2014; Juhász *et al.*, 2020). 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.

Third, by its setting, this paper adds to the literature that studies technical change and the diffusion of novel technologies during the Age of Mass Migration. A growing number of papers examines the short-run (Arkolakis *et al.*, 2020; Moser and San, 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. Moreover, we contribute to studies examining the archetypal Italian case of mass migration. Among those, Hatton and Williamson (1998) study the aggregate determinants of Italian emigration. Spitzer *et al.* (2020) validate the Gould (1980) 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. Methodologically, we present newly digitized district-level data from Italian population and industrial censuses. In terms of new findings, we show that the mass migration was unlikely to have hampered the structural shift toward manufacturing. Our results suggest that the opposite impact prevailed: immigration *restriction* likely hampered economic modernization in Italy.

Outline of the paper We structure the paper as follows. Section I describes the Italian mass migration, the policies that shaped it, and the fundamental economic characteristics of early 20th-century Italy. Section II discusses our data-collection contribution and the sources. In Section III, we detail the empirical strategy and present our three sets of results. Section IV concludes.

I Historical Background

I.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 returns 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).

I.1.1 A Short History of the Italian 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 and other Southern and Eastern European countries didn't start experiencing mass emigration until the 1880s. The country's migration patterns over 1870–1925 display substantial time variation. Until the 1880s, its emigration rate remained relatively modest, and most migrants hailed from Northern regions. Prohibitively high transportation costs and prevailing poverty in rural Southern areas largely inhibited migration from the *Mezzogiorno*. During the 1880s, Northerners chiefly moved to neighbor 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). By the 1890s, Italy had become the global leader in sheer numbers of emigrants and in emigration rate, which grew from 5‰ in 1880 to a peak of 25‰ in 1913.

Italian emigration collapsed during World War 1 (WW1) but quickly regained momentum in the years immediately following the war. The epoch ended in the early 1920s when the U.S. Congress enacted 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 Online Appendix Table B.2, we tabulate data from official statistics on regional out-migration to the United States and other international destinations to gauge the geographical evolution of the Italian mass migration over time.

Internal migration is one last, largely overlooked component of labor mobility in Italy during the Age of Mass 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. We provide a quantitative assessment of this claim in Online Appendix Table B.3, which confirms that the number of international emigrants is one order of magnitude larger than that of internal migrants. Second, internal mobility was largely temporary and seasonal, inherently different from transoceanic migration (Gallo, 2012). Third, internal migrations reflected historically deep-rooted, persistent economic relationships between regions unlikely to influence our results on economic modernization in the 1930s.

I.1.2 Composition and Determinants of the Migratory Movements

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 provided our empirical strategy's backbone. Until the early 1880s, most 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.

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 information diffusion. Information about emigration opportunities required time to spread across the country, and this diffusion accelerated as the volume of emigration increased. 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). Spitzer and Zimran (2023) further provide evidence consistent with Gould's diffusion hypothesis. They show that emigration began in a few districts in the 1870s and 1880s and subsequently spread to nearby districts over time through immigrants' social networks. In Online Appendix A.3.3, we present some evidence that points in the same direction.

I.2 Migration Policy in Italy and the United States

An appealing feature of this context is that migratory flows from Europe to the United States remained essentially unregulated until 1921 (Abramitzky and Boustan, 2017). This section describes the key features of the historical Italian and U.S. systems of migration laws.

I.2.1 Italian Emigration Policy

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). The first such attempt was the 1888 Crispi-De Zerbi law, which introduced and regulated the emigration contract between the migrant and the migration agency. However, the law was manifestly inadequate to deal with the waves of migration that unfolded starting in the 1890s, and it effectively failed (Foerster, 1919).

As emigration to the United States gained momentum, a new emigration law was passed in 1901 under the renewed understanding that emigration was no artificial phenomenon and could positively affect Italy (Foerster, 1919). 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).

I.2.2 American Immigration Policy

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 and Boustan, 2017). 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, immigrants' economic and social conditions. The Commission's 41-volume report favored "old" immigration countries such as England and Germany over "new" ones, mainly Southern and Eastern European.

When immigration ramped up again after WWI, nativist demands for restrictions surged, and the Emergency Quota Act was passed in 1921. It was modified by the 1924 Immigration Act, which further tightened immigration restrictions on second-wave countries. The 1921 Emergency Quota Act envisaged a (temporary) annual quota of 360,000 immigrants from Europe.[5] 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, targets Southern and Eastern European countries that took part in the Mass Migration starting in the late 1890s, as advised by the report of the Dillingham Commission. Abramitzky *et al.* (2023) note that the impact of the 1924 Immigration Act on immigration was highly heterogeneous across 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 very modest. Since the 1890s, the United States had been absorbing 30% to 40% of all Italian emigration, so the Quota Acts represented a significant policy shock for Italy.

I.3 Technology Adoption and Economic Growth in Italy

Italy entered the Age of Mass Migration in the 1880s. The country was amid an agrarian crisis that followed two decades of stagnation (Toniolo, 2014). The period from 1895 to 1913 was the only time until the 1950s "economic miracle" in which Italy outperformed and narrowed the income gap with the leading industrial nations. Still in the 1920s and 1930s, however, during the Fascist period, Italy remained a mainly agricultural country, featuring low income per capita and languishing productivity growth (Cohen and Federico, 2001). 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 agrarian policies resulted in neither substantial intervention nor sizeable progress (Zamagni, 1993). Growth slowed after 1925, and regional disparities widened (Cohen and Federico, 2001). Historical evidence

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

is thus consistent with our finding that following the 1921–1924 U.S. emigration restrictions, Italy underwent a period of economic distress and rising North-South economic inequality.

We relate the migration shock to diminished investment in capital goods, especially technologically advanced ones, and a shift to labor-intensive production routines. Italy operated far from the technological frontier throughout the period, and skill premia *declined* from the 1890s onward (Vasta, 1999). Similarly to contemporary developing countries, Italy lagged behind advanced industrial nations in research-and-development expenditures, and it imported substantial amounts of foreign technology, both patents and machinery. Italian firms bundled different vintages of capital whenever possible, adding new machines to existing ones instead of renovating the whole stock (Cohen and Federico, 2001). 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 modernization process they had been undergoing before the Quota Acts.

II Data

Our analysis spans the years 1881 to 1936. We collect data from a number of sources; data are organized by census years and analyzed at the *circondario* (henceforth, "district") level if available. Otherwise, the units of observation are provinces.[6] In 1921, there were 216 districts and 69 provinces, each consisting of a variable number of municipalities (see Online Appendix section A.1 for a complete description of the data). Because districts were abolished in 1927, all subsequent district-level data are collected at the municipality level and aggregated at the 1921-district boundaries. We adopt the geographical cross-walk procedure described by Eckert *et al.* (2020) to ensure that we work with consistent 1921-border district and province geographies. Table 1 reports summary statistics for the variables in our final dataset. Online Appendix Table A.1 lists the source and coverage of each variable in the final dataset.

II.1 Emigration Data

Italian official emigration statistics provide insufficient information because out-migration flows by destination were recorded at the province level (Hatton and Williamson, 1998). This poses a major challenge because we would rather work with more granular geographical units than provinces whenever possible. This limitation

[6]Population censuses were taken in 1881, 1901, 1911, 1921, 1931, and 1936. Manufacturing censuses were taken in 1911, 1927, and 1936. 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. Districts were organized in provinces, which encompassed between one and five districts. In the Online Appendix section A.1, we discuss in more detail the sources that we digitized and present a visual summary of all the variables we analyze. In Online Appendix A.2, we describe how the final estimation samples are constructed.

precludes the use of official statistics for a spatially-detailed econometric exercise. In addition, official statistics typically relied on issued passports, which were nonetheless not required to emigrate to the US. We nonetheless digitize province-level emigration outflows from official statistics and use them to construct overall emigration rates and validate the series we derive from the individual dataset that we assemble.[7]

To overcome these issues, we collect administrative records of Italians who entered the country between 1892 and 1930 through the Ellis Island immigration station.[8] This was by far the largest, though not the only, immigration gateway during this period.[9] Administrative records report, for most migrants, name and surname, year of arrival, age, municipality of origin, and sailing ship. This study concentrates on the migration year and the municipality of origin. Ultimately, we collected approximately 2.7 million individual observations from 1890 to 1930.

The municipality variable displays frequent coding errors. We perform manual and automated trimming of the raw data and geo-code each municipality to precise geographical coordinates. We can match 1.6 million migrants to their municipality of origin. Among those, 800,000 municipalities are coded with no error. We then map each municipality to the district it belonged to in 1921 through historical GIS boundary files and compute the resulting aggregate U.S. emigration.[10] To the best of our knowledge, the Ellis Island repository is the most comprehensive source spanning the whole Age of Mass Migration for Italy at this level of aggregation.[11] Similar data have been assembled by Gray *et al.* (2019) and Spitzer and Zimran (2023).[12] Spitzer and Zimran (2018) relied on a small subset of the overall Ellis Island data in their seminal contribution. We validate this dataset with coarser data available from official statistics as detailed in Online Appendix A.3.2. There is a positive, statistically significant, and large correlation between the Ellis Island data and official statistics, which remains stable throughout the estimation period.

To conduct a Bartik-type instrumental variable analysis, we link the individual-level Ellis Island immi-

[7] We construct district-level emigration rates from official statistics. To compute a district's emigration rate, we multiply the province-level number of emigrants by the share of that district's population in the province. In other words, we assume that the emigration rate is constant among districts in the same province.

[8] These records are freely available at `heritage.statueofliberty.org`. We run queries over a comprehensive pool of 20,000 Italian surnames over the 1890–1930 period. In Online Appendix section A.3.2, we document that our constructed series correlates well with existing—albeit less granular—emigration data from official statistics.

[9] 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.

[10] Appendix A.3.1 discusses the methodological details, including the frequency of missing data on immigrants' municipality of origin.

[11] Brum (2019) produced a similar dataset for the pre-1900 period.

[12] Compared to Gray *et al.* (2019), who list 4.8 million observations, we recover a smaller proportion of migrants because we do not allow for fuzzy matches. This choice is motivated by the fact that the fuzzy match tool provided by the database search engine returns very distant and only vaguely similar results for the given queried surname. By only including exact matches, we ensure that our database does not comprise these (potentially) spurious observations.

gration records for the period 1895–1900 with the full-count US population census (Ruggles *et al.*, 2021). To the best of our knowledge, ours represents the first attempt at linking census records with the Ellis Island administrative data. Concretely, we link individuals with records of Italians who appear in the 1900 census on their name, surname, and immigration year. We detail the procedure in Online Appendix section A.3.4. Using this linked sample, we attach a municipality of origin to Italian migrants, which in turn allows us to compute migration flows from Italian districts to US counties.

Figure 1 plots the overall country-level yearly inflow of immigrants who landed in Ellis Island from 1892 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, and was definitively shut down by the Quota Acts in 1921 and 1924. Our data are consistent with comprehensive U.S. immigration data and overall Italian migration patterns (Brum, 2019; Sequeira *et al.*, 2020). In Figure 2, we plot the geographical distribution of migrants across districts. Panel 2a displays the variation in the emigrants-to-population ratio, i.e., the emigration rate. Panel 2b reports the unconditional variation in the U.S. emigrants-to-overall emigrants ratio, which is the baseline measure for treatment exposure. Online Appendix A.3.3 presents some further stylized facts that the data allows documenting.

II.2 Population and Manufacture Censuses

The population of each municipality was compiled by the Italian Statistical Office (ISTAT). We aggregate it by district and province for each census between 1881 and 1936.[13] We compute the k-urbanization rates as (i) the share of the population residing in cities with at least k-thousand inhabitants and (ii) the share of cities with at least k-thousand inhabitants. These tables also contain the altitude and area of each municipality and an indicator variable returning value one for towns near the sea. We collapse the first two at the district and province levels, weighting them by municipality population and tagging districts and provinces with access to the sea.

We construct manufacturing and agriculture employment series from disaggregated data listed in census records. These are available at the district level before 1921 and the municipality level afterwards. For consistency, we recast them at the district level throughout the sample period. Agriculture employment is not reported in the 1931 population census. Until 1921, population censuses contain sector-level employment data for manufacturing firms at the district level. After that, the same variables are reported in the manufacturing census at the province level. Thus, we can observe sector-level manufacturing employment at the province level throughout the sample period.

Manufacturing censuses contain detailed information on the quantity and quality of capital employed in

[13] Popoulation censuses were taken in 1881, 1901, 1911, 1921, 1927, and 1936. Manufacture censuses were taken in 1911, 1927, and 1936. Data in manufacturing censuses are only available at the province level.

each province by manufacturing firms in 1911, 1927, and 1937. We collect within-manufacturing sector-level 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 (David, 1990). This allows us to disentangle the possibly differential impact of the labor supply shock induced by the migration shock on different technology vintages.

II.3 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.[14] The dataset provides rich information on Italian military personnel who died during WWI. Importantly for our analysis, it includes the municipality of origin of each soldier. Because we conduct the analysis at the district level, we collapse the dataset from municipalities to 1921 districts, and we measure the war's severity in a given district as the ratio between deaths and population in 1910.

To construct the instrumental variable, we use aggregate information on US GDP and industrial production. These are constructed by Maddison (2007) and Davis (2004), respectively.

To account for changes in market access, we digitize the entire Italian railway network over the 1839–1926 period.[15] We know all the stations it is connected to for each railway section. 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, the distance it covered, and the train's traction system. We use these data to construct the Italian railway network. To capture its evolution over time, we take snapshots of the network at decade frequency.

[14] Death records were compiled 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 it in the economics literature.

[15] 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 our knowledge, this is the first paper to use these data. Compared to Ciccarelli *et al.* (2021), we do not have access to the geography of historical railway routes as we only digitize the list of stations and the year when each trunk was opened. The advantage of our data is that we can reconstruct the network until 1924, which marks the end of the transatlantic emigration, while previous studies focus on the period until 1913.

III Empirical Strategy and Results

This section presents the empirical strategy that we adopt to identify the effects of the Quota Acts and provides evidence that supports the validity of the research design. Then, we present and discuss the key findings of the paper and the possible limitations of the analysis.

III.1 Empirical Approach

The empirical analysis hinges on the observation that areas whose emigrants were more likely to settle in the United States before 1921 would be more exposed to the Quota Acts. This implicitly assumes that emigrants do not perfectly substitute the United States with other—internal or international—destinations. The first step of the analysis validates this assumption. This section, however, takes it as valid and explains how to leverage it to causally estimate the labor market effects of immigration restriction policies.

III.1.1 Measuring Exposure to the Quota Acts

In principle, districts with a larger share of emigrants headed toward the United States before 1921 would be relatively more exposed to the Quota Acts shock. Formally, let i and t denote a district and a year. Exposure to the Quota Acts would read out as follows:

$$\text{Gross Quota Exposure}_i \equiv \frac{\text{US Emigrants}_i}{\text{Population}_{i,1880}} \tag{1}$$

where US Emigrants_i and $\text{Population}_{i,1880}$ denote, respectively, the number of emigrants who settled in the US and who originated in district i between 1900 and 1914, and population before the mass migration begun (in 1880). The baseline definition (1) includes the years 1900-1914 because municipality data before 1900 and after 1921 are less comprehensive (see Appendix A). We also exclude the years of WW1 because the large drop in emigration rates during the war years could be correlated with district characteristics, such as distance from emigration ports. In the robustness checks, we enlarge the period used to construct the treatment.

However, if we were to leverage variation in (1), we would be comparing districts with very different emigration rates. If the decision to emigrate were correlated with other—possibly unobserved—characteristics, then the resulting estimates would conflate this underlying spurious association. We thus decompose (1) in two margins as follows:

$$\text{Gross Quota Exposure}_i = \underbrace{\frac{\text{US Emigrants}_i}{\text{Emigrants}_i}}_{\substack{\text{Quota Exposure} \\ (\equiv \text{QE}_i)}} \times \underbrace{\frac{\text{Emigrants}_i}{\text{Population}_{i,1880}}}_{\substack{\text{Extensive Margin} \\ (\equiv \text{EM}_i)}} \tag{2}$$

The first term denotes US emigrants as a share of the total number of emigrants between 1900 and 1914; the second term, in turn, expresses the overall number of emigrants as a share of the pre-mass migration population over the same time period.

Evidently, (1) and (2) are entirely analogous. However, recasting the definition of Quota exposure in (2) is useful for illustrating the empirical strategy. Our key idea is to exploit variation in the intensive margin of US emigration (QE_i), holding fixed the extensive margin (EM_i). In other words, the identifying variation arises by comparing districts with similar emigration rates, whose emigrants settle in different destination countries.

III.1.2 Baseline Difference-in-Differences Model

Throughout the paper, we estimate variations of the following double-differences (DiD) Poisson regression:

$$\ln \mathbb{E}\left(y_{it} \,|\, \mathbf{X}_{it}\right) = \alpha_i + \beta_t + \gamma \times \text{EM}_i \times \text{I}(t \geq 1921) + \sum_{\tau \neq -1} \delta_\tau \times \text{QE}_i \times \text{I}(t = 1921 - \tau) \tag{3}$$

where $\mathbb{E}(\cdot)$ denotes the expected value of a generic outcome y conditional on a set of baseline controls $\mathbf{X}$; terms α_i and β_t capture, respectively, unit and time fixed effects, and $\text{I}(\cdot)$ are indicators for each period since the Quota shock. The term (QE) is the cumulative net Quota Exposure metric defined in (2). We estimate (3) as a Poisson regression to account for the non-normality of the outcome and the presence of zeros, particularly when looking at capital investment outcomes. Importantly, equation (3) controls for an interaction term between the cumulative emigration rate (EM) and a post-Quota indicator variable. This ensures that the DiD estimators $\{\delta_\tau\}$ compare districts with similar emigration rates. In the baseline analysis, $\mathbf{X}$ contains an indicator variable for southern regions interacted with a post-Quota indicator to account for potential diverging North-South trends after the Quotas. In several robustness, however, we enrich the set of included controls.

The identification assumption that model (3) requires can be stated in terms of conventional parallel trends. Absent the Quotas, units with conditionally more US emigrants would not have displayed different patterns in y compared with districts with fewer US emigrants. While this assumption is not testable, we estimate pre-treatment coefficients that are never statistically different from zero throughout the paper. This supports the parallel trends assumption.

In Table 2, however, we provide one additional exercise to gauge the validity of the research design. In each line, we report the correlation between district-level variables from population censuses (Panel A), province-level manufacturing employment by sector (Panel B), province-level capital variables (Panel C), the Quota Exposure measure (in columns 1–6), and the instrumental variable defined in (4). In columns (1–2) and (7–8), the variables are measured in 1901; in columns (3–4) and (9–10), they are measured in 1911. Importantly, identification in a difference-in-differences setting requires that the treatment does not correlate with *changes* in pre-determined characteristics, which may impact the outcomes of interest. In columns (5–6) and (11–12), we thus compute the correlation between the growth rate of each variable between 1901 and 1911 and the two treatment indicators. These indicate that observation units appear comparable along essentially all observed variables in growth rates, despite substantial differences in levels. Since we do not observe two pre-treatment values for capital variables, in Panel C, we document the absence of any statistically significant

correlation between their level in 1911 and the treatments. Although imperfect, data limitations preclude a more appropriate comparison in changes.

To avoid the pitfalls of the DiD estimator with continuous treatments reported by Callaway *et al.* (2021), we always report estimates of model (3) where the intensive margin is coded as a binary variable returning value one for districts above the median value, and zero for districts below the median.

III.1.3 Instrumental Variable for Quota Exposure

To provide more solid evidence in favor of a causal interpretation of our estimates, we construct a simple shift-share instrument to predict actual US emigration. Our approach mirrors a widely employed methodology to estimate the causal impact of immigration (Card, 2001; Tabellini, 2020). The instrument predicts district-level out-migration flows to the United States by interacting the 1895–1900 migration ties between Italian districts and US counties with subsequent county-level immigrant inflows from countries other than Italy. Formally, the actual US emigration is instrumented with

$$\widehat{\text{US Emigrants}}_{it} \equiv \sum_j \omega_{ij} \times \text{Immigrants}_{jt}^{-\text{Italy}} \tag{4}$$

where ω_{ij} denotes the number of immigrants from district i into county j between 1895 and 1900, while the variable $(\text{Immigrants}_{jt}^{-\text{Italy}})$ is the number of non-Italian immigrants who settle in county j in year t, expressed as a share of the overall number of non-Italian immigrants who enter the US in year t. To compute the exposure share terms (ω_{ij}), we rely on our novel dataset that links Ellis Island immigrant records with the full-count US population census (Ruggles *et al.*, 2021). We construct bilateral flows between districts and counties by attaching a municipality of origin to Italian immigrants recorded in the US census. In regression (3), we substitute the baseline treatment indicator (QE_i) with the ratio between the predicted stock of US emigrants from (4) and the overall number of emigrants to obtain a predicted measure of Quota Exposure $(\widehat{\text{QE}}_i)$.

Our key identification assumption when using the instrumental variable in the double differences setup is that the initial sorting of emigrants across US counties, weighted by the pull shocks represented by non-Italian immigration inflows, does not correlate with unobserved factors that may impact the *changes* in the outcome variables in the same periods. In contrast to a standard instrumental-variable estimation setting, we do not require that the instrument does not correlate with the *levels* of such variables, as in Anelli *et al.* (2023). While this assumption cannot be tested, in Table 2, we show that the instrument does not systematically correlate with changes in the outcome variables before the Quotas. This imperfect test supports the validity of this research design.

