Outmigration and Long-Run Economic Development: Evidence from the Italian Mass Migration*

Nicola Fontana† Marco Manacorda‡ Gianluca Russo§
Marco Tabellini¶

August 2025

Abstract

In this paper, we study the long run effects of emigration on economic development. We consider the case of historical mass migration from Italy between 1890 and 1920, when more than 10 million people left the country. We exploit variation in access to information about opportunities abroad to derive an instrument for outmigration at the municipality level. We find that areas with higher historical emigration are poorer, less educated, and less densely populated at the turn of the 21st century. These effects emerged early and persisted, as emigration led to sustained depopulation that, combined with declining fertility and lower human capital investment, constrained the structural transformation from agriculture to manufacturing and services.

JEL Codes: F22, N33, O15.
Keywords: Outmigration, long-run economic development.

*We thank Silvia Farina and Yannay Spitzer for useful comments. Jacopo Ballabio, Ludovica Ciasullo, Matteo Giugovaz, Matteo Grigoletto, Marta Leva, and Elena Verzellesi provided outstanding research assistance. All remaining errors are ours.

†Dept. of Economics & TIME & TRiSS, Trinity College Dublin; CEP, LSE. Email: nicola.fontana@tcd.ie
‡Queen Mary University of London, CEP (LSE), CEPR, and IZA. Email: m.manacorda@qmul.ac.uk
§Universitat Pompeu Fabra and Barcelona School of Economics. Email: gianluca.russo@upf.edu
¶Harvard Business School, NBER, CEPR, and IZA. Email: mtabellini@hbs.edu

1 Introduction

Between 1850 and 1920, more than 40 million people left Europe for the US, Canada, and Latin America during what is known as the Age of Mass Migration (Hatton and Williamson, 1998; Abramitzky and Boustan, 2017). European immigrants have been shown to increase income per capita and fuel economic growth overseas, with effects that persist until today (Rocha et al., 2017; Sequeira et al., 2020). Recent work has also begun to document the long-run consequences of emigration for sending countries in Europe: emigration raised support for political reform and promoted innovation in Sweden (Karadja and Prawitz, 2019; Andersson et al., 2022) and increased educational attainment in Galicia, Spain (Fernández Sánchez, 2022). Yet, despite these important advances, we still lack a comprehensive understanding of what may be the most fundamental question: how did emigration affect long-run economic development in the places migrants left behind?

In this paper, we answer this question focusing on the Italian experience between 1880 and 1920, when more than 10 million people—or, 30% of the 1900 Italian population—left the country. The relationship between emigration and long-run development is *ex ante* ambiguous. On the one hand, emigration might have released pressure on slack labor markets, improved living standards through remittances, and increased incentives to innovate and adopt labor-saving technologies. Sending countries may also have benefited from emigration via return migration and the flow of new ideas. On the other hand, emigration may have hindered economic development by reducing the size of the labor force and setting in motion a process of "brain drain" that depleted the stock of human capital, slowed down innovation, and depressed economic growth.

Figure 1 summarizes our main results, presenting the correlation between income per capita in 2001 and the cumulated number of migrants from 1884 to 1920, scaled by 1881 population, across Italian municipalities.[1] The figure shows that places that sent more migrants during the Age of Mass Migration are today poorer. To attach a causal interpretation to these motivating patterns, we derive an instrument for emigration that exploits variation across municipalities in access to information about opportunities abroad at the onset of the Italian mass migration. We retrieved and digitized historical documents from a special inquiry conducted by the Italian government aimed at understanding the factors responsible for emigration between 1882 and 1884. These documents record whether residents emigrated in response to information about economic opportunities abroad, which we use as a proxy for exposure to pull factors. Based on this, we calculated the distance between each municipality and the nearest city or town with such exposure.

[1] The figure partials out province fixed effects. Italian provinces are administrative units corresponding to NUTS-3 European regions and comparable to US counties. As of 1901, Italy had 69 provinces.

The existing historical evidence indicates that access to information about migration opportunities was more important than local push factors, such as economic downturns or natural disasters, in driving Italian emigration (Gould, 1980a; Spitzer et al., 2025). Moreover, such information spread gradually across space, through a process of contagion (Spitzer and Zimran, 2023). As for many other countries at the time (Hatton and Williamson, 1998; Karadja and Prawitz, 2019), also Italian migration was highly path dependent (Gould, 1980b). Confirming this idea, distance from places with access to information about opportunities abroad early on strongly predicts emigration throughout the entire 1884-1920 period.

The validity of the instrument relies on the assumption that proximity to municipalities with early exposure to information about migration opportunities influenced long-run economic development only through its effect on emigration. This assumption would be violated if the distance to such areas also affected economic growth via other channels. On the one hand, exposure to information early on may be correlated with better market access and with proximity to more dynamic neighboring local economies. On the other hand, smaller and more remote areas may be both less exposed to migration-related information and less likely to grow.

To address these concerns, our preferred specification includes 1881 population as well as a large vector of pre-determined or time-invariant controls related to geography (area; latitude and longitude; altitude; slope; ruggedness; presence of lakes and rivers; crop suitability) and market access (e.g., distance to coastline, border, province capital, main ports, and railroads). In addition, we control for long-term precipitation, temperature, and the number of violent earthquakes within 100 km—factors that may independently affect development and be spuriously correlated with distance to areas with migration-related information.

Estimating 2SLS regressions that include these controls as well as province fixed effects, we find that historical emigration reduces economic development in the long-run, confirming the patterns displayed in Figure 1. We begin by documenting a strong negative impact on population: moving from the 25th to the 75th percentile of the emigration distribution is associated with a 54% lower population in 2001. We then show that these areas are also poorer and less educated. Income per capita and the share of individuals with at least a high school degree are 13.7% and 15.9% lower in a high (75th percentile) than in a low (25th percentile) migration municipality. In all cases, 2SLS coefficients are larger (in absolute value) than OLS ones, suggesting that emigration was higher from municipalities (within the same province) where the population was growing faster. This is consistent with recent findings in Blanc and Wacziarg (2025) that migration from Europe during this time period was fueled by high fertility rates, and thus fast population growth.

These results are robust to controlling for a range of additional variables—including World War I (WWI) mortality rates, proximity to early migration hubs, early emigration activity, and characteristics of neighboring municipalities presenting information on migration opportunities. These controls are designed to address the concern that local economic development may have been influenced by factors correlated with—but not caused by—historical emigration. We further confirm that our findings are robust to alternative inference methods, including Conley standard errors, and to restricting the sample by excluding outliers or early migration epicenters. Results also remain stable when sequentially dropping each Italian region, and when exploring alternative constructions of the instrument.

Having established the long-run effect of emigration on economic development, we turn to its dynamics. We begin by analyzing population, for which consistent data are available from 1871 onward. Providing support for the validity of our empirical strategy, instrumented emigration is uncorrelated with population growth prior to 1880. Starting in 1901, the estimates become negative and statistically significant, consistent with emigration reducing population in the short run. These negative effects persist well beyond the end of the Age of Mass Migration and intensify over time, pointing to lasting demographic divergence across municipalities.

One channel through which these effects may have deepened over time is fertility. If historical emigration reduced local economic opportunities, it may also have weakened the incentives or capacity of younger generations to form families (Monras et al., 2023; Carlana and Tabellini, 2025). To test this hypothesis, we trace out the effects of emigration on the share of the population under age six—a proxy for fertility—throughout the 20th century. While the estimates are positive and imprecise in the short run—reflecting the fact that prime-age men were more likely to migrate—they become negative and statistically significant after 1951. This pattern suggests that fertility decline likely contributed to the persistent demographic divergence associated with historical emigration.

Another potential contributor to long-run population decline is continued outmigration in the postwar period. Between the 1950s and early 1970s, many Italians left for Northern and Western Europe under bilateral labor agreements (Faini and Venturini, 1994; Colucci et al., 2015), and internal migration from the rural South to the industrializing North followed a similar pattern (Pugliese, 2002; Mocetti and Porello, 2010). Although we lack detailed data on these mid-century flows, we use information on domestic and international migration from 2002 to 2015 to examine whether municipalities with high historical emigration continued to experience greater outflows. We find no evidence of such a relationship, suggesting that continued migration is unlikely to fully account for the persistent demographic divergence associated with historical emigration.

Next, we examine whether historical emigration hindered long-run development through its impact on human capital. Spitzer and Zimran (2018) show that Italian emigrants were positively selected within provinces, suggesting that emigration may have reduced not only the quantity but also the quality of labor. In line with Spitzer and Zimran (2018), we document a negative relationship between emigration and literacy in 1911 and 1921—the only proxy for education systematically available before 1950. However, the effects are small and imprecise, indicating that selection is unlikely to be the main driver of our results. Instead, emigration may have affected long-run development by lowering returns to schooling and reducing incentives to invest in education over time. Supporting this interpretation, we find that emigration is associated with lower rates of high school completion starting in 1951, the first year for which this measure is available. These effects are initially small and imprecise but grow over time and become statistically significant.

Finally, we test whether the negative long-run effects of emigration on economic development partly come from delayed structural transformation. Over the 20th century, Italy transitioned from an economy primarily based on agriculture to one centered on manufacturing and, more recently, services. In 1936, agriculture accounted for 66% of total employment; by 2001, the end of our sample period, its share had fallen below 10%, while manufacturing and services made up roughly 40% and 50%, respectively. We find that historical emigration slowed this reallocation process. Municipalities with more emigration experienced lower employment in manufacturing and services, starting as early as 1936. The number of manufacturing firms was also smaller, suggesting that emigration not only reduced labor supply but also hindered the scale and complexity of local production.

In contrast, we find no relationship between emigration and agricultural employment until after 1980. Using data from the 1929 Agricultural Census—the earliest available year—we also detect no effect on a wide range of agricultural productivity indicators, including yields, farm size, and capital intensity. This suggests that emigration did not lead to a reorganization of agricultural production or generate efficiency gains in the sector. Depressed human capital and stalled sectoral change likely reinforced the negative effects of historical emigration on long-run development.

Our results speak to the literature on the economic and political effects that immigrants have in receiving countries (Burchardi et al., 2018, 2020; Sequeira et al., 2020; Tabellini, 2020; Peters, 2022). We complement these studies by focusing on the consequences of outmigration for sending regions, contributing to a recent literature that has examined the effects of outmigration on innovation (Waldinger, 2010, 2012, 2016; Andersson et al., 2022; Coluccia and Dossi, 2025), technology adoption (Hornbeck and Naidu, 2014; Coluccia and Spadavecchia, 2021), human capital accumulation (Fernández Sánchez, 2022), and political change

(Karadja and Prawitz, 2019). We expand this line of work by providing systematic evidence on the developmental consequences of one of the largest episodes of voluntary migration in history. Within the outmigration literature, our findings are especially connected to those from Fernández Sánchez (2022) and Coluccia and Spadavecchia (2021).

Fernández Sánchez (2022) examines mass emigration from Galicia (Spain) between 1900 and 1930 and finds that short-run human capital losses were eventually offset by return migration and migrant-funded educational investments. In contrast, we document persistent declines in population and education, suggesting that when migration is not accompanied by reinvestment or return flows, it can entrench long-run divergence. Coluccia and Spadavecchia (2021) study the effects of the 1920s U.S. immigration quotas, which abruptly curtailed Italian emigration. They find that the resulting reduction in outflows increased local population and reduced the pace of labor-saving technological adoption. Our work complements theirs by studying the broader and longer-term consequences of outmigration itself, rather than the short-run adjustment to an external policy shock.

More broadly, our results speak to the long-standing debate in development economics on the consequences of emigration. Some scholars argue that brain drain and high-skill emigration might depress growth in origin countries by depleting human capital, distorting labor markets, and widening global inequality (Bhagwati and Hamada, 1974; McCulloch and Yellen, 1977; Miyagiwa, 1991). Others have pointed to offsetting mechanisms such as remittances, return migration, diaspora investment, and the diffusion of ideas and norms (Beine et al., 2001; Docquier and Rapoport, 2012; Batista et al., 2025). Our findings add nuance to this debate by showing that when such channels are absent or weak, emigration can hinder local development through demographic decline, slower structural transformation, and diminished human capital accumulation.

These longer-run effects are consistent with findings from the urban and regional economics literature on agglomeration economies (Glaeser and Gottlieb, 2009; Bleakley and Lin, 2012; Kline and Moretti, 2014; Leonardi and Moretti, 2023). This literature has shown that population density can sustain higher productivity, wages, and innovation—and that these effects can persist even after the underlying conditions have changed. Related insights emerge from recent macroeconomic models of endogenous growth and structural transformation, where population growth fuels firm entry, innovation, and sectoral reallocation; when population growth slows, so does economic dynamism (Peters, 2022; Eckert and Peters, 2022). In our context, the sustained population losses triggered by historical emigration may have pushed municipalities below the critical thresholds needed to support dynamic, higher-productivity sectors. This interpretation helps explain why our estimated effects persist—and even grow—long after the end of the Age of Mass Migration, reinforcing the view that outmigration can have lasting developmental consequences even in the absence of continued demographic shocks.

Our findings also relate to those of Chaney and Hornbeck (2016), who examine the long-run effects of the Morisco expulsion in Spain. While they find that towns that lost more people remained smaller for centuries—a pattern similar to our findings on population—their results suggest no adverse effects on income per capita, and possibly even gains due to slower population growth. In contrast, we find that emigration lowered both population and income per capita, driven by declines in manufacturing, services, and educational attainment. This divergence underscores the importance of context: in Italy, historical emigration appears to have weakened the forces of structural transformation and human capital investment, leading to slower long-run development.

Finally, our paper contributes to the literature on the Age of Mass Migration (Hatton and Williamson, 1998; Abramitzky and Boustan, 2017). It is especially related to work on the Italian experience, including Spitzer and Zimran (2018) and Spitzer and Zimran (2023), who examine migrant selection and the spatial diffusion of migration, and Pérez (2021), who compares the assimilation trajectories of Italian immigrants in Argentina and the United States.

2 Historical Background

2.1 The Italian Mass Migration

Between 1880 and 1920, Italy experienced one of the largest episodes of voluntary migration in recorded history. Although Italy was a relative latecomer to the Age of Mass Migration (1850-1920), emigration from the country grew rapidly at the turn of the century. By 1913, Italy had surpassed most other European countries in the absolute number of migrants departing each year, with per capita rates rivaling those of Ireland and Norway (Foerster, 1924). In total, more than 10 million Italians—roughly 30% of the 1900 population—left the country during this period (Spitzer and Zimran, 2023).

The United States was the most common destination, absorbing close to 40% of the flow, followed by Argentina and Brazil (Livi-Bacci, 1961). Italian migration followed a distinctive geographic and temporal pattern: it originated in the more urbanized and industrialized northern regions and spread gradually to the agrarian South. By the early twentieth century, emigration rates in southern regions were among the highest in Europe (Spitzer and Zimran, 2023). Figure 2 plots the number of emigrants between 1884 and 1920, scaled by the 1881 population, based on the official Italian Migration Statistics. While the figure confirms that southern regions accounted for a large share of total emigration, it also reveals substantial spatial variation in migration intensity across the country.

Italian migrants were overwhelmingly young men from rural areas, typically traveling alone with the intent of working abroad temporarily. Migration was often circular: between 30% and 50% of Italian emigrants ultimately returned home, with return rates especially high from the United States and parts of Latin America (Gould, 1980a; Bandiera et al., 2013). Migrants tended to concentrate in low-skilled occupations both at home and abroad, often working in agriculture, construction, and urban services. Despite this, recent evidence has shown that migrants were positively selected relative to their local populations—taller, and likely better educated or wealthier—especially in poorer provinces (Spitzer and Zimran, 2018).

A rich historical literature has examined the economic, social, and institutional forces shaping Italian emigration during this period (Foerster, 1924; Faini and Venturini, 1994; Gráda and O'Rourke, 1997; Hatton and Williamson, 1998). On the one hand, rural poverty, demographic pressure, land fragmentation, and sluggish industrial development—particularly in the South—are often cited as important contributors to emigration. Agricultural crises, poor harvests, and local shocks such as earthquakes or disease outbreaks occasionally played a role, though their effects were typically short-lived and geographically limited. On the other hand, expanding labor demand abroad—especially in the United States and South America—offered migrants the prospect of higher wages, land access, and greater social mobility.

While these accounts highlight key enabling conditions, more recent evidence suggests that variation in migration across space and time was shaped less by local economic distress and more by the diffusion of information through social and kinship networks (Spitzer and Zimran, 2023). Migration rarely began in isolation; early movers reduced the costs and uncertainty of migration for others by offering support and information about life abroad. As noted by Gould (1980b), a process of demonstration and imitation caused migration to radiate gradually from initial epicenters to neighboring areas. Spitzer and Zimran (2023) formalize this diffusion mechanism, showing that emigration followed an S-shaped temporal pattern and expanded spatially. In this setting, access to migration-related information—rather than local shocks—was often the key determinant of when a municipality entered the migration stream (Spitzer et al., 2025).

The Age of Mass Migration came to an abrupt halt in the early 1920s, when the United States—by then the dominant destination for Italian emigrants—passed a series of restrictive immigration laws (Abramitzky and Boustan, 2017). These restrictions effectively ended the era of mass transatlantic mobility for Italians. In the decades following World War

II (WWII), hundreds of thousands of Italians moved to Northern and Western Europe—especially to Germany, Switzerland, and France—while large-scale internal migration also reshaped Italy's demographic geography, as millions moved from the rural South to the industrial cities of the North (Pugliese, 2002; Colucci, 2008).

2.2 The Italian Economy in the 20th Century

At the time of the Mass Migration, Italy was a predominantly agrarian country marked by deep regional disparities. The North had begun to industrialize in the late 19th century, especially in the so-called "industrial triangle" of Milan, Turin, and Genoa, while the South remained largely rural and underdeveloped. National unification in 1861 had done little to integrate the economies of different regions, and poverty, entrenched land inequality, and low productivity continued to characterize much of the South. Large estates (*latifundia*) dominated the agrarian structure, especially in regions like Sicily, where the conditions of agricultural workers and peasants were particularly dismal (Acemoglu et al., 2020). Despite some modest gains in infrastructure and literacy, economic growth in the decades following unification was uneven and slow, and Italy remained one of the poorest countries in Europe by the turn of the century (Cafagna, 1989; Toniolo, 2013).

Following WWI, Italy faced high inflation, political instability, and a crisis of public finance. The Fascist regime that came to power in 1922 implemented a series of state-led interventions aimed at achieving autarky and restoring national pride. Chief among them was the Battle for Grain (1925–1939), a wide-ranging campaign to achieve self-sufficiency in wheat production. While the policy was intended to strengthen the agricultural sector, it had unintended positive effects on industrialization and long-run economic growth in areas more exposed to wheat-suitable land, largely through technological change and human capital accumulation (Carillo, 2021). Nevertheless, overall economic progress under Fascism remained limited, and Italy entered WWII still lagging behind the more industrialized countries of Western Europe (Federico, 1994).

