On the same boat

Three essays on the causes of pre-WWI Italian international migration

A thesis submitted to the University of Dublin, Trinity College
in application for the degree of Doctor of Philosophy in Economics

Trinity College Dublin
Project funded by the Irish Research Council

Candidate*: Gaspare Tortorici*
Supervisor*: Prof Gaia Narciso*

March 2019

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I declare that I was the lead author for each of the co-authored chapters in the thesis.

Open access electronic theses are freely available over the World Wide Web for users to read, copy, download, and distribute subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement. Withheld access will be implemented, as applicable, according to the terms in Calendar Part III section 1.38.14.

A digital repository is an online, searchable, web-accessible database containing works of research deposited by scholars, the purpose of which is both increased access to scholarship and long-term preservation.

Gaspare Tortorici

Summary

This thesis analyzes the causes of Italian mass international migration before the outbreak of WWI. While each chapter is a stand-alone paper that explores different motives and dynamics, there is a common quantitative core based on Ellis Island administrative records (1892-1912) and official Italian out-migration statistics (1876-1912).

Chapter I is centered on the reconstruction of US-bound out-migration flows from Italian municipalities between 1892 and 1912. The main findings can be summarized as follows. First, international migration followed a process of path-dependency and progressive spatial diffusion at the local level. Second, the further away from a port a given municipality of origin was, the more passengers tended to travel together therefore creating ex-ante networks. Third, cyclical migration was a sizable and widespread phenomenon.

Chapter II - joint with Prof Rowena Gray and Prof Gaia Narciso - investigates the local migration response to international agricultural commodity price shocks during the first globalization era. The paper revolves around the idea that price shocks affect different areas according to their agricultural production structure. We conduct a Fixed-Effects analysis at the province level and find that higher agricultural commodity prices released liquidity constraints thus promoting out-migration. On the other hand, higher volatility translated itself into higher incentives to out-migrate. Results are statistically significant across a battery of robustness checks.

Chapter III - joint with Prof Yannay Spitzer and Ariell Zimran - studies the local migration response to the 1908 Messina and Reggio Calabria earthquake, one of the deadliest and most disruptive natural disasters in European history. The paper is based on a Difference-in-Difference approach coupled with a series of event studies that allow to test the pre-trends hypothesis. We find that the quake only produced a short-lived decline of out-migration rates from affected municipalities - as opposed to non-affected ones - in Calabria. We test two

mechanisms: migration networks and damage accumulation. In the former case, we find that there is no statistically significant effect of networks; in the latter, we show how migration decreased more in municipality which had been affected by another earthquake in the recent past.

Ackowledgments

I thank the Irish Research Council for having funded this project,
the administrative staff at Trinity College Dublin for their helpfulness,
all who shared their ideas with me and challenged my beliefs,
all the students that I had the chance to teach to,
all those who encouraged me and were kind enough to listen,
my co-authors for having given me the opportunity to grow intellectually,
my real friends, in Dublin and abroad,
my supervisor, Prof Gaia Narciso, for her gentle guidance, trust and positiveness,
my family for having supported me anytime, anyhow, anyway.

To the ones who stood by my side
and those who still do

The pure products of America
go crazy
. . .

WILLIAM CARLOS WILLIAMS
To Elsie, 1923

Table of Contents

Introduction

He must have been five when he first saw a classical guitar. For years, he kept asking his father to get him one at the local flea market. But the man was reluctant: there was no need for musicians in his family. He agreed at last, surrendering to the resourceful insistence of the boy and the decent bargain offered by the merchant. The child was pretty skilled at school and had no intention to work as a laborer in the family-owned tile industry. As time went by, the young man became known as *Canta-la-notte* – the one who sings at night – a specialized guitarist who got by playing serenades to young women chased by penniless and often unrequited lovers. Despite the abundance of clients, he struggled to make ends meet in the Sicilian village he hailed from. By the time he turned 27, he decided that those dog days had to stop, took a train to Palermo and set sail toward the United States on a steamship named after the queen of Italy. He set foot on Ellis Island about ten days afterward – on September 1st 1907 – and never went back. He spoke no English, had no networks to rely on, no pre-paid ticket and next-to-no money. He never sent remittances, letters or paid any visits. He had just disappeared, swallowed whole by the New World.[1]

About 55 million individuals left Europe between the early 1820s and the outbreak of the First World War.[2] Favored by technological progress in steamship,[3] this continuous ebb and flow of people changed the traits of societies and labor markets across the Oceans. Destination-wise, the United States absorbed more than 30 million people while the remaining two fifths of the wave opted for Latina America and other overseas destinations. While major origin countries varied over time, this event was endemic to the Old Continent.[4] During the first decades of the 19th century, migrants from the British Isles accounted for the lion's share

[1] I am grateful to Joan Sewall, the granddaughter of *Canta-la-notte* – Giuseppe Tortorici according to the registry office of Pietraperzia (Enna), a first cousin of my grandfather – for having shared bits and pieces of her past with me.

[2] This paragraph is mostly based on the first two chapters of Hatton and Williamson (1998).

[3] Steamships substituted sailing ships as major transportation means since the early 1850s, determining a rapid decline of fare prices.