In Figure 3, we check that actual and predicted US emigration rates are positively correlated. From (4), we construct a yearly district-level series of the predicted US emigration rate. In the baseline regression (3), however, we exploit cross-sectional variation arising from the overall US emigration rate over time. We

thus check the instrument's relevance in a panel setting, where each district is observed yearly, and in the cross-section, where each district is observed once. In Panel (3a), we report the first exercise: red dots only include the baseline controls included in (3), whereas the blue dots include district fixed effects. Instead, panel (3b) refers to the second exercise. Here, we cannot include district-level fixed effects because the underlying variation is cross-sectional. Hence, the blue dots include region-fixed effects. Both exercises reveal a strong, positive, and significant correlation between the instrument and observed US emigration rates. We tabulate this exercise and additional checks in Online Appendix Table B.1.

III.2 Population Increases in Districts More Exposed to the Quota Acts

The first step of our argument maintains that areas more exposed to the US Quota Acts experienced an unexpected population increase. Implicitly, this requires that not all those who would have migrated to the United States had the Quotas not been promulgated would settle in a different country. This "imperfect substitution" argument can be tested and quantified empirically. If the US and other countries were perfectly substitutable, we would expect to find no effect of exposure to the Quotas on population.

To assess the validity of the imperfect substitution hypothesis, we estimate model (3) using population as the outcome variable. Table 3 reports the results. We find that districts more exposed to the IRP shock display a larger population after 1921. The effect remains quantitatively unchanged if we focus on the subsample of Southern districts, where exposure to the Quotas was generally higher (column 2).[16] Importantly, in columns (3–4), we repeat the estimation using a binary treatment for districts below and above the median US emigration rate to avoid issues related to continuous treatment DiD designs. Finally, in columns (5) and (6), we include region-by-decade and province-by-decade fixed effects to account for time-varying unobserved heterogeneity non-parametrically. All these specifications yield quantitatively consistent results. For instance, as shown in column (3), districts above-median Quota exposure display a 6.2% increase in population compared to districts below-median exposure. To put it differently, we estimate that approximately 70% of foregone migrants remained in more severely affected areas.[17]

The validity of the DiD design hinges on a standard parallel trends assumption. This requires that if the US had not promulgated the Quota Acts, the population would have evolved similarly in districts with relatively higher or lower US emigration rates. To gauge the sensibility of this assumption, we estimate model (3) as a generalized DiD using the binary treatment interacted with time dummies. Figure 4 reports the estimated dynamic treatment effects. We find that districts with above- and below-median exposure to the Quotas did not

[16] By "Southern" districts, we mean areas in the NUTS-2 ITF and ITG regions, which correspond to the former Kingdom of the Two Sicilies.

[17] To obtain this figure, we first compute the average yearly number of foregone emigrants by multiplying the population in the treatment group by the estimated average treatment effect and normalizing by the number of post-treatment years. We divide this figure by the number of US emigrants in treatment districts in the pre-Quota period, normalizing by the number of pre-treatment periods.

statistically differ before 1921. This supports the parallel trends assumption. We estimate a positive effect of the Quotas already in 1921, which persists for at least 15 years. The immediacy of the impact of the Quotas is not surprising. As highlighted in Figure 1, US emigration collapsed following WW1. Moreover, the provisions of the Act became strictly and immediately enforced (Abramitzky and Boustan, 2017).

We repeat the exercise using the predicted US emigration rates from (4) to reduce further concerns about treatment non-excludability. Columns (7–8) report the estimated treatment effects, which confirm the baseline results.

III.3 Capital Investments Decrease in Districts More Exposed to the Quota Acts

How did the increased population interact with technology adoption and investment? In this section, we show that investment in labor-saving and possibly productivity-enhancing technologies decreased in areas more exposed to the shock. To do so, we estimate model (3) using several proxies of capital investment as outcome variables. We distinguish between the overall number of firms, the number of firms with at least one installed engine, mechanical and electrical engines, and the mechanical and electrical horsepower generated by the respective installed engines. The data cover manufacturing firms only.

Table 4 reports the estimated effect of exposure to the Quotas on capital investment. Panel A refers to the observed US emigration rate, whereas in Panel B, the treatment is the predicted values from (4). We find that investment in physical capital decreased in areas more exposed to the Quotas. This finding is confirmed across all the imperfect proxies available. In particular, we estimate larger treatment effects for electrical engines and horsepower. This is particularly striking, as David (1990) and Mokyr (1998) note that electrical engines stood as a major defining technology of the Second Industrial Revolution, which could yield sizable productivity advantages. Gaggl *et al.* (2021), moreover, show that electrification in the US was a catalyst for urbanization and structural transformation. In terms of magnitude, we estimate that, relative to areas below-median exposure to the Quota Acts, districts above-median display a 23% decrease in the number of active firms, a 18% drop in the number of firms using at least one engine, a 21% decrease in mechanical and 62% less electrical engines, and a 16% and 36% drop in horsepower energy produced by mechanical and electrical engines, respectively.

Figure 5 displays the associated dynamic DiD coefficients. Unfortunately, the data are available for three points in time only. Of these, 1911 is the only pre-treatment observation. We thus cannot estimate pre-treatment coefficients. Instead, in Table 2, we show that treated and control provinces were similar along all outcome variables except mechanical engines in 1911. All variables decreased in 1927: the effect remains persistent for electrical engines, while it appears to be shorter-lived for mechanical ones.

Thus far, we grouped all manufacturing sectors. Nonetheless, the data allows us to undertake a more disaggregated analysis. We report the results in Figure 6. Each "Capital" panel reports the estimated effect of the Quotas on each capital indicator for a given manufacturing sector. For readability, the coefficients are

all standardized. The reaction to the population shock does not exhibit sizable heterogeneity across industries. Capital investment decreases in all sectors except for chemical manufacturing.

III.4 The Quota Acts as Passive Labor Market Policies

Why did investments in capital and technology decrease in areas more exposed to the Quota Acts? We advance and validate the hypothesis that the immigration restriction shock triggered directed technical adoption mechanics *à la* Zeira (1998) and Acemoglu (2002). Under the imperfect substitution hypothesis, which we validated in section III.2, areas that had been sending relatively more migrants to the United States are, in fact, subject to a larger labor supply shock because a relatively larger fraction of people who would have emigrated are prevented from doing so. This positive labor supply shock decreases wages, which triggers firms' incentive to substitute capital with labor. We formalize this argument in Online Appendix section D.

To test this mechanism more formally, we study how manufacturing and agriculture employment reacted to the shock. We look at manufacturing and agriculture because, together, they account for more than 80% of overall employment. In section I, we noted that Italian agriculture in this period was labor-intensive and not mechanized (Cohen and Federico, 2001). The directed technical adoption narrative would thus predict that the labor supply shock generated by the Quota Acts would primarily flow into increased manufacturing employment.

We test these hypotheses in Table 5, which reports the results of the baseline DiD regression using employment in manufacture (Panel A) and agriculture (Panel B) as the dependent variables. Manufacturing employment increases in areas more exposed to the immigration restrictions shock. This result holds within the sub-sample of Southern districts (column 2), re-coding the treatment as a binary indicator (columns 3–4) and including region-by-decade and province-by-decade fixed effects. Importantly, these hold when using the predicted exposure to the Quota Acts instead of the measured treatment. In contrast, in Panel B, we find evidence indicating that agriculture employment decreased in response to the Quota Acts. Therefore, it appears that the labor supply shock generated by the Quotas prompted a reallocation of labor from agriculture, a largely labor-intensive activity, to manufacturing. Overall, the evidence presented in the table provides solid evidence that manufacturing employment increased in areas more exposed to the Quotas. From a quantitative perspective, manufacturing employment in districts above-median treatment exposure increases by 21%, while agriculture employment decreases by approximately 9%.

As usual, we report the flexible difference-in-differences estimates in Figure 7. We confirm that treated and control districts were not statistically different before 1921 in terms of either manufacturing or agriculture employment. Manufacturing employment increased after the Quotas. The effect is persistent as it lasts at least until 1936. In this specification, agriculture employment decreases in districts more exposed to the treatment. Because manufacturing employment increases proportionally more than population, our estimates may indicate

some reallocation of the workforce out of the agriculture sector into manufacturing.

We can break up manufacturing employment by sector, although at the province aggregation level. As in section III.3, we do not uncover substantial heterogeneity across sectors. Employment increases in all industries, although marginally less so in textiles. Overall, this confirms that a positive labor supply shock is associated with a decrease in capital investment in all manufacturing sectors.

III.5 Robustness Checks

This section summarizes the robustness checks, which we presented in Appendix C. We provide evidence that the baseline estimates are robust to alternative definitions of quota exposure, to the inclusion of several covariates controlling for both push and pull factors, to the exclusion of specific parts of the sample, and to the use of different methods to estimate the standard error.

Since measured quota exposure may be subject to mismeasurement, in Tables C.1, C.2, C.3, and C.4 we show that the results are robust to alternative definitions. The first concern is that our results might be driven either by remote migration patterns or by more recent migration closer to the introduction of the quotas. We address these concerns by constructing two measures of quota exposure which assign increasing or decreasing weights on more recent out-migration flows to the US. In addition, we also exclude all migrants to the US that left before 1900. As a further test, we construct our measured quota exposures using migration to the US that happened before the first Quota Act in 1921. Last, as discussed in Section II, though emigration collapsed during WW1, it did not completely dry out. During the war, districts closer to emigration ports are, in fact, disproportionately represented relative to previous shares, possibly because of their geographic position. To control for this, we restrict the sample of US emigration to the years before the outbreak of WWI. In all cases, we find that all our baseline results hold.

In Tables C.5, C.6, C.8, and C.9 we show that our baseline results are robust to the inclusion of a large set of covariates measured before the Acts, interacted with a post-Quota indicator variable, as further controls. These unit-specific control variables are the literacy rate in 1901, measured as the share of people who could read relative to the district's overall population; the urbanization rate in 1901, measured as the share of people living in towns with more than 5'000 inhabitants in the district; the altitude at which the district is located; an indicator variable that returns a value of one if the district has access to the railway network before 1901;[18] the number of deaths due to World War One relative to population in 1901.[19] In order to control for pull factors, in some specifications, we explicitly control for an interaction between US GDP, which serves as an indicator of the business cycle, and quota exposure. All results remain quantitatively unchanged.

[18] For province-level analyses, we use the share of municipalities in the province that had access to the railway network before 1901

[19] Shares are expressed relative to the population in 1911 for province-level regression, as 1911 is the first decade of observation for analyses at that level.

In the above-mentioned Tables and in Table C.7, we also show the robustness of our results to these specifications when using the shift-share instrument as treatment. In addition, we test that our results are robust to the exclusion of specific parts of the sample. In Figures C.2 and C.4, we show that the baseline estimates are not driven by any specific region, as regression coefficients remain stable irrespective of the region that we exclude from the estimation sample.

In Figures C.1 and C.3, we compare the baseline estimated standard errors with a battery of alternative estimators. Among others, we employ the correction suggested by Conley (1999) to allow for time and spatial autocorrelation. The significance of the baseline results remains largely unaltered.

III.6 Discussion and Limitations of the Analysis

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 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 estimate employment gains and capital investment losses in First Industrial Revolution, traditionally low-skilled and labor-intensive sectors. 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 C.5, C.6, C.8 and C.9. 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 and given data limitations, we abstracted from including remittances in our analysis. Remittances 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., Fernández, 2022). 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.

One 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 against this interpretation throughout this paper. Historical scholarship, however, notes that assimilation patterns of Italian immigrants in the United States and Argentina during this period substantially differed (Klein, 1983).[20] 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.

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

[20] 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.

poor countries, whereas it is starkly different from emigration from rich countries (e.g., Gibson *et al.*, 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.

IV Conclusions

The adoption of foreign technology is a crucial driver of economic growth for developing countries, which typically operate far from the technology frontier (Eaton and Kortum, 1999; Suri, 2011). This paper explores the relationship between technology adoption and out-migration, a common feature of the development process.

We study the Italian mass migration to the United States between 1892 and 1936 using individual-level data on Italian immigrants and newly digitized census data. Between 1921 and 1924, the U.S. promulgated two immigration restriction policies—the "Quota Acts"—which completely halted the inflow of Italian immigrants. 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. Moreover, we produce a novel sample of Ellis Island immigrants linked to the US full-count population census to construct a shift-share instrumental variable, which confirms the baseline results.

We document three facts. First, we find that the population increased in areas that were comparatively more exposed to the Quota Acts. This finding supports an "imperfect substitution" narrative whereby immigration restriction policies generate foregone emigration because those who would have migrated do not—or cannot—substitute the restricted location with alternative destination countries. We thus interpret immigration restriction policies as "passive" labor market policies that increase the supply of labor in the countries they target. Second, we show that firms in treated locations decreased capital investment and technology adoption, particularly in relatively more advanced technology vintages. How can we reconcile a positive labor supply shock with depressed investment in productivity-enhancing capital? We argue that the immigration restriction policy triggered directed technical change incentives (Zeira, 1998; Acemoglu, 2002). Firms face an incentive to substitute capital with more abundant, hence cheaper labor. To validate this hypothesis, we show that manufacturing employment increased in the districts more exposed to the Quotas. These dynamics operated in each manufacturing sector.

This paper presents novel evidence that out-migration can act as a driver of technology adoption, thus potentially empowering long-run economic growth. We emphasize that this channel operates plausibly independently on "brain drain" effects. It also suggests that immigration restriction policies enacted by immigration—typically developed—countries may hamper the modernization of emigration—typically developing—ones. Our findings thus inform policymakers about the potential long-term consequences of such policies.

References

ABRAMITZKY, R., P. AGER, L. BOUSTAN, E. COHEN and C. W. HANSEN (2023). "The Effect of Immigration Restrictions on Local Labor Markets: Lessons from the 1920s Border Closure." *American Economic Journal: Applied Economics*, 15(1): 164–191. (Cited on p. 2, 8)

ABRAMITZKY, R. and L. BOUSTAN (2017). "Immigration in American Economic History." *Journal of Economic Literature*, 55(4): 1311–1345. (Cited on p. 2, 7, 8, 17)

ACEMOGLU, D. (2002). "Directed Technical Change." *The Review of Economic Studies*, 69(4): 781–809. (Cited on p. 3, 18, 22, D29)

——— (2007). "Equilibrium Bias of Technology." *Econometrica*, 75(5): 1371–1409. (Cited on p. 3, D29)

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. (Cited on p. 12)

AKCIGIT, U., J. GRIGSBY and T. NICHOLAS (2017). "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age." (Cited on p. 5)

ALVAREZ, F. E., F. J. BUERA and J. LUCAS, ROBERT E. (2013). "Idea Flows, Economic Growth, and Trade." (Cited on p. 1)

ANDERSSON, D., M. KARADJA and E. PRAWITZ (2022). "Mass Migration and Technological Change." *Journal of the European Economic Association*, 20(5): 1859–1896. (Cited on p. 4, 5)

ANELLI, M., G. BASSO, G. IPPEDICO and G. PERI (2023). "Emigration and Entrepreneurial Drain." *American Economic Journal: Applied Economics*, 15(2): 218–252. (Cited on p. 4, 15)

ARKOLAKIS, C., M. PETERS and S. K. LEE (2020). "European Immigrants and the United States' Rise to the Technological Frontier." (Cited on p. 5)

BANDIERA, O., I. RASUL and M. VIARENGO (2013). "The Making of Modern America: Migratory Flows in the Age of Mass Migration." *Journal of Development Economics*, 102: 23–47. (Cited on p. 5)

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. (Cited on p. 4)

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. (Cited on p. 4)

BERTOLI, S. and F. MARCHETTA (2015). "Bringing It All Back Home – Return Migration and Fertility Choices." *World Development*, 65: 27–40. (Cited on p. 4)

BRUM, M. (2019). "Italian Migration to the Unites States: The Role of Pioneers' Locations." (Cited on p. 3, 10, 11)

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. (Cited on p. 1, 5)

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

(Cited on p. 1)

BURCHARDI, K. B., T. CHANEY, T. A. HASSAN, L. TARQUINIO and S. J. TERRY (2020). "Immigration, Innovation, and Growth." (Cited on p. 5)

CALLAWAY, B., A. GOODMAN-BACON and P. H. C. SANT'ANNA (2021). "Difference-in-Differences with a Continuous Treatment." (Cited on p. 15)

CARD, D. (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64. (Cited on p. 2, 15)

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. (Cited on p. 4)

CHOATE, M. I. (2008). "Emigrant Nation: The Making of Italy Abroad." In "Emigrant Nation," Harvard University Press. (Cited on p. 1, 5)

CICCARELLI, C., C. MAGAZZINO and E. MARCUCCI (2021). "Early Development of Italian Railways and Industrial Growth: A Regional Analysis." *Research in Transportation Economics*, 88: 100916. (Cited on p. 12)

CLEMENS, M. A. (2011). "Economics and Emigration: Trillion-Dollar Bills on the Sidewalk?" *Journal of Economic Perspectives*, 25(3): 83–106. (Cited on p. 4, 20)

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): 1468–1487. (Cited on p. 4)

COHEN, J. and G. FEDERICO (2001). *The Growth of the Italian Economy, 1820-1960*. Cambridge University Press. (Cited on p. 2, 4, 8, 9, 18)

COLUCCIA, D. M. and G. DOSSI (2023). "Return Innovation: The Knowledge Spillovers of the British Migration to the United States, 1870–1940." (Cited on p. 4, A3)

COMIN, D. and B. HOBIJN (2011). "Technology Diffusion and Postwar Growth." *NBER Macroeconomics Annual*, 25: 209–246. (Cited on p. 1)

CONLEY, T. G. (1999). "GMM estimation with cross sectional dependence." *Journal of Econometrics*, 92(1): 1–45. (Cited on p. 20)

DAVID, P. A. (1990). "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox." *American Economic Review*, 80(2): 355–361. (Cited on p. 12, 17)

DAVIS, J. H. (2004). "An Annual Index of U. S. Industrial Production, 1790–1915." *The Quarterly Journal of Economics*, 119(4): 1177–1215. (Cited on p. 12)

DE HAAS, H., K. NATTER and S. VEZZOLI (2015). "Conceptualizing and measuring migration policy change." *Comparative Migration Studies*, 3(1): 15. (Cited on p. 1)

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. (Cited on p. 5)

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. (Cited on p. 2, 4)

EATON, J. and S. KORTUM (1999). "International Technology Diffusion: Theory and Measurement." *International Economic Review*, 40(3): 537–570. (Cited on p. 1, 22)

ECKERT, F., A. GVIRTZ, J. LIANG and M. PETERS (2020). "A Method to Construct Geographical Crosswalks with an Application to US Counties since 1790." (Cited on p. 9, A2)

FERNÁNDEZ, M. (2022). "Mass Emigration and Human Capital over a Century: Evidence from the Galician Diaspora." (Cited on p. 21)

FOERSTER, R. F. (1919). *The Italian Emigration of Our Times*. Harvard University Press. (Cited on p. 7)

GABACCIA, D. R. (2013). *Italy's Many Diasporas: Elites, Exiles, and the Workers of the World*. Routledge. (Cited on p. 7)

GAGGL, P., R. GRAY, I. MARINESCU and M. MORIN (2021). "Does electricity drive structural transformation? Evidence from the United States." *Labour Economics*, 68: 101944. (Cited on p. 17)

GALLO, S. (2012). *Senza attraversare le frontiere: Le migrazioni interne dall'Unità a oggi*. Gius.Laterza & Figli Spa. (Cited on p. 6)

GERSCHENKRON, A. (1962). *Economic Backwardness in Historical Perspective:*. Cambridge, MA: Belknap Press. (Cited on p. 1)

GIBSON, J. and D. MCKENZIE (2011). "Eight Questions about Brain Drain." *Journal of Economic Perspectives*, 25(3): 107–128. (Cited on p. 1)

GIBSON, J., D. MCKENZIE and S. STILLMAN (2011). "The Impacts of International Migration on Remaining Household Members: Omnibus Results from a Migration Lottery Program." *The Review of Economics and Statistics*, 93(4): 1297–1318. (Cited on p. 22)

GOULD, J. D. (1980a). "European inter-continental emigration. The road home: return migration from the U.S.A." *The Journal of European Economic History*. (Cited on p. 5, A4, A5, A11)

——— (1980b). "European inter-continental emigration: the role of "diffusion" and "feedback"." *The Journal of European Economic History*. (Cited on p. 3, 7)

GRAY, R., G. NARCISO and G. TORTORICI (2019). "Globalization, Agricultural Markets and Mass Migration: Italy, 1881–1912." *Explorations in Economic History*, 74: 101276. (Cited on p. 10)

GRIFFITH, R., R. HARRISON and J. VAN REENEN (2006). "How Special Is the Special Relationship? Using the Impact of U.S. R&D Spillovers on U.K. Firms as a Test of Technology Sourcing." *American Economic Review*, 96(5): 1859–1875. (Cited on p. 1)

HABAKKUK, H. J. (1962). *American and British Technology in the Nineteenth Century: The Search for Labour Saving Inventions*. Cambridge University Press. (Cited on p. 1, 4)

HANLON, W. W. (2015). "Necessity Is the Mother of Invention: Input Supplies and Directed Technical Change." *Econometrica*, 83(1): 67–100. (Cited on p. 4)