Italy's postwar recovery was dramatic but uneven (Croce et al., 2025). Between the late 1940s and early 1970s, the country experienced rapid economic growth and structural transformation. Industrial production boomed, especially in the North, driven by exports, urbanization, and public infrastructure investment. Per capita income rose sharply, and living standards improved. However, the South failed to converge. Despite targeted development programs, including massive public transfers and infrastructure projects, southern Italy remained characterized by high unemployment, low productivity, and sustained migration to the northern urban areas (Pugliese, 2002; Colucci, 2008). The divergence in regional development would become one of the defining features of Italy's postwar economic history (Daniele and Malanima, 2011; Toniolo, 2013).

The postwar period also marked a profound shift in Italy's economic structure. As shown in Figure A.1, the share of employment in agriculture declined steadily from 1936 onward, while manufacturing and services absorbed growing portions of the labor force. Until the 1980s, structural transformation was driven largely by the expansion of manufacturing; thereafter, services became the dominant engine of employment growth (Toniolo, 2013). Figure A.2 illustrates that this transition was markedly uneven across regions. Northern Italy experienced an earlier and more pronounced shift into manufacturing and services, while Southern and Central regions remained more reliant on agriculture for much longer. These trends underscore the importance of labor reallocation in shaping Italy's postwar development and highlight the regional disparities that persisted throughout the century. As we discuss in Section 6.3, historical emigration may have played a role in this process by limiting the capacity of sending areas to diversify out of agriculture and participate fully in the country's broader structural transformation.

3 Data

This section describes the main variables used in the analysis; further details are provided in Appendix B. The unit of analysis is the municipality.[2] Because municipal boundaries changed over time, due to the creation, dissolution, or merging of municipalities, as detailed in Appendix B, we construct spatial units by grouping together all municipalities that have ever shared or exchanged territory, forming the largest stable agglomeration. This leaves us with 5,803 units, which we will refer to as municipalities throughout the paper. Tables A.1 and A.2 present summary statistics.

Migration statistics. We digitized yearly data at the municipality level on Italian migration between 1884 and 1920 from the *Statistica della Emigrazione Italiana per l'Estero* (see Figure A.3 for an example).[3] These publications were produced by the Italian government and report official statistics on international migration based on passport issuances, which became compulsory for overseas travel in 1901. The data record the number of emigrants from each municipality in a given year and includes departures to all international destinations (see also Spitzer and Zimran, 2023, for more details). In some years, the statistics report separately temporary and permanent migrants, which we always aggregate to derive the total number of emigrants from each municipality and year.

[2] In 1861, there were 7,720 municipalities, while in 2001 there were 8,101.

[3] We could not retrieve data for 7 out of 37 years (1888, 1889, 1890, 1894, 1895, 1916 and 1917).

We also rely on the 1884-1885 volume of the *Statistica della Emigrazione Italiana per l'Estero* to construct a measure of early exposure to information about overseas migration, which we then use to derive an instrument for emigration (see Section 4.2 for more details). This volume includes a government survey of municipal authorities—specifically, mayors and prefects—conducted by the Ministry of Agriculture, Industry, and Commerce to investigate the presence and influence of migration promoters. Officials were asked whether there were emigration agents or intermediaries in the municipality, whether outside individuals had visited to encourage migration, and what means of persuasion (such as pamphlets, subsidies, or public advertisements) were employed. Responses were given in free-text form and vary in detail and completeness (see Figure A.4 for an example). We manually coded a municipality as having been exposed to a migration driver if at least one specific channel of persuasion was mentioned.

We complement the historical migration statistics with a comprehensive database of modern migration data from *Anagrafe Comunali* (Istat), which includes the number of in- and out-migrants for each municipality between 2002 and 2015, by origin and destination (domestic and international) and by educational attainment.

Socio-economic variables. We assembled a panel of demographic and economic indicators at the municipal level from a range of official sources. Resident population figures for each census year from 1871 to 2001 are drawn from newly digitized historical population censuses conducted by ISTAT every 10 years.[4] For the postwar period, we supplement these data with information from the 8mila Census database (ISTAT). We obtained literacy rates for the resident population aged six and above in 1911 and 1921 from Fontana et al. (2023), the share of population aged six or above with at least a high school diploma for years 1951, 1961, 1971, 1981, 1991, and 2001 from 8mila Census), the number of individuals employed in agriculture, manufacturing and services for years 1936, 1951, 1961, 1971, 1981, 1991, and 2001 from Colussi et al. (2020) and 8mila Census, and income per capita for 2001 from the Italian Ministry of Finance. We also collected data on the number of firms and workers by sector from the Industrial Censuses for the years 1927, 1951, 1961, 1971, 1981, 1991, and 2001.[5] Finally, we digitized the 1929 Census of Agriculture to obtain several variables, including cultivated area, number of workers, agricultural output, average farm size, and the number of horses—the main form of agricultural capital at the time.

Additional variables. To account for potential confounders of long-run development, we collect a range of additional controls at the municipality level, including elevation, surface water features, agro-climatic conditions, and number of earthquakes. Additional variables

[4]Except for 1891 and 1941, when no census was held. In 1936 an additional Census was conducted.

[5]The 1951-2001 Industrial Censuses data are available on ISTAT website; we digitized the 1927 Industrial Census.

are described below, when relevant (see Tables B.1 and B.2 for the complete list).

4 Empirical Strategy

In this section, we introduce the baseline estimating equation (Section 4.1), construct the instrument for outmigration (Section 4.2), and present first stage estimates (Section 4.3).

4.1 Baseline Estimating Equation

To study the long-run effects of outmigration on economic development, we consider the sample of Italian municipalities described in Section 3, and estimate:

$$y_{mp} = \beta M_{mp} + \alpha_p + X_{mp} + \epsilon_{mp} \tag{1}$$

where y_{mp} is the outcome for municipality m in province p (e.g., the log of income per capita or the share of the population with at least a high school diploma in 2001); M_{mp} is the total number of emigrants from 1884 to 1920 from municipality m, scaled by 1881 population; α_p are province fixed effects; and, X_{mp} is a vector of time-invariant or historical municipality-level controls. We cluster standard errors at the province level.

In the preferred specification, X_{mp} includes the log of 1881 population and controls for three sets of variables that may be correlated with historical emigration and may have differential effects on economic development. First, we include a vector of geographic controls: area, latitude and longitude, altitude, slope, ruggedness, presence of lakes and rivers, and crop suitability. Second, we control for various measures of market access, such as distance to coastline, border, province capital, main ports, and railroads. Third, we control for deviations from long-term precipitation and temperature as well as for the number of violent earthquakes within 100 km.[6]

4.2 Instrument for Historical Outmigration

The main challenge in estimating equation (1) using OLS is that emigration may be correlated with other factors that independently shape long-run economic development. Recent work has shown that during this historical period, emigration was higher in European countries where fertility rates were higher (Blanc and Wacziarg, 2025). This suggests that emigration may be spuriously correlated with population growth, which may in turn drive long-run economic development. OLS estimates may be upward biased also if emigration

[6]See Table B.2 for the detailed description of each variable, with the corresponding source.

was stronger in municipalities where income was higher and more households could afford the cost of transatlantic travel (Foerster, 1924; Faini and Venturini, 1994). Conversely, if migration was driven by poverty or adverse economic and environmental shocks (Gráda and O'Rourke, 1997; Karadja and Prawitz, 2019), OLS estimates would be biased downward.

To address these and similar concerns, we derive an instrument for historical emigration that leverages variation in access to information about opportunities abroad in the mid-1880s, at the onset of the Italian Mass Migration. As detailed in Section 3, we draw on a special inquiry conducted in 1884-1885 by the Italian Ministry of Agriculture, Industry, and Commerce, which surveyed municipal authorities about the presence and activities of migration promoters. Based on manual coding of the free-text responses, we classify a municipality as exposed to a migration source of information if at least one specific channel—such as an agent, pamphlet, or recruiter—was mentioned. To capture the spatial diffusion of information, we compute the straight-line distance from each municipality to the nearest municipality reporting the presence of a migration driver. This variable—distance to migration information source—serves as the basis of our instrumental variable strategy.

The instrument exploits variation that aligns with historical accounts: emigration from Italy during the Age of Mass Migration was shaped primarily by social networks, rather than by economic distress (Spitzer and Zimran, 2023). Access to migration-related information early on may have played a key role in triggering emigration, which then spread through a process of chain migration (Gould, 1980b; Spitzer and Zimran, 2018). As a result, we expect municipalities with early exposure to migration information sources to experience persistent outflows throughout the period—consistent with historical evidence from other settings (Karadja and Prawitz, 2019; Andersson et al., 2022).

The key identifying assumption is that proximity to municipalities with early exposure to migration-related information affected long-run economic development only through its impact on emigration. This assumption would be violated if proximity also captured other determinants of development. For instance, areas closer to early nodes of migration activity may also have had better market access or more dynamic neighboring economies. Similarly, remote and sparsely populated areas may have been both less exposed to information and less likely to grow, even in the absence of emigration.

To tackle these concerns, our preferred specification controls for the log of 1881 population as well as for different proxies for market access. We also verify below that 1871 population is uncorrelated with the instrument, suggesting that early exposure to migration-related information was not systematically related to pre-existing trends in population growth. In addition, we show that our results are robust to a variety of alternative specifications. These include controlling for: *i*) distance to the nearest early migration epicenter or border municipality, as defined by Spitzer and Zimran (2023); *ii*) whether the municipality sent at least ten migrants in any year between 1882 and 1884; *iii*) the full vector of baseline covariates for the nearest municipality which is a migration information source; *iv*) average baseline characteristics of all municipalities within a 100 km radius; and, *v*) WWI mortality rates. We present these and other robustness checks in detail below, following the main results.

4.3 First Stage Estimates

Table 1 presents first stage results, where we regress cumulative emigration between 1884 and 1920, scaled by 1881 population, against the instrument described in Section 4.2.[7] Column 1 estimates a parsimonious specification that includes province fixed effects, the log of 1881 population, and altitude to account for the unique features of Italian territory. The negative and statistically significant coefficient indicates that emigration was higher in municipalities that were located closer to early sources of migration-related information.

Column 2 adds the full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for the main crops) and controls for the (straight-line) distance between the municipality and the coastline, the border, the province capital, and the closest main port.[8] Column 3 controls for average distance to the closest rail station and rail line, addressing concerns that the contemporaneous expansion of the rail network (Facchini et al., 2021) may be correlated with information diffusion. Column 4 adds exposure to weather shocks and violent earthquakes.[9]

In all cases, the coefficient on distance to migration information source remains negative, statistically significant, and close to that reported in column 1. According to our preferred specification, reported in column 4, increasing the distance to the closest location with information about opportunities abroad by 10 km reduces cumulated migration by 3.9 percentage point (or, 6.3% relative to the mean). Figure A.5 shows the corresponding residualized binscatter plot, visually confirming the strong negative relationship between proximity to early migration information and subsequent emigration.

In Figure A.6, we replicate the preferred first stage specification, binning the instrument by 10 km of distance to the nearest municipality with early migration-related information. The plot shows a monotonic decline in cumulative emigration as distance increases, with municipalities in the closest distance bin exhibiting significantly higher emigration rates than those further away. The slope of the relationship flattens at greater distances, consistent with

[7]The instrument is scaled by 10, meaning that the reported coefficients correspond to the effect of a 10 km increase in distance.

[8]The three main ports for international departures were Genova, Palermo, and Naples.

[9]We calculate exposure to weather shocks by taking the average of the yearly deviation of temperature and precipitation from the long-run (1800-1920) means. See Table B.2 for more details.

the idea that information diffused locally and attenuated with geographic separation.

Finally, in Figure A.7, we estimate the first stage regression separately for five-year periods to assess whether the effects of early exposure persist over time or fade quickly. Consistent with a process of chain migration, the coefficients remain stable and statistically significant throughout the period, closely matching the estimate from the full sample (plotted as the first dot on the left).[10]

5 The Long-Run Effects of Outmigration

This section studies the long-run effects of outmigration on economic development. Section 5.1 establishes that municipalities with higher historical outmigration are poorer, less educated, and less densely populated at the turn of the 21st century. Section 5.2 presents several robustness checks supporting these findings.

5.1 Main Results

We begin by examining whether municipalities that experienced more outmigration during the Age of Mass Migration continue to have smaller populations today. Table 2 presents 2SLS (Panel A) and OLS (Panel B) estimates for the effect of historical emigration on the log of resident population in 2001. Column 1 includes province fixed effects, log population in 1881, and altitude. The 2SLS coefficient is negative and statistically significant. Results remain stable when sequentially adding controls for distance to ports and other access points (column 2), rail-based market access (column 3), and exposure to weather shocks and earthquakes (column 4). According to our preferred specification (column 4), moving from a a low (25th percentile) to a high (75th percentile) historical emigration municipality reduces population in 2001 by 53.8%.[11] This result indicates that emigration caused long-lasting demographic divergence across municipalities.

Then, we ask whether this population loss also translated into weaker economic outcomes. Tables 3 and 4 show that municipalities with higher historical emigration are not only smaller today, but also poorer and less educated. Specifically, we estimate the effects of emigration on the log of income per capita and on the share of the population with at least a high school diploma in 2001.[12] The 2SLS coefficients are negative, statistically significant, and stable

[10]While the coefficient for the 1910–1914 period is smaller in magnitude, it remains precisely estimated and is not statistically different from the others.

[11]To get at these numbers, note that moving from a municipality at the 25th to a municipality at the 75th percentile of the historical emigration distribution increases outmigration by 0.592. This implies: $(e^{-1.303\times0.592}) - 1) \times 100$=53.8%.

[12]ISTAT reports the number of individuals with at least a high school diploma relative to the population aged six and older, which includes many who are too young to have completed high school. Results are similar when using the log of the number of high school graduates or when restricting the denominator to those aged 15 and older.

across columns. According to our preferred specification (column 4), increasing emigration from a low to a high migration municipality reduces income per capita and the share of the population with at least a high school degree by 13.6% and 4.3 percentage points (or, 15.7% relative to the mean), respectively.

Comparing coefficients in Panels A and B of Tables 2 to 4 reveals that OLS estimates are systematically smaller (in absolute value) than 2SLS ones. This discrepancy may be partly explained by attenuation bias due to measurement error in the emigration data. More substantively, it suggests that municipalities with higher historical emigration may have been on faster economic trajectories prior to migration, biasing OLS estimates toward zero. These patterns align with recent findings by Blanc and Wacziarg (2025), who document that population pressure was a key driver of emigration in late 19th-century Europe. It is also in line with the idea that credit constraints limited poorer households' ability to migrate, as discussed by Foerster (1924) and Faini and Venturini (1994).

5.2 Robustness Checks

In this section, we assess the robustness of our main results across a range of alternative specifications, inference strategies, and sample definitions. In Table A.3, we report 2SLS estimates analogous to our baseline specification, sequentially adding controls to address concerns about omitted variables related to proximity, historical migration exposure, and spatial spillovers. We consider each of the three main outcomes—the log of population, the log of income per capita, and the share of individuals with at least a high school degree—in Panels A to C, respectively.

Column 1 reproduces the baseline results to ease comparisons. In column 2, we control for the distance to the nearest migration epicenter or to the national border, following Spitzer and Zimran (2023). This addresses the concern that proximity to early-sending hubs or border regions may have directly influenced subsequent economic development through channels unrelated to emigration (e.g., earlier industrialization, trade exposure). In column 3, we include a dummy for whether the municipality appears in the early emigration tables—defined as having sent at least ten migrants in any year between 1882 and 1884—to account for possible unobserved characteristics of municipalities with early emigration activity.

Columns 4 and 5 control for the characteristics of neighboring municipalities with migration information source. Column 4 includes controls for the nearest municipality with early exposure to migration-related information: log population in 1881, altitude, distance to the provincial capital, and an indicator for early emigration. Column 5 extends this approach by including the average characteristics of all municipalities within 100 km. These controls are designed to address the concern that the characteristics of municipalities with migration information sources—rather than their role in information diffusion—may have directly affected the development trajectories of nearby areas, thus violating the exclusion restriction. By conditioning on both the attributes of the nearest migration-source municipalities and their spatial averages, we isolate the identifying variation in access to migration-related information from broader regional confounders.

A separate concern is that the distance to early migration information source may be spuriously correlated with WWI casualties. If municipalities closer to early migration hubs also experienced fewer wartime deaths, our estimates could be confounded by the demographic effects of the war rather than reflecting the impact of emigration. To address this issue, column 6 replicates our baseline specification while controlling for the number of WWI deaths per capita, measured as the number of casualties relative to municipal population in 1911. The coefficient on emigration remains negative, statistically significant, and—if anything—becomes slightly larger in absolute value, reinforcing the conclusion that our results are not driven by differential exposure to wartime mortality.

We then examine the robustness of our results to excluding selected subsets of the sample. In columns 2 to 4 of Table A.4, we drop municipalities: in the bottom and top 1% of the emigration distribution; reporting migration information source in the 1884–85 inquiry within their territory; and, those identified by Spitzer and Zimran (2023) as early migration epicenters. Reassuringly, coefficients remain stable and similar to the baseline (reported in column 1). In addition, Figure A.8 presents coefficient plots from re-estimating the baseline specification while sequentially excluding each of the 20 Italian regions one at a time. Each panel corresponds to a different outcome. In all cases, the estimated effects remain in line with the baseline, indicating that no single region is driving the results.

One may also be concerned that the results are entirely driven by southern and central Italian regions, which sent many migrants but may have followed divergent development paths for reasons unrelated to historical emigration. To address this, columns 5 and 6 of Table A.4 report estimates from separate regressions for the North and the Center-South.[13] In both samples, the coefficients remain negative and are not statistically different from each other; if anything, they are somewhat larger in magnitude and more precisely estimated for the North. These estimates are less precise, likely due to reduced sample size, but they provide further reassurance that the results are not specific to one macro-region.

Finally, Table A.5 replicates the analysis: excluding pamphlets from the set of information about opportunities abroad early on (column 2); considering any possible source

[13]Following National Statistic Office classification, regions in the North are: Emilia-Romagna, Friuli-Venezia Giulia, Liguria, Lombardia, Piemonte, Veneto; and the Center-South regions are: Abruzzo, Basilicata, Calabria, Campania, Lazio, Marche, Puglia, Sardegna, Sicilia, Toscana, Umbria.

of information, including previous migrants, news from neighboring cities, and unspecified drivers of emigration (column 3); splitting the sample in 10 km bins (column 4); and, estimating standard errors following Conley (1999), which allows for spatial autocorrelation (column 5). Even if at times the F-stat drops below 10, especially in column 4, coefficients remain statistically significant, stable, and similar to the baseline (reported in column 1).

6 Mechanisms

This section examines the mechanisms through which emigration hindered long-run economic development. First, we show that emigration triggered early population losses that persisted throughout the 20th century, driven in part by lower fertility and reduced in-migration (Section 6.1). Second, we find negative, albeit small and imprecisely estimated, effects on literacy in the short run, but document a sustained decline in human capital accumulation over time, likely due to lower returns to schooling (Section 6.2). Third, we present evidence that emigration slowed structural transformation by constraining the growth of manufacturing and services, while failing to improve agricultural productivity (Section 6.3).

6.1 Population Dynamics

We begin our analysis of the mechanisms by studying how historical emigration shaped population dynamics over time. While outmigration may have triggered immediate demographic changes, we show that these effects not only appeared early on but also intensified during the 20th century, leading to persistent population divergence.

Early and persistent population decline. Figure 3 plots the coefficients from a series of 2SLS regressions of log population on instrumented emigration, estimated separately for each census year from 1871 to 2001.[14] Importantly, the coefficient for 1871 is quantitatively small and indistinguishable from zero, indicating no relationship between emigration and population prior to the Age of Mass Migration. This supports the validity of our identification strategy and alleviates concerns that the instrument may be predicting larger outmigration from places that were already on distinct demographic paths.

Starting in 1901—the first census after the beginning of the Italian Mass Migration—the coefficient turns negative and, except for 1911, becomes statistically significant at the 5% level, consistent with emigration reducing population in the short run. Between 1901 and 1921, the coefficient ranges from –0.16 to –0.28. This implies that comparing a low to a

[14]Since we always condition on the log of 1881 population, this year is not included in the analysis. Data on population in 1891 and 1941 are missing, since no censuses were conducted in those years. The corresponding 2SLS and OLS estimates in tabular form are reported in Table A.8.

high migration municipality reduces population by approximately 9% to 15.3% over this period. The fact that the relationship is less than one-for-one likely reflects the circular nature of migration at the time: estimates suggest that between 30% and 50% of Italian migrants eventually returned home, particularly from destinations such as the United States and Argentina (Gould, 1980a; Bandiera et al., 2013). Over time, the negative effects remain statistically significant and become larger (in absolute value), with the coefficient reaching -1.303 by 2001.

These patterns suggest that emigration set municipalities on divergent demographic trajectories, with population losses compounding well beyond the original migration episode. This resonates with findings from Chaney and Hornbeck (2016), who show that the expulsion of the Moriscos from Spain in 1609 led to prolonged population declines, which lasted for over two centuries. These long-run population dynamics are also consistent with the literature that has documented the role of agglomeration economies in driving local economic growth. Declines in population density can reduce local productivity, wages, and innovation through weaker labor market matching, lower knowledge spillovers, and diminished demand for infrastructure and amenities—and these effects can persist well beyond the initial shock, even after the underlying conditions have changed (Bleakley and Lin, 2012; Kline and Moretti, 2014). In our context, the demographic divergence triggered by historical emigration may have compounded economic losses over time by weakening the forces that sustain growth in densely populated areas.

Consistent with historical records, which document that emigration from Italy during this period was disproportionately male (Foerster, 1924), we also examine the effects of emigration on sex ratios (defined as the number of men relative to the number of women). Figure A.9 shows that emigration reduced the number of men relative to women in sending municipalities in the early decades of the twentieth century.[15] These effects are statistically significant and economically meaningful in the 1911 and 1921 censuses. However, the relationship weakens and disappears after WWII, suggesting that gender imbalances were a transitory demographic consequence of early outflows rather than a long-term legacy of emigration.

The role of fertility. The persistence of population decline raises the question of whether fertility responses contributed to the long-run demographic effects of emigration. Several studies have documented a positive relationship between fertility and economic stability (Sobotka et al., 2011; Monras et al., 2023; Carlana and Tabellini, 2025). If emigration reduced local growth prospects, it may have also discouraged younger generations from forming families and having children.

[15] Data on the population by gender at the municipality level before 1951 are only available for 1911 and 1921, explaining why we have only two point estimates before 1950 in Figure A.9. See Table A.9 for the corresponding 2SLS (Panel A) and OLS (Panel B) estimates in tabular form.

To examine this channel, we estimate the effect of historical emigration on the number of individuals below age six—a proxy for fertility—available for 1911 and 1921 and consistently from 1951 onwards. Ideally, we would scale this count by the number of women of childbearing age (e.g., 18–40), but such data are unavailable. Instead, we report two sets of estimates: one scaled by total female population (grey diamonds), available for all years, and another scaled by population 15-64 (black circles), available only from 1951 onward.

Figure 4 shows that in the short run (1911 and 1921), the coefficient is positive, though not statistically significant at conventional levels. This is consistent with the fact that most emigrants were prime-age men, whose departure would not have directly reduced the number of young children in the short run (Foerster, 1924). Starting in 1951, coefficients become negative and statistically significant, suggesting that fertility decline emerged as a key contributor to long-run demographic divergence. These effects persist through the 1970s but appear to fade by 1991.[16]

Migration persistence. Another explanation for the persistent population gap is continued out-migration or reduced in-migration until today. Between the 1950s and early 1970s, Italy experienced a new wave of international emigration—this time primarily to Northern and Western Europe—and a surge in internal migration from the rural South to the industrializing North (Colucci et al., 2015; Croce et al., 2025). Although we lack municipality-level migration data for this period, we test for potential persistence in outmigration and inmigration by examining domestic and international migration flows from 2002 to 2015.

Table 6, Panel A, presents 2SLS estimates for in-migration (columns 1 to 4) and outmigration (columns 5 to 8).[17] We scale migration by 2001 population, and separately examine domestic (columns 1 and 5) and international flows (columns 2 and 6). We also distinguish between migrants with and without a college degree (columns 3–4 and 7–8, respectively). Columns 1 to 4 show a consistently negative relationship between historical emigration and modern in-migration across all categories—domestic, international, skilled, and unskilled—with the estimates statistically significant in the case of international flows. These results suggest that high-emigration municipalities have become less attractive destinations over time and struggle to draw in new residents.

Columns 5 to 8 show a similarly negative relationship between historical emigration and modern out-migration, with the exception of international migration, where the coefficient is slightly positive but close to zero. The effect is statistically significant only for domestic outmigration, and only at the 10% level. This weighs against the idea that persistent outflows have driven long-run population decline. Instead, these areas appear to have become less

[16] See Tables A.12 and A.13 for the corresponding 2SLS (Panel A) and OLS (Panel B) estimates in tabular form.

[17] Panel B reports the corresponding OLS coefficients.

dynamic: rather than continuing to lose people through out-migration, they now fail to attract or retain residents.

We interpret these results as consistent with the idea that historical emigration triggered long-run depopulation, which in turn reduced local economic dynamism. The decline in in-migration is particularly relevant: not only does it reinforce population divergence, but it also reflects a persistent lack of pull factors, such as employment opportunities (as discussed in Section 6.3) or urban amenities, that might otherwise attract new residents and offset earlier losses. While these findings do not rule out the possibility that out-migration continued during the postwar decades—particularly between 1950 and 1970—they suggest that more recent emigration is not the driver of long-run population divergence. Instead, the evidence points toward reduced attractiveness and declining mobility in historically high-emigration areas.

6.2 The Role of Human Capital

A second channel through which historical emigration may have hindered long-run development is through its impact on human capital accumulation. This channel may operate through two complementary mechanisms. First, if emigrants were positively selected on education or skills, their departure could have directly reduced the average quality of the local labor force. Second, over time, emigration may have lowered incentives to invest in education, especially if it led to weaker labor market opportunities and lower returns to schooling.

Migrant selection. We begin by exploring the selection channel. Using height as a proxy for selection, Spitzer and Zimran (2018) show that while Italian emigrants were negatively selected at the national level, they were positively selected within provinces—especially in poorer municipalities. To assess whether similar patterns hold in our setting, we examine literacy rates—the only proxy for education prior to 1951. We calculate the ratio of literate to illiterate individuals aged six years or older within each municipality, and plot the 2SLS coefficients on historical emigration in Figure 5, Panel A.[18]

The estimates are negative, implying that emigration may have increased the number of illiterate relative to literate individuals. However, these effects are small and statistically insignificant. The patterns are consistent with the direction of selection suggested by Spitzer and Zimran (2018), but the imprecision and limited variation in the literacy data make it difficult to draw firm conclusions. After 1971, the coefficients become larger (in absolute value), but remain imprecise—likely reflecting the fact that literacy rates were already above

[18]Ideally, we would scale the number of literate individuals by the adult population, but such data are not consistently available prior to 1951. Table A.10 reports the corresponding 2SLS (Panel A) and OLS (Panel B) estimates in tabular form.

90% by that time, rendering literacy an uninformative proxy for human capital accumulation in the modern period.

Human capital investment. We next turn to the second mechanism: the idea that historical emigration may have reduced incentives for human capital investment over time. Panel B of Figure 5 plots the estimated effects of emigration on the share of individuals with at least a high school diploma from 1951 to 2001.[19] The coefficients are consistently negative, and become larger over time. The point estimate is statistically significant at the 10% level in 1991 and at the 5% level in 2001. These effects emerge more than four decades after the end of the Age of Mass Migration, suggesting that they cannot be explained by direct migrant selection. These findings help explain why these areas remain poorer today, and are consistent with work stressing the role of human capital accumulation as a key driver of long-run economic development (Rocha et al., 2017; Valencia Caicedo, 2019; Althoff and Reichardt, 2024).

It is instructive to compare our results with those of Fernández Sánchez (2022), who examines the long-run effects of mass emigration from Galicia, Spain, during the period 1900–1930. In Galicia, the initial loss of educated individuals was reversed within a decade after the migration wave ended. Over time, high-emigration municipalities became more educated than their counterparts. Fernández Sánchez (2022) attributes this long-run educational gain to two main mechanisms: the establishment of migrant-funded schools and the diffusion of pro-education norms via return migrants and diaspora networks. In contrast, we find no evidence of a similar reversal in the Italian context. The negative effects of emigration on educational attainment persist over the long run, even decades after the Age of Mass Migration ended. While data limitations prevent us from directly testing the role of return migration, migrant-funded schooling, or diaspora-led institutional investments, the persistence and magnitude of the negative effects on educational attainment suggest that such channels were much weaker in our setting.

This comparison highlights that the long-run consequences of emigration for human capital formation hinge not only on the characteristics of those who leave, but also on the institutional and social dynamics set in motion afterward. When emigration is accompanied by reinvestment and norm diffusion—as in Galicia—it can ultimately foster educational upgrading. When it leads instead to depopulation, reduced in-migration, and weakened local institutions—as in our setting—it may entrench educational and economic divergence.

[19] Table A.11 reports the corresponding 2SLS (Panel A) and OLS (Panel B) estimates in tabular form.

6.3 Structural Transformation and Sectoral Employment

A third channel through which emigration may have affected long-run development is by slowing the structural transformation of the local economy. As discussed in Section 2.2, between the early 20th century and the postwar decades, Italy underwent a profound transition from agriculture to manufacturing and later to services. This reallocation of labor was central to the country's growth: until 1980, most economic expansion was driven by the rise of manufacturing; in subsequent decades, the growth engine shifted increasingly to services. If emigration depleted the local labor force or drew away individuals with higher skills and initiative, it may have hindered this reallocation and, in turn, limited local development.

Structural transformation. We begin by documenting the long-run effects of historical emigration on total and sectoral employment patterns. Figure 6, Panel A, plots the estimated effect of emigration on the log of total employment across census years from 1936 to 2001. The coefficients are negative, statistically significant, and persistent over time, indicating that high-emigration municipalities experienced lower employment overall. Panel B disaggregates these effects by sector. Emigration had no discernible impact on agricultural employment (black dots) until after 1980, when the coefficients turn negative but remain statistically insignificant. By contrast, manufacturing employment (grey diamonds) shows a persistent decline, though the estimates are imprecise and not statistically significant. The strongest effects appear in services (grey squares), where emigration is associated with steep and sustained declines that are statistically significant throughout the period and grow more negative over time. These results suggest that emigration primarily reduced employment in non-agricultural sectors, and only later began to affect agriculture.[20]

To better capture the evolution of sectoral composition, Figure 7 plots the effect of emigration on the employment share of each sector, measured relative to total employment.[21] The results reveal clear disruptions to structural transformation. The agricultural employment share increases modestly in the medium run—though the effect is not statistically significant at conventional levels—and declines only gradually thereafter. The coefficient on emigration for the manufacturing share initially falls, returns to zero, and becomes weakly positive after 1971, but remains statistically insignificant throughout. By contrast, the service sector share shows a persistent and statistically significant decline beginning in 1936. These patterns point to a dual distortion: a delayed and weakened transition out of agriculture and into manufacturing, and a long-run drag on the expansion of services—the sector that has been central to economic growth in recent decades. Together, these findings suggest that emigration undermined the local reallocation of labor toward higher-productivity

[20]See Tables A.14 to A.17 for the corresponding 2SLS (Panel A) and OLS (Panel B) estimates in tabular form.

[21]See Tables A.18 to A.20 for the corresponding 2SLS (Panel A) and OLS (Panel B) estimates in tabular form.

activities.

Evidence from manufacturing. We complement these findings with data from the Census of Manufacturing, which provides information on employment and number of establishments beginning in 1927. Figure 8 presents 2SLS estimates for the log number of manufacturing firms (black dots) and workers (grey diamonds).[22] The figure confirms the patterns observed in the population census data (Figure 6): the negative effect of historical emigration on manufacturing employment is already evident in 1927 and remains relatively stable over time. Municipalities with higher emigration also consistently host fewer manufacturing firms.

Evidence from agriculture. Finally, we turn to agriculture, using data that we digitized from the 1929 Agricultural Census. In Table 5, we report estimates for a wide range of outcomes: cultivated area, number of farms, number of agricultural workers, the number of horses (the main form of agricultural capital at the time) per farm or per worker, agricultural output, crop yield (output per acre), and an index of land concentration. We find no statistically significant effects on any of these outcomes. Emigration did not lead to changes in land use, farm size, capital intensity, or productivity. This lack of effect suggests that emigration did not trigger a reorganization of agricultural production or an improvement in agricultural efficiency. Instead, it appears to have simply depleted the labor force without inducing compensating gains in productivity.

Taken together, these findings indicate that emigration slowed the reallocation of labor from agriculture into higher-productivity sectors. In doing so, it hindered both the timing and the depth of structural transformation, with long-run consequences for economic development. These results are consistent with a large literature on structural transformation and development, which emphasizes that sustained growth requires shifting labor out of agriculture and into more productive sectors such as manufacturing and services (Kongsamut et al., 2001; Herrendorf et al., 2014; Eckert and Peters, 2022).

7 Conclusion

In this paper, we study the long-run economic consequences of mass emigration, focusing on the case of Italy between 1880 and 1920, when over 10 million people—roughly 30% of the population—left the country. We instrument historical emigration across Italian municipalities exploiting variation in the distance from places with information about opportunities abroad in the early 1880s. We find that emigration led to long-lasting negative effects on economic development: at the turn of the 21st century, municipalities with higher historical emigration are poorer, less educated, and have lower levels of population and employment.

[22]The corresponding 2SLS (Panel A) and OLS (Panel B) estimates are reported in Tables A.21 and A.22.

These effects emerged early on and grew over time. Exploring the mechanisms, we find that emigration led to early and persistent population losses, further compounded by declines in fertility. Declining population, combined with lower human capital investment, impeded structural transformation, slowing the shift from agriculture to manufacturing and services.

Although our findings may be specific to the historical context of early 20th-century Italy, they offer valuable insights for today's developing countries, many of which are facing large-scale emigration. Our results highlight how emigration—through depopulation, lower investment in human capital, and a slowdown in structural transformation—can have long-lasting consequences for local economies. While the outflow of skilled or educated individuals might contribute to human capital depletion, emigration also hampers the transition to more productive sectors, such as manufacturing and services, and limits overall economic dynamism. Comparing our results with those in Fernández Sánchez (2022) can offer insights on the relevance that mechanisms such as reinvestment, return migration, and the exchange of ideas and resources between emigrants and their home communities can shape long run economic growth.

Our findings also raise intriguing questions for future research. To what extent do emigration-driven losses in human capital also reflect a decline in social capital and the weakening of local institutions? How might emigration shape political preferences and ideologies in the long run, particularly in societies experiencing high levels of outward migration? These questions, which go beyond the scope of our analysis, offer fascinating avenues for exploring the broader societal impacts of emigration.

References

Abramitzky, R. and L. Boustan (2017). Immigration in american economic history. *Journal of Economic Literature 55*(4), 1311–45.

Acemoglu, D., G. De Feo, and G. D. De Luca (2020). Weak states: Causes and consequences of the sicilian mafia. *The Review of Economic Studies 87*(2), 537–581.

Althoff, L. and H. Reichardt (2024). Jim crow and black economic progress after slavery. *The Quarterly Journal of Economics 139*(4), 2279–2330.

Andersson, D., M. Karadja, and E. Prawitz (2022). Mass Migration and Technological Change. *Journal of the European Economic Association 20*(5), 1859–1896.

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. Migration and Development.

Batista, C., D. Han, J. Haushofer, G. Khanna, D. McKenzie, A. M. Mobarak, C. Theoharides, and D. Yang (2025). Brain drain or brain gain? effects of high-skilled international emigration on origin countries. *Science 388*(6749), eadr8861.

Beine, M., F. Docquier, and H. Rapoport (2001). Brain drain and economic growth: theory and evidence. *Journal of Development Economics 64*(1), 275–289.

Bhagwati, J. and K. Hamada (1974). The brain drain, international integration of markets for professionals and unemployment: a theoretical analysis. *Journal of Development Economics 1*(1), 19–42.

Blanc, G. and R. T. Wacziarg (2025). Malthusian migrations. NBER Working Paper w33542, NBER and University of Manchester / UCLA.

Bleakley, H. and J. Lin (2012). Portage and path dependence *. *The Quarterly Journal of Economics 127*(2), 587–644.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2018). Migrants, ancestors, and foreign investments. *The Review of Economic Studies 86*(4), 1448–1486.

Burchardi, K. B., T. Chaney, T. A. Hassan, L. Tarquinio, and S. J. Terry (2020). Immigration, innovation, and growth. Technical report, National Bureau of Economic Research.

Cafagna, L. (1989). Dualismo e sviluppo nella storia d'italia. In P. Saraceno (Ed.), *Dualismo e sviluppo nella storia d'Italia*, pp. 21–50. Roma-Bari: Laterza.

Carillo, M. F. (2021). Agricultural policy and long-run development: Evidence from mussolini's battle for grain. *The Economic Journal 131*(635), 2346–2379.

Carlana, M. and M. Tabellini (2025). Happily ever after: Immigration, natives' marriage, and fertility. *Journal of Economic History*. Forthcoming.

Chaney, E. and R. Hornbeck (2016). Economic dynamics in the malthusian era: Evidence from the 1609 spanish expulsion of the moriscos. *The Economic Journal 126*(594), 1404–1440.

Colucci, M. (2008). *Lavoro in movimento: L'emigrazione italiana in Europa, 1945–57.* Roma: Donzelli Editore.

Colucci, M., S. Gallo, et al. (2015). *L'emigrazione italiana. Storia e documenti.* Brescia, Italy: Morcelliana.

Coluccia, D. M. and G. Dossi (2025). Return innovation: The knowledge spillovers of the british migration to the united states, 1870-1940. CEP Discussion Papers dp2069, Centre for Economic Performance, LSE.

Coluccia, D. M. and L. Spadavecchia (2021). The economic effects of immigration restriction policies - evidence from the italian mass migration to the us. (9361).

Colussi, T., G. Lecce, M. Manacorda, and M. G. Onorato (2020). The politics and economics of government aid: Evidence from the italian cassa per il mezzogiorno. Technical report, Working paper.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1–45.

Croce, P., M. Filippi, P. Piselli, and A. Ramazzotti (2025). Internal migration, local development and structural change: Evidence from the italian golden age.