HATTON, T. and J. G. WILLIAMSON (1998). "The Age of Mass Migration: Causes and Economic Impact." OUP Catalogue, Oxford University Press. (Cited on p. 1, 5, 6, 7, 9)

HICKS, J. (1963). "The Theory of Wages." Palgrave Macmillan Books, Palgrave Macmillan. (Cited on p. 1, 4)

HIGHAM, J. (2002). *Strangers in the Land: Patterns of American Nativism, 1860-1925*. Rutgers University Press. (Cited on p. 2)

HORNBECK, R. and S. NAIDU (2014). "When the Levee Breaks: Black Migration and Economic Development in the American South." *American Economic Review*, 104(3): 963–990. (Cited on p. 4)

JUHÁSZ, R., M. P. SQUICCIARINI and N. VOIGTLÄNDER (2020). "Technology Adoption and Productivity Growth: Evidence from Industrialization in France." (Cited on p. 5)

KAPUR, D. (2014). "Political Effects of International Migration." *Annual Review of Political Science*, 17(1): 479–502. (Cited on p. 4)

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. (Cited on p. 4)

KEELING, D. (1999). "The Transportation Revolution and Transatlantic Migration, 1850–1914." *Research in Economic History*, 19: 39–74. (Cited on p. 6)

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. (Cited on p. 21)

KWOK, V. and H. LELAND (1982). "An Economic Model of the Brain Drain." *American Economic Review*, 72(1): 91–100. (Cited on p. 1)

LEWIS, E. (2011). "Immigration, Skill Mix, and Capital Skill Complementarity." *The Quarterly Journal of Economics*, 126(2): 1029–1069. (Cited on p. 4)

MADDISON, A. (2007). *Contours of the World Economy 1-2030 AD: Essays in Macro-Economic History*. OUP Oxford. (Cited on p. 12, A7)

MOKYR, J. (1998). "The Second Industrial Revolution, 1870-1914." In "Storia dell'Economia Mondiale," pp. 219–245. Rome (Italy): Laterza. (Cited on p. 2, 17)

MOSER, P. and S. SAN (2020). "Immigration, Science, and Invention. Lessons from the Quota Acts." (Cited on p. 5)

OTTINGER, S. and L. ROSENBERGER (2023). "The American Origin of the French Revolution." (Cited on p. 4)

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. (Cited on p. 5, 21)

RAPOPORT, H. and F. DOCQUIER (2006). "The Economics of Migrants' Remittances." In S.-C. Kolm and J. M. Ythier, eds., "Handbook of the Economics of Giving, Altruism and Reciprocity," volume 2 of *Applications*, pp. 1135–1198. Elsevier. (Cited on p. 20)

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

(Cited on p. 1)

ROSOLI, G. (1998). *La political migratoria dell'Italia dall'Unità al Fascismo*, volume 32. Annali della Fondazione Luidi Einaudi. (Cited on p. 7)

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." (Cited on p. 11, 15, A3, A5)

SAN, S. (2023). "Labor Supply and Directed Technical Change: Evidence from the Termination of the Bracero Program in 1964." *American Economic Journal: Applied Economics*, 15(1): 136–163. (Cited on p. 3, D29)

SEQUEIRA, S., N. NUNN and N. QIAN (2020). "Immigrants and the Making of America." *The Review of Economic Studies*, 87(1): 382–419. (Cited on p. 5, 11, 21)

SMITH, D. M. (1997). *Modern Italy: A Political History*. Yale University Press. (Cited on p. 6)

SORI, E. (1979). *L'emigrazione italiana dall'Unità d'Italia alla Seconda Guerra Mondiale*. Bologna (IT): Il Mulino. (Cited on p. 6)

SPILIMBERGO, A. (2009). "Democracy and Foreign Education." *American Economic Review*, 99(1): 528–543. (Cited on p. 4)

SPITZER, Y., G. TORTORICI and A. ZIMRAN (2020). "International Migration Responses to Natural Disasters: Evidence from Modern Europe's Deadliest Earthquake." (Cited on p. 3, 5)

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. (Cited on p. 10, 20)

——— (2023). "Like an Ink Blot on Paper: Testing the Diffusion Hypothesis of Mass Migration, Italy 1876-1920." (Cited on p. 7, 10, A4, A5)

SURI, T. (2011). "Selection and Comparative Advantage in Technology Adoption." *Econometrica*, 79(1): 159–209. (Cited on p. 1, 5, 22)

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. (Cited on p. 15)

TAYLOR, A. M. and J. G. WILLIAMSON (1997). "Convergence in the age of mass migration." *European Review of Economic History*, 1(1): 27–63. (Cited on p. 5)

TONIOLO, G. (2014). *An Economic History of Liberal Italy: 1850-1918*. Routledge. (Cited on p. 8)

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. (Cited on p. 4)

UNRAU, H. D. (1984). *Ellis Island, Statue of Liberty National Monument, New York-New Jersey*. U.S. Department of the Interior, National Park Service. (Cited on p. 10)

VAN PATTEN, D. (2023). "International Diffusion of Technology: Accounting for Heterogeneous Learning Abilities." (Cited on p. 1)

VASTA, M. (1999). *Innovazione tecnologica e capitale umano in Italia (1880-1914)*. Bologna (IT): Il Mulino.

(Cited on p. 9)

YANG, D. (2011). "Migrant Remittances." *Journal of Economic Perspectives*, 25(3): 129–152. (Cited on p. 21)

ZAMAGNI, V. (1993). *The Economic History of Italy 1860-1990*. Clarendon Press. (Cited on p. 8)

ZEIRA, J. (1998). "Workers, Machines, and Economic Growth." *The Quarterly Journal of Economics*, 113(4): 1091–1117. (Cited on p. 3, 18, 22, D29, D30)

Tables

TABLE 1. DESCRIPTIVE STATISTICS

	Mean	Std. Dev.	Median	Min.	Max.	Units	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. District-Level Variables from Population Censuses (in 10,000 units unless otherwise specified)							
Population	1.704	1.544	1.278	0.250	15.041	211	1235
US Emigrants	0.067	0.075	0.045	0.001	0.668	211	1235
Emigrants	1.458	0.941	1.325	0.106	5.921	211	1235
Manufacturing Employment Share (%)	10.810	6.191	9.170	0.278	38.287	211	1235
Agriculture Employment Share (%)	27.209	8.786	26.803	0.953	77.549	211	1024
Altitude	3.402	2.143	3.302	0.013	9.669	211	1235
Area	1.335	0.831	1.097	0.105	4.982	211	1235
Share of Cities Above 20,000 (%)	0.000	0.000	0.000	0.000	0.001	211	1235
Share of Population in Cities Above 20,000 (%)	20.771	23.277	16.499	0.000	100.000	211	1235
Panel B. Province-Level Manufacturing Employment by Sector (in 10,000 units)							
Agriculture	1.542	1.569	1.116	0.131	11.856	69	345
Chemicals	0.178	0.493	0.051	0.000	5.186	69	345
Construction	1.359	1.262	0.944	0.162	8.906	69	345
Metalworking	0.924	1.757	0.403	0.065	17.608	69	345
Mining	0.153	0.265	0.066	0.000	2.137	69	345
Textiles	2.128	2.924	1.189	0.070	21.097	69	345
Panel C. Province-Level Capital Variables (in 1,000 units)							
N. of Firms	8.231	7.817	6.162	0.000	57.109	71	208
N. of Firms with Engine	1.503	2.066	0.913	0.000	18.511	71	208
N. of Electrical Engines	5.267	15.082	1.555	0.000	168.288	71	208
Electrical Horsepower	39.576	94.605	10.408	0.000	893.338	71	208
N. of Mechanical Engines	0.572	0.416	0.430	0.000	2.079	71	208
Mechanical Horsepower	15.770	18.595	8.636	0.000	104.668	71	208

Notes. This Table reports summary statistics for the main variables used in the paper. Data in Panel A are tabulated from population censuses; data in Panel B and C are digitized from manufacturing censuses. Variables in Panels A and B are reported in 10,000 units unless otherwise specified; variables in Panel C are reported in 1,000 units. All variables are cross-walked to consistent 1921 district (Panel A) and province (Panels B and C) borders. Referenced on page 9.

TABLE 2. CORRELATION BETWEEN TREATMENT AND PRE-PERIOD OBSERVABLE VARIABLES

	Observed Quota Exposure						Shift-Share Instrumental Variable					
	1901		1911		Growth Rate		1901		1911		Growth Rate	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A. District-Level Variables from Population Census												
Population	0.612***	(0.104)	0.670***	(0.122)	0.002	(0.005)	2.452***	(0.641)	2.790***	(0.749)	0.005	(0.030)
Manufacturing Employment Share	0.003	(0.003)	0.003	(0.005)	-0.009	(0.018)	0.008	(0.017)	0.016	(0.027)	0.009	(0.107)
Agriculture Employment Share	-0.013***	(0.004)	-0.013*	(0.007)	0.009	(0.019)	-0.055**	(0.025)	-0.084**	(0.040)	-0.054	(0.111)
Share of Cities Above 20,000 (%)	-0.186	(0.952)	-0.278	(0.985)	0.030	(0.040)	3.175	(4.160)	2.691	(4.444)	0.040	(0.251)
Share of Population in Cities Above 20,000 (%)	6.743***	(1.956)	6.545***	(2.015)	0.005	(0.029)	26.797**	(11.276)	27.534**	(11.587)	0.011	(0.181)
Panel B. Province-Level Manufacturing Employment by Sector												
Agriculture	-1.707	(3.658)	-3.269	(5.848)	-0.485	(0.419)	4.857*	(2.767)	4.974	(4.508)	-0.397	(0.325)
Chemicals	-0.114	(0.261)	-0.690	(0.639)	1.606	(2.245)	0.421**	(0.194)	1.319***	(0.467)	-0.886	(1.750)
Construction	-1.665	(2.061)	-2.805	(3.185)	0.111	(0.645)	1.085	(1.606)	4.968**	(2.395)	1.035**	(0.480)
Metalworking	-1.943	(1.840)	-3.398	(3.544)	0.132	(0.652)	3.357***	(1.367)	6.413**	(2.630)	0.825*	(0.494)
Mining	0.888	(0.974)	1.094	(0.763)	0.348	(1.812)	0.355	(0.762)	0.703	(0.597)	-0.737	(1.403)
Textiles	-2.117	(7.871)	-2.104	(7.410)	-0.124	(0.638)	9.134	(5.988)	10.008*	(5.592)	0.551	(0.490)
Panel C. Province-Level Capital Indicators												
N. of Firms			-6.506	(7.406)					7.210	(5.714)		
N. of Firms with Engine			-1.428	(2.436)					2.322	(1.873)		
N. of Electrical Engines			-1.156	(7.065)					5.413	(5.443)		
Electrical Horsepower			-36.534	(41.462)					50.347	(31.706)		
N. of Mechanical Engines			-1.124	(0.992)					0.990	(0.769)		
Mechanical Horsepower			-61.149*	(34.144)					50.478*	(26.447)		

Notes. This Table displays the correlation between the observable district- and province-level variables, the baseline metric of quota exposure (columns 1–6), and the instrumental variable for quota exposure (columns 7–12). In columns (1–2) and (7–8), the correlation is between the treatments and the variables in 1901; in columns (3–4) and (9–10), the variables are recorded in 1911. In columns (5–6) and (11-12), we report the correlation between the two treatments and the growth rate of each variable between 1901 and 1911. The absence of any statistically significant correlation between the growth rates and the treatment variables is the crucial test in support of the validity of the research design. In Panel A, the units of observation are districts; in Panel B and C, the units of observation are provinces. Capital variables, displayed in Panel C, are observed in 1911, 1927, and 1937, so we cannot compute the growth rate between two consecutive pre-treatment periods. Each regression controls for the emigration rate and region-fixed effects. Standard errors are displayed in parentheses. Referenced on pages 14, 15, 17.

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

TABLE 3. THE RESPONSE OF POPULATION

	Dependent Variable: Population							
	Difference-in-Differences						Instrumented DiD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Quota Exposure × Post	0.033***	0.021***			0.019***	0.036***		
	(0.011)	(0.006)			(0.007)	(0.008)		
I(Quota Exposure) × Post			0.061***	0.017				
			(0.022)	(0.016)				
$\widehat{\text{Quota Exposure}}$ × Post							0.117***	
							(0.038)	
I($\widehat{\text{Quota Exposure}}$) × Post								0.046**
								(0.019)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	–	–	Yes	Yes
Region-Decade FE	No	No	No	No	Yes	–	No	No
Province-Decade FE	No	No	No	No	No	Yes	No	No
Regions in Sample	All	South	All	South	All	All	All	All
N. of Districts	202	88	202	88	201	188	192	192
N. of Observations	1198	518	1198	518	1192	1115	1140	1140
R^2	0.672	0.574	0.671	0.574	0.674	0.662	0.669	0.670
Mean Dep. Var.	17.342	15.683	17.342	15.683	17.233	16.237	17.877	17.877
Std. Beta Coef.	0.025	0.020	0.026	0.008	0.015	0.028	0.012	0.020

Notes. This Table reports the estimated effect of district-level exposure to the US Quota Acts on population. The unit of observation is a district observed at a census-decade frequency between 1881 and 1936. In columns (1–2) and (5–6), the treatment is an interaction between a post-Quota (1921) indicator variable the Quota Exposure metric. In columns (3–4), we interact the post-Quota indicator with a dummy returning value one if district-level exposure to the Quota Acts (measured as the ratio between US and overall emigrants) is above the median and zero otherwise. Columns (7–8) substitute measured Quota exposure with the shift-share instrument. In columns (2) and (4), the sample includes only regions in the South (former Kingdom of Two Sicilies). Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. Robust standard errors are reported in parentheses. Referenced on page 16, C20, C20, C26, C26.

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

TABLE 4. THE RESPONSE OF CAPITAL AND TECHNOLOGY ADOPTION

	Dependent Variable: Province-Level Number of... (in 1,000 units)					
	(1) Firms	(2) Firms with Engines	(3) Mechanical Engines	(4) Electrical Engines	(5) Mechanical Horsepower	(6) Electrical Horsepower
Panel A. Difference-in-Differences Estimates (Treatment: Measured Quota Exposure)						
I(Quota Exposure) × Post	-0.273**	-0.178	-0.214***	-0.969**	-0.185	-0.457
	(0.130)	(0.163)	(0.079)	(0.405)	(0.135)	(0.328)
R^2	0.485	0.287	0.126	0.665	0.318	0.662
Panel B. Instrumented Difference-in-Differences Estimates (Treatment: Predicted Quota Exposure)						
I($\widehat{\text{Quota Exposure}}$) × Post	-0.199*	-0.292***	-0.097	-1.009**	-0.267**	-0.654**
	(0.113)	(0.112)	(0.069)	(0.402)	(0.124)	(0.258)
R^2	0.484	0.288	0.126	0.665	0.319	0.663
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
N. of Provinces	69	69	69	69	69	69
N. of Observations	206	206	206	206	206	206
Mean Dep. Var.	7.723	1.305	0.569	3.589	1.472	3.092

Notes. This Table reports the estimated effect of exposure to the US Quota Acts on capital investment. The unit of observation is a province observed at a census-decade frequency between 1911 and 1936. The dependent variables are the number of firms (column 1), the number of firms with at least one engine (column 2), the number of mechanical engines (column 3), the number of electrical engines (column 4), the horsepower generated by mechanical (column 5) and electrical (column 6) engines. In Panel A, the treatment is an interaction between a post-Quota (1921) indicator variable and a dummy for districts with above-median exposure to the Quota Acts (measured as the ratio between US and overall emigrants). In Panel B, the treatment substitutes measured Quota exposure with the shift-share instrument. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. Regressions include province and decade-fixed effects. Robust standard errors are reported in parentheses. Referenced on page 17, C21, C21, C22, C22, C27, C27.

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

TABLE 5. THE RESPONSE OF MANUFACTURING AND AGRICULTURE EMPLOYMENT

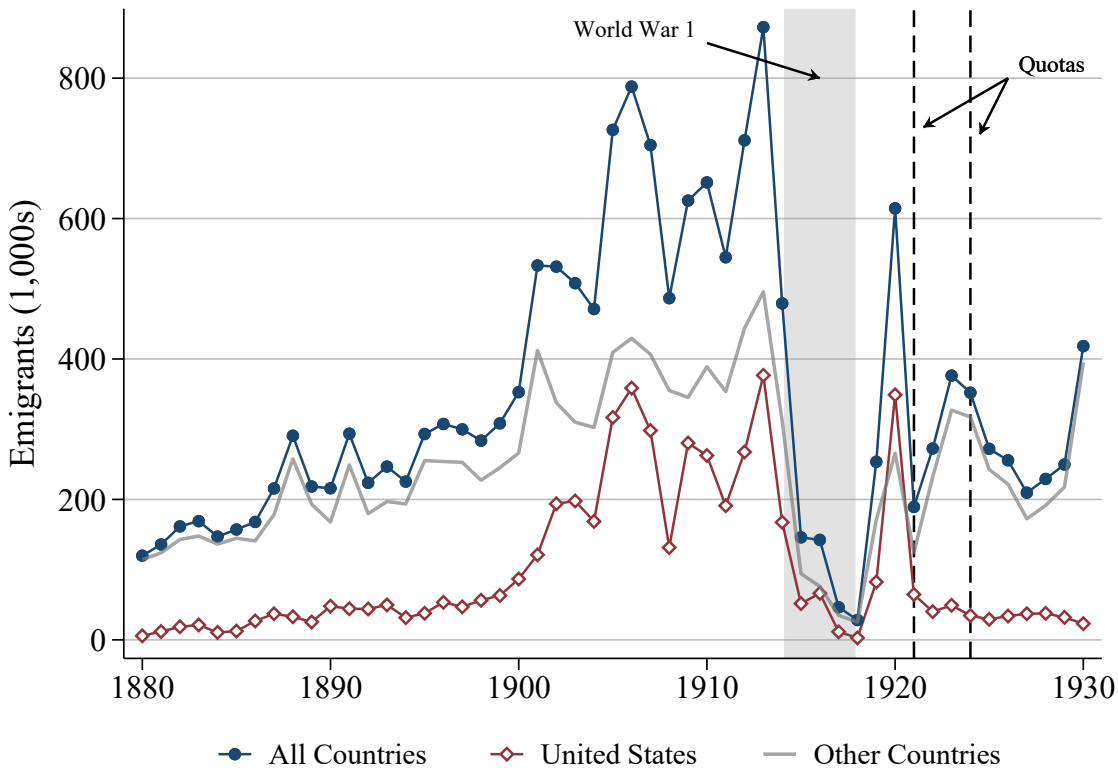
	Difference-in-Differences						Instrumented DiD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Dependent variable: Manufacture Employment (% of Pre-Treatment Population)								
Quota Exposure × Post	0.071**	0.072			0.051*	0.082***		
	(0.036)	(0.049)			(0.030)	(0.017)		
I(Quota Exposure) × Post			0.193***	0.103				
			(0.046)	(0.064)				
$\widehat{\text{Quota Exposure}}$ × Post							0.372***	
							(0.120)	
I($\widehat{\text{Quota Exposure}}$) × Post								0.148***
								(0.042)
R^2	0.775	0.665	0.776	0.664	0.792	0.808	0.775	0.776
Mean Dep. Var.	15.231	10.752	15.231	10.752	15.118	14.550	15.745	15.745
Std. Beta Coef.	0.054	0.068	0.193	0.103	0.040	0.063	0.038	0.148
Panel B. Dependent Variable: Agriculture Employment (% of Pre-Treatment Population)								
Quota Exposure × Post	-0.047**	-0.056			-0.041*	-0.010		
	(0.023)	(0.035)			(0.022)	(0.008)		
I(Quota Exposure) × Post			-0.085***	-0.011				
			(0.023)	(0.031)				
$\widehat{\text{Quota Exposure}}$ × Post							-0.220**	
							(0.103)	
I($\widehat{\text{Quota Exposure}}$) × Post								-0.079***
								(0.022)
R^2	0.628	0.480	0.628	0.478	0.636	0.598	0.615	0.616
Mean Dep. Var.	29.456	26.604	29.456	26.604	29.126	27.282	30.368	30.368
Std. Beta Coef.	-0.033	-0.049	-0.034	-0.005	-0.029	-0.007	-0.020	-0.031
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	–	–	Yes	Yes
Region-Decade FE	No	No	No	No	Yes	–	No	No
Province-Decade FE	No	No	No	No	No	Yes	No	No
Regions in Sample	All	South	All	South	All	All	All	All
N. of Districts	202	88	202	88	201	188	192	192
N. of Observations	996	430	996	430	991	927	948	948

Notes. This Table reports the estimated effect of exposure to the US Quota Acts on manufacturing (Panel A) and agriculture (Panel B) employment. The unit of observation is a district observed at a census-decade frequency between 1881 and 1936.In columns (1–2) and (5–6), the treatment is an interaction between a post-Quota (1921) indicator variable the Quota Exposure metric. In columns (3–4), we interact the post-Quota indicator with a dummy returning value one if district-level exposure to the Quota Acts (measured as the ratio between US and overall emigrants) is above the median, and zero otherwise. Columns (7–8) substitute measured Quota exposure with the shift-share instrument. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. Robust standard errors are reported in parentheses. Referenced on page 18, C23, C23, C24, C24, C26, C26.