Daniele, V. and P. Malanima (2011). Il divario nord–sud in italia 1861–2011. *Rivista di Politica Economica 3*, 267–315.

Docquier, F. and H. Rapoport (2012). Globalization, brain drain, and development. *Journal of Economic Literature 50*(3), 681–730.

Eckert, F. and M. Peters (2022). Spatial structural change. Technical report, National Bureau of Economic Research.

Facchini, G., R. Bonfatti, A. Tarasov, G. L. Tedeschi, and C. Testa (2021). Pork, infrastructure and growth: Evidence from the italian railway expansion. CEPR Discussion Papers 16462, C.E.P.R. Discussion Papers.

Faini, R. and A. Venturini (1994). Migration and growth: The experience of southern europe. Cepr discussion paper no. 964, Centre for Economic Policy Research.

Federico, G. (1994). *The Economic Development of Italy Since 1870.* Cheltenham: Edward Elgar.

Fernández Sánchez, M. (2022). Mass emigration and human capital over a century: Evidence from the galician diaspora. *Available at SSRN 4267207*.

Foerster, R. F. (1924). *The Italian Emigration of Our Times*, Volume 20. Harvard University Press.

Fontana, N., T. Nannicini, and G. Tabellini (2023). Historical roots of political extremism: The effects of nazi occupation of italy. *Journal of Comparative Economics 51*(3), 723–743.

Glaeser, E. L. and J. D. Gottlieb (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature 47*(4), 983–1028.

Gould, J. (1980a). European inter-continental emigration: The road home: Return migration from the u.s.a. *Journal of European Economic History 9*(1), 41–74.

Gould, J. (1980b). European international emigration: The role of diffusion and feedback. *Journal of European Economic History 9*(2), 267–315.

Gráda, C. O. and K. H. O'Rourke (1997). Migration as disaster relief: Lessons from the great irish famine. *European Review of Economic History 1*(1), 3–25.

Hatton, T. and J. G. Williamson (1998). *The Age of Mass Migration: Causes and Economic Impact*. Oxford University Press.

Herrendorf, B., R. Rogerson, and Á. Valentinyi (2014). Growth and structural transformation. *Handbook of Economic Growth 2*, 855–941.

Hornbeck, R. and S. Naidu (2014). When the levee breaks: Black migration and economic development in the american south. *American Economic Review 104*(3), 963–90.

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.

Kline, P. and E. Moretti (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *The Quarterly Journal of Economics 129*(1), 275–332.

Kongsamut, P., S. Rebelo, and D. Xie (2001). Beyond balanced growth. *Review of Economic Studies 68*(4), 869–882.

Leonardi, M. and E. Moretti (2023). The agglomeration of urban amenities: Evidence from milan restaurants. *American Economic Review: Insights 5*(2), 141–57.

Livi-Bacci, M. (1961). *L'Immigrazione e l'Assimilazione degli Italiani negli Stati Uniti secondo le Statistiche Demografiche Americane*. A. Giuffrè.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner (2004). European seasonal and annual temperature variability, trends, and extremes since 1500. *Science 303*(5663), 1499–1503.

McCulloch, R. and J. L. Yellen (1977). Factor mobility, regional development, and the distribution of income. *Journal of Political Economy 85*(1), 79–96.

Miyagiwa, K. (1991). Scale economies in education and the brain drain problem. *International Economic Review 32*(3), 743–759.

Mocetti, S. and C. Porello (2010). How does immigration affect native internal mobility? new evidence from italy. *Regional Science and Urban Economics 40*(6), 427–439.

Monras, J., E. Polo-Muro, and J. Vázquez-Grenno (2023). Labor market stability and fertility decisions: Evidence from a large immigration reform. Working Paper 2023-36, Federal Reserve Bank of San Francisco.

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in africa. *Review of Economics and Statistics 94*(1), 20–36.

Nunn, N. and N. Qian (2011). The potato's contribution to population and urbanization: evidence from a historical experiment. *The quarterly journal of economics 126*(2), 593–650.

Pauling, A., J. Luterbacher, C. Casty, and H. Wanner (2006). Five hundred years of gridded high-resolution precipitation reconstructions over europe and the connection to large-scale circulation. *Climate dynamics 26*, 387–405.

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.

Peters, M. (2022). Market size and spatial growth—evidence from germany's post-war population expulsions. *Econometrica 90*(5), 2357–2396.

Pugliese, E. (2002). *L'Italia tra migrazioni internazionali e migrazioni interne*. Bologna, Italy: Il Mulino.

Rocha, R., C. Ferraz, and R. R. Soares (2017). Human capital persistence and development. *American Economic Journal: Applied Economics 9*(4), 105–136.

Rovida, A. N., M. Locati, R. D. Camassi, B. Lolli, and P. Gasperini (2019). Catalogo parametrico dei terremoti italiani cpti15, versione 2.0.

Sequeira, S., N. Nunn, and N. Qian (2020). Immigrants and the making of america. *The Review of Economic Studies 87*(1), 382–419.

Sobotka, T., V. Skirbekk, and D. Philipov (2011). Economic recession and fertility in the developed world. *Population and Development Review 37*(2), 267–306.

Spitzer, Y., G. Tortorici, and A. Zimran (2025). International migration responses to modern europe's most destructive earthquake: Messina and reggio calabria, 1908. *The Journal of Economic History*. Forthcoming.

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.

Spitzer, Y. and A. Zimran (2023). Like an ink blot on paper: Testing the diffusion hypothesis of mass migration, italy 1876-1920. Technical report, National Bureau of Economic Research.

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.

Toniolo, G. (2013). *The Oxford Handbook of the Italian Economy Since Unification*. Oxford: Oxford University Press.

Valencia Caicedo, F. (2019). The mission: Human capital transmission, economic persistence, and culture in south america. *The Quarterly Journal of Economics 134*(1), 507–556.

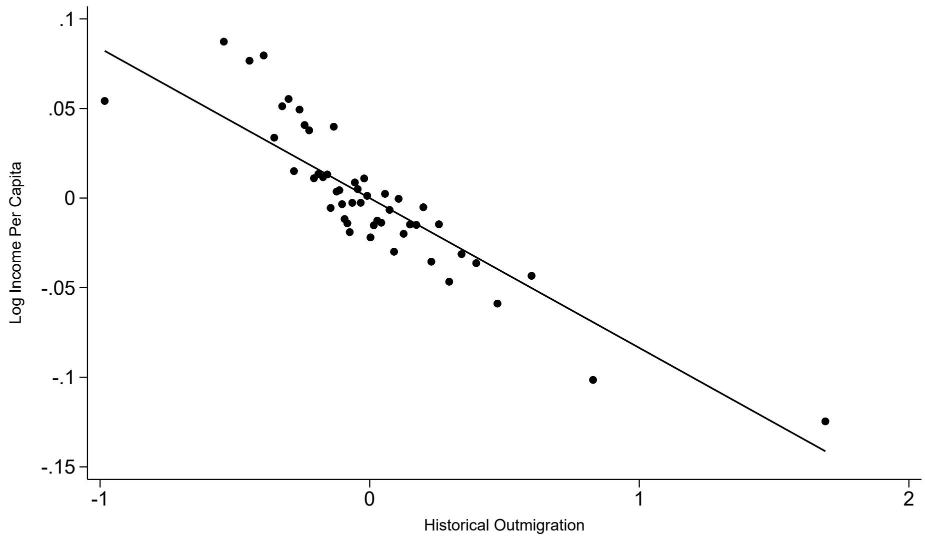
Waldinger, F. (2010). Quality matters: The expulsion of professors and the consequences for phd student outcomes in nazi germany. *Journal of political economy 118*(4), 787–831.

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in nazi germany. *The review of economic studies 79*(2), 838–861.

Waldinger, F. (2016). Bombs, brains, and science: The role of human and physical capital for the creation of scientific knowledge. *Review of Economics and Statistics 98*(5), 811–831.

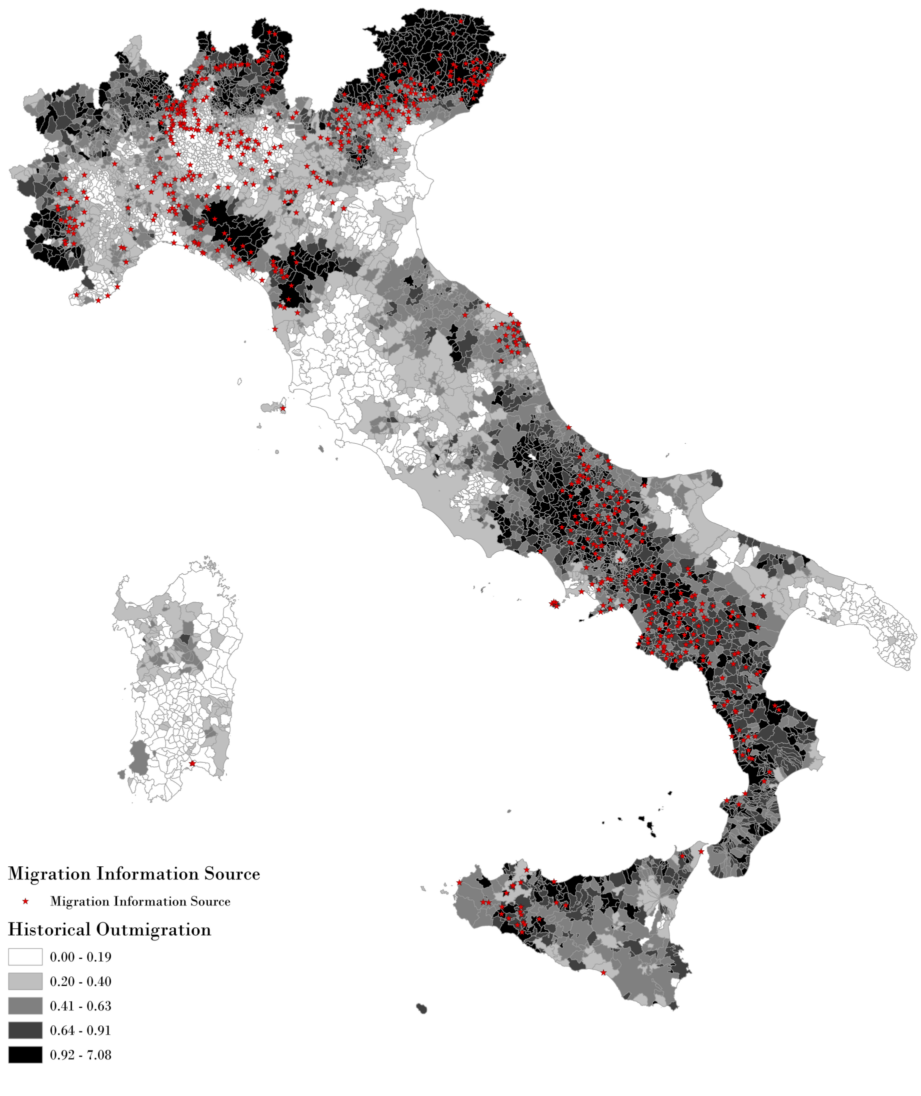
Figures and Tables

Figure 1. Historical Outmigration and Income per Capita (2001)



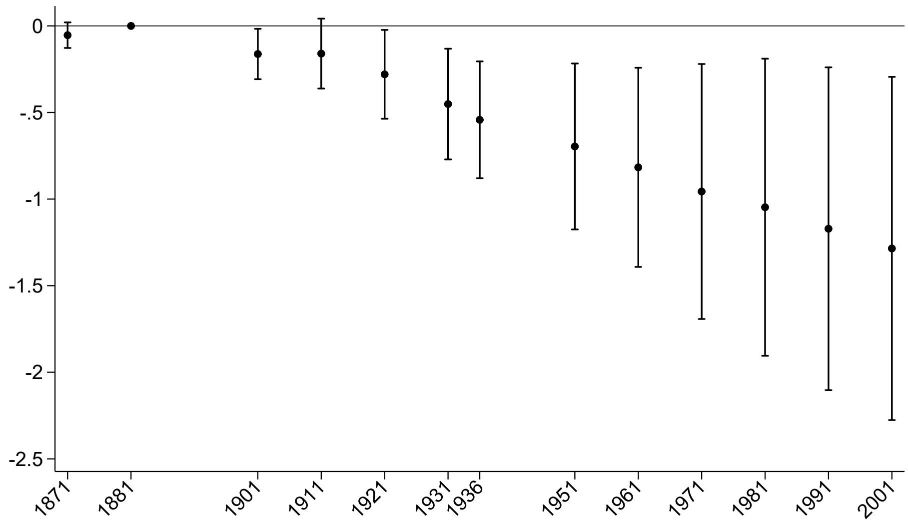
Notes: The figure displays the relationship between log income per capita in 2001 (y-axis) and historical outmigration from 1884 to 1920 scaled by 1881 population (x-axis), across Italian municipalities, after partialling out province fixed effects. The scatterplot pools observations into 50 bins. The point estimate is -0.084, and the associated standard errors, clustered at the province level, are 0.017.

Figure 2. Historical Outmigration Rate



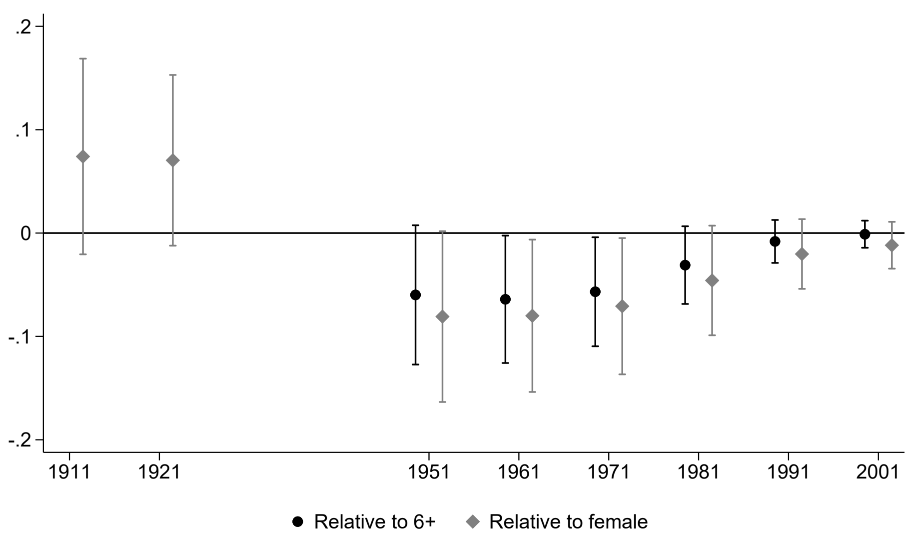
Notes: The figure displays the cumulated number of out-migrants between 1884 and 1920, relative to 1881 population, recorded in official Italian Migration Statistics. Red dots depict locations where at least one migration driver was reported in the 1884-1885 volume of the *Statistica della Emigrazione Italiana per l'Estero*. See Section 3 for more details.

Figure 3. Effects of Historical Outmigration on Resident Population



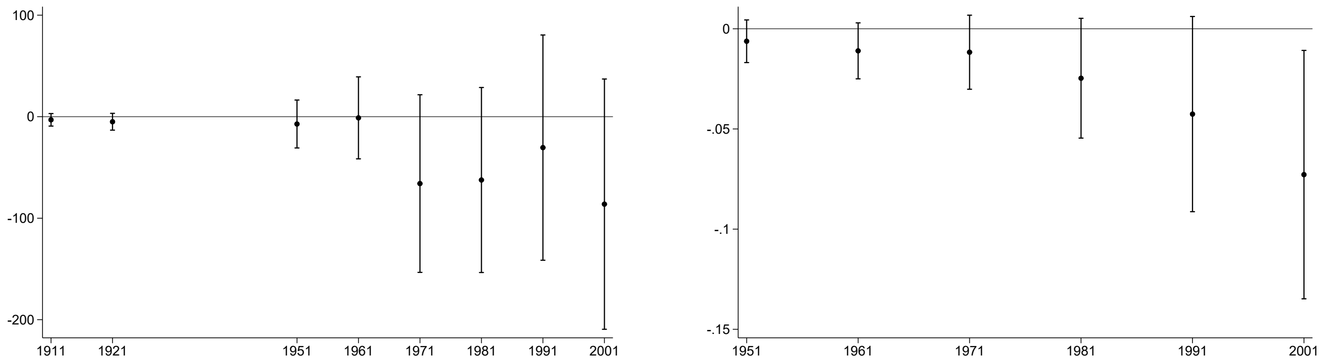
Notes: The figure reports 2SLS coefficients, with corresponding 95% confidence intervals, on historical outmigration from 1884 to 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. The dependent variable is the log of municipal resident population in each of the years indicated on the x-axis. All regressions include province fixed effects, the log of 1881 population, altitude, the full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops), distance between the municipality and several access points (coastline, border, province capital, and the closest main port), distance to the closest rail station and rail line, temperature and precipitation shocks (measured in deviation from the long-run mean), and exposure to violent earthquakes. Standard errors are clustered at the province level. For the corresponding estimates in tabular form, see Table A.8.

Figure 4. Effects of Historical Outmigration on Population Below Age of 6



Notes: The figure reports 2SLS coefficients, with corresponding 95% confidence intervals, on historical outmigration from 1884 to 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. The dependent variable is the share of municipal population under the age of 6 in each of the years indicated on the x-axis, computed relative to population aged 15-64 and female population. All regressions include province fixed effects, the log of 1881 population, altitude, the full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops), distance between the municipality and several access points (coastline, border, province capital, and the closest main port), distance to the closest rail station and rail line, temperature and precipitation shocks (measured in deviation from the long-run mean), and exposure to violent earthquakes. Standard errors are clustered at the province level. For the corresponding estimates in tabular form, see Table A.12 and Table A.13, respectively.

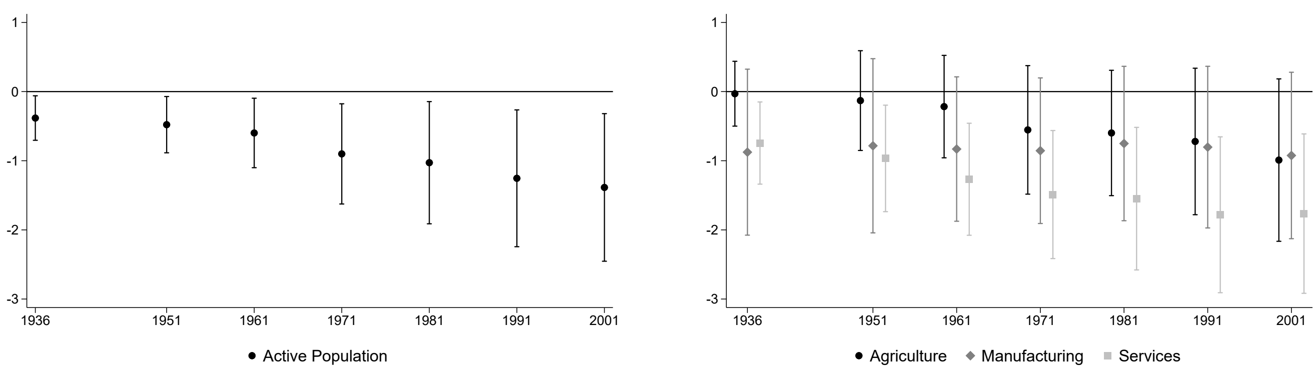
Figure 5. Effects of Historical Outmigration on on Literacy and Educational Attainment



(a) Literate vs. Illiterate Population (b) Population w/ at Least High School Degree

Notes: The figure reports 2SLS coefficients, with corresponding 95% confidence intervals, on historical outmigration from 1884 to 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. In panel (a), the dependent variable is the share of municipal literate population relative to municipal illiterate population in each of the years indicated in the x-axis. In panel (b), the dependent variable is the share of municipal population with at least a high school degree in each of the years indicated in the x-axis. All regressions include province fixed effects, the log of 1881 population, altitude, and a full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops); distance to multiple access points (coastline, border, province capital, and the closest main port); proximity to the nearest railway station and rail line; weather shocks (temperature and precipitation deviations from long-run means); and exposure to violent earthquakes. Standard errors are clustered at the province level. For the corresponding estimates in tabular form, see Table A.10 and Table A.11, respectively.

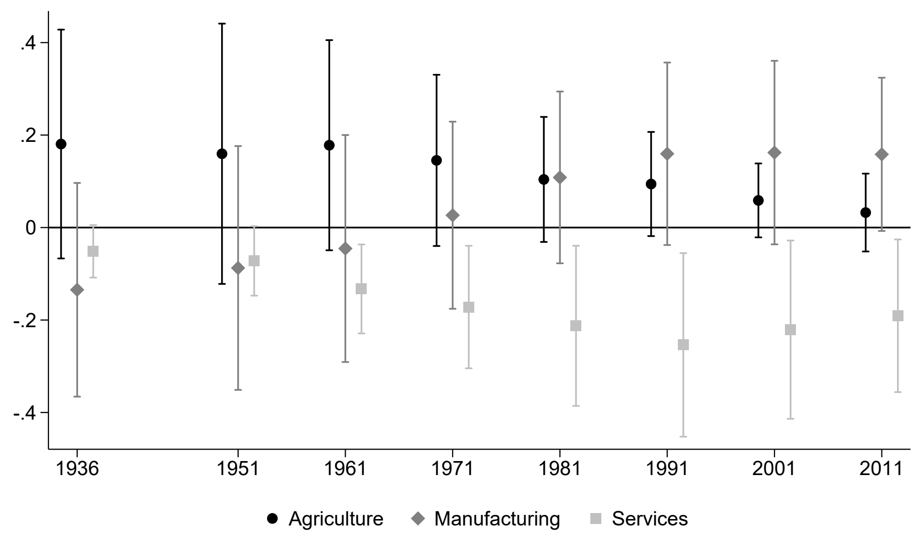
Figure 6. Effects of Historical Outmigration on Logged Employment



(a) Active Population (b) Employment by Sector

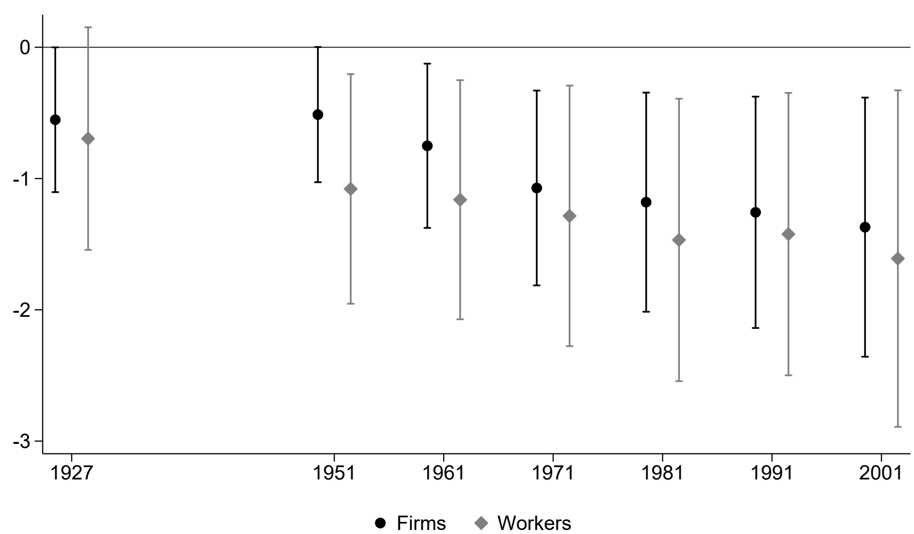
Notes: The figure reports 2SLS coefficients, with corresponding 95% confidence intervals, on historical outmigration from 1884 to 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. In panel (a), the dependent variable is the log of the municipal active population in each of the years indicated in the x-axis. In panel (b), the dependent variable is the log of the municipal employed population, disaggregated by sector (agriculture, manufacturing, and services), in each of the years indicated in the x-axis. All regressions include province fixed effects, the log of 1881 population, altitude, and a full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops); distance to multiple access points (coastline, border, province capital, and the closest main port); proximity to the nearest railway station and rail line; weather shocks (temperature and precipitation deviations from long-run means); and exposure to violent earthquakes. Standard errors are clustered at the province level. For the corresponding estimates in tabular form, see Table A.14, Table A.15, Table A.16, and Table A.17, respectively.

Figure 7. Effects of Historical Outmigration on Employment Shares



Notes: The figure reports 2SLS coefficients, with corresponding 95% confidence intervals, on historical outmigration from 1884 to 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. The dependent variable is the share of municipal employment in agriculture, manufacturing, and services for each of the years indicated on the x-axis. All regressions include province fixed effects, the log of 1881 population, altitude, and a full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops); distance to multiple access points (coastline, border, province capital, and the closest main port); proximity to the nearest railway station and rail line; weather shocks (temperature and precipitation deviations from long-run means); and exposure to violent earthquakes. Standard errors are clustered at the province level. For the corresponding estimates in tabular form, see Table A.18, Table A.19, and Table A.20, respectively.

Figure 8. Effects of Historical Outmigration on Number of Firms and Number of Workers



Notes: The figure reports 2SLS coefficients, with corresponding 95% confidence intervals, on historical outmigration from 1884 to 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. The dependent variables are the log number of municipal firms and the log number of municipal workers for each of the years indicated on the x-axis. All regressions include province fixed effects, the log of 1881 population, altitude, and a full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops); distance to multiple access points (coastline, border, province capital, and the closest main port); proximity to the nearest railway station and rail line; weather shocks (temperature and precipitation deviations from long-run means); and exposure to violent earthquakes. Standard errors are clustered at the province level. For the corresponding estimates in tabular form, see Table A.22 and Table A.21.

Table 1. First Stage Estimates: Distance to Migration Information Source

Dep. Variable:	Historical Outmigration			
	(1)	(2)	(3)	(4)
Distance to Migration Information	-0.022***	-0.040***	-0.040***	-0.039***
	(0.007)	(0.010)	(0.010)	(0.009)
KP F-stat	10.51	17.18	17.31	19.36
Observations	5,803	5,803	5,803	5,803
Province FE	Yes	Yes	Yes	Yes
Log 1881 Population	Yes	Yes	Yes	Yes
Geography	Altitude	Yes	Yes	Yes
Crop Suitability		Yes	Yes	Yes
Distance to Access Points		Yes	Yes	Yes
Distance to Railway			Yes	Yes
Weather Shocks				Yes
Violent Earthquakes				Yes

Notes: The table presents first stage regressions of cumulated historical emigration from 1884 to 1920, scaled by 1881 population, on the distance to migration information source presented in Section 4.2. The unit of observation is the municipality (see Section 3 and Appendix B for more details). All specifications include province fixed effects and the log of 1881 population. Column 1 also controls for altitude. Column 2 adds the full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, and area) along with indicators of suitability for main crops and distances from the municipality to key access points (coastline, border, province capital, and the nearest main port). Column 3 additionally includes distances to the closest railway station and railway line. Column 4 adds temperature and precipitation shocks (measured as deviations from the long-run mean) and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 2. Effects of Historical Outmigration on Resident Population, 2001

Dep. Variable:	Log Resident Population			
	(1)	(2)	(3)	(4)
Panel A: 2SLS				
Historical Outmigration	-1.201*	-1.129**	-1.152**	-1.303***
	(0.673)	(0.442)	(0.455)	(0.483)
KP F-stat	10.51	17.18	17.31	19.36
Panel B: OLS				
Historical Outmigration	-0.119***	-0.125***	-0.123***	-0.135***
	(0.043)	(0.044)	(0.043)	(0.045)
Mean Dep. Var.	9,587	9,587	9,587	9,587
Observations	5,803	5,803	5,803	5,803
Province FE	Yes	Yes	Yes	Yes
Log 1881 Population	Yes	Yes	Yes	Yes
Geography	Altitude	Yes	Yes	Yes
Crop Suitability		Yes	Yes	Yes
Distance to Access Points		Yes	Yes	Yes
Distance to Railway			Yes	Yes
Weather Shocks				Yes
Violent Earthquakes				Yes

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. The unit of observation is the municipality (see Section 3 and Appendix B for more details). The dependent variable is the log of resident population in 2001. Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects and the log of 1881 population. Column 1 also controls for altitude. Column 2 adds the full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, and area) along with indicators of suitability for main crops and distances from the municipality to key access points (coastline, border, province capital, and the nearest main port). Column 3 additionally includes distances to the closest railway station and railway line. Column 4 adds temperature and precipitation shocks (measured as deviations from the long-run mean) and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3. Effects of Historical Outmigration on Income per Capita, 2001

Dep. Variable:	Log Income per Capita			
	(1)	(2)	(3)	(4)
Panel A: 2SLS				
Historical Outmigration	-0.213	-0.217**	-0.227**	-0.246***
	(0.129)	(0.085)	(0.087)	(0.089)
KP F-stat	10.51	17.18	17.31	19.36
Panel B: OLS				
Historical Outmigration	-0.039***	-0.039***	-0.039***	-0.038***
	(0.013)	(0.013)	(0.013)	(0.012)
Mean Dep. Var.	12,656	12,656	12,656	12,656
Observations	5,803	5,803	5,803	5,803
Province FE	Yes	Yes	Yes	Yes
Log 1881 Population	Yes	Yes	Yes	Yes
Geography	Altitude	Yes	Yes	Yes
Crop Suitability		Yes	Yes	Yes
Distance to Access Points		Yes	Yes	Yes
Distance to Railway			Yes	Yes
Weather Shocks				Yes
Violent Earthquakes				Yes

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. The unit of observation is the municipality (see Section 3 and Appendix B for more details). The dependent variable is the log of income per capita in 2001. Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects and the log of 1881 population. Column 1 also controls for altitude. Column 2 adds the full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, and area) along with indicators of suitability for main crops and distances from the municipality to key access points (coastline, border, province capital, and the nearest main port). Column 3 additionally includes distances to the closest railway station and railway line. Column 4 adds temperature and precipitation shocks (measured as deviations from the long-run mean) and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Effects of Historical Outmigration on Educational Attainment, 2001

Dep. Variable:	Share w/ at Least High School Degree			
	(1)	(2)	(3)	(4)
Panel A: 2SLS				
Historical Outmigration	-0.100*	-0.061*	-0.066**	-0.072**
	(0.053)	(0.031)	(0.030)	(0.030)
KP F-stat	10.51	17.17	17.31	19.36
Panel B: OLS				
Historical Outmigration	-0.006**	-0.006**	-0.006**	-0.006*
	(0.003)	(0.003)	(0.003)	(0.003)
Mean Dep. Var.	0.272	0.272	0.272	0.272
Observations	5,803	5,803	5,803	5,803
Province FE	Yes	Yes	Yes	Yes
Log 1881 Population	Yes	Yes	Yes	Yes
Geography	Altitude	Yes	Yes	Yes
Crop Suitability		Yes	Yes	Yes
Distance to Access Points		Yes	Yes	Yes
Distance to Railway			Yes	Yes
Weather Shocks				Yes
Violent Earthquakes				Yes

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. The unit of observation is the municipality (see Section 3 and Appendix B for more details). The dependent variable is the share of population with at least a high school degree, relative to total population aged six or more in 2001. Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects and the log of 1881 population. Column 1 also controls for altitude. Column 2 adds the full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, and area) along with indicators of suitability for main crops and distances from the municipality to key access points (coastline, border, province capital, and the nearest main port). Column 3 additionally includes distances to the closest railway station and railway line. Column 4 adds temperature and precipitation shocks (measured as deviations from the long-run mean) and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5. Effects of Historical Outmigration on Agriculture, 1929

Dep. Variables:	Log Cultivated Area	Log Number of Farms	Log Number of Workers	Horses per Farm	Horses per Worker	Log Output	Crop Yield	Land Concentration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 2SLS								
Historical Outmigration	-0.592	0.024	-0.315	-0.827	-0.254	-0.305	0.560	0.127
	(0.440)	(0.276)	(0.408)	(0.544)	(0.182)	(0.730)	(3.444)	(0.085)
KP F-stat	19.55	19.67	19.71	19.68	19.72	18.29	18.29	19.67
Panel B: OLS								
Historical Outmigration	-0.077***	0.098***	-0.116**	-0.163**	-0.027*	-0.236***	-0.617	-0.008
	(0.022)	(0.030)	(0.053)	(0.062)	(0.014)	(0.056)	(0.467)	(0.006)
Mean Dep. Var.	4,552	537.5	1,850	0.751	0.256	18,880	15.89	0.084
Observations	5,794	5,790	5,788	5,788	5,786	5,693	5,693	5,790

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. The unit of observation is the municipality (see Section 3 and Appendix B for more details). The dependent variable is the log of cultivated area (column 1), the log of the number of farms (column 2), the log of the number of agricultural workers (column 3), the number of horses per farm (column 4) and per worker (column 5), the log of agricultural output (column 6), crop yield, defined as output per acre (column 7), and an index of land concentration, calculated as the Herfindahl-Hirschman Index (HHI) (column 8). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Effects of Historical Outmigration on Modern Migration, 2002–2015

Dep. Variables:	In-Migration Rate				Out-Migration Rate			
	Domestic	International	No College	College Degree	Domestic	International	No College	College Degree
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 2SLS								
Historical Outmigration	-0.154	-0.046**	-0.103	-0.104	-0.136*	0.009	-0.080	-0.088
	(0.096)	(0.022)	(0.075)	(0.075)	(0.081)	(0.010)	(0.061)	(0.061)
KP F-stat	19.36	19.36	19.36	19.36	19.36	19.36	19.36	19.36
Panel B: OLS								
Historical Outmigration	-0.010	-0.005**	0.002	0.000	-0.012	0.001	-0.001	-0.002
	(0.009)	(0.002)	(0.008)	(0.008)	(0.007)	(0.001)	(0.007)	(0.007)
Mean Dep. Var.	0.368	0.078	0.292	0.284	0.369	0.019	0.287	0.279
Observations	5,803	5,803	5,803	5,803	5,803	5,803	5,803	5,803

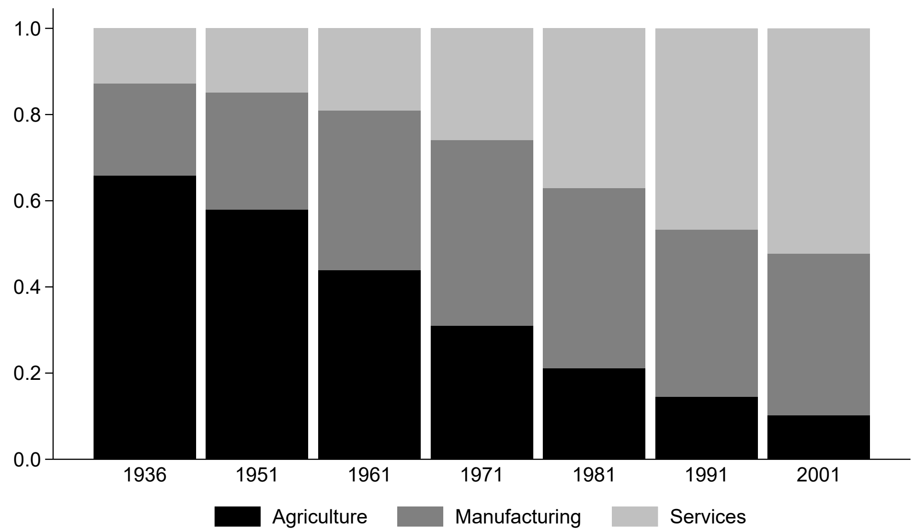
Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. The unit of observation is the municipality (see Section 3 and Appendix B for more details). The dependent variable is the share of in- (columns 1 to 4) and out- (columns 5 to 8) migrants from 2002 to 2015, relative to 2001 municipality population. In both cases, we distinguish between domestic (columns 1 and 5) and international (columns 2 and 6) migrants, and between individuals without (columns 3 and 7) and with (columns 4 and 8) a college degree. Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix: Additional material

Table of Contents

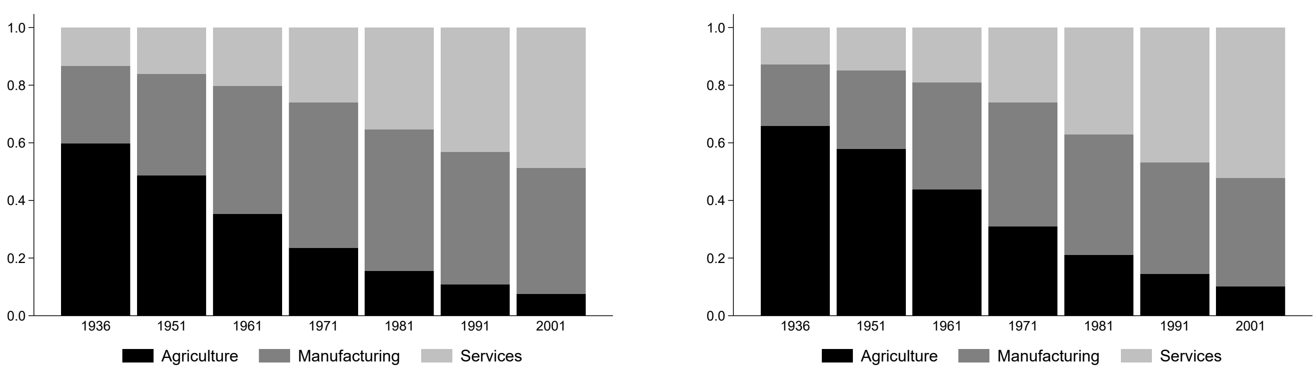
A Online Appendix – Additional Figures and Tables

Figure A.1. Sectoral Composition of Employment over Time



Notes: The figure displays the share of municipal employment in agriculture, manufacturing, and services between 1936 and 2001. Data come from the 8mila Census database (ISTAT).

Figure A.2. Sectoral Composition of Employment over Time



(a) Northern Regions

(b) Southern and Central Regions

Notes: Panel A displays the share of municipal employment in agriculture, manufacturing, and services between 1936 and 2001 for Northern Italian regions: Emilia-Romagna, Friuli-Venezia Giulia, Liguria, Lombardia, Piemonte, Veneto. Data come from the 8mila Census database (ISTAT). Panel B displays the share of municipal employment in agriculture, manufacturing, and services between 1936 and 2001 for Southern and Central Italian regions. We define Southern regions: Abruzzo, Basilicata, Calabria, Campania, Puglia, Sardegna, Sicilia. We define Central regions: Lazio, Marche, Toscana, Umbria. Data come from the 8mila Census database (ISTAT).