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

Figures

FIGURE 1. INTERNATIONAL MIGRANTS, 1880–1930

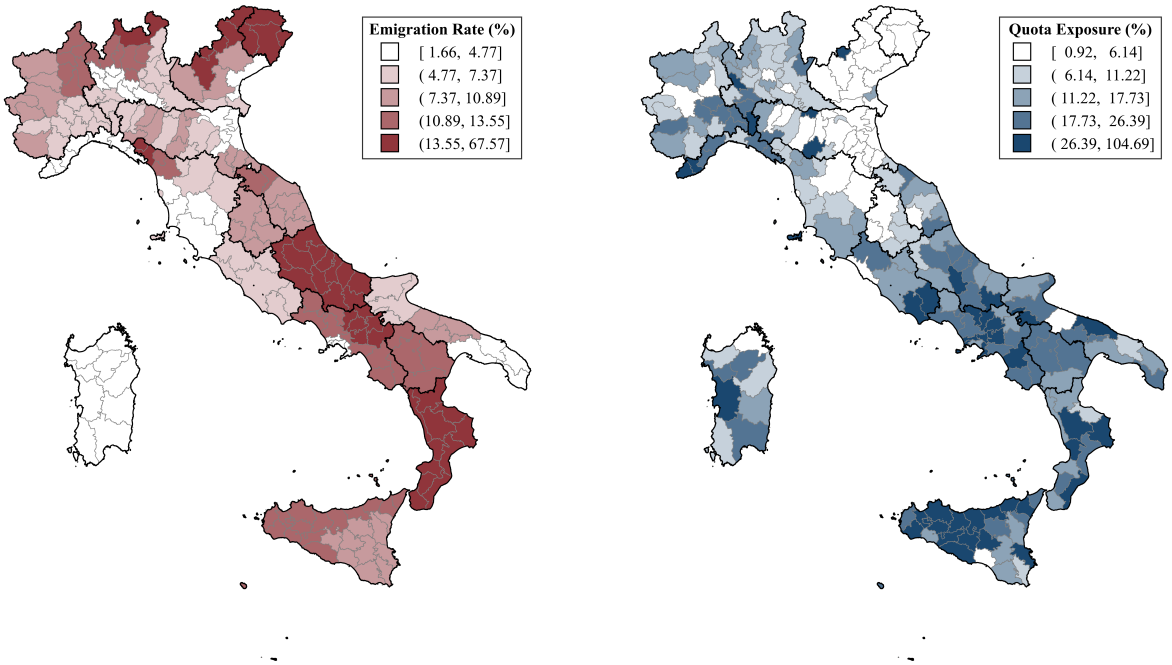


Notes. This figure reports the yearly outflow of international migrants from Italy between 1880 and 1930. Data are digitized from the *Annuario statistico dell'emigrazione italiana dal 1876 al 1925: con notizie sull'emigrazione negli anni 1869-1875* edited by the Italian Statistical Office (1926). The blue line reports the overall number of international migrants; the red line reports the number of migrants to the United States, the single most important destination country over this period; the gray line reports emigrants to every other country. The shaded gray area marks the 1914–1918 war years; the dashed vertical black lines mark the 1921 and 1924 Emergency and (Johnson-Reed) Immigration Quota Acts, respectively. Referenced on pages 1, 3, 11, 16.

FIGURE 2. SPATIAL DISTRIBUTION OF INTERNATIONAL EMIGRATION, 1892–1925

(A) Share of Emigrants Relative to 1881 Population

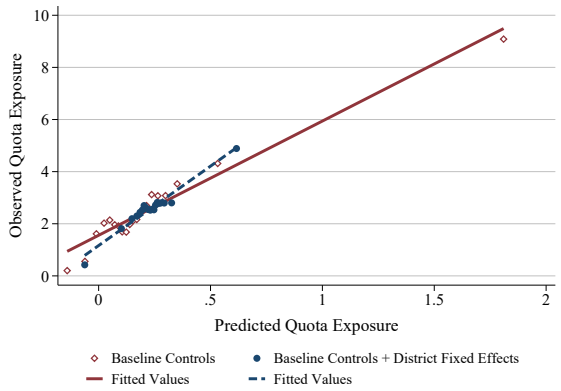
(B) Share of US Migrants Relative to Total Emigrants



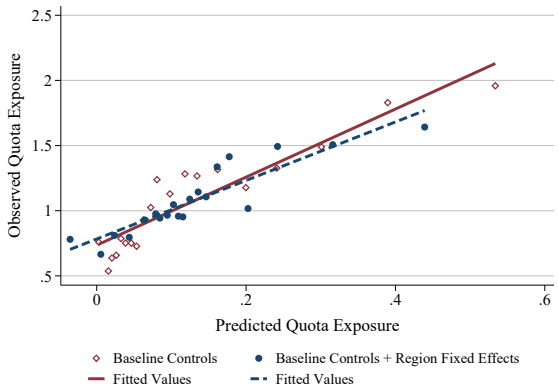
Notes. These figures report the spatial distribution of the emigration rate (Figure 2a) and of quota exposure (Figure 2b) across districts. In each map, the unit of observation is a district. Black lines superimpose region boundaries. The geography is at 1921 borders. We exclude Trentino-Alto Adige, Venezia-Giulia, Istria, and Dalmatia because districts in those areas are acquired after the Treaty of Versailles (1918). Panel 2a reports the share of international emigrants relative to the district population in 1881. This is computed from province-level data using district-level population, assuming no within-province variation in emigration rates. Panel 2b reports the share of migrants to the United States relative to the number of international migrants. Data refer to the period 1892–1925 and are collapsed at the district level for visualization. Referenced on page 11.

FIGURE 3. CORRELATION BETWEEN MEASURED AND PREDICTED QUOTA EXPOSURE

(A) Panel Correlation

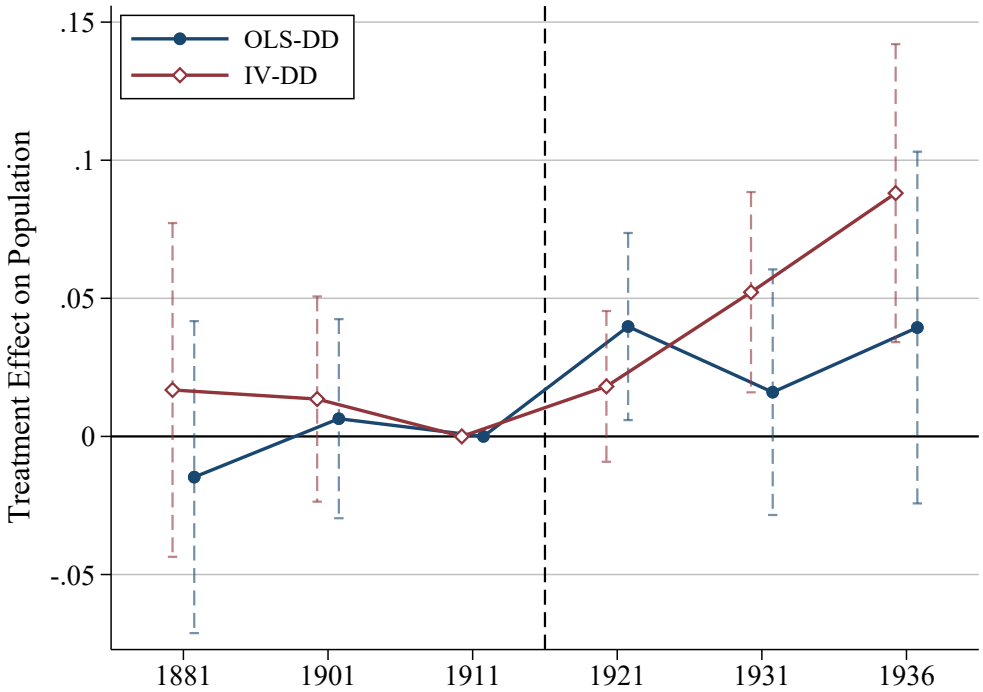


(B) Cross-Sectional Correlation



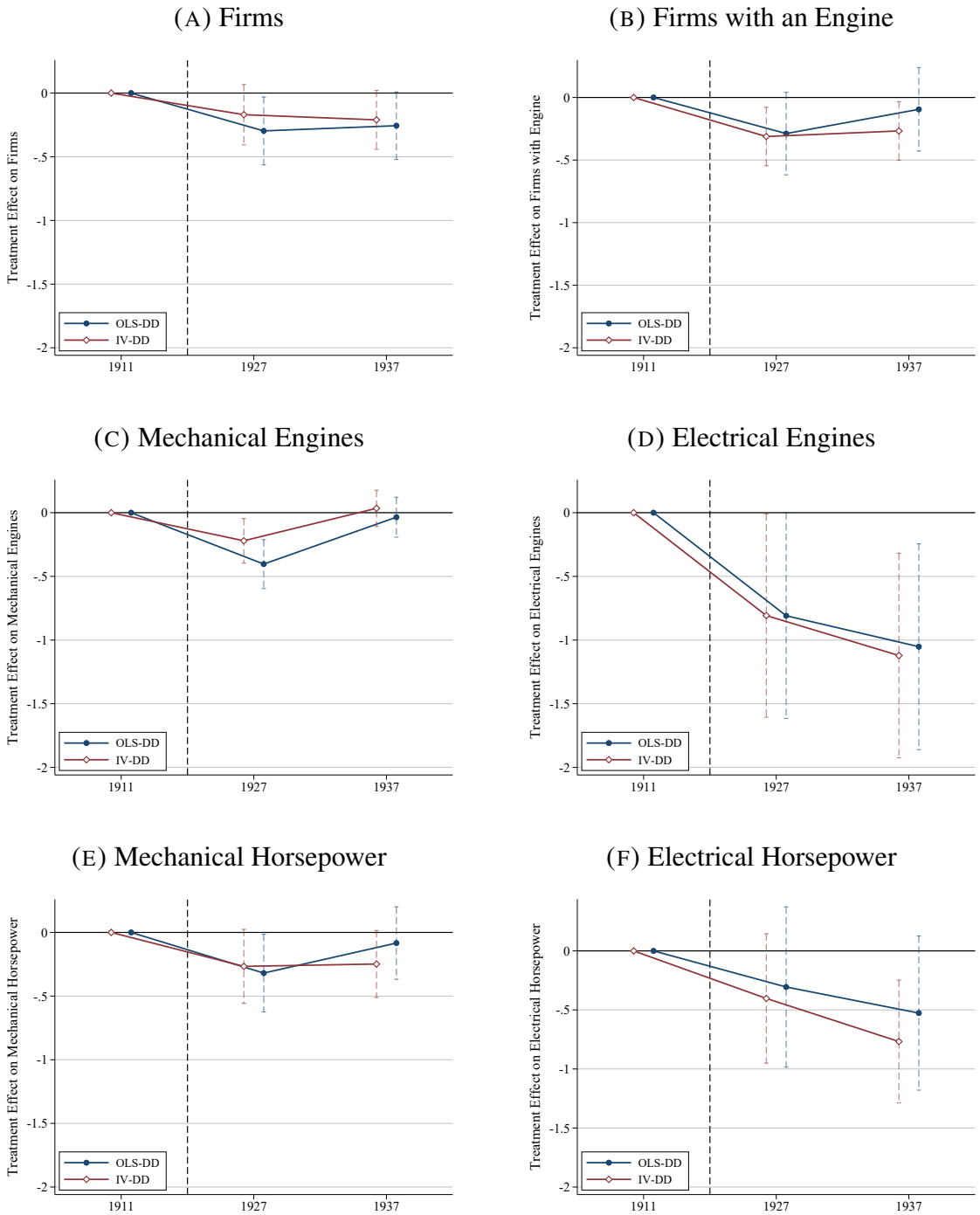
Notes. These figures report binned scatterplot between actual quota exposure and the shift-share instrument. In Panel 3a, districts are observed yearly between 1900 and 1914. Red dots report the baseline correlation controlling for year-fixed effects and the emigration rate; the blue dots add district-level fixed effects. In Panel 3b, both variables are collapsed at the district level so that each district is observed once. This reflects the treatment variation used in the difference-in-differences analysis. Red dots report the baseline correlation controlling for the emigration rate; the blue dots add region-fixed effects. Referenced on pages 15, A5.

FIGURE 4. EFFECT OF EXPOSURE TO THE US QUOTA ACTS ON POPULATION



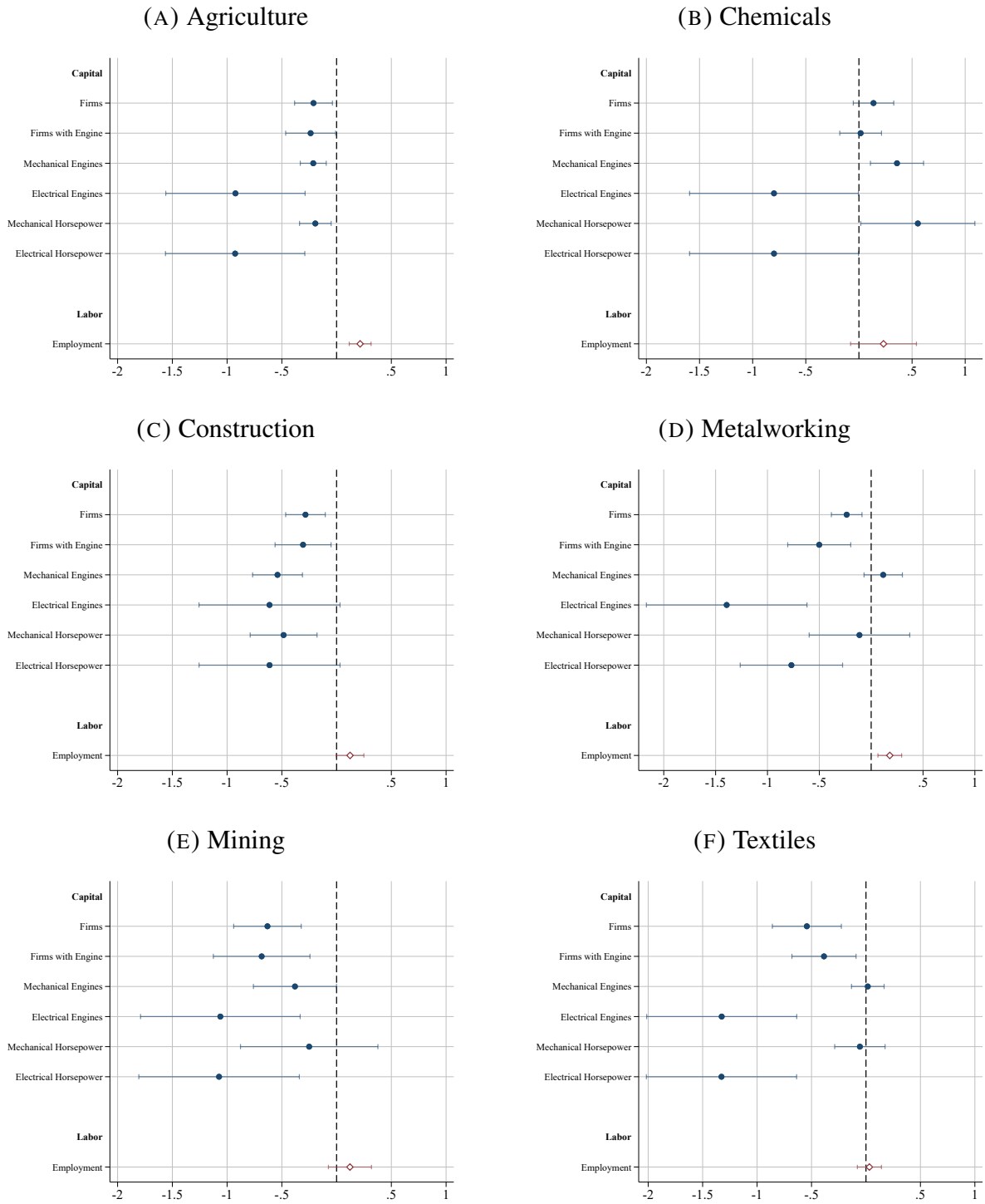
Notes. These figures report the estimated effect of district-level exposure to the US Quota Acts on population. The unit of observation is a district observed at a census-decade frequency between 1881 and 1936. The outcome variable is population. The blue dots report the estimated treatment effects of the baseline Quota Exposure; the red dots report the estimates using the shift-share instrumental variable. In both cases, the treatment is an indicator returning value one if district-level exposure is above the median and zero otherwise. Each regression includes district and decade-fixed effects and controls for the emigration rate, and an indicator variable for Southern regions, both interacted with a post-1921 indicator. The bands report 95% confidence intervals from robust standard errors. Referenced on page 16.

FIGURE 5. EFFECT OF EXPOSURE TO THE US QUOTA ACTS ON CAPITAL INVESTMENT



Notes. These figures report the estimated effect of exposure to the US Quota Acts on capital investment over time. The unit of observation is a province observed at a census-decade frequency between 1911 and 1936. The outcome variables are firms (panel 5a), firms with an engine (panel 5b), mechanical (panel 5c) and electrical (panel 5d) engines, mechanical (panel 5e) and electrical (panel 5f) horsepower. The blue dots report the estimated treatment effects of the baseline Quota Exposure; the red dots report the estimates using the shift-share instrumental variable. In both cases, the treatment is an indicator returning value one if province-level exposure is above the median and zero otherwise. Each regression includes province and decade-fixed effects and controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. Bands report 95% confidence intervals from robust standard errors. Referenced on page 17.

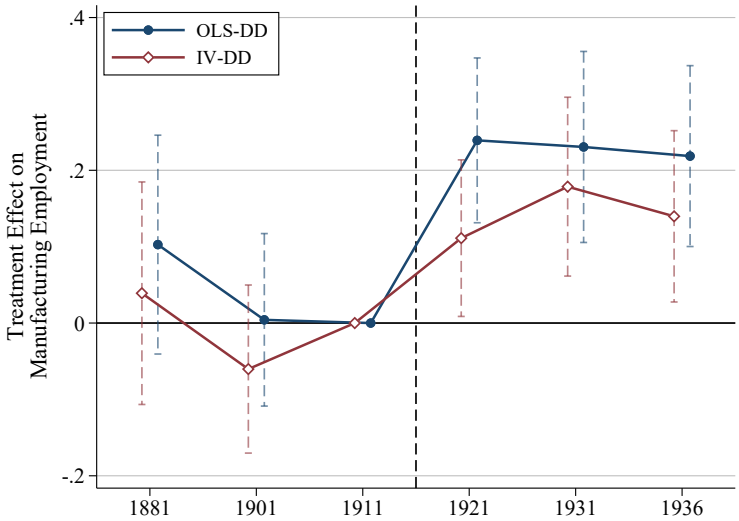
FIGURE 6. EFFECT OF THE US QUOTA ACTS ON CAPITAL AND EMPLOYMENT BY SECTOR



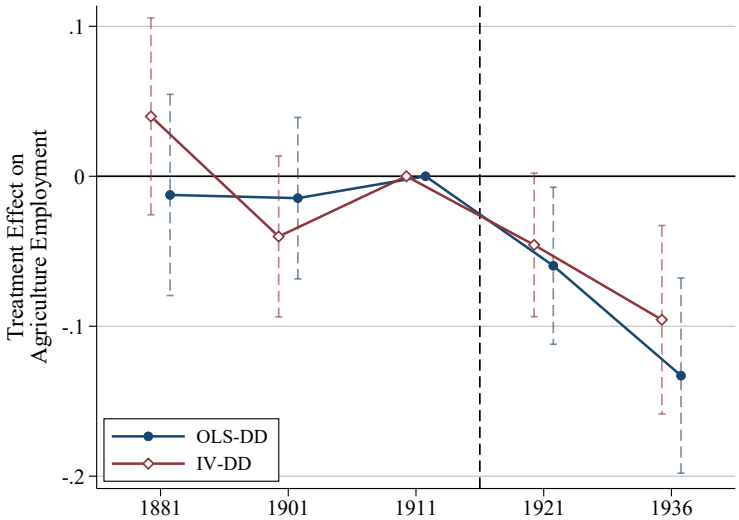
Notes. These figures report the estimated effect of exposure to the US Quota Acts on capital and employment by sector over time. The unit of observation is a district observed at a census-decade frequency between 1901 and 1936. Each panel reports the results for a specific manufacturing sector. The outcome variables are listed by row. The treatment is an indicator returning value one if exposure to the Quota Acts is above the median and zero otherwise. Each regression includes district and decade-fixed effects and controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. Bands report 90% confidence intervals from robust standard errors. Referenced on page 17.

FIGURE 7. EFFECT OF EXPOSURE TO THE US QUOTA ACTS ON EMPLOYMENT BY SECTOR

(A) Manufacture Employment



(B) Agriculture Employment



Notes. These figures report the estimated effect of district-level exposure to the US Quota Acts on manufacturing (Panel 7a) and agriculture (Panel 7b) employment over time. Both outcomes are normalized by the 1881 district population. The unit of observation is a district observed at a census-decade frequency between 1881 and 1936. In panel 7a, the outcome variable is manufacturing employment; in panel 7b, the outcome variable is agriculture employment. The blue dots report the estimated treatment effects of the baseline Quota Exposure; the red dots report the estimates using the shift-share instrumental variable. In both cases, the treatment is an indicator returning value one if district-level exposure is above the median and zero otherwise. Each regression includes district and decade-fixed effects and controls for the emigration rate, and an indicator variable for Southern regions, both interacted with a post-1921 indicator. The bands report 95% confidence intervals from robust standard errors. Referenced on page 18.

Online Appendix

Emigration Restrictions and Economics Development

Evidence from the Italian Mass Migration to the United States, 1892–1936

Davide M. Coluccia & Lorenzo Spadavecchia

February, 2024

A Data Appendix

A.1 Data Sources

This section lists the sources, methodology, and coverage of the data that we assemble. We defer a more detailed description of the emigration data to section A.3. Table A.1 summarizes the content of this section.

Population Censuses. We extract information from population censuses in 1881, 1901, 1911, 1921, 1931, and 1936. No census was taken in 1891. We digitize district-level information on the number of people employed in all major sectors in 1881–1921 censuses. The same information is available at the municipality level for 1931 and 1936. Hence we aggregate it at the district level. To assign municipalities to districts, we geo-code each municipality and overlay the coordinates to the 1921 district shape file. From population censuses, we also extract information on employment by industry: this is available by districts until 1921 but only by provinces in 1931 and 1936. We thus aggregate industry employment to provinces. Population data have been tabulated by ISTAT for each municipality. We aggregate them at the district and province levels. We also code two types of urbanization variables: the share of the population living in districts with at least k-thousand municipalities and the share of municipalities with at least k-thousand inhabitants. The ISTAT population tables report information on the area, altitude, and access to the sea for each municipality.

Manufacture Censuses. Manufacture censuses were taken in 1911, 1927, and 1937. They report province-level information on the universe of firms. We digitize data on a set of proxies for capital investment: the number of firms, the number of firms with at least one installed engine, the number of installed mechanical and electrical engines, and the number of installed mechanical and electrical horsepower.

World War One Deaths. WW1 deaths are from `cadutigrandeguerra.it`, a dataset maintained by ISTORECO, a branch of the *Associatione Nazionale Partigiani d'Italia* (ANPI). The underlying data were collected by the Fascist regime for propaganda purposes. The honor roll call contains information on 570,355 Italian soldiers who lost their lives during the war. The data appears to be comprehensive since most estimates place the total death toll at 650,000. For each individual, we know the name and surname, the birth year, the municipality of origin—which we map to the municipalities listed in the ISTAT data—, the military rank at death, the regiment, when, where, and why they died, and the decoration, if any. Deaths span the war years (1915–1918). We aggregate them at the district and province levels to have a cross-sectional indicator of mortality.

Railway Data. We reconstruct the network of Italian railways between 1839 and 1926 from a volume curated by the Italian Statistical Office and edited in 1927. This is the first paper using these data. The unit of observation in the data is truck lines connecting two stations. We have information on the precise date when each line opened and the distance it covered, and the name of the station it connected. We geo-code the location of each station and generate a dummy variable that indicates whether a given municipality had at least one station,

as well as a count variable returning the number of stations. We also define a variable that returns the number of kilometers that separate every given municipality from the closest transatlantic migration port. Calabrese (2017) suggests that railway access to Genoa, Palermo, or Naples—the only ports with ships sailing toward the United States—was a crucial condition to ignite migration movements. We thus compute the shortest path connecting each municipality with each transatlantic port given the state of the railway network in every given year and take the minimum among the three. We aggregate this variable by district and province, taking a population-weighted average of the shortest railway distances to emigration ports.

A.2 Construction of the Sample

All variables that are not computed from geo-coded data are cross-walked to consistent 1921 district and border geographies using the method described by Eckert *et al.* (2020). To do so, we use GIS boundary files publicly provided by ISTAT for each census year. Even though this yields quantitatively minor corrections, it is important to ensure that geographies remain consistent because the Fascist regime undertook an extensive revision of local government divisions, which ultimately abolished districts in 1927.

We forcibly exclude areas that were annexed as a consequence of World War One in 1918—Trento and Trieste, Südtirol, Istria, and Zara—because we do not observe pre-Quota outcomes in those regions.

Unlike for US emigrants, we do not have district-level data on overall emigration. To assign province-level emigration to districts, we assume that emigration rates were constant within each province. We then impute district emigration by multiplying province emigration by the share of inhabitants in each district compared to the province population before the mass migration (in 1881).

We assemble three distinct samples (Samples 1, 2, and 3). Sample 1, which comprises population and employment-by-sector data, is at the district level and covers the years 1881, 1901, 1911, 1921, 1931, and 1936. Agriculture employment is not available in 1931. All data in Sample 1 are either digitized from population censuses or are aggregated from data tabulated by ISTAT. Sample 2 runs at the province level, covers capital variables and is available in 1911, 1927, and 1937. Capital variables are retrieved from manufacturing censuses. Sample 3 runs at the province level, covers employment-by-industry variables and spans the years 1901, 1911, 1921, 1927, and 1936. For the first three decades, the data are from population censuses; for the latter two, we digitize them from manufacturing censuses. Since capital variables exhibit substantial excess kurtosis, we winsorize the top and bottom 5% of their distribution.

A.3 Details on the Emigration Data

In this section, we document in detail the emigration data that we collect. The raw data can be found at `https://heritage.statueofliberty.org/`. First, we describe the methodology that we adopt to assemble the data. Second, we show how to validate this dataset with external sources. Last, we present some stylized

facts that the new data allow us to document.

This dataset responds to a key limitation of commonly used US census data (Ruggles *et al.*, 2021). These list the country of origin of immigrants residing in the US, but they do not report where immigrants originated from within their home country. This issue applies to all countries. Hence separate papers developed strategies to reconstruct such information for, among others, Norway (Abramitzky *et al.*, 2014) and England (Coluccia and Dossi, 2023). This paper looks at emigrants from Italy, a major emigration country among the so-called "second-wave" nations.

A.3.1 Methodology

We run queries over a comprehensive set of the most common 20,000 Italian surnames between 1890 and 1930. We collect individual-level information on the name and surname of immigrants, their municipality of origin, their immigration year, and whether they can read or write.

The municipality of origin is recorded consistently only between 1892 and 1924. Names, surnames, and municipalities are frequently coded with spelling errors, possibly because they were recorded by American enumerators. In this paper, we are interested in the municipality of origin of immigrants. We tackle this data quality issue in two steps. First, we pick the 1,000 most common origin municipalities in the data, and we correct eventual coding errors in those manually. We also discard entries that are too coarse, such as "Italy" or "Sicily." Then, we geo-code the remaining entries using Google Maps' auto-correction algorithm. Then, we manually checked that the return geo-coded locations are reliable for a subset of 200 municipalities. The algorithm successfully matches 189 out of 200 municipalities. The remaining 11 are impossible to match even by hand. We assess the plausibility of this matching exercise in section A.3.2.

The municipality of origin is missing for a non-negligible sub-sample of immigrants. In Figure A.1, we report graphical evidence. The top panel reports the absolute number of recorded immigrants (in blue) and those which we match to a municipality. The bottom panel reports the share of immigrants with at least one listed origin (in blue) and the share that we match to a municipality. Throughout the years 1900–1914—the key years of the mass migration and the ones on which we compute our treatment variable—the sample of immigrants with a matched municipality always exceeds 80% of the total records. This share is lower at the beginning of the sample—perhaps because origins started being recorded in 1892—and after World War Two. Municipalities after the 1924 Immigration Act, in particular, were seldom recorded, but we never use this sample period in this paper.

In the analysis, the dataset is aggregated by district or province depending on the part of the analysis using boundaries in 1921 from historical shape files provided by ISTAT.

A.3.2 Validation

The granular nature of the dataset implies that we cannot validate it with existing data at similar levels of aggregation. Our strategy, instead, is to aggregate it at the regional level and compare it with data from official statistics on Italian emigration to the United States collected by the Commissioner General for Emigration. These span the period 1877–1925 and are available by region.

In Table A.2, we report the correlation between the Ellis Island dataset and US emigration as recorded in official statistics. In columns (1–5), we report the correlation between the raw series, while in columns (6–7), we take logs. We find a positive and large unconditional correlation between the two (columns 1 and 6). In particular, Ellis Island migration explains more than 80% of the region-level variation in US emigration as measured in official statistics. This correlation remains conditioning on year (2 and 7), region (3 and 8), and year and region (4 and 9) fixed effects. Importantly it holds within the sub-sample period we use to compute the treatment (5 and 10). Figure A.2 displays the unconditional correlation between the two series (A.2a) and the one absorbing region and year fixed effects (A.2b). These exercises document a positive, large, and statistically significant correlation between the Ellis Island US emigration and data from official statistics. Finally, in Figure A.3, we check that the correlation remains high in each year of the observation sample. Each dot in the figure reports the correlation between the two series in one year between 1892 and 1924. The correlation remains stable, positive, and statistically significant throughout the sample period.

A.3.3 Stylized Facts

Dissecting the specific features of mass emigration to the United States is beyond the scope of this paper. Instead, we present two suggestive facts.

First, in Table A.3, we list the districts that were relatively more exposed to the US migration. In columns (1–3), districts are ranked by the absolute number of emigrants. In columns (4–6), we rank them by the emigration rate, expressed as the ratio between overall US emigrants and the 1921 population. The vast majority of top-ranked districts are located in Southern regions. Emigration rates are higher in Sicily and Campania. The district of Palermo alone accounts for almost 90,000 emigrants out of a population of 850,000.

We then provide evidence supporting the S-hypothesis advanced by Gould (1980) and recently analyzed by Spitzer and Zimran (2023). This maintains that local out-migration patterns followed a logistic-type dynamic, with initially low uptake, large increases in a relatively short time period, and subsequent plateau. Gould (1980) connects these dynamics to information diffusion within the population. To test this hypothesis, we mark the beginning of the mass migration in each district when the US emigration rate exceeded 0.1‰. This generates a setting akin to a staggered treatment roll-out. We then use the method of Borusyak *et al.* (2022) to estimate the dynamics of US out-migration. Importantly, this approach ensures that we compare emigration districts with areas where the migration had not already begun. We find that emigration followed Gould's S-shaped pattern

as documented by Spitzer and Zimran (2023). The event-study figures associated with the resulting model are listed in A.4 and show that out-migration follows an S-shaped pattern as argued by Gould (1980). We interpret this finding as additional evidence supporting the quality of the underlying data.

A.3.4 Linked Sample

To produce the instrumental variable for Quota Exposure, we require information on the origin district and province of Italian immigrants by US county. This information is not available in the US census or reported in the Ellis Island data. To circumvent this limitation, we link the full-count non-anonymized US population census (Ruggles *et al.*, 2021) and the Ellis Island records. To the best of our knowledge, ours is the first attempt to produce a linked sample between these two uniquely rich sources.

The algorithm builds on similar automated linking procedures (e.g., Abramitzky *et al.*, 2021). First, we translate the names of Italian-born individuals recorded in the US census to their Italian version.[21] For each record in the Ellis Island dataset i who immigrates in year t_i, we pick the set of Italian-born individuals J in the 1900 US census whose initial name and surname Soundex-adjusted letters are the same as i's and whose immigration year recorded in the US census is in the window $[t_i - 1, t_i + 1]$.[22] We then compute the Monge-Elkan similarity with Jaro-Winkler inner word distance between the name and surname of i and those of the individuals in J. Among those, we pick the j's with the highest name and surname similarity as potential matches. If both the maximal name and surname similarities are above a given quality threshold, which in the baseline exercise is set at .9, the match(es) is (are) accepted; otherwise, they are discarded.

In Figure A.5, we report the distribution of name and surname similarities for the sample of individuals with at least one potential match. There is substantial mass at 1, where matches are literal. In Figure A.6, we report the share of Ellis Island records with at least one match, in blue, and one *accepted* match, in red. The gross matching rate remains constant throughout the sample at approximately 50%. There are several reasons why someone recorded at Ellis Island may not appear in the 1900 census. First, that person may have left before 1900. Second, women could change their surname. Alternatively, Italians could choose to change their surname as an assimilation effort. If the name change did not simply consist of a translation, we would fail to detect this practice. Finally, the immigration year in the census may be coded with errors. Of the 50% fraction with at least one match, between one-third and one-half presents at least one accepted match with sufficiently high quality. The resulting 22% matching rate is not very distant from benchmark rates for intergenerational linked samples using US census data (Abramitzky *et al.*, 2021).

A key concern for the empirical strategy is that the probability of matching individuals from the Ellis

[21] For instance, we convert "Peter" to "Pietro." This procedure ensures that Anglicizations of Italian names that occur in the US census but not in the Ellis Island records do not artificially deflate our matching rate.

[22] We use the Soundex-adjusted algorithm to ensure that different spellings with similar phonetics are treated in the same way. For instance, the Soundex-adjusted initial of "Katherine" and "Catherine" is encoded as the same hard "c."

Island data is correlated with their area of origin. This possibility would induce selection in the resulting linked sample, thus ultimately invalidating the relevance of the shift-share instrument. While we provide quantitative evidence to support the relevance of the first stage, we can check whether any systematic selection pattern emerges directly from the linked data. In Figure A.7, we report the correlation between the probability of matching and the origin region (panel A.2a) and province (panel A.2b) of immigrants. There appears to be no systematic selection pattern. Individuals from Calabria and Sicily are marginally more likely to be matched, even though this difference likely reflects a larger sample size of emigrants from those regions. Overall, as evidenced in Figure 3, the resulting instrumental variable retains considerable correlation with the observed emigration outflows.

A.4 Tables

TABLE A.1. SUMMARY OF THE DATA SOURCES AND COVERAGE

Variable	Observation Unit	Source	Observed Years
(1)	(2)	(3)	(4)
Panel A. Demographics			
Population	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Area	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Urbanization	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Literacy	Municipality	Population Censuses	1881-1936, excl.1891
Panel B. Employment, by Sector			
Manufacture	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Agriculture	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Trade	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Liberal Professions	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Public Administration	District (1881-1921), Municipality (1931-1936)	Population Censuses	1881-1936, excl.1891
Panel C. Capital			
Firms	Province	Manufacture Censuses	1911, 1927, 1937
Firms with Engine	Province	Manufacture Censuses	1911, 1927, 1937
Mechanical Engines	Province	Manufacture Censuses	1911, 1927, 1937
Electrical Engines	Province	Manufacture Censuses	1911, 1927, 1937
Mechanical Horsepower	Province	Manufacture Censuses	1911, 1927, 1937
Electrical Horsepower	Province	Manufacture Censuses	1911, 1927, 1937
Panel D. Manufacturing Employment, by Industry			
Agriculture	District (1901–1921), Province (1927–1936)	Population Censuses, Manufacture Census	1901–1936
Chemicals	District (1901–1921), Province (1927–1936)	Population Censuses, Manufacture Census	1901–1936
Construction	District (1901–1921), Province (1927–1936)	Population Censuses, Manufacture Census	1901–1936
Metalworking	District (1901–1921), Province (1927–1936)	Population Censuses, Manufacture Census	1901–1936
Mining	District (1901–1921), Province (1927–1936)	Population Censuses, Manufacture Census	1901–1936
Textiles	District (1901–1921), Province (1927–1936)	Population Censuses, Manufacture Census	1901–1936
Panel E. Emigration			
US Emigration	Municipality	Ellis Island Data	1892-1924
Overall Emigration	Province, imputed to Districts	Official Statistics of the Commissioner General	1877-1925
Panel F. Other			
WW1 deaths	Municipality	ISTORECO, ANPI	1915-1918
Railways	Municipality	ISTAT	1839-1926
US GDP	National	Maddison (2007)	
Panel G. GIS Files			
Shapefiles	District, Provinces	ISTAT	1881–1936, excl. 1891

Notes. This table reports all variables used in the paper. Column (2) returns the level of aggregation at which the variable is measured. Column (3) displays the type of source the raw data are extracted from. Further references to original sources can be found in the text's main body. Column (4) reports the years when the raw data is available. ISTAT: Italian Statistical Office or previous denominations. ISTORECO: Istituto per la storia della Resistenza e della società contemporanea, part of the Associazione Nazionale Partigiani Italiani (ANPI). Referenced on pages 9, A1.

TABLE A.2. CORRELATION BETWEEN ELLIS ISLAND AND OFFICIAL STATISTICS US EMIGRATION

	Official Statistics Emigrants					ln(1+Official Statistics Emigrants)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ellis Island Migrants	0.906***	0.906***	0.591***	0.591***	0.578***					
	(0.108)	(0.112)	(0.067)	(0.070)	(0.110)					
ln(1+Ellis Island Migrants)						0.957***	1.034***	0.835***	0.618***	0.722***
						(0.048)	(0.087)	(0.045)	(0.059)	(0.143)
Region FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Year FE	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes
Sample Years	All	All	All	All	1900–1914	All	All	All	All	1900–1914
N. of Regiones	16	16	16	16	16	16	16	16	16	16
N. of Observations	527	527	527	527	240	527	527	527	527	240
R^2	0.820	0.820	0.971	0.971	0.982	0.820	0.882	0.894	0.955	0.972
Std. Beta Coef.	0.906	0.906	0.591	0.591	0.578	0.906	0.978	0.790	0.584	0.669

Notes. This Table compares the number of US emigrants recorded in Italian official statistics with data from the Ellis Island Foundation dataset. The unit of observation is a region at a yearly frequency. The sample period spans 1892 to 1925. In columns (1–5), the dependent and independent variables are the number of emigrants recorded in official statistics and at Ellis Island, respectively. In columns (6–10), both variables are taken as log(1+). Columns (1) and (6) display the unconditional correlation; in columns (2) and (7), (3) and (8), and (4) and (9), we include year, region, and year and region fixed effects. In columns (5) and (10), we restrict the observation sample to the years 1900–1914, which is the period that we use to construct the baseline treatment in the analysis. Standard errors are always clustered at the region level and are displayed in parentheses. Referenced on page A4.

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

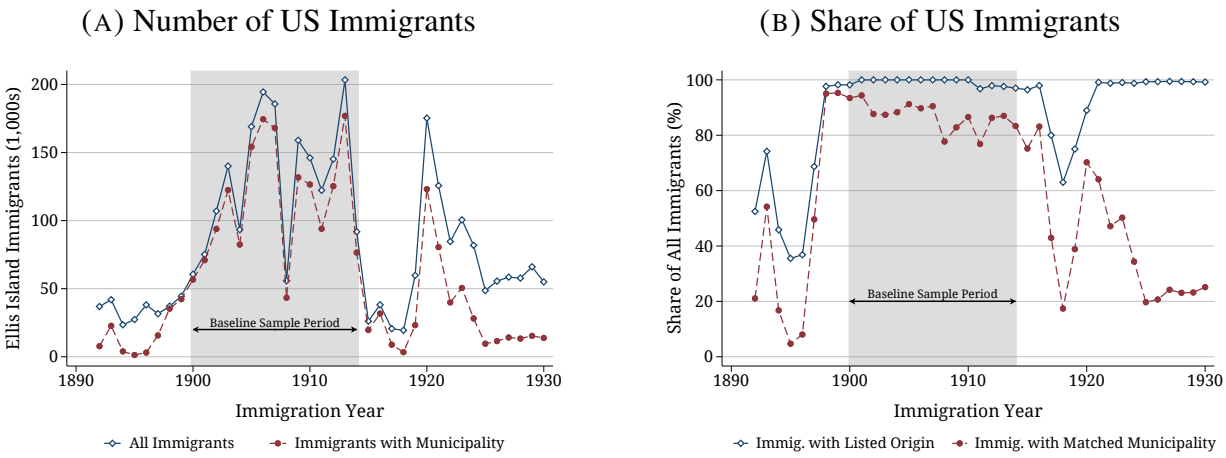
TABLE A.3. MOST COMMON EMIGRATION DISTRICTS

	Absolute Number of Emigrants			Emigration Rate		
	(1) District	(2) Emigrants	(3) Emigration Rate (‰)	(4) District	(5) Emigrants	(6) Emigration Rate (‰)
1	Palermo	89546	14.097	Termini Imerese	26069	26.358
2	Caserta	40586	11.970	Piedimonte d'Alife	11256	22.794
3	Cosenza	38821	16.673	Campobasso	26965	21.390
4	Bari delle Puglie	37918	8.485	Avellino	37422	18.932
5	Avellino	37422	18.932	Sulmona	18928	18.294
6	Girgenti	34467	12.135	Cefalu	19185	17.604
7	Salerno	33096	10.414	Cosenza	38821	16.673
8	Frosinone	29422	13.111	Isernia	22900	16.466
9	Campobasso	26965	21.390	Asiago	5214	16.429
10	Termini Imerese	26069	26.358	Sant'Angelo de' Lombardi	21404	16.020
11	Messina	25765	8.735	Cerreto Sannita	13676	15.582
12	Napoli	24625	2.479	Nola	17583	15.566
13	Isernia	22900	16.466	Corleone	9192	15.482
14	Sant'Angelo de' Lombardi	21404	16.020	Nicastro	19631	15.360
15	Gerace Marina	21094	13.711	Benevento	18385	14.410
16	Catanzaro	20535	11.606	Palermo	89546	14.096
17	Gaeta	20329	11.228	Gerace Marina	21094	13.711
18	Potenza	19792	12.651	Mistretta	8763	13.612
19	Nicastro	19631	15.360	Cotrone	11811	13.593
20	Cefalu	19185	17.604	Campagna	14273	13.572
21	Sulmona	18928	18.294	Bivona	10908	13.390
22	Aquila degli Abruzzi	18562	12.411	Melfi	13898	13.231
23	Benevento	18385	14.410	Sala Consilina	10458	13.143
24	Caltanissetta	17715	10.747	Frosinone	29422	13.111
25	Lucca	17626	4.748	Castroreale	15883	13.047
26	Nola	17583	15.566	Potenza	19792	12.650
27	Oristano	17183	12.268	Aquila degli Abruzzi	18562	12.411
28	Castellammare di Stabia	16420	7.290	Oristano	17183	12.268
29	Castroreale	15883	13.047	Bovino	6784	12.224
30	Roma	15831	1.616	Patti	15739	12.185

Notes. This Table reports the districts with the largest number of US emigrants (columns 1–3) and US emigration rates relative to the 1921 population (columns 4–6). We list the top 30 origin districts in each category. Emigration rates are expressed in per-thousand units. Referenced on page A4.