Figure A.3. Sample of 1891-1892 Emigration Statistics

Tavola I. — Emigrazione per Comuni

Piemonte. **Alessandria.**

Circondari Distretti e Comuni	Popolazione (Censimento al 31 Dicembre 1881)	1891 Emigrazione Propria	1891 Emigrazione Temporanea	1892 Emigrazione Propria	1892 Emigrazione Temporanea	Circondari Distretti e Comuni	Popolazione (Censimento al 31 Dicembre 1881)	1891 Emigrazione Propria	1891 Emigrazione Temporanea	1892 Emigrazione Propria	1892 Emigrazione Temporanea	Circondari Distretti e Comuni	Popolazione (Censimento al 31 Dicembre 1881)	1891 Emigrazione Propria	1891 Emigrazione Temporanea	1892 Emigrazione Propria	1892 Emigrazione Temporanea

PROVINCIA DI ALESSANDRIA.

Circondario di Acqui.

Bruno	953	2	..	10	7	Mombaruzzo	3008	..	6	..	28	33 Comuni senza emigrazione	51110	..	..	..	..
Calamandrana	2052	11	..	4	..	Orsara Bormida	1312	..	..	..	11						
Castelnuovo Belbo	1877	..	..	13	1	Ponzone	3756	..	60	..	75						
Castelnuovo Bormida	1793	4	..	25	..	Spigno Monferr.	3204	2	26	..	..						
Melazzo	2082	1	12	..	14	21 altri comuni	35079	19	5	39	41	Totale N. 63.	106226	39	109	91	177

Circondario di Alessandria.

Alessandria	62464	60	..	..	..	Quattordio	1885	1	..	13	..	17 comuni senza emigrazione	33534	..	..	..	..
Bassignana	3604	5	10	23	..	Solero	3994	..	..	15	..						
Cassine	5283	17	..	68	..	Valenza	9835	22	..	65	..						
Pecetto di Valenza	2099	..	1	10	1	10 altri comuni	30276	16	2	28	14	Totale N. 34.	152974	121	13	222	15

Circondario di Asti.

Asti	33233	33	10	37	12	Isola d'Asti	2860	..	10	..	..	50 comuni senza emigrazione	63351	..	..	..	..
Baldichieri	787	..	..	..	12	Mombercelli	3653	4	..	1	30						
Canelli	5461	8	..	28	..	Moransengo	523	..	..	..	12						
Cocconato	2694	..	..	..	28	Rocchetta Tanaro	3410	..	1	..	14						
Frinco	1565	1	..	..	15	27 altri comuni	51648	34	58	25	75	Totale N. 86.	169185	80	79	91	198

Circondario di Casale Monferrato.

Balzola	3295	8	..	44	..	Mombello Monferrato	3174	9	..	18	..	Villanova Monferrato	3403	11	..	16	..
Calliano	3247	1	..	42	..	Moncestino	885	5	..	11	..	34 altri comuni	64875	51	..	80	..
Casale Monferrato	28711	25	..	67	..	Montalero	649	2	..	16	..						
Cerrina	1398	2	..	10	..	Morano sul Po	2956	2	..	18	..	21 comuni senza emigrazione	30763	·			
Fubine	3609	6	..	48	..	Varengo	751	4	..	33	..						
Gabiano	2664	4	..	50	..	Villamiroglio	1546	2	..	11	..	Totale N. 71.	151926	132	..	464	..

Circondario di Novi Ligure.

Arquata Scrivia	2817	1	16	..	21	Francavilla Bisio	678	..	28	..	5	4 comuni senza emigrazione	5109	..	..	..	..
Borghetto di Borbera	749	3	8	..	10	Mongiardino Ligure	1715	1	18	..	36						
Cabella Ligure	1598	2	8	..	26	Novi Ligure	13783	23	21	..	37						
Cantalupo Ligure	1243	2	9	..	17	Roccaforte Ligure	1167	..	1	..	16						
Capriata d'Orba	3072	9	2	..	30	Rocchetta Ligure	1142	2	12	..	8						
Carrega	2673	1	18	..	10	23 altri comuni	47952	17	40	..	66	Totale N. 38.	83698	61	181	..	282

Figure A.4. 1884-1885 Emigration Statistics: Information on Opportunities Abroad

Sugli agenti di emigrazione.

6. Esistevano nel comune vere e proprie agenzie di emigrazione o, quanto meno, tra le persone ivi stabilite ve ne erano alcune incaricate di dare informazioni su uno od altro Stato di oltremare a chi desiderava di emigrare?

7. In difetto di veri agenti e di altri incaricati stabili, o in aggiunta a questi, consta che nel corso dell'anno siansi recate temporaneamente nel comune persone che eccitavano la popolazione ad emigrare verso un dato paese?

8. Quali mezzi di eccitamento (sussidi, anticipazioni, opuscoli, ecc.), vennero più specialmente adoperati per persuadere gli abitanti alla emigrazione?

9. Furono messi in opera altri mezzi di propaganda, come sarebbe l'invio agli osti, ai caffettieri, ecc. di manifesti da affiggere?

(a) Survey to the Prefects and Mayors

SUNTO DELLE RISPOSTE DATE DAI SINDACI RIGUARDO AGLI AGENTI D'EMIGRAZIONE ED AI MEZZI ADOPERATI PER FOMENTARE L'EMIGRAZIONE

Avvertenze. I numeri qui sotto segnati ricordano i quesiti della circolare ministeriale 21 marzo 1884, ai quali furono date dai sindaci le risposte relative agli agenti d'emigrazione.

PIEMONTE.

Provincia di Alessandria.

6. Nel comune di Spigno-Monferrato (Acqui), fin dal settembre 1883, un agente autorizzato dava informazioni a chi desiderava di emigrare. In Castelnuovo-Scrivia e Montacuto (Tortona) esistono degl'incaricati stabili. A Tortona poi sonvi due agenzie autorizzate di emigrazione.

7. A S. Marzano-Oliveto (Asti) un tale, reduce dall'Argentina, eccitava ad emigrare in America.

A Fubine (Casal Monferrato) speculatori interessati nei trasporti e agenti pagati dai Governi di paesi di immigrazione o da Società intrapenditrici di colonizzazione tentarono nel 1883 di indurre gli abitanti ad emigrare; ma a poco valsero i loro tentativi.

A Silvano d'Orba (Novi Ligure) trovavasi un incaricato per conto di imprese private di colonizzazione. Infine a Garbagna e Sale (Tortona) si ebbero degl'incaricati che eccitavano la popolazione ad emigrare per l'America. Quello presentatosi nel comune di Sale sembra che agisse per conto di qualche compagnia di trasporti marittimi, poichè s'assumeva di provvedere all'imbarco.

8. I mezzi di eccitamento più specialmente adoperati per persuadere gli abitanti all'emigrazione nei comuni di S. Marzano-Oliveto (Asti), Fubine (Casal Monferrato), Silvano d'Orba (Novi Ligure) Sale, Garbagna e Tortona furono le promesse di grossi guadagni. Nei due ultimi comuni si anticiparono denari.

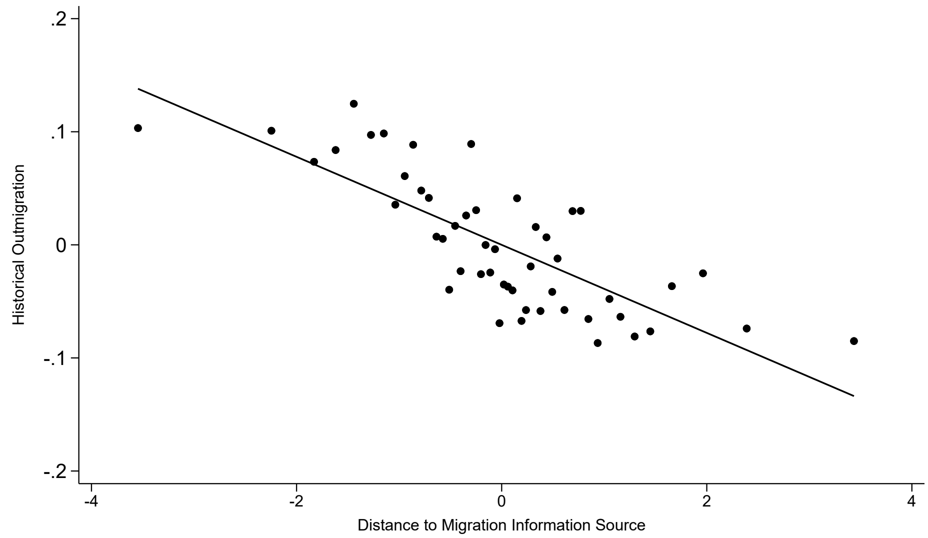
Nei comuni di Grondona e Ovada (Novi Ligure), Castelnuovo Scrivia, Cerreto-Grue e Viguzzolo (Tortona) influirono i consigli e gli eccitamenti dei parenti ed amici già all'estero.

9. In qualche comune della provincia di Alessandria furono inviati agli osti, ai caffettieri, ecc. manifesti da affiggere. Tali manifesti si riferivano alla partenza dei piroscafi, ai prezzi d'imbarco, ecc.

Nel comune di Valenza (Alessandria) furono inviati ai pubblici esercenti degli opuscoli, che vantavano la fertilità delle terre d'America e promettevano o lasciavano sperare alle famiglie di abili agricol.

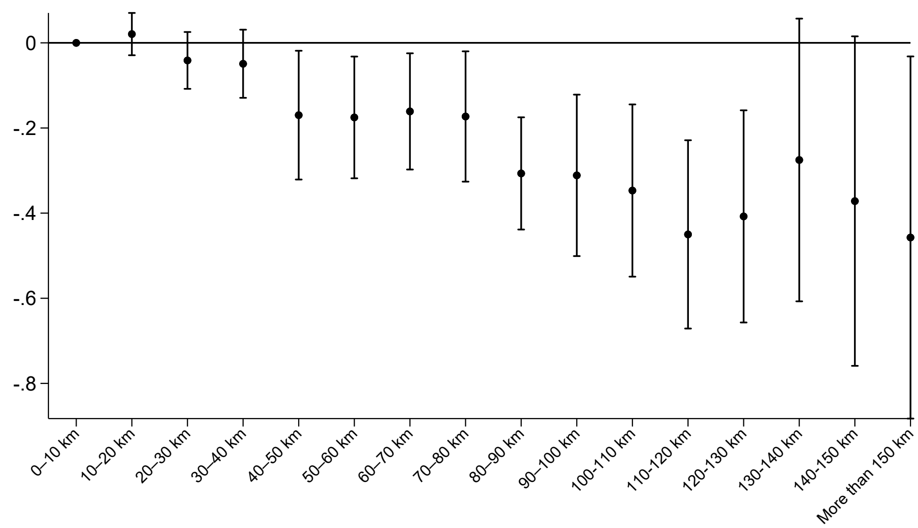
(b) Sample of Answers

Figure A.5. First Stage: Residualized Binscatter



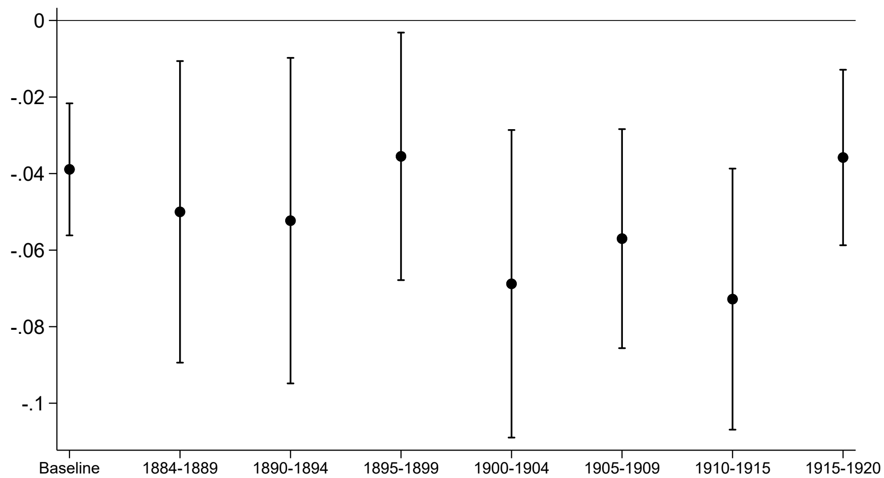
Notes: The figure displays the relationship between historical outmigration from 1884 to 1920, scaled by 1881 population (y-axis), and the distance to migration information source (x-axis), across Italian municipalities. Both variables are residualized with respect to the full set of controls from the preferred specification (see column 4 of Table 1). The scatterplot pools observations into 50 bins. The point estimate is -0.039, and the associated standard errors, clustered at the province level, are 0.009.

Figure A.6. First Stage Robustness: Distance to Migration Information Source (10 km bins)



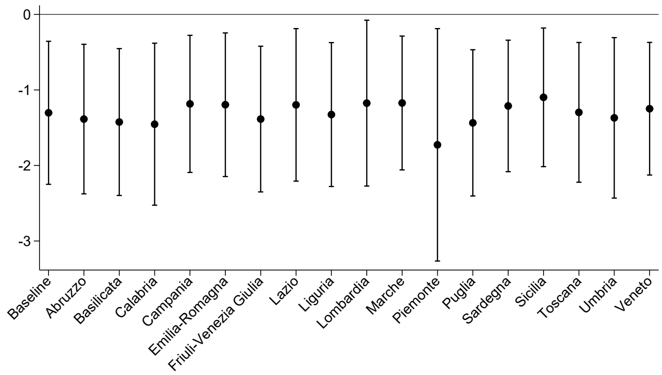
Notes: The figure reports the coefficients from the preferred first stage specification, with corresponding 95% confidence intervals, where the instrument is discretized into 10 km intervals. The instrument is the distance to migration information source presented in Section 4.2. The dependent variable is historical outmigration from 1884 to 1920, scaled by 1881 population. The specification include province fixed effects, the log of 1881 population, altitude, and a full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops); distance to multiple access points (coastline, border, province capital, and the closest main port); proximity to the nearest railway station and rail line; weather shocks (temperature and precipitation deviations from long-run means); and exposure to violent earthquakes. Standard errors are clustered at the province level.

Figure A.7. First Stage Robustness: Estimates by Five-Year Periods

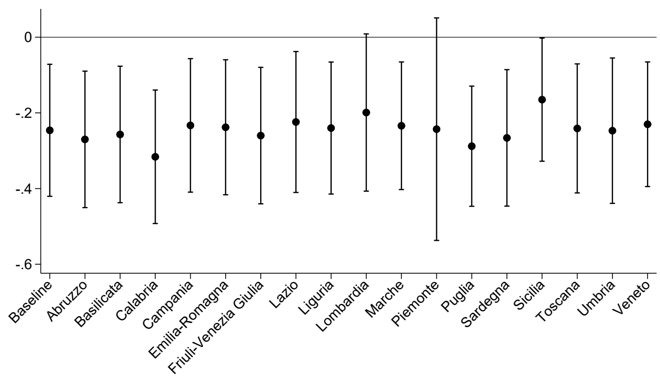


Notes: The figure reports the coefficients from the preferred first stage specification, estimated for separate year bins, with corresponding 95% confidence intervals. The instrument is the distance to migration information source presented in Section 4.2. The dependent variable is historical outmigration from 1884 to 1920, scaled by 1881 population. Each specification includes province fixed effects, the log of 1881 population, altitude, and a full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops); distance to multiple access points (coastline, border, province capital, and the closest main port); proximity to the nearest railway station and rail line; weather shocks (temperature and precipitation deviations from long-run means); and exposure to violent earthquakes. Standard errors are clustered at the province level.For the corresponding estimates in tabular form, see Table A.7.

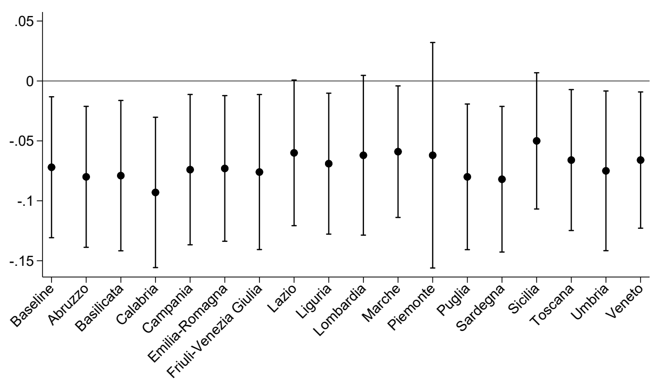
Figure A.8. Robustness Checks: Excluding One Region at the Time



(a) Resident Population



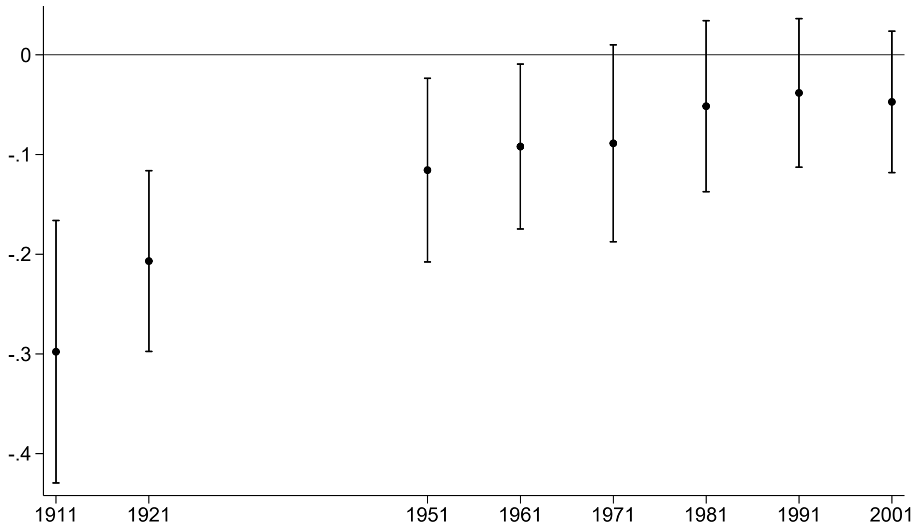
(b) Income per Capita



(c) Population w/ at least High School Degree

Notes: The figure reports 2SLS coefficients, with corresponding 95% confidence intervals, on historical outmigration from 1884 to 1920, scaled by 1881 population. The figure shows robustness checks where the baseline specification is re-estimated sequentially excluding each of the 20 Italian regions one at a time. The instrument is the distance to migration information source presented in Section 4.2. In panel (a), the dependent variable is the the log of municipality population. In panel (b), the dependent variable is the log of income per capita. In panel (c), the dependent variable is the share of municipal population with at least a high school degree (Panel C). All regressions include province fixed effects, the log of 1881 population, altitude, a full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops), distance to multiple access points (coastline, border, province capital, and the closest main port), proximity to the nearest railway station and rail line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. Standard errors are clustered at the province level.