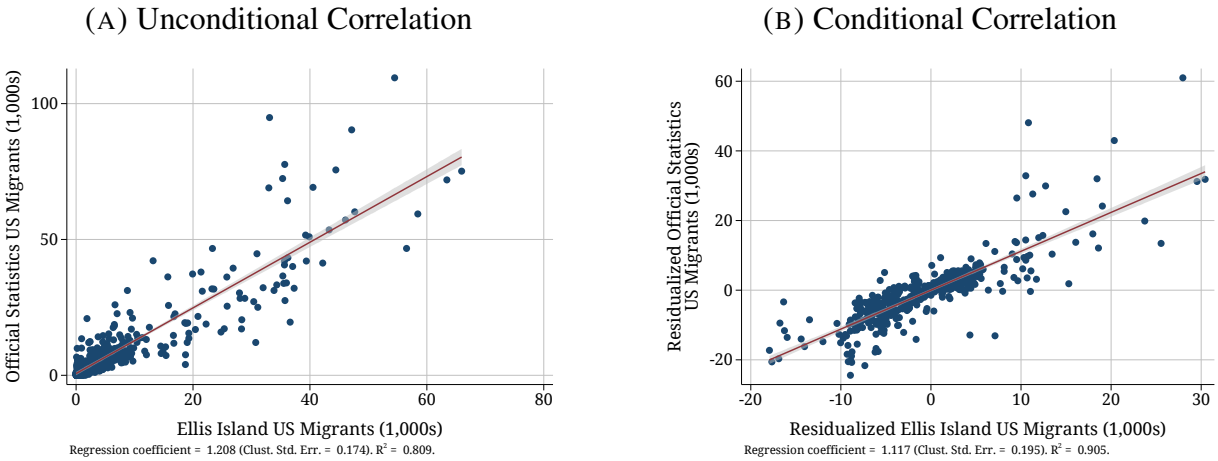
A.5 Figures

FIGURE A.1. MISSING ORIGIN IN THE ELLIS ISLAND DATASET



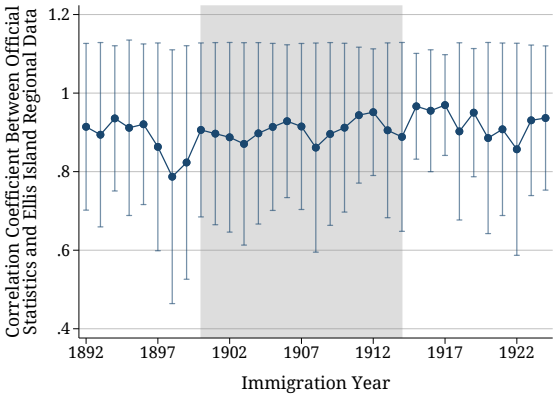
Notes. This Figure reports data on the missing municipality of origin in the Ellis Island dataset. In Panel A.1a, we display the absolute number of immigrants (in blue) and immigrants for whom we can assign a municipality of origin (in red). In Panel A.1b, we display the share of immigrants with at least one listed place of origin (in blue) and those for whom we can assign the listed origin to a municipality. Referenced on page A3.

FIGURE A.2. CORRELATION BETWEEN OFFICIAL STATISTICS AND ELLIS ISLAND US EMIGRATION



Notes. This Figure reports the correlation between the number of US emigrants recorded in the Italian official statistics and the number of Italian immigrants recorded in the Ellis Island dataset. The unit of observation is a region at a yearly frequency between 1892 and 1925. In Panel A.2a, we report the unconditional correlation. Panel A.2b displays the correlation between the variables after residualizing region and year-fixed effects. Both graphs report the fitted values of a linear regression between the variables. Referenced on page A4.

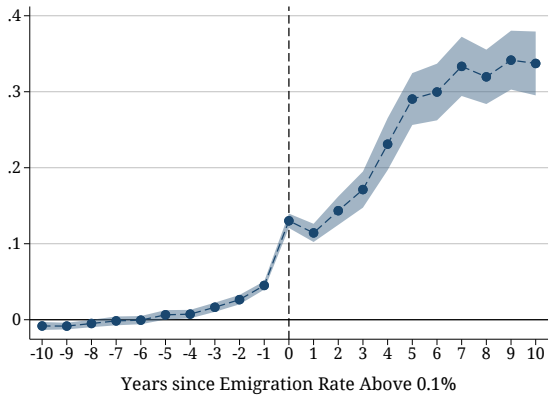
FIGURE A.3. YEAR-BY-YEAR CORRELATION BETWEEN ELLIS ISLAND AND OFFICIAL STATISTICS US EMIGRANTS



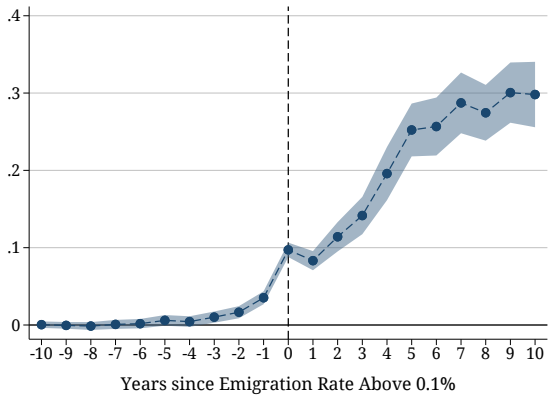
Notes. This Figure reports the correlation between the number of US emigrants recorded in the Italian official statistics and the number of Italian immigrants recorded in the Ellis Island dataset. Each dot reports the correlation between the two variables in one specific year between 1892 and 1925. In each dot, the unit of observation is a region. Both variables are standardized to have zero mean and unitary standard deviation in each year for readability. The bars report 95% confidence intervals from standard errors clustered at the regional level. The grey area marks the years that we used to construct the baseline treatment of exposure to the Quota Acts. Referenced on page A4.

FIGURE A.4. TESTING THE S-SHAPED HYPOTHESIS

(A) Unconditional Emigration

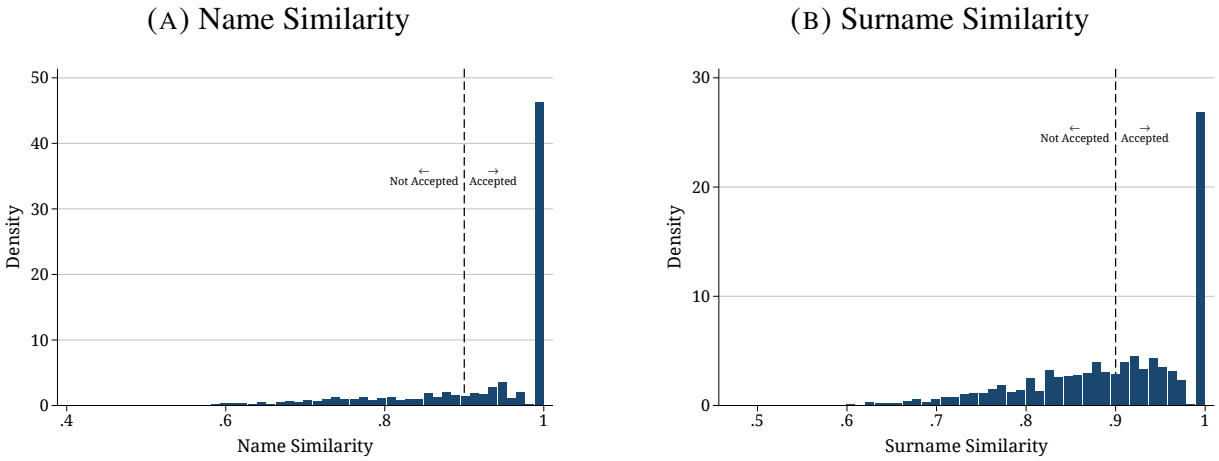


(B) Emigration Net of Fixed Effects



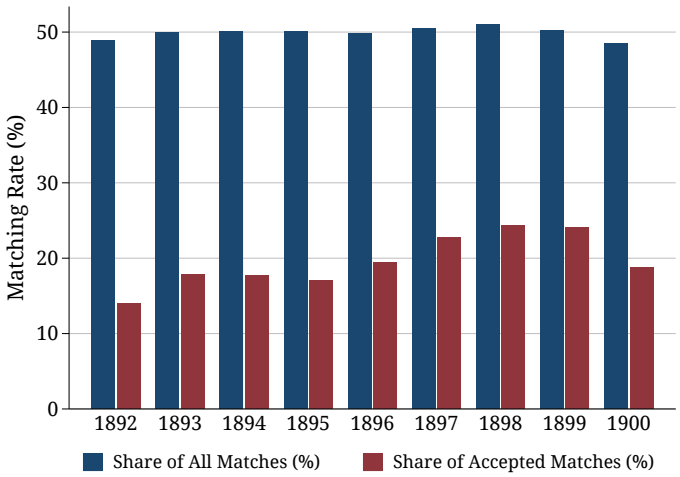
Notes. These figure test the S-shape emigration hypothesis of Gould (1980). The dependent variable is US emigration, observed at the district level every year between 1892 and 1924. For each district, we define an indicator variable that returns the number of years since the ratio between US emigrants and population in 1921 exceeds 0.1%. We report the estimated difference-in-differences coefficients obtained using the method of Borusyak *et al.* (2022) around this threshold. Panel A.4a does not include any fixed effect in the regression; in Panel A.4b, we include district and year fixed effects. Standard errors are clustered at the district level; bands report 99% confidence intervals. Referenced on page A4.

FIGURE A.5. NAME AND SURNAME SIMILARITY IN THE LINKED ELLIS ISLAND–US CENSUS SAMPLE



Notes. These figures reports the name (panel A.5a) and surname (panel A.5b) similarities between the records of Ellis Island and those that appear in the US census. To compare name and surname matches, we adopt the Monge-Elkan method with the embedded Jaro-Winkler word measure. The resulting values are normalized between zero and one, with values closer to one indicating closer comparisons. For each name and surname recorded in the Ellis Island data, we select the individual in the US census with the highest name and surname similarity among those whose initial Soundex-adjusted letter is the same as the Ellis Island record to be matched. The dashed lines mark the quality thresholds below which we reject the matches as uncertain. Referenced on page A5.

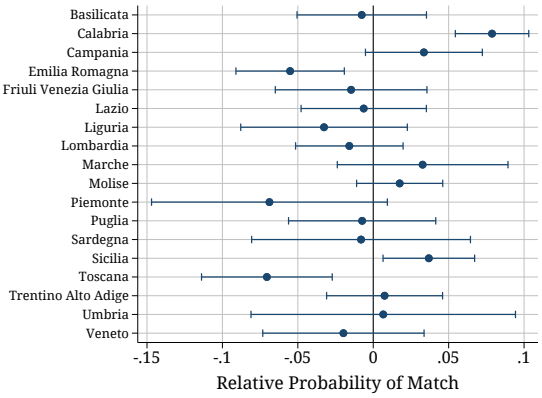
FIGURE A.6. MATCHING RATE IN THE LINKED ELLIS ISLAND–US CENSUS SAMPLE



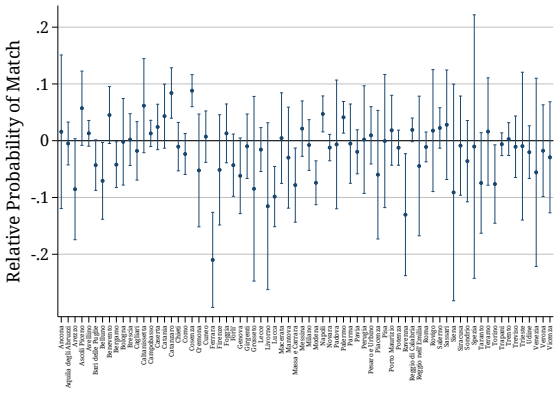
Notes. This Figure reports the matching rate in the linked sample between the Ellis Island data and the records of Italians living in the US as recorded in the population census. The blue bars report the share of individuals with at least one match in the US census. The red bars report the share of individuals with at least one acceptable match—i.e., with name and surname similarity above 0.9—in the US census. Matching rates are expressed in percentage terms and by immigration year. Referenced on page A5.

FIGURE A.7. PROBABILITY OF MATCH: BREAKDOWN BY ORIGIN OF ELLIS ISLAND IMMIGRANTS

(A) Correlation by Region of Origin



(B) Correlation by Province of Origin



Notes. These figures report the correlation between the matching status and the region (panel A.7a) and the province (panel A.7b) of origin in the Ellis Island–US Census linked sample. The sample comprises all individuals who appear in the Ellis Island dataset who immigrated between 1892 and 1900. The dependent variable is an indicator equal to one if the individual has at least one accepted match in the linked sample and zero otherwise. The right-hand side consists of a series of indicator variables which tag regions, in panel A.7a, or provinces, in panel A.7b, of origin. Standard errors are clustered at the year level, and the bands report the associated 95% confidence intervals. Referenced on page A5.

B Additional Facts and Results

B.1 Tables

TABLE B.1. FIRST-STAGE REGRESSIONS

	Dependent Variable: Measured Quota Exposure							
	Panel Regressions				Cross-Sectional Regressions			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Quota Exposure	3.743***	3.684***	6.088***	3.361**	2.632***	2.125***	2.201***	2.430***
	(0.688)	(0.701)	(0.781)	(1.513)	(0.257)	(0.266)	(0.363)	(0.406)
Controls	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Region FE	No	Yes	–	–	No	No	Yes	–
Province FE	No	No	–	–	No	No	No	Yes
District FE	No	No	Yes	Yes	–	–	–	–
Decade FE	Yes	Yes	Yes	–	–	–	–	–
Region-Decade FE	No	No	No	Yes	–	–	–	–
N. of Districts	201	201	201	201	201	201	201	190
N. of Observations	3014	3014	3014	3014	201	201	201	190
R^2	0.326	0.348	0.586	0.697	0.457	0.553	0.636	0.719
Mean Dep. Var.	2.567	2.567	2.567	2.567	1.077	1.077	1.077	1.102
Std. Beta Coef.	0.482	0.475	0.784	0.433	0.676	0.545	0.565	0.630

Notes. This Table reports the correlation between the US emigration rate and the predicted US emigration rate constructed from equation (4). The units of observation are districts. In columns (1–4), each district is observed yearly between 1900 and 1914. In columns (5–8), each district is observed once, and the outcome and the dependent variables are aggregated over time. Columns (1) and (5) report the unconditional correlations; in columns (2), (3), and (4), we additionally include region, district, and region-by-year fixed effects. In columns (6), (7), and (8), we include the baseline controls, region, and province fixed effects. Robust standard errors are displayed in parentheses. Referenced on page 15.

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

TABLE B.2. REGIONAL US AND OVERALL EMIGRATION, 1876–1925

	Emigrants to US					Emigrants to all destinations					Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	76-87	88-99	00-12	13-25	Total	76-87	88-99	00-12	13-25	Total	
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
Basilicata	28.4	53.3	108.1	38.5	228.3	74.1	106.5	179.8	70.5	431.0	53.0
Calabria	15.0	58.5	457.7	125.1	656.3	74.1	178.5	539.8	253.6	1046.1	62.7
Campania	44.3	157.5	637.8	241.5	1081.2	131.3	339.6	871.0	360.7	1702485	63.5
Emilia Romagna	1.3	8.4	62.0	24.0	95.8	60.5	137.7	422.4	178.7	799.2	12.0
Lazio	0.02	2.3	109.4	50.1	161.9	0.4	14.0	151.4	72.9	238.6	67.8
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
Marche	0.2	2.0	62.0	30.6	94.8	12.7	48.0	280.6	131.1	472.3	20.1
Piemonte	5.2	12.3	109.8	43.4	170.8	353.3	332.5	697.2	527.9	1910.8	8.9
Puglie	1.3	12.9	164.7	107.9	286.9	8.1	37.2	283.4	172.4	501.2	57.2
Sardegna	0.01	0.03	8.5	5.7	14.2	1.3	6.2	72.8	43.9	124.1	11.5
Sicilia	12.6	117.2	687.7	356.1	1173.6	26.8	170.9	946.5	516.4	1660.6	70.7
Toscana	3.3	12.9	89.6	42.0	147.8	110.7	157.5	412.4	230.6	911.2	16.2
Umbria	0.1	0.5	24.1	11.8	36.6	0.5	6.0	129.9	59.4	195.7	18.7
Veneto	1.0	6.0	52.7	48.4	108.1	486.3	1197.6	1298.2	651.0	3633.1	3.0
Total	152.1	533.9	3029.1	1326.0	5041.3	1699.3	3206.9	7635.8	4045.4	16587.4	30.4

Notes. This Table reports regional emigration towards the US and total emigration from 1876 to 1925. Figures are in thousands. Column (11) indicates the percentage of total emigrants towards US relative to all emigrants from the given region in the whole period 1876-1925. Referenced on page 6.

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 B.3. INTERNAL AND INTERNATIONAL MIGRATIONS, 1921–1931

	Absolute numbers			Share over Population	
	(1)	(2)	(3)	(4)	(5)
	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 (1) reports the population in 1881. Column (2) 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 (3) reports the number of international emigrants. Figures are in thousands. Columns (4–5) report net internal and international migration figures relative to the 1881 population. Figures are in percentage terms. Referenced on page 6.

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.

C Robustness Checks

C.1 Tables

TABLE C.1. EXPOSURE TO THE US QUOTA ACTS AND POPULATION

	Difference-in-Differences					Instrumented DiD				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Quota Exposure × Post - increasing weight	0.145***					0.172***				
	(0.029)					(0.056)				
Quota Exposure × Post - decreasing weight		0.136***					0.183***			
		(0.020)					(0.060)			
Quota Exposure × Post - migrants betweem 1900 and 1924			0.033***					0.070***		
			(0.006)					(0.023)		
Quota Exposure × Post - migrants between 1900 and 1920				0.035***					0.081***	
				(0.007)					(0.026)	
Quota Exposure × Post - migrants between 1900 and 1914					0.036***					0.117***
					(0.008)					(0.038)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Decade FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Regions in Sample	All	All	All	All	All	All	All	All	All	All
N. of Districts	188	188	188	188	188	192	192	192	192	192
N. of Observations	1115	1115	1115	1115	1115	1140	1140	1140	1140	1140
R^2	0.662	0.662	0.662	0.662	0.662	0.669	0.669	0.669	0.669	0.669
Mean Dep. Var.	16.237	16.237	16.237	16.237	16.237	17.877	17.877	17.877	17.877	17.877
Std. Beta Coef.	0.026	0.037	0.032	0.030	0.028	0.012	0.012	0.012	0.012	0.012

Notes. This Table reports the estimated effect of district-level exposure to the US Quota Acts on population. The unit of observation is a district observed at a census-decade frequency between 1881 and 1936. Each column represent the estimates of a regression where the treatment is an interaction between a post-Quota (1921) indicator variable and the district-level exposure to the Quota Acts (measured as the ratio between US and overall emigrants). We use five different measures of Quota Exposure, constructed using different weighting of the number of emigrants to the US or computing the number of US emigrants over different time spans. In column 1 and 2 we respectively assign a higher weight to emigrants departed closer to the Quota Acts date and much earlier in time: for both measures we adopt an exponential weighting of factor 0.9. In columns 3, 4 and 5 we respectively construct the Quota Exposure using the number emigrants to the US departed, respectively, between 1900 and 1924, between 1900 and 1920, and between 1900 and 1914. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. In columns 6-10 we replicate the analysis substituting the measured Quota exposure with the shift-share instrument. Robust standard errors are reported in parentheses. Referenced on page 19.

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

TABLE C.2. EXPOSURE TO THE US QUOTA ACTS AND CAPITAL INVESTMENT

	Dependent Variable: Province-Level Number of... (in 1,000 units)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Firms	Firms with Engines	Mechanical Engines	Electrical Engines	Mechanical Horsepower	Electrical Horsepower
Panel A. Increasing weight (Treatment: Measured Quota Exposure)						
I(Quota Exposure) × Post - increasing weight	-0.279**	-0.169	-0.161**	-0.966**	-0.194	-0.452
	(0.132)	(0.162)	(0.079)	(0.402)	(0.134)	(0.326)
Panel B. Decreasing weight (Treatment: Measured Quota Exposure)						
I(Quota Exposure) × Post - decreasing weight	-0.274**	-0.176	-0.217***	-0.971**	-0.187	-0.458
	(0.132)	(0.164)	(0.080)	(0.405)	(0.135)	(0.328)
Panel C. US emigrants between 1900 and 1924 (Treatment: Measured Quota Exposure)						
I(Quota Exposure) × Post - migrants between 1900 and 1924	-0.274**	-0.176	-0.217***	-0.971**	-0.187	-0.458
	(0.132)	(0.164)	(0.080)	(0.405)	(0.135)	(0.328)
Panel D. US emigrants between 1900 and 1920 (Treatment: Measured Quota Exposure)						
I(Quota Exposure) × Post - migrants between 1900 and 1920	-0.274**	-0.176	-0.217***	-0.971**	-0.187	-0.458
	(0.132)	(0.164)	(0.080)	(0.405)	(0.135)	(0.328)
Panel E. US emigrants between 1900 and 1914 (Treatment: Measured Quota Exposure)						
I(Quota Exposure) × Post - migrants between 1900 and 1914	-0.274**	-0.176	-0.217***	-0.971**	-0.187	-0.458
	(0.132)	(0.164)	(0.080)	(0.405)	(0.135)	(0.328)
Panel F. Any sample period (Treatment: Predicted Quota Exposure)						
I($\widehat{\text{Quota Exposure}}$) × Post - any measure	-0.210*	-0.292***	-0.094	-1.011**	-0.268**	-0.655**
	(0.114)	(0.112)	(0.070)	(0.402)	(0.124)	(0.258)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
N. of Provinces	69	69	69	69	69	69
N. of Observations	206	206	206	206	206	206
Mean Dep. Var.	7.714	1.304	0.567	3.589	1.470	3.092

Notes. This Table reports the estimated effect of exposure to the US Quota Acts on capital investment. The unit of observation is a province observed at a census-decade frequency between 1911 and 1936. The dependent variables are the number of firms (column 1), the number of firms with at least one engine (column 2), the number of mechanical engines (column 3), the number of electrical engines (column 4), the horsepower generated by mechanical (column 5) and electrical (column 6) engines. The treatment is an interaction between a post-Quota (1921) indicator variable and a dummy for districts with above-median exposure to the Quota Acts (measured as the ratio between US and overall emigrants). In the five panels we use different measures of Quota Exposure, constructed using different values for the number of US emigrants. In Panel A and B we respectively assign a higher weight to emigrants departed closer to the Quota Acts date and much earlier in time: for both measures we adopt an exponential weighting of factor 0.9. In Panel C, D and E we respectively construct the Quota Exposure using the number of emigrants to the US departed, respectively, between 1900 and 1924, between 1900 and 1920, and between 1900 and 1914. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. In Panel F, we substitute our measured quota exposure with the shift share instrument: we use an indicator variable for province with above-median (predicted) exposure to the Quota Acts. We report the results in Panel F only, since the indicator variable takes the same value, regardless of the way we construct it using different sample periods. Robust standard errors are reported in parentheses. Referenced on page 19.

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

TABLE C.3. EXPOSURE TO THE US QUOTA ACTS AND MANUFACTURE EMPLOYMENT

	Difference-in-Differences					Instrumented DiD				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Quota Exposure × Post - increasing weight	0.151***					0.547***				
	(0.033)					(0.176)				
Quota Exposure × Post - decreasing weight		0.219***					0.583***			
		(0.041)					(0.190)			
Quota Exposure × Post - migrants betweem 1900 and 1924			0.072***					0.221***		
			(0.015)					(0.072)		
Quota Exposure × Post - migrants between 1900 and 1920				0.075***					0.258***	
				(0.015)					(0.083)	
Quota Exposure × Post - migrants between 1900 and 1914					0.082***					0.372***
					(0.017)					(0.120)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Decade FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Regions in Sample	All	All	All	All	All	All	All	All	All	All
N. of Districts	188	188	188	188	188	192	192	192	192	192
N. of Observations	1115	1115	1115	1115	1115	1140	1140	1140	1140	1140
R^2	0.808	0.808	0.808	0.808	0.808	0.775	0.775	0.775	0.775	0.775
Mean Dep. Var.	14.550	14.550	14.550	14.550	14.550	15.745	15.745	15.745	15.745	15.745
Std. Beta Coef.	0.062	0.069	0.066	0.065	0.063	0.038	0.038	0.038	0.038	0.038

Notes. This Table reports the estimated effect of exposure to the US Quota Acts on manufacturing employment. The unit of observation is a district observed at a census-decade frequency between 1881 and 1936. Each column represent the estimates of a regression where the treatment is an interaction between a post-Quota (1921) indicator variable and the district-level exposure to the Quota Acts (measured as the ratio between US and overall emigrants). We use five different measures of Quota Exposure, constructed using different weighting of the number of emigrants to the US or computing the number of US emigrants over different time spans. In column 1 and 2 we respectively assign a higher weight to emigrants departed closer to the Quota Acts date and much earlier in time: for both measures we adopt an exponential weighting of factor 0.9. In columns 3, 4 and 5 we respectively construct the Quota Exposure using the number emigrants to the US departed, respectively, between 1900 and 1924, between 1900 and 1920, and between 1900 and 1914. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. In columns 6-10 we replicate the analysis substituting the measured Quota exposure with the shift-share instrument. Robust standard errors are reported in parentheses. Referenced on page 19.

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

TABLE C.4. EXPOSURE TO THE QUOTA ACTS AND AGRICULTURE EMPLOYMENT

	Difference-in-Differences					Instrumented DiD				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Quota Exposure × Post - increasing weight	-0.020					-0.325**				
	(0.014)					(0.153)				
Quota Exposure × Post - decreasing weight		-0.025					-0.347**			
		(0.019)					(0.163)			
Quota Exposure × Post - migrants betweem 1900 and 1924			-0.009					-0.132**		
			(0.007)					(0.062)		
Quota Exposure × Post - migrants between 1900 and 1920				-0.009					-0.153**	
				(0.007)					(0.072)	
Quota Exposure × Post - migrants between 1900 and 1914					-0.010					-0.220**
					(0.008)					(0.103)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Decade FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Regions in Sample	All	All	All	All	All	All	All	All	All	All
N. of Districts	188	188	188	188	188	192	192	192	192	192
N. of Observations	927	927	927	927	927	948	948	948	948	948
R^2	0.598	0.598	0.598	0.598	0.598	0.615	0.615	0.615	0.615	0.615
Mean Dep. Var.	27.282	27.282	27.282	27.282	27.282	30.368	30.368	30.368	30.368	30.368
Std. Beta Coef.	-0.007	-0.007	-0.007	-0.007	-0.007	-0.020	-0.020	-0.020	-0.020	-0.020

Notes. This Table reports the estimated effect of exposure to the US Quota Acts on agriculture employment. The unit of observation is a district observed at a census-decade frequency between 1881 and 1936. Each column represent the estimates of a regression where the treatment is an interaction between a post-Quota (1921) indicator variable and the district-level exposure to the Quota Acts (measured as the ratio between US and overall emigrants). We use five different measures of Quota Exposure, constructed using different weighting of the number of emigrants to the US or computing the number of US emigrants over different time spans. In column 1 and 2 we respectively assign a higher weight to emigrants departed closer to the Quota Acts date and much earlier in time: for both measures we adopt an exponential weighting of factor 0.9. In columns 3, 4 and 5 we respectively construct the Quota Exposure using the number emigrants to the US departed, respectively, between 1900 and 1924, between 1900 and 1920 and between 1900 and 1914. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. In columns 6-10 we replicate the analysis substituting the measured Quota exposure with the shift-share instrument. Robust standard errors are reported in parentheses. Referenced on page 19.

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

TABLE C.5. EXPOSURE TO THE US QUOTA ACTS AND POPULATION

	Difference-in-Differences						Instrumented DiD					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Quota Exposure × Post	0.018***	0.021***	0.031***	0.034***	0.031***	0.036***						
	(0.005)	(0.007)	(0.008)	(0.008)	(0.007)	(0.008)						
$\widehat{\text{Quota Exposure}}$ × Post							0.119***	0.082***	0.077**	0.111***	0.066**	0.115***
							(0.037)	(0.032)	(0.037)	(0.038)	(0.033)	(0.039)
Literacy × Post	1.102***						-0.006					
	(0.095)						(0.057)					
Urbanization × Post		0.361***						0.239***				
		(0.035)						(0.027)				
Altitude × Post			-0.000***						-0.000***			
			(0.000)						(0.000)			
Railway × Post				0.048***						0.068***		
				(0.011)						(0.014)		
WW1 deaths × Post					-13.777***						-7.721***	
					(1.957)						(2.750)	
US GDP growth × Quota Exposure × Post						0.000						
						(0.001)						
US GDP growth × $\widehat{\text{Quota Exposure}}$ × Post												0.003
												(0.002)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Regions in Sample	All	All	All	All	All	All	All	All	All	All	All	All
N. of Districts	188	188	188	188	188	188	192	192	192	192	192	192
N. of Observations	1115	1115	1115	1115	1115	1115	1140	1140	1140	1140	1140	1140
R^2	0.663	0.662	0.662	0.662	0.662	0.662	0.669	0.671	0.670	0.670	0.670	0.669
Mean Dep. Var.	16.237	16.237	16.237	16.237	16.237	16.237	17.877	17.877	17.877	17.877	17.877	17.877
Std. Beta Coef.	0.014	0.016	0.024	0.027	0.024	0.028	0.012	0.008	0.008	0.011	0.007	0.012

Notes. This Table reports the estimated effect of district-level exposure to the US Quota Acts on population. We replicate the results obtained in column 6 of Table 3. In each column we control for the interaction between the post-Quota indicator variable and a district- or country-specific variable. In column 1 we use the district's literacy rate in 1901; in column 2 the urbanization rate in 1901, measured as the share of people living in towns with more than 5'000 inhabitants in the district; in column 3 the altitude at which the district is located; in column 4 an indicator variable specifying whether the district was connected to the railway network before 1901; in column 5 district's number of deaths caused by WW1, measured as share of the district's population. In column 6 we use the US GDP growth: being this a time-varying variable common to all districts, we further interact it with the district's Quota Exposure. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. In columns 7-12 we replicate the analysis substituting the measured Quota exposure with the shift-share instrument. Robust standard errors are reported in parentheses. Referenced on page 19, 20.

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

TABLE C.6. EXPOSURE TO THE US QUOTA ACTS AND CAPITAL INVESTMENT

	Dependent Variable: Province-Level Number of... (in 1,000 units)					
	(1) Firms	(2) Firms with Engines	(3) Mechanical Engines	(4) Electrical Engines	(5) Mechanical Horsepower	(6) Electrical Horsepower
Panel A. Control for: Literacy × Post						
I(Quota Exposure) × Post	-0.219*	-0.100	-0.180**	-0.803*	-0.122	-0.342
	(0.131)	(0.161)	(0.088)	(0.411)	(0.126)	(0.323)
Panel B. Control for: Urbanization × Post						
I(Quota Exposure) × Post	-0.324***	-0.170	-0.187**	-1.085**	-0.201	-0.478
	(0.123)	(0.173)	(0.079)	(0.439)	(0.130)	(0.337)
Panel C. Control for: Altitude × Post						
I(Quota Exposure) × Post	-0.287**	-0.167	-0.210***	-1.021**	-0.165	-0.458
	(0.136)	(0.175)	(0.079)	(0.409)	(0.144)	(0.332)
Panel D. Control for: Railway access × Post						
I(Quota Exposure) × Post	-0.275**	-0.195	-0.217***	-0.963**	-0.157	-0.529
	(0.131)	(0.163)	(0.080)	(0.421)	(0.139)	(0.342)
Panel E. Control for: WW1 death rate × Post						
I(Quota Exposure) × Post	-0.188*	-0.111	-0.213***	-0.873**	-0.067	-0.335
	(0.109)	(0.158)	(0.081)	(0.389)	(0.135)	(0.331)
Panel F. Control for: US GDP growth × I(Quota Exposure) × Post						
I(Quota Exposure) × Post	-0.300**	-0.300*	-0.438***	-0.797*	-0.339**	-0.297
	(0.138)	(0.169)	(0.103)	(0.409)	(0.160)	(0.348)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
N. of Provinces	68	68	68	68	68	68
N. of Observations	204	204	204	204	204	204
Mean Dep. Var.	7.665	1.297	0.567	3.528	1.455	3.028

Notes. This Table replicates Panel A of Table 4. In the five panels we add a different control, given by the interaction between the post-Quota indicator variable and a province-specific variable. We use respectively: the literacy rate in 1901; the urbanization rate in 1901 (share of people in towns with more than 5'000 inhabitants); the altitude; the share of municipalities connected to the railway network before 1901; the number of deaths caused by WW1, measured as population's share; the US GDP growth interacted with the province's Quota Exposure. Referenced on page 19, 20.
***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.10$

TABLE C.7. EXPOSURE TO THE US QUOTA ACTS AND CAPITAL INVESTMENT: IV-DID

	Dependent Variable: Province-Level Number of... (in 1,000 units)					
	(1) Firms	(2) Firms with Engines	(3) Mechanical Engines	(4) Electrical Engines	(5) Mechanical Horsepower	(6) Electrical Horsepower
Panel A. Control for: Literacy × Post						
$\widehat{\text{I(Quota Exposure)}} \times$ Post	-0.125	-0.220*	-0.013	-0.776	-0.168	-0.458
	(0.126)	(0.130)	(0.094)	(0.482)	(0.142)	(0.319)
Panel B. Control for: Urbanization × Post						
$\widehat{\text{I(Quota Exposure)}} \times$ Post	-0.301***	-0.299**	-0.050	-1.215**	-0.322***	-0.726**
	(0.107)	(0.124)	(0.071)	(0.472)	(0.121)	(0.295)
Panel C. Control for: Altitude × Post						
$\widehat{\text{I(Quota Exposure)}} \times$ Post	-0.243**	-0.284**	-0.080	-1.268***	-0.212	-0.682**
	(0.124)	(0.135)	(0.071)	(0.436)	(0.155)	(0.312)
Panel D. Control for: Railway access × Post						
$\widehat{\text{I(Quota Exposure)}} \times$ Post	-0.211*	-0.306***	-0.093	-0.993**	-0.257**	-0.675***
	(0.112)	(0.110)	(0.069)	(0.402)	(0.123)	(0.251)
Panel E. Control for: WW1 death rate × Post						
$\widehat{\text{I(Quota Exposure)}} \times$ Post	-0.088	-0.220**	-0.084	-0.878**	-0.141	-0.503*
	(0.097)	(0.110)	(0.080)	(0.393)	(0.142)	(0.269)
Panel F. Control for: US GDP growth × I(Quota Exposure) × Post						
$\widehat{\text{I(Quota Exposure)}} \times$ Post	-0.180	-0.320***	-0.240**	-0.785*	-0.278*	-0.387
	(0.121)	(0.119)	(0.094)	(0.404)	(0.151)	(0.277)
Province FE						
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
N. of Provinces	68	68	68	68	68	68
N. of Observations	204	204	204	204	204	204
Mean Dep. Var.	7.665	1.297	0.567	3.528	1.455	3.028

Notes. This Table replicates Panel B of Table 4, where we substitute quota exposure with the shift share instrument. In the five panels we add a different control, given by the interaction between the post-Quota indicator variable and a province-specific variable. We use respectively: the literacy rate in 1901; the urbanization rate in 1901 (share of people in towns with more than 5'000 inhabitants); the altitude; the share of municipalities connected to the railway network before 1901; the number of deaths caused by WW1, measured as population's share; the US GDP growth interacted with the province's Quota Exposure. Referenced on page 19.

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

TABLE C.8. EXPOSURE TO THE US QUOTA ACTS AND MANUFACTURE EMPLOYMENT

	Difference-in-Differences						Instrumented DiD					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Quota Exposure × Post	0.048***	0.077***	0.085***	0.080***	0.073***	0.082***						
	(0.015)	(0.017)	(0.018)	(0.018)	(0.016)	(0.017)						
$\widehat{\text{Quota Exposure}}$ × Post							0.151	0.368***	0.350***	0.358***	0.283**	0.383***
							(0.124)	(0.116)	(0.119)	(0.121)	(0.123)	(0.119)
Literacy × Post	1.088***						0.714***					
	(0.255)						(0.131)					
Urbanization × Post		0.068						0.013				
		(0.085)						(0.078)				
Altitude × Post			0.000						-0.000			
			(0.000)						(0.000)			
Railway × Post				0.040						0.211***		
				(0.033)						(0.051)		
WW1 deaths × Post					-8.286*						-9.983**	
					(4.757)						(5.029)	
US GDP growth × Quota Exposure × Post						-0.000						
						(0.002)						
US GDP growth × $\widehat{\text{Quota Exposure}}$ × Post												-0.016
												(0.011)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Regions in Sample	All	All	All	All	All	All	All	All	All	All	All	All
N. of Districts	188	188	188	188	188	188	192	192	192	192	192	192
N. of Observations	1115	1115	1115	1115	1115	1115	1140	1140	1140	1140	1140	1140
R^2	0.808	0.808	0.808	0.808	0.808	0.808	0.777	0.775	0.776	0.776	0.776	0.776
Mean Dep. Var.	14.550	14.550	14.550	14.550	14.550	14.550	15.745	15.745	15.745	15.745	15.745	15.745
Std. Beta Coef.	0.038	0.060	0.066	0.062	0.056	0.063	0.016	0.038	0.036	0.037	0.029	0.039

Notes. This Table reports the estimated effect of district-level exposure to the US Quota Acts on population. We replicate the results obtained in column 6 of Panel A in Table 5. In each column we control for the interaction between the post-Quota indicator variable and a district- or country-specific variable. In column 1 we use the district's literacy rate in 1901; in column 2 the urbanization rate in 1901, measured as the share of people living in towns with more than 5'000 inhabitants in the district; in column 3 the altitude at which the district is located; in column 4 an indicator variable specifying whether the district was connected to the railway network before 1901; in column 5 district's number of deaths caused by WW1, measured as share of the district's population. In column 6 we use the US GDP growth: being this a time-varying variable common to all districts, we further interact it with the district's Quota Exposure. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. In columns 7-12 we replicate the analysis substituting the measured Quota exposure with the shift-share instrument. Robust standard errors are reported in parentheses. Referenced on page 19, 20.

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

TABLE C.9. EXPOSURE TO THE QUOTA ACTS AND AGRICULTURE EMPLOYMENT

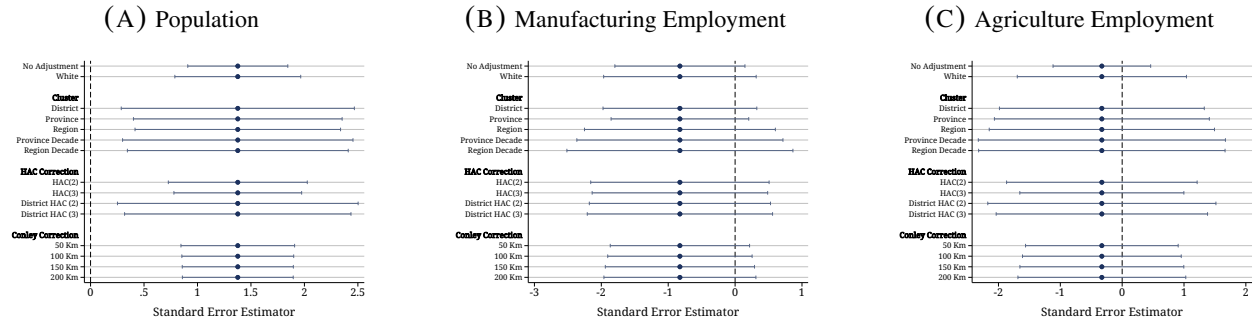
	Difference-in-Differences						Instrumented DiD					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Quota Exposure × Post	-0.008	-0.010	-0.012	-0.011	-0.009	-0.008						
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)						
$\widehat{\text{Quota Exposure}}$ × Post							-0.021	-0.203*	-0.215**	-0.215**	-0.181*	-0.187**
							(0.088)	(0.108)	(0.100)	(0.103)	(0.097)	(0.079)
Literacy × Post	-0.179						-0.797***					
	(0.115)						(0.064)					
Urbanization × Post		0.003						0.082*				
		(0.032)						(0.046)				
Altitude × Post			-0.000***						-0.000***			
			(0.000)						(0.000)			
Railway × Post				0.032*						-0.045		
				(0.018)						(0.028)		
WW1 deaths × Post					3.357**						8.389***	
					(1.625)						(2.486)	
US GDP growth × Quota Exposure × Post						-0.001						
						(0.001)						
US GDP growth × $\widehat{\text{Quota Exposure}}$ × Post												-0.007
												(0.015)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Regions in Sample	All	All	All	All	All	All	All	All	All	All	All	All
N. of Districts	188	188	188	188	188	188	192	192	192	192	192	192
N. of Observations	927	927	927	927	927	927	948	948	948	948	948	948
R^2	0.598	0.598	0.598	0.598	0.598	0.598	0.621	0.616	0.616	0.615	0.616	0.615
Mean Dep. Var.	27.282	27.282	27.282	27.282	27.282	27.282	30.368	30.368	30.368	30.368	30.368	30.368
Std. Beta Coef.	-0.006	-0.007	-0.009	-0.008	-0.007	-0.005	-0.002	-0.018	-0.019	-0.019	-0.016	-0.017

Notes. This Table reports the estimated effect of district-level exposure to the US Quota Acts on population. We replicate the results obtained in column 6 of Panel B in Table 5. In each column we control for the interaction between the post-Quota indicator variable and a district- or country-specific variable. In column 1 we use the district's literacy rate in 1901; in column 2 the urbanization rate in 1901, measured as the share of people living in towns with more than 5'000 inhabitants in the district; in column 3 the altitude at which the district is located; in column 4 an indicator variable specifying whether the district was connected to the railway network before 1901; in column 5 district's number of deaths caused by WW1, measured as share of the district's population. In column 6 we use the US GDP growth: being this a time-varying variable common to all districts, we further interact it with the district's Quota Exposure. Each regression controls for the emigration rate and an indicator variable for Southern regions, both interacted with a post-1921 indicator. In columns 7-12 we replicate the analysis substituting the measured Quota exposure with the shift-share instrument. Robust standard errors are reported in parentheses. Referenced on page 19, 20.