Figure A.9. Effects of Historical Outmigration on Sex Ratios



Notes: The figure reports 2SLS coefficients, with corresponding 95% confidence intervals, on historical outmigration from 1884 to 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. The dependent variable is the municipality sex ratio (defined as the number of men relative to the number of women) in each of the years indicated on the x-axis. All regressions include province fixed effects, the log of 1881 population, altitude, the full set of geographic controls (slope, ruggedness, lakes and rivers, coordinates, area, and crop suitability for main crops), distance between the municipality and several access points (coastline, border, province capital, and the closest main port), distance to the closest rail station and rail line, temperature and precipitation shocks (measured in deviation from the long-run mean), and exposure to violent earthquakes. Standard errors are clustered at the province level. For the corresponding estimates in tabular form, see Table A.9.

Table A.1. Summary Statistics: Main Variables

Variables:	Mean	SD	Min	P25	Median	P75	Max	Obs	Year
Panel A: Dynamics									
Resident Population	7,201	42,470	31.00	1,505	2,757	5,430	3,641,018	82,288	1871-2001
Sex Ratio	0.969	0.084	0.322	0.930	0.970	1.009	4.070	47,871	1936-2001
Share of Literate/Illiterate	76.77	152.3	0.074	5.311	18.91	85.82	4,643	47,195	1911-2001
Share of High School Graduates	0.103	0.097	0	0.025	0.063	0.158	0.721	36,060	1951-2001
Share of Population Below 6 (15-64)	0.121	0.049	0.032	0.083	0.110	0.150	0.392	35,767	1951-2001
Share of Population Below 6 (Female)	0.201	0.108	0.041	0.119	0.172	0.259	0.730	46,903	1911-2001
Active Population	3,825	24,421	10.31	640.5	1,286	2,751	1,748,573	42,067	1936-2001
Workers in Agriculture	794.8	1,947	0	103.3	311.2	825.7	106,276	42,067	1936-2001
Workers in Manufacturing	1,310	8,087	0	154.2	375.9	953.2	528,452	42,059	1936-2001
Workers in Services	1,479	14,287	0	122.0	276.3	712.3	1,162,512	42,067	1936-2001
Agriculture Employment Share	0.349	0.271	0	0.104	0.284	0.581	1.000	42,069	1936-2001
Manufacturing Employment Share	0.353	0.181	0	0.212	0.335	0.479	0.996	42,061	1936-2001
Services Employment Share	0.298	0.178	0	0.145	0.262	0.430	0.953	42,069	1936-2001
Total Number of Workers	1,603	14,056	0	104.0	275.0	842.0	885,007	46,476	1927-2001
Number of Firms	380.7	2,491	0	48.00	109.0	259.0	208,983	46,476	1927-2001
Panel B: Cross Section									
Historical Outmigration Rate	0.621	0.597	0	0.239	0.519	0.831	7.080	5,803	1884–1920
Distance to Migration Information Source	3.755	6.823	0.009	0.669	1.350	3.459	42.13	5,803	1882-1884
Income per Capita	12,656	3,163	5,104	10,122	12,536	14,856	39,874	5,803	2001
Cultivated Area	4,552	9,830	81.00	1,260	2,438	4,802	452,096	5,794	1929
Number of Farms	632.9	1,153	0	216.0	380.0	684.0	45,781	5,803	1929
Number of Agricultural Workers	1,849	3,055	0	578.0	1,092	2,045	73,782	5,790	1929
Number of Horses per Farm	0.812	1.472	0	0.289	0.590	0.933	83.47	5,802	1929
Number of Horses per Worker	0.256	0.305	0	0.073	0.166	0.325	4.479	5,786	1929
Agricultural Output	18,564	39,038	0	2,999	8,603	19,912	1,088,960	5,790	1929
Crop Yield	15.89	7.869	0.770	10.04	13.82	19.94	53.94	5,693	1929
Land Concentration Index	0.095	0.159	0	0.011	0.029	0.096	1.000	5,802	1929
In-Migration, Domestic	3,090	16,891	8.000	338.0	911.0	2,558	968,249	5,803	2002–2015
In-Migration, International	848.8	8,032	0	74.00	189.0	540.0	475,478	5,803	2002–2015
In-Migration, No Degree	2,380	12,593	7.000	281.0	730.0	1,966	751,463	5,803	2002–2015
In-Migration, College Degree	2,312	12,206	7.000	266.0	703.0	1,921	726,590	5,803	2002–2015
Out-Migration, Domestic	3,097	17,753	23.00	408.0	932.0	2,427	930,134	5,803	2002–2015
Out-Migration, International	176.9	1,328	0	16.00	46.00	125.0	74,160	5,803	2002–2015
Out-Migration, No Degree	2,404	13,899	15.00	320.0	721.0	1,842	738,524	5,803	2002–2015
Out-Migration, College Degree	2,327	13,385	15.00	310.0	695.0	1,776	704,411	5,803	2002–2015

Table A.2. Summary Statistics: Geographic Controls

Variables:	Mean	SD	Min	P25	Median	P75	Max	Obs
Altitude	414.0	419.8	0	110.0	287.0	582.0	2,687	5,803
Ruggedness Index	2.080	2.121	0	0.294	1.479	3.118	12.31	5,803
Slope	6.280	6.591	0	0.844	4.335	9.391	39.81	5,803
Number of Large Lakes	0.022	0.150	0	0	0	0	2.000	5,803
Number of Large Rivers	0.034	0.183	0	0	0	0	2.000	5,803
Longitude	11.64	2.854	6.731	9.106	11.46	14.03	18.46	5,803
Latitude	43.25	2.554	35.59	41.15	44.33	45.45	46.63	5,803
Area	49.44	102.5	1.010	13.75	26.14	52.62	4,511	5,803
Distance to Coastline	62.84	54.04	0	16.68	48.59	100.5	224.2	5,803
Distance to Border	276.6	261.8	0	53.62	161.3	458.2	989.6	5,803
Distance to Provincial Capital	31.00	18.00	0.173	17.99	27.13	40.85	219.0	5,803
Distance to Port	172.5	95.85	0	104.0	154.0	229.4	466.6	5,803
Suitability Index – Barley	29.65	16.95	0	17.52	28.14	39.29	86.55	5,803
Suitability Index – Rye	20.86	11.76	0	13.70	20.13	27.90	66.81	5,803
Suitability Index – Wheat	21.14	12.06	0	13.44	19.65	26.70	66.04	5,803
Suitability Index – Dry Rice	20.79	13.21	0	12.19	19.57	28.67	65.16	5,803
Suitability Index – Foxtail Millet	21.28	15.11	0	11.87	17.47	27.31	71.76	5,803
Suitability Index – Maize	21.28	15.11	0	11.87	17.47	27.31	71.76	5,803
Suitability Index – Pearl Millet	29.34	16.71	0	17.47	28.15	38.37	86.68	5,803
Suitability Index – Wet Rice	30.39	17.22	0	18.34	28.53	40.68	86.68	5,803
Suitability Index – White Potatoes	21.14	12.06	0	13.44	19.65	26.70	66.04	5,803
Suitability Index – Sorghum	4.963	8.111	0	0	0	9.738	44.46	5,803
Suitability Index – Sweet Potato	4.711	11.55	0	0	0.048	2.730	63.51	5,803
Mean Annual Precipitation	246.1	79.54	54.80	189.6	217.7	286.8	548.3	5,803
Mean Annual Temperature	11.57	3.122	-1.182	10.31	11.86	13.57	17.79	5,803
Mean Number of Violent Earthquakes	0.207	0.189	0	0.081	0.135	0.324	0.838	5,803
Mean Distance to Railway Line	6.862	7.842	0	0.338	5.307	9.976	216.0	5,803
Mean Distance to Railway Station	11.26	7.896	0	6.270	10.18	15.13	216.1	5,803

Table A.3. Robustness Checks: Additional Controls

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dep. Variable: Log Resident Population						
Historical Outmigration	-1.303***	-1.925**	-1.360***	-1.341**	-1.169**	-1.497***
	(0.483)	(0.777)	(0.507)	(0.526)	(0.450)	(0.521)
Mean Dep. Var.	9,587	9,587	9,587	9,587	9,588	9,590
Panel B: Dep. Variable: Log Income per Capita						
Historical Outmigration	-0.246***	-0.333***	-0.250***	-0.246**	-0.161*	-0.284***
	(0.089)	(0.123)	(0.093)	(0.095)	(0.088)	(0.090)
Mean Dep. Var.	12,656	12,656	12,656	12,656	12,657	12,656
Panel C: Dep. Variable: Share w/ at Least High School Degree						
Historical Outmigration	-0.072**	-0.099**	-0.074**	-0.077**	-0.053	-0.087***
	(0.030)	(0.042)	(0.032)	(0.035)	(0.033)	(0.032)
Mean Dep. Var.	0.272	0.272	0.272	0.272	0.272	0.272
KP F-stat	19.36	12.37	18.10	15.26	20.15	17.05
Observations	5,803	5,803	5,803	5,803	5,801	5,800
Specification	Baseline	Migration Epicenter	Early Migration	Muni Factor	Muni 100km	WWI Casualties

Notes: The table replicates the 2SLS regressions presented in column 4 of Tables 2 to 4, adding further controls. In panel A, the dependent variable is the log of municipality population. In panel B, the dependent variable is the log of income per capita. In panel C, the dependent variable is the share of municipal population with at least a high school degree. Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. Column 1 reports the baseline specification presented in column 4 of Tables 2, 3 and 4, respectively. Column 2 adds the distance to the nearest early migration epicenter or border municipality. Column 3 controls for a dummy indicating whether the municipality sent at least ten migrants in any year between 1882 and 1884. Column 4 controls for the nearest municipality with early exposure to migration-related information: log population in 1881, altitude, distance to the provincial capital, and an indicator for early emigration. Column 5 extends this approach by including the average characteristics of all municipalities within 100 km. Column 6 adds a control for WWI casualties per 1911 population. Standard errors, clustered at the province level, are reported in parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.4. Robustness Checks: Alternative Samples

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dep. Variable: Log Resident Population						
Historical Outmigration	-1.303***	-1.363***	-1.346***	-1.234**	-0.681*	-1.481
	(0.483)	(0.476)	(0.506)	(0.481)	(0.349)	(1.004)
Mean Dep. Var.	9,587	9,718	7,838	9,485	7,838	11,873
Panel B: Dep. Variable: Log Income per Capita						
Historical Outmigration	-0.246***	-0.257***	-0.252***	-0.232**	-0.254***	-0.194
	(0.089)	(0.090)	(0.094)	(0.092)	(0.083)	(0.124)
Mean Dep. Var.	12,656	12,637	12,589	12,657	14,505	14,021
Panel C: Dep. Variable: Share w/ at Least High School Degree						
Historical Outmigration	-0.072**	-0.076**	-0.075**	-0.062**	-0.072**	-0.012
	(0.030)	(0.032)	(0.032)	(0.030)	(0.032)	(0.060)
Mean Dep. Var.	0.272	0.272	0.272	0.272	0.279	0.265
KP F-stat	19.36	20.28	17.58	20.75	28.22	14.12
Observations	5,803	5,686	5,360	5,666	3,086	2,717
Specification	Baseline	Drop 1/99 pct.	No Factor	No Epicenter	North	Center/ South

Notes: The table replicates the 2SLS regressions presented in column 4 of Tables 2 to 4, using alternative samples. In panel A, the dependent variable is the log of municipality population. In panel B, the dependent variable is the log of income per capita. In panel C, the dependent variable is the share of municipal population with at least a high school degree. Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument is the distance to migration information source presented in Section 4.2. Column 1 reports the baseline specification presented in column 4 of Tables 2, 3 and 4, respectively. Column 2 excludes outliers by dropping municipalities in the bottom and top 1st and 99th percentiles of the outmigration distribution. Column 3 excludes municipalities that hosted an emigration factor. Column 4 excludes municipalities classified as either early migration epicenters or border municipalities. Columns 5 and 6 estimate the model separately for northern and central/southern municipalities. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.5. Robustness Checks: Alternative Specifications and Standard Errors

	(1)	(2)	(3)	(4)	(5)
Panel A: Dep. Variable: Log Resident Population					
Historical Outmigration	-1.303***	-1.436*	-0.727*	-0.722***	-1.303***
	(0.483)	(0.779)	(0.399)	(0.221)	(0.264)
Mean Dep. Var.	9,587	9,587	9,587	9,587	9,484
Panel B: Dep. Variable: Log Income per Capita					
Historical Outmigration	-0.246***	-0.369**	-0.262***	-0.205***	-0.246***
	(0.089)	(0.173)	(0.078)	(0.049)	(0.051)
Mean Dep. Var.	12,656	12,656	12,656	12,656	12,707
Panel C: Dep. Variable: Share w/ at Least High School Degree					
Historical Outmigration	-0.072**	-0.109*	-0.088***	-0.040*	-0.072***
	(0.030)	(0.056)	(0.029)	(0.023)	(0.016)
Mean Dep. Var.	0.272	0.272	0.272	0.272	0.273
KP F-stat	19.36	6.48	19.40	3.69	20.74
Observations	5,803	5,803	5,803	5,803	5,803
Specification	Baseline	Without pamphlets	Any factor	Split10	Conley SE

Notes: The table replicates the 2SLS regressions presented in column 4 of Tables 2 to 4, using alterative instrument and standard errors' definitions. In panel A, the dependent variable is the log of municipality population. In panel B, the dependent variable is the log of income per capita. In panel C, the dependent variable is the share of municipal population with at least a high school degree. Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. Column 1 reports the baseline specification presented in column 4 of Tables 2, 3 and 4, respectively. Column 2 excludes pamphlets from the set of migration incentives. Column 3 includes all available factors mentioned in historical sources, including references to past migrants and unspecified or unknown drivers. Column 4 transforms the distance variable into deciles of proximity to migration factors. Column 5 reproduces the baseline regression using Conley standard errors (spatial kernel 100 km). The Kleibergen-Paap F-statistic for the excluded instruments is reported in the last row of each panel. All regressions cluster standard errors at the province level (except in Column 5, where Conley SEs are reported). Standard errors are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.6. First Stage Robustness: Distance to Migration Information Source (10km bins)

Dep. Variable:	Historical Outmigration	
	(1)	(2)
Distance to Migration Information	-0.039*** (0.009)	
10–20 km		0.021 (0.025)
20–30 km		-0.041 (0.033)
30–40 km		-0.049 (0.040)
40–50 km		-0.170* (0.076)
50–60 km		-0.175* (0.072)
60–70 km		-0.161* (0.068)
70–80 km		-0.173* (0.077)
80–90 km		-0.307*** (0.066)
90–100 km		-0.311** (0.095)
100–110 km		-0.347** (0.101)
110–120 km		-0.450*** (0.111)
120–130 km		-0.408** (0.125)
130–140 km		-0.275 (0.166)
140–150 km		-0.372 (0.194)
More than 150 km		-0.457* (0.213)
Observations	5,803	5,803
KP F-stat	19.36	3.688

Notes: Column 1 reports the first stage regression presented in column 4 of Table 1. Column 2 reports the robustness exercise with 10km bins. Standard errors clustered at the province level in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table A.7. First Stage Robustness: Estimates by Five-Year Periods

Dep. Variable:	Historical Outmigration							
	Baseline	1884–89	1890–94	1895–99	1900–04	1905–09	1910–14	1915–20
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance to Migration Information	-0.039***	-0.050**	-0.052**	-0.036**	-0.069***	-0.057***	-0.073***	-0.036***
	(0.009)	(0.020)	(0.022)	(0.017)	(0.021)	(0.015)	(0.017)	(0.012)
Observations	5,803	5,803	5,803	5,803	5,803	5,803	5,803	5,803
KP F-stat	19.36	6.199	5.832	4.615	11.30	15.31	17.43	9.344

Notes: The table replicates the first stage regression presented in column 4 of Table 1, estimated separate year bins. Column 1 reports results for the full sample. All the other columns report results the years indicated at the top of the column. All regressions cluster standard errors at the province level. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.8. Effects of Historical Outmigration on Resident Population

Dep. Variable:	Log Resident Population											
	1871	1901	1911	1921	1931	1936	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: 2SLS												
Historical Outmigration	-0.054	-0.162**	-0.160	-0.280**	-0.451***	-0.542***	-0.696***	-0.817***	-0.956**	-1.047**	-1.171**	-1.285**
	(0.038)	(0.074)	(0.103)	(0.131)	(0.163)	(0.172)	(0.244)	(0.293)	(0.376)	(0.438)	(0.475)	(0.505)
KP F-stat	19.96	19.96	19.96	19.96	19.96	19.96	19.93	19.96	19.96	19.96	19.96	19.96
Panel B: OLS												
Historical Outmigration	-0.015***	0.023***	0.024***	0.017*	-0.024	-0.053**	-0.057**	-0.073***	-0.106***	-0.124***	-0.137***	-0.152***
	(0.005)	(0.007)	(0.007)	(0.010)	(0.018)	(0.023)	(0.025)	(0.027)	(0.035)	(0.040)	(0.044)	(0.047)
Mean Dep. Var.	4,618	5,685	6,177	6,633	6,903	7,119	7,982	8,505	9,099	9,514	9,554	9,587
Observations	5,803	5,803	5,803	5,803	5,803	5,803	5,802	5,803	5,803	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the log of municipal resident population in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.9. Effects of Historical Outmigration on Sex Ratio

Dep. Variable:	Sex Ratio							
	1911	1921	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 2SLS								
Historical Outmigration	-0.298***	-0.207***	-0.116**	-0.092**	-0.089*	-0.052	-0.038	-0.047
	(0.067)	(0.046)	(0.047)	(0.042)	(0.050)	(0.044)	(0.038)	(0.036)
KP F-stat	19.96	19.97	19.96	19.96	19.96	19.96	19.96	19.96
Panel B: OLS								
Historical Outmigration	-0.087***	-0.040***	-0.017***	-0.014**	-0.014***	-0.011***	-0.009***	-0.005**
	(0.015)	(0.012)	(0.006)	(0.005)	(0.005)	(0.003)	(0.002)	(0.002)
Mean Dep. Var.	0.941	0.953	0.985	0.987	0.983	0.969	0.963	0.969
Observations	5,803	5,802	5,803	5,803	5,803	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the municipality sex ratio (number of men relative to women) in the year at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. All specifications include province fixed effects, the log of 1881 population, geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, area), crop suitability indicators, distances to key access points (coastline, border, province capital, nearest main port), proximity to the closest railway station and line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.10. Effects of Historical Outmigration on Literacy Attainment

Dep. Variable:	Share of Literate vs. Illiterate Population							
	1911	1921	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 2SLS								
Historical Outmigration	-3.089	-4.978	-7.196	-1.147	-65.89	-62.37	-30.46	-86.15
	(3.143)	(4.220)	(12.04)	(20.61)	(44.65)	(46.50)	(56.62)	(62.91)
KP F-stat	19.92	19.93	20.03	20.19	20.70	20.41	20.66	21.08
Panel B: OLS								
Historical Outmigration	-0.538	-1.035	3.413	-3.886	-0.126	-9.782	-17.77**	-7.782
	(0.391)	(0.710)	(3.157)	(2.769)	(8.174)	(7.384)	(8.656)	(7.790)
Mean Dep. Var.	4.511	7.871	23.392	40.562	74.072	107.7	137.4	168.2
Observations	5,801	5,798	5,791	5,780	5,704	5,671	5,646	5,599