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

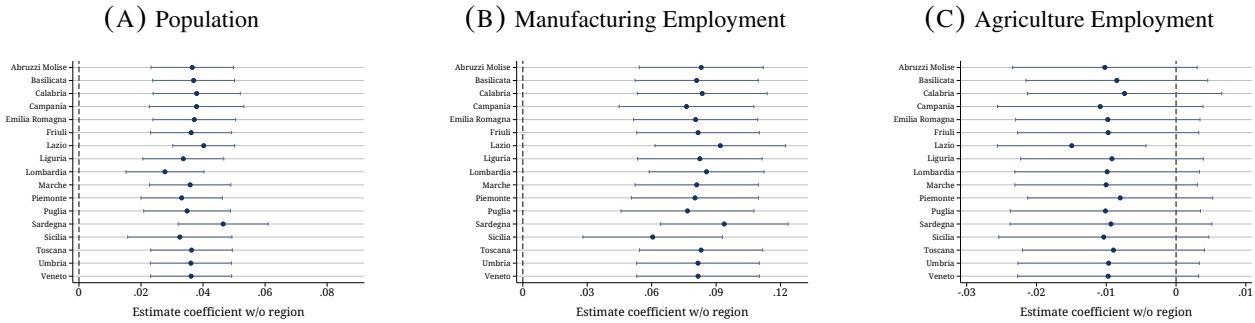
C.2 Figures

FIGURE C.1. ALTERNATIVE STANDARD ERRORS: DISTRICT-LEVEL LINEAR PROBABILITY MODELS



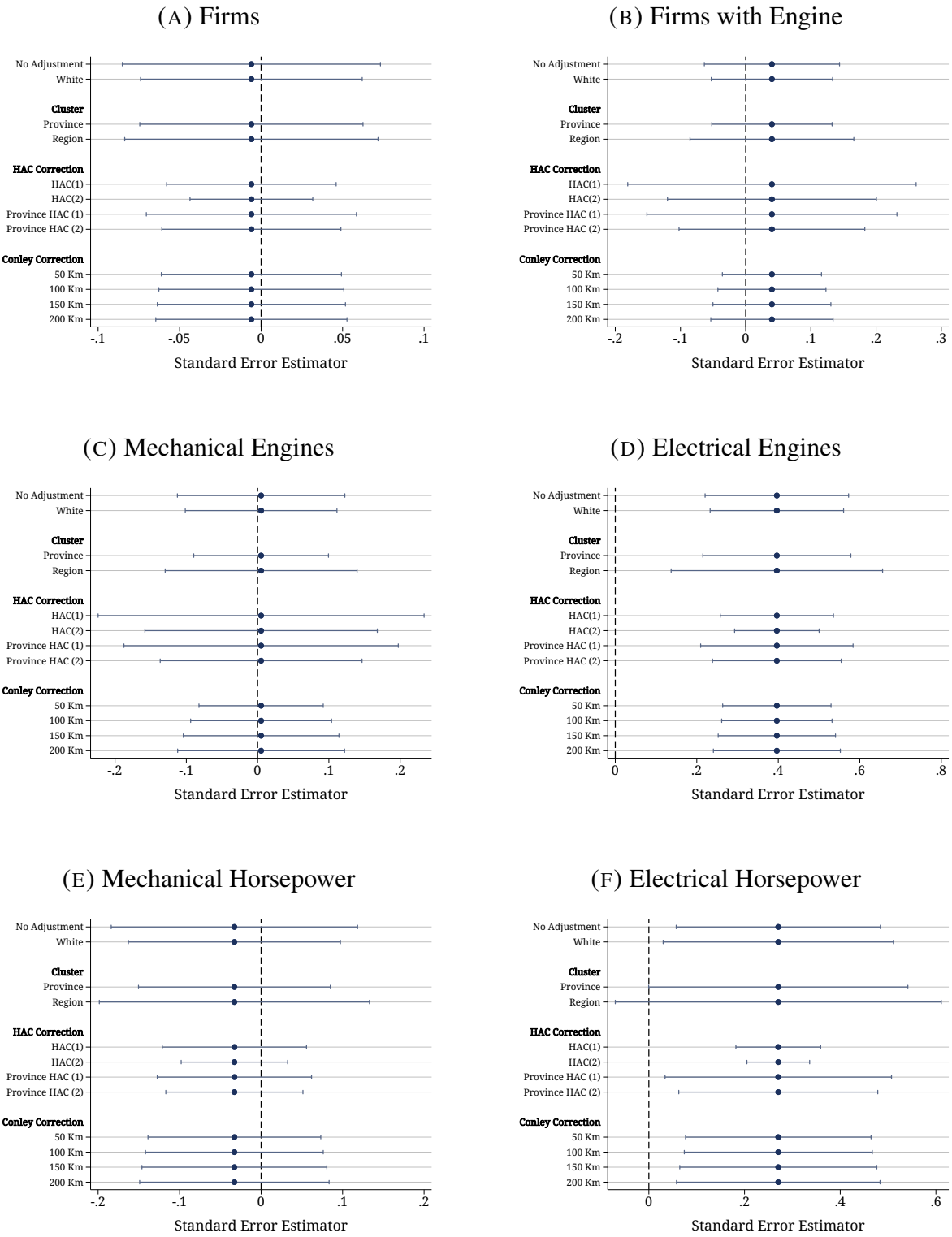
Notes. This Figure reports alternative estimators of the standard errors in the baseline district-level regressions. The unit of observation is a district, at a decade frequency between 1881 and 1936. The dependent variable is population (Panel C.1a), manufacturing employment (Panel C.1b), and agriculture employment (Panel C.1c). Regressions are estimated through OLS and include district and decade-fixed effects. Bands report 90% confidence intervals. Referenced on page 20.

FIGURE C.2. EXCLUDING ONE REGION AT THE TIME: DISTRICT-LEVEL REGRESSIONS



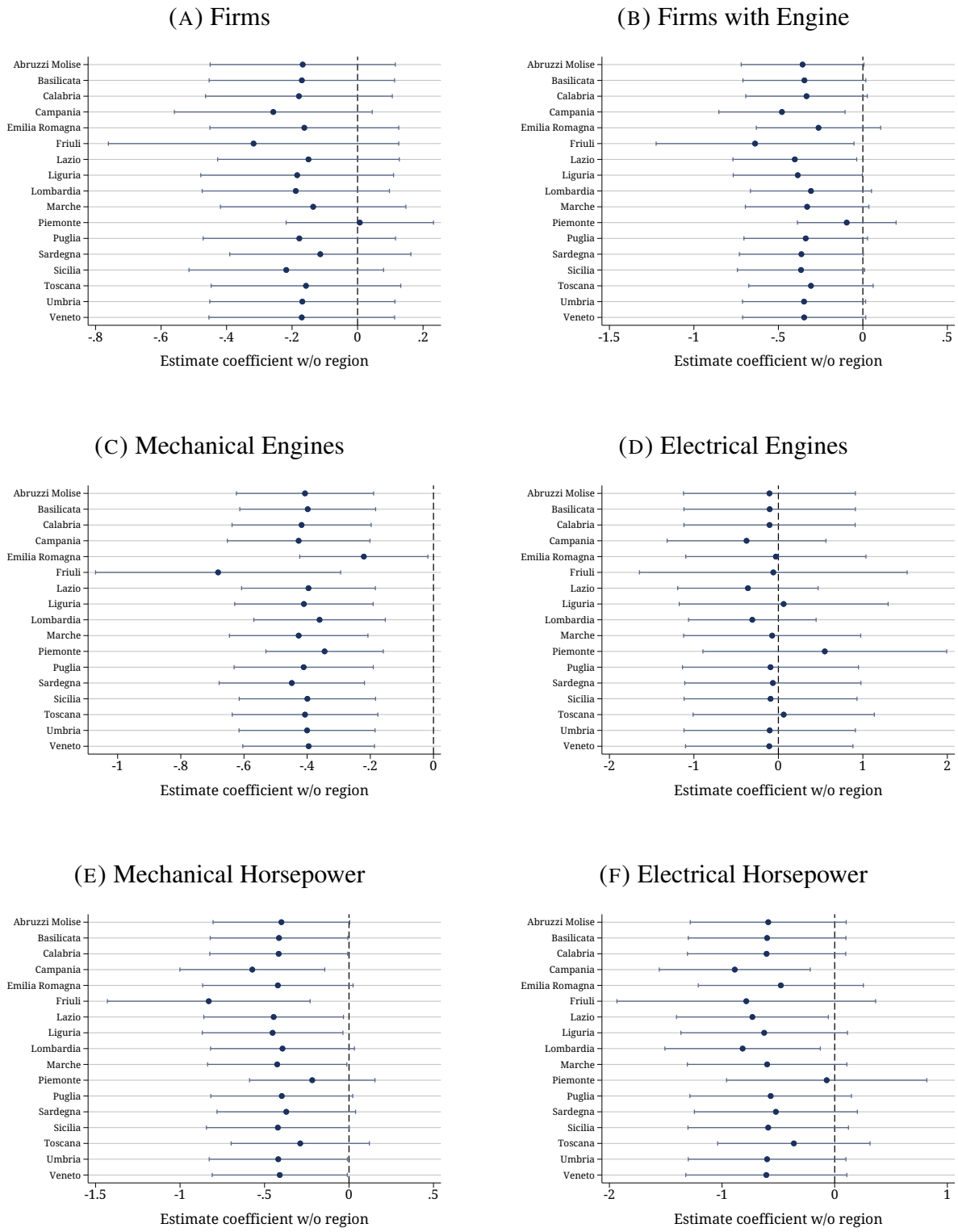
Notes. This Figure reports the estimates of the baseline district-level regressions when excluding one region. The unit of observation is a district, at a decade frequency between 1881 and 1936. The dependent variable is population (Panel C.2a), manufacturing employment (Panel C.2b), and agriculture employment (Panel C.2c). Regressions are estimated through Poisson regression and include district and province-decade fixed effects: refer to column 6 of Table 3 and 5. The left axis indicates the excluded region. Bands report 90% confidence intervals. Referenced on page 19.

FIGURE C.3. ALTERNATIVE STANDARD ERRORS: PROVINCE-LEVEL REGRESSIONS



Notes. This Figure reports alternative estimators of the standard errors in the baseline province-level regressions. The unit of observation is a province, at a decade frequency between 1911 and 1937. Regressions are estimated through OLS on the $\log(1+\cdot)$-transformed variable and include province and decade-fixed effects. Bands report 90% confidence intervals. Referenced on page 20.

FIGURE C.4. EXCLUDING ONE REGION AT THE TIME: PROVINCE-LEVEL REGRESSIONS



Notes. This Figure reports the estimates of the baseline province-level regressions when excluding one region. The unit of observation is a province, at a decade frequency between 1911 and 1937. Regressions are estimated through a double-difference (DiD) Poisson regression and include province and decade-fixed effects: refer to Panel A of Table 4. The left axis indicates the excluded region. Bands report 90% confidence intervals. Referenced on page 19.

D Theoretical Framework

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 D.3.

D.1 Theoretical Framework

In this section, we develop a simple analytical framework inspired to Zeira (1998) and San (2023) 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, as in 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 **?** 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:

$$Y = A\left[\int_0^{\iota} mx(j)^{\alpha}\,dj + \int_{\iota}^{1} e(j)^{\alpha}\,dj\right] \tag{D.1}$$

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 (2023), 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

$$\max_{\iota,\{x(j),e(j)\}_{j\in[0,1]}} Y - \int_0^{\iota} p(j)x(j)\,dj - w\int_{\iota}^{1} e(j)\,dj \tag{D.2}$$

where $p(j)$ is the price of a machine for task j, w is the nominal wage, subject to the technology constraint (D.1). Note that the price of the consumption good is implicitly normalized to one. In section D.3, we formally show that the demand schedules for machines and labor are given by the following demand schedules:

$$x(j) = p(j)^{-\frac{1}{1-\alpha}}(\alpha A m)^{\frac{1}{1-\alpha}} \quad \forall j \in [0,\iota] \tag{D.3a}$$

$$e(j) = w^{-\frac{1}{1-\alpha}}(\alpha A)^{\frac{1}{1-\alpha}} \qquad \forall j \in [\iota,1] \tag{D.3b}$$

Combining (D.3a)-(D.3b) with the first order condition for the industrialization rate, it follows that in the equilibrium ι^* is pinned down by the following:

$$m = \left[\frac{p(\iota^*)}{w}\right]^{\alpha} \tag{D.4}$$

The economic intuition behind condition (D.4) 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 a demand for machines (D.3a). 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.[23] 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

$$\max_{p(j)} [p(j) - \psi(j)]\,x(j) \tag{D.5}$$

subject to (D.3a). In section D.3, we show that the first-order conditions imply

$$p(j) = \min\left\{mw, \frac{\psi(j)}{\alpha}\right\} \tag{D.6}$$

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 the 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 D.3.

[23] 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.

Lemma 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 D.1. *There exists one and only one $\iota^* \in [0,1]$ which solves the problem of the final good producer (D.3a)-(D.3b)-(D.4) as well as the problem of the machine producers (D.6) and verifies labor market clearing. In particular, the equilibrium industrialization ι^* is the solution to the following:*

$$\psi(\iota^*) = L^{\alpha-1}(1-\iota^*)^{1-\alpha}\alpha^2 A m^{1/\alpha}.$$

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.

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 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 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 (D.3b) 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 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 that were more exposed to the Quota Acts decreased investment in machinery—section III.3—and increased employment—-section III.4. These findings are fully in line with the empirical predictions D.2 of the model and hence provide evidence in favor of labor-saving directed technical adoption.

D.3 Proofs of Analytical Results

Solution of the problem of the final good producer. Plugging the technology constraint into problem (D.2), the problem of the final good producer reads out as follows:

$$\max_{\iota,\{x(j),e(j)\}_{j\in[0,1]}} A\left[\int_0^{\iota} mx(j)^{\alpha}\,dj+\int_{\iota}^{1} e(j)^{\alpha}\,dj\right]-\int_0^{\iota} p(j)x(j)\,dj-w\int_{\iota}^{1} e(j)\,dj$$

The—necessary and sufficient—first-order conditions with respect to labor and capital in the generic task j are

$$x(j)=p(j)^{-\frac{1}{1-\alpha}}(\alpha Am)^{\frac{1}{\alpha}}\ \forall j\in[0,\iota]$$
$$e(j)=w^{-\frac{1}{1-\alpha}}(\alpha A)^{\frac{1}{\alpha}}\quad \forall j\in[\iota,1]$$

To obtain the first-order condition for the optimal industrialization rate, apply the Leibniz integral rule with respect to ι to get:

$$x(\iota^*)\left[mx(\iota^*)^{\alpha-1}-p(\iota^*)\right]=e(\iota^*)\left[e(\iota^*)^{\alpha-1}-w\right]$$

Plugging (D.3a)-(D.3b) into the expression above we get $m=(p(\iota^*)/w)^{\alpha}$. □

Solution of the problem of the monopolist. The solution is trivial upon plugging (D.3a) into the objective function (D.5). □

Proof of Lemma D.1. From (D.6) and (D.4), it is

$$p(\iota^*)=\min\left\{\frac{\psi(\iota^*)}{\alpha},mw\right\}$$
$$p(\iota^*)=m^{1/\alpha}w$$

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 $\psi(\iota^*)/\alpha=wm^{1/\alpha}$. Because $m\in(0,1)$, $m^{1/\alpha}<m$ since $\alpha\in(0,1)$, and therefore $\psi(\iota^*)/\alpha=wm^{1/\alpha}<wm$. This implies that the solution is acceptable. Hence, $p(\iota^*)=\psi(\iota^*)/\alpha$ and this concludes the proof. □

Proof of Proposition D.1. Because $w(j) = w$ for all $j \in [0,1]$, from (D.3b) 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 D.1, it is $w = \psi(\iota^*)/(\alpha m^{1/\alpha})$. Plugging this into the previous equation, we get

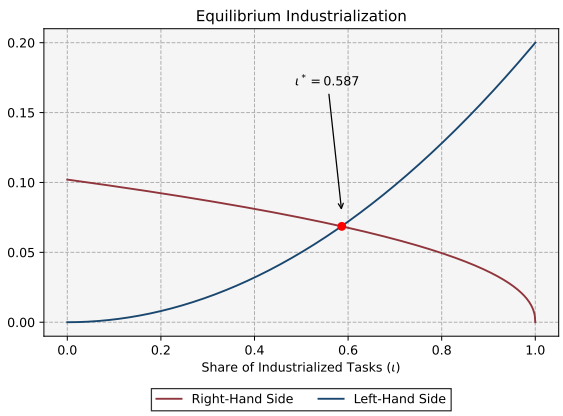
$$\left(\frac{\psi(\iota^*)}{\alpha m^{1/\alpha}}\right)^{-\frac{1}{1-\alpha}}(\alpha\beta)^{\frac{1}{1-\alpha}} = \frac{L}{1-\iota^*}$$

$$\frac{\psi(\iota^*)}{\alpha m^{1/\alpha}}(\alpha\beta)^{-1} = \left(\frac{L}{1-\iota^*}\right)^{-1+\alpha}$$

$$\psi(\iota^*)L^{1-\beta} = (1-\iota^*)^{1-\alpha}\alpha^2 A m^{1/\alpha}$$

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. □

FIGURE D.1. EQUILIBRIUM OF THE MODEL



Notes. This figure plots the equilibrium of the model. The blue and red lines, respectively, display the left and right-hand sides of the final equation of the proof of Proposition D.1. We assume $\psi(j) = \gamma j^2$ even though quadratic costs do not verify the Inada conditions.
Parametrization: $\alpha = .55$, $\beta = .45$, $\gamma = .2$, $A = .5$, $L = 1$, $m = .5$.

Proof of Implication D.1. From Lemma 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 D.2. First note that because w is invariant across tasks, then by (D.3b) $e(j) = e$ 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 (D.3b), yielding

$$w^{-\frac{1}{1-\alpha}}(\alpha A)^{\frac{1}{1-\alpha}} = \frac{L}{1-\iota^*}$$

Using Lemma D.1 into the previous equation we get

$$\frac{\psi(\iota^*)}{\alpha m^{1/\alpha}} = \left(\frac{L}{1-\iota^*}\right)^{\alpha-1} \alpha A$$
$$L^{1-\alpha} = \frac{(1-\iota^*)^{1-\alpha}}{\psi(\iota^*)} \alpha^2 A m^{1/\alpha}$$

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. □

Appendix References

ABRAMITZKY, R., L. BOUSTAN, K. ERIKSSON, J. FEIGENBAUM and S. PÉREZ (2021). "Automated Linking of Historical Data." *Journal of Economic Literature*, 59(3): 865–918. (Cited on p. A5)

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. (Cited on p. A3)

ACEMOGLU, D. (2002). "Directed Technical Change." *The Review of Economic Studies*, 69(4): 781–809. (Cited on p. 3, 18, 22, D29)

——— (2007). "Equilibrium Bias of Technology." *Econometrica*, 75(5): 1371–1409. (Cited on p. 3, D29)

BORUSYAK, K., X. JARAVEL and J. SPIESS (2022). "Revisiting Event Study Designs: Robust and Efficient Estimation." (Cited on p. A4, A11)

CALABRESE, V. (2017). "Land of Women: Basilicata, Emigration, and the Women Who Remained Behind, 1880-1914." *Dissertations, Theses, and Capstone Projects*. (Cited on p. A2)

COLUCCIA, D. M. and G. DOSSI (2023). "Return Innovation: The Knowledge Spillovers of the British Migration to the United States, 1870–1940." (Cited on p. 4, A3)

ECKERT, F., A. GVIRTZ, J. LIANG and M. PETERS (2020). "A Method to Construct Geographical Crosswalks with an Application to US Counties since 1790." (Cited on p. 9, A2)

GOULD, J. D. (1980). "European inter-continental emigration. The road home: return migration from the U.S.A." *The Journal of European Economic History*. (Cited on p. 5, A4, A5, A11)

MADDISON, A. (2007). *Contours of the World Economy 1-2030 AD: Essays in Macro-Economic History*. OUP Oxford. (Cited on p. 12, A7)

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." (Cited on p. 11, 15, A3, A5)

SAN, S. (2023). "Labor Supply and Directed Technical Change: Evidence from the Termination of the Bracero Program in 1964." *American Economic Journal: Applied Economics*, 15(1): 136–163. (Cited on p. 3, D29)

SPITZER, Y. and A. ZIMRAN (2023). "Like an Ink Blot on Paper: Testing the Diffusion Hypothesis of Mass Migration, Italy 1876-1920." (Cited on p. 7, 10, A4, A5)

ZEIRA, J. (1998). "Workers, Machines, and Economic Growth." *The Quarterly Journal of Economics*, 113(4): 1091–1117. (Cited on p. 3, 18, 22, D29, D30)