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the share of municipal literate population relative to municipal illiterate population in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.11. Effects of Historical Outmigration on Educational Attainment

Dep. Variable:	Share w/ at Least High School Degree					
	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2SLS						
Historical Outmigration	-0.006	-0.011	-0.012	-0.025	-0.043*	-0.073**
	(0.005)	(0.007)	(0.009)	(0.015)	(0.025)	(0.032)
KP F-stat	19.96	19.96	19.96	19.96	19.96	19.96
Panel B: OLS						
Historical Outmigration	0.000	-0.001	-0.001	-0.003*	-0.005*	-0.006*
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Mean Dep. Var.	0.020	0.026	0.048	0.092	0.158	0.272
Observations	5,803	5,803	5,803	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the share of municipal population with at least a high school degree in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.12. Effects of Historical Outmigration on Population Below Age of 6 Relative to Population 15–64

Dep. Variable:	Share of Resident Population Below Age of 6 Relative to Population aged 15–64					
	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2SLS						
Historical Outmigration	-0.060*	-0.064**	-0.057**	-0.031	-0.008	-0.001
	(0.034)	(0.031)	(0.027)	(0.019)	(0.011)	(0.007)
KP F-stat	19.93	19.99	19.86	19.82	20.77	19.99
Panel B: OLS						
Historical Outmigration	-0.005***	-0.008***	-0.007***	-0.002*	0.000	-0.002**
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Mean Dep. Var.	0.165	0.142	0.141	0.109	0.084	0.080
Observations	5,802	5,798	5,787	5,752	5,695	5,693

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the share of municipal population under the age of 6 relative to municipal population aged 15-64 in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.13. Effects of Historical Outmigration on Population Below Age of 6 Relative to Female Population

Dep. Variable:	Share of Resident Population Below Age of 6 Relative to Female Population							
	1911	1921	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 2SLS								
Historical Outmigration	0.074	0.070*	-0.081*	-0.080**	-0.071**	-0.046*	-0.020	-0.012
	(0.048)	(0.042)	(0.042)	(0.038)	(0.034)	(0.027)	(0.017)	(0.012)
KP F-stat	16.86	19.89	19.93	19.99	19.86	19.82	20.77	19.99
Panel B: OLS								
Historical Outmigration	0.053***	0.047***	-0.008***	-0.011***	-0.011***	-0.005***	-0.002*	-0.004***
	(0.009)	(0.010)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
Mean Dep. Var.	0.380	0.318	0.219	0.186	0.176	0.136	0.110	0.102
Observations	5,347	5,592	5,802	5,798	5,787	5,752	5,695	5,693

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the share of municipal population under the age of 6 relative to municipal female population in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table A.14. Effects of Historical Outmigration on Active Population

Dep. Variable:	Log Active Population						
	1936	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	-0.383***	-0.478**	-0.598**	-0.901**	-1.910**	-1.254**	-1.386**
	(0.164)	(0.208)	(0.256)	(0.370)	(0.407)	(0.505)	(0.545)
KP F-stat	19.97	19.93	19.96	19.96	19.96	19.96	19.96
Panel B: OLS							
Historical Outmigration	-0.059***	-0.060**	-0.078***	-0.133***	-0.123***	-0.168***	-0.179***
	(0.019)	(0.024)	(0.028)	(0.035)	(0.044)	(0.047)	(0.049)
Mean Dep. Var.	3,027	3,734	3,636	3,950	4,494	4,494	4,387
Observations	5,802	5,802	5,803	5,803	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the log of municipal active population in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.15. Effects of Historical Outmigration on Workers in Agriculture

Dep. Variable:	Logged Number of Workers in Agriculture						
	1936	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	-0.031	-0.130	-0.217	-0.554	-0.598	-0.721	-0.991
	(0.239)	(0.368)	(0.378)	(0.474)	(0.462)	(0.540)	(0.599)
KP F-stat	19.97	19.93	19.96	19.96	19.92	19.91	19.98
Panel B: OLS							
Historical Outmigration	-0.064**	-0.139***	-0.177***	-0.211***	-0.266***	-0.216***	-0.259***
	(0.026)	(0.046)	(0.061)	(0.052)	(0.045)	(0.052)	(0.056)
Mean Dep. Var.	1,465	1,569	1,063	690.6	379.0	266.3	214.1
Observations	5,802	5,802	5,803	5,798	5,798	5,791	5,789

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the log of municipal population employed in agriculture in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.16. Effects of Historical Outmigration on Workers in Manufacturing

Dep. Variable:	Logged Number of Workers in Manufacturing						
	1936	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	-0.876	-0.784	-0.831	-0.855	-0.751	-0.803	-0.924
	(0.612)	(0.643)	(0.533)	(0.537)	(0.570)	(0.596)	(0.614)
KP F-stat	20.05	19.96	19.99	19.96	19.96	19.96	19.96
Panel B: OLS							
Historical Outmigration	-0.010	0.029	0.011	-0.083**	-0.119***	-0.141***	-0.158***
	(0.058)	(0.034)	(0.029)	(0.035)	(0.037)	(0.040)	(0.043)
Mean Dep. Var.	889.8	1,210	1,480	1,749	1,351	1,299	1,299
Observations	5,794	5,799	5,800	5,803	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the log of municipal population employed in manufacturing in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.17. Effects of Historical Outmigration on Workers in Services

Dep. Variable:	Logged Number of Workers in Services						
	1936	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	-0.744**	-0.966**	-1.268***	-1.489***	-1.548***	-1.782***	-1.765***
	(0.303)	(0.393)	(0.413)	(0.472)	(0.526)	(0.575)	(0.588)
KP F-stat	19.97	19.93	19.96	19.99	19.96	19.96	19.96
Panel B: OLS							
Historical Outmigration	-0.061*	-0.049	-0.076**	-0.122***	-0.135***	-0.159***	-0.163***
	(0.032)	(0.032)	(0.034)	(0.045)	(0.049)	(0.054)	(0.054)
Mean Dep. Var.	672.7	955.6	1,095	1,512	1,669	2,113	2,351
Observations	5,802	5,801	5,803	5,802	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the log of municipal population employed in services in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.18. Effects of Historical Outmigration on Agricultural Employment Share

Dep. Variable:	Share of Employment in the Agricultural Sector						
	1936	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	0.181*	0.160	0.178	0.145	0.104	0.094	0.059
	(0.126)	(0.144)	(0.116)	(0.094)	(0.069)	(0.057)	(0.041)
KP F-stat	19.97	19.96	19.96	19.96	19.96	19.96	19.96
Panel B: OLS							
Historical Outmigration	-0.007	-0.026**	-0.027**	-0.018***	-0.015***	-0.010**	-0.010***
	(0.012)	(0.012)	(0.010)	(0.007)	(0.005)	(0.004)	(0.003)
Mean Dep. Var.	0.660	0.582	0.442	0.312	0.213	0.146	0.103
Observations	5,802	5,803	5,803	5,803	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the share of municipal employment in agriculture in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.19. Effects of Historical Outmigration on Manufacturing Employment Share

Dep. Variable:	Share of Employment in the Manufacturing Sector						
	1936	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	-0.135	-0.087	-0.045	0.027	0.108	0.160	0.162
	(0.118)	(0.135)	(0.125)	(0.103)	(0.095)	(0.101)	(0.101)
KP F-stat	20.05	19.96	19.96	19.96	19.96	19.96	19.96
Panel B: OLS							
Historical Outmigration	0.010	0.028***	0.031***	0.018***	0.011**	0.009	0.005
	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.007)	(0.007)
Mean Dep. Var.	0.214	0.272	0.371	0.433	0.421	0.390	0.377
Observations	5,794	5,803	5,803	5,803	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the share of municipal employment in manufacturing in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A.20. Effects of Historical Outmigration on Services Employment Share

Dep. Variable:	Share of Employment in the Service Sector						
	1936	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	-0.051*	-0.072*	-0.133***	-0.172**	-0.212**	-0.254**	-0.221**
	(0.029)	(0.038)	(0.049)	(0.068)	(0.088)	(0.101)	(0.098)
KP F-stat	19.97	19.96	19.96	19.96	19.96	19.96	19.96
Panel B: OLS							
Historical Outmigration	-0.003	-0.002	-0.004	0.000	0.004	0.001	0.005
	(0.003)	(0.004)	(0.005)	(0.006)	(0.007)	(0.008)	(0.007)
Mean Dep. Var.	0.126	0.146	0.187	0.255	0.366	0.464	0.520
Observations	5,802	5,803	5,803	5,803	5,803	5,803	5,803

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the share of municipal employment in services in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.21. Effects of Historical Outmigration on Number of Workers

Dep. Variable:	Log Number of Workers						
	1927	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	-0.696	-1.079**	-1.161**	-1.284**	-1.468***	-1.423**	-1.610**
	(0.433)	(0.446)	(0.465)	(0.506)	(0.549)	(0.549)	(0.654)
KP F-stat	19.93	20.25	20.06	20.06	20.07	20.07	20.07
Panel B: OLS							
Historical Outmigration	-0.078**	-0.139***	-0.152***	-0.184***	-0.182***	-0.213***	-0.235***
	(0.038)	(0.043)	(0.040)	(0.049)	(0.054)	(0.053)	(0.058)
Mean Dep. Var.	915.1	1,140	1,575	1,856	2,243	2,269	2,338
Observations	5,799	5,669	5,669	5,669	5,670	5,670	5,670

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the log number of municipal workers in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table A.22. Effects of Historical Outmigration on Number of Firms

Dep. Variable:	Log Number of Firms						
	1927	1951	1961	1971	1981	1991	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Historical Outmigration	-0.552*	-0.513*	-0.750**	-1.072***	-1.180***	-1.257***	-1.370***
	(0.281)	(0.263)	(0.319)	(0.379)	(0.426)	(0.450)	(0.504)
KP F-stat	19.93	20.25	20.06	20.06	20.07	20.07	20.07
Panel B: OLS							
Historical Outmigration	0.001	-0.076***	-0.086***	-0.108***	-0.116***	-0.146***	-0.173***
	(0.021)	(0.027)	(0.028)	(0.032)	(0.038)	(0.041)	(0.047)
Mean Dep. Var.	296.6	275.4	344.9	405.3	515.5	536.4	606.9
Observations	5,799	5,669	5,669	5,669	5,670	5,670	5,670

Notes: The table presents 2SLS and OLS estimates from equation (1) in Panels A and B, respectively. Each column reports results from a separate regression, where the dependent variable is the log number of municipal firms in the year indicated at the top of the column. The unit of observation is the municipality (see Section 3 and Appendix B for more details). Historical outmigration is the cumulated number of emigrants between 1884 and 1920, scaled by 1881 population. The instrument in Panel A is the distance to migration information source presented in Section 4.2. All specifications include province fixed effects, the log of 1881 population, a full set of geographic controls (altitude, slope, ruggedness, lakes and rivers, coordinates, and area), indicators of suitability for main crops, distances from the municipality to key access points (coastline, border, province capital, and the nearest main port), proximity to the closest railway station and railway line, weather shocks (temperature and precipitation deviations from long-run means), and exposure to violent earthquakes. KP F-stat is the Kleibergen-Paap F-stat for weak instruments. Standard errors, clustered at the province level, are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

B Online Appendix – Data Sources and Description

B.1 Unit of analysis

The administrative structure of Italian municipalities changed substantially over our study period. In 1871, there were 8,383 municipalities, compared to 8,101 in 2001. The number fluctuated, reaching a peak of 9,194 in the 1921 Census and a low of 7,311 in the 1931 Census. Over the almost 140 years we consider, municipalities were created, dissolved, merged, or divided, with associated changes in their territorial boundaries. Using SISTAT (Sistema Informativo Storico delle Amministrazioni Territoriali), we identify all municipalities involved in boundary changes. To the best of our knowledge, no historical dataset provides the exact share of territory affected by each change. ISTAT shapefiles allow reconstruction of municipal boundaries only from the 1991 Census onward.

To ensure a time-consistent unit of analysis, we construct *aggregated municipality*, defined as the largest set of municipalities that have ever exchanged territory with each other at any point during our period of interest. In the remainder of the paper, we refer to these simply as "municipalities," except in this section, where we use the term *aggregated municipality* for clarity.

All underlying data are first collected at the original (historical) municipality level and then aggregated to the *aggregated municipality.* For ratio variables (e.g., the share of the population over age 6 with at least a high school diploma), we compute population-weighted averages across constituent municipalities, where weights reflect each municipality's share of the total*aggregated municipality* population.

All geographical variables are computed using custom shapefiles based on *aggregated municipality* boundaries.

For fixed effects, we reconstruct historical provincial boundaries at different points in time. Our default specification uses provinces and regions as defined in 1881, but results are robust to using boundaries from other dates. When an *aggregated municipality* spans multiple provinces, we assign it to the province containing the largest share of its population in each year for which we reconstruct provincial borders.

Table B.1. Variable Dictionary: Main Variables (Table A.1)

Variable Name	Description	Source
Panel A: Dynamics		
Resident Population	Resident population	Istat - Population Census
Sex Ratio	Male-to-female population ratio	Istat - Population Census
Share of Literate/Illiterate	Literate-to-Illiterate population aged 6 and above ratio	Istat - Population Census
Share of High School Graduates	Share of the population aged 6 and above with at least a high school diploma	Istat - Population Census
Share of Population Below 6 (15-64)	Ratio of the population aged 6 or younger to the population aged 15–64	Istat - Population Census
Share of Population Below 6 (Female)	Ratio of the population aged 6 or younger to the female population	Istat - Population Census
Active Population	Economically active resident population	Istat - Population Census
Workers in Agriculture	Resident population employed in the agricultural sector[a]	Istat - Population Census
Workers in Manufacturing	Resident population employed in the manufacturing sector[a]	Istat - Population Census
Workers in Services	Resident population employed in the services sector[a]	Istat - Population Census
Agriculture Employment Share	Share of total employment in agriculture[a]	Istat - Population Census
Manufacturing Employment Share	Share of total employment in manufacturing[a]	Istat - Population Census
Services Employment Share	Share of total employment in services[a]	Istat - Population Census
Number of Workers	Employed persons in manufacturing by place of work	Istat - Manufacturing Census
Number of Firms	Number of manufacturing firms	Istat - Manufacturing Census
Panel B: Cross Section		
Historical Outmigration Rate	Ratio of total recorded migrants 1884-1920 to the 1881 population[b]	Statistica della Emigrazione Italiana per l'Estero
Distance to Migration Information Source	Straight-line distance from each municipality to the nearest municipality reporting the presence of a migration driver (agency, agent, pamphlet, or recruiter)	Statistica della Emigrazione Italiana per l'Estero
Income per Capita	Total income per capita	Ministry of Economy and Finance
Cultivated Area	Utilized agricultural area	Istat - Agricultural Census
Number of Farms	Number of agricultural farms	Istat - Agricultural Census
Number of Agricultural Workers	Employed persons in agriculture (primary occupation) by place of work	Istat - Agricultural Census
Number of Horses per Farm	Number of horses per farm	Istat - Agricultural Census
Number of Horses per Worker	Number of horses per agricultural workers	Istat - Agricultural Census
Agricultural Output	Total production (only wheat, rice, maize, other cereals)	Istat - Agricultural Census
Crop Yield	Total production over total cultivated land (only wheat, rice, maize, other cereals)	Istat - Agricultural Census
Land Concentration Index	Sum of squared land shares by farm size class (HHI)	Istat - Agricultural Census
In-Migration, Domestic	Number of recorded in-migrations from other Italian municipalities	Istat
In-Migration, International	Number of recorded in-migrations from abroad	Istat
In-Migration, No Degree	Number of recorded in-migrations, individuals without a College Degree	Istat
In-Migration, College Degree	Number of recorded in-migrations, individuals with a College Degree	Istat
Out-Migration, Domestic	Number of recorded out-migrations to other Italian municipalities	Istat
Out-Migration, International	Number of recorded out-migrations to abroad	Istat
Out-Migration, No Degree	Number of recorded out-migrations, individuals without a College Degree	Istat
Out-Migration, College Degree	Number of recorded out-migrations, individuals with a College Degree	Istat

Notes: [a] Until 1971 inclusive, ISTAT reports the number of active population by sector; from 1981 onwards, it reports the number of employed population by sector. Employment shares are calculated relative to the corresponding denominator, *i.e.*, until 1971 inclusive, relative to total active population; from 1981 onwards, relative to total employment.
[b] We could not retrieve data for 7 of the 37 years between 1884 and 1920 (1888, 1889, 1890, 1894, 1895, 1916, and 1917). For these years, we assume the outmigration rate equals the average outmigration rate over the entire period.

Table B.2. Data Description (Table A.2)

Variable Name	Description	Source
Altitude	Median altitude in the municipality	Nunn and Puga (2012)
Ruggedness Index	Median terrain ruggdness index (TRI) in the municipality	Nunn and Puga (2012)
Slope	Median slope in the municipality	Nunn and Puga (2012)
Number of Large Lakes	Number of large lakes	HydroSHEDS
Number of Large Rivers	Number of large rivers	HydroSHEDS
Longitude	Longitude	Istat
Latitude	Latitude	Istat
Area	Total area	Istat
Distance to Coastline	Straight-line distance to coastline	Istat
Distance to Border	Straight-line distance to national border	Istat
Distance to Provincial Capital	Straight-line distance to the closest 1881 Provincial Capital City	Istat
Distance to Port	Straight-line distance to the closest main port (Genova, Napoli or Palermo)	Istat
Suitability Index – Barley	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Barley	FAO
Suitability Index – Rye	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Rye	FAO
Suitability Index – Wheat	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Wheat	FAO
Suitability Index – Dry Rice	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Dry Rice	FAO
Suitability Index – Foxtail Millet	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Foxtail Millet	FAO
Suitability Index – Maize	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Maize	FAO
Suitability Index – Pearl Millet	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Pearl Millet	FAO
Suitability Index – Wet Rice	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Wet Rice	FAO
Suitability Index – White Potatoes	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - White Potatoes	FAO
Suitability Index – Sorghum	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Sorghum	FAO
Suitability Index – Sweet Potato	Rain-fed FAO's GAEZ suitability index, following Nunn and Qian (2011) - Sweet Potato	FAO
Mean Annual Precipitation	Average over 1884–1920 of the ratio between annual precipitation and the period's long-term average precipitation	Luterbacher et al. (2004)
Mean Annual Temperature	Average over 1884–1920 of the ratio between annual temperature and the period's long-term average temperature	Pauling et al. (2006)
Mean Number of Violent Earthquakes	Average over 1884-1920 of the number of earthquakes within 100 km with Mercalli intensity greater than or equal to 7	Rovida et al. (2019)
Mean Distance to Railway Line	Average over 1884-1920 of the straight-line distance to the closest railway line	Sviluppo delle ferrovie italiane dal 1839 al 31 dicembre 1926
Mean Distance to Railway Station	Average over 1884-1920 of the straight-line distance to the closest railway station	Sviluppo delle ferrovie italiane dal 1839 al 31 dicembre 1926