[4] France is a notable exception, possibly because it experienced an early fertility transition accompanied by declining death rates before any other European country.

of immigration to the United States. During the 1850s German laborers flocked overseas, followed by Scandinavians. Migration flows from less developed areas in Southern and Easter Europe initially remained limited and eventually incremented drastically from the 1890s - with an estimated 1.5 million Russian Jews leaving the Pale of Settlement only (Spitzer, 2016). An estimated 11 million individuals emigrated from Italy between the 1870s and WWI, temporarily or permanently.[5]

Why did all these people move? Are there any common variables driving them? What does explain the timing of each wave and its volatility? Why did not poor peripheral economies go with the flow earlier? These questions – alongside many others – were already of primal interest to contemporary policy makers[6] and certainly constitute a focal issue in the current debate on international migration. Drawing a parallel with the present, a contemporary observer might wonder what sort of policies were in place if so many individuals were allowed to resettle. Indeed, this era was characterized by the implementation of virtually open-borders and unrestrictive immigration laws.[7] It is therefore a particularly valuable historical prior that allows to analyze the dynamics of migration in the absence of legal constraints – for the vast majority of immigrants were driven by economic motives.[8]

The story I open the introduction with is clearly a splinter among others and yet it is interesting to pose a question: how representative is it? Ginzburg (1978) argued that the reconstruction of micro-histories could offer precious insights on general norms. The strand of literature initiated by the Italian historian rehabilitate the value of details concealed by aggregate depictions while recognizing the necessity of a broader account that might serve as a frame of reference. It is noteworthy how, on a different scale, the recent economic literature

[5]Note that both cyclical and return migration were very widespread. See Bandiera, Rasul and Viarengo (2013) for more details on return migration flows from Ellis Island.

[6]See, for instance, the reports produced by the Dillingham Commission in 1911.

[7]The Unites States implemented a virtually open-border policy until the introduction of the 1917, 1921 and 1924 Immigration Acts (Goldin, 1994). Canada introduced a literacy test in 1919. Brazil kept its border open until WWI; so did Argentina - with minor occasional limitations (Sánchez-Alonso, 2013).

[8]There were refugees, too. O Grada and O'Rourke (1997) and Collins and Zimran (2018) - among others - focus on the Great Irish Famine of the 1840s.

on international migration during the first globalization era has advocated a similar shift toward finely grained analyses that could account for the extraordinary heterogeneity of this experience across geographical areas.

The literature on overseas Italian out-migration is immense. It has accrued and stratified over time since contemporaries noticed the relevance of the event. Foerster (1919) minutely pinpointed the traits of the phenomenon ten years before Ferenczi and Willcox (1929) quantitatively documented the magnitude and evolution of flows. On the European side of the Atlantic, hundreds of social scientists have written extensively on the topic for decades - see for example the early contributions of Sori (1979) or the more recent overarching contribution of Fauri (2015) and Bevilacqua (2001). How does this project contribute to the debate? Taken as a whole, these essays constitute a long-run analysis of the causes that motivated Italian international migration from the early 1880s to the outbreak of World War I. In doing so, I adopt a systematic approach that hinges on two key concepts: spatial dis-aggregation and quantitative evidence based on recently digitized large demographic collections.

Why is it so crucial to analyze the causes of international out-migration at the local level? Recent contributions have highlighted how Italian migration patterns are extraordinarily heterogeneous along multiple dimensions. Out-migration rates differed widely across provinces, so did timing and self-selection patterns (Spitzer and Zimran, 2018). This issue translates itself into a severe aggregation bias whenever the analysis is conducted at the national level.[9] Only a more spatially grained approach could mitigate these concerns. Given that Italy as a whole had been exposed to the same international environment,[10] differences in migration rates are likely to be rooted in intrinsic features of each geographical area.

While each chapter of this thesis tests different hypotheses, there is a common quantitative

[9]This point has been made, among others, by Faini and Venturini (1994) and Gomellini and O'Grada (2013).

[10]This is, to a certain extent, a simplification. Geography and access to different international labor markets were not the same everywhere.

core throughout it. Most of the descriptive and econometric analyses rely on a recently digitized archive: Ellis Island administrative records. This data set contains individual-level information on all the passengers that entered the United States through the port of New York between 1892 and 1924. I restrict the limelight to pre-WWI Italian entries, a grand total of about 3.5 million observations. The characterizing feature of this collection is the availability of migrants' municipalities of origin. However, the deluge of spelling mistakes – largely due to the stratification of transcription and digitization errors – prevented me and my co-authors from immediately employing these data. In order to overcome this issue, I devoted a significant effort to correctly identifying place names. In particular, I purposely developed a matching script – all the details are in Section 2.1, Chapter 1 – that eventually allows to retrieve more than half the available information. Then I simply tally how many passengers hailed from a certain municipality, per year. The result of is a rich panel data set that traces local migration patterns from 1892 to 1912.

The first chapter of the thesis represents a descriptive account of unexplored migration patterns and mechanisms. Entirely centered on Ellis Island administrative records, the paper makes its major contribution to the literature through the reconstruction of municipal US-bound migration flows. It can be broken down into three main arguments.

First, it confirms that US-bound migration from Italy was indeed an extremely heterogeneous phenomenon, even when considering a relatively confined geographical unit such as a province. There is evidence of migration rates swinging widely across municipalities. Recognizing such variability represents a stepping stone toward a subtler analysis of, for example, networks: while economic conditions might be relatively homogeneous province-wise, migration networks might only be active at a sub-provincial – certainly sub-national – level. When considering the temporal dimension of such flows, path-dependency and convergence seem to have played a central role – in line with the diffusionist approach pioneered by Gould (1979). For instance, in the case of Sicily, there was a sustained flow of passengers hailing from municipalities

where US-bound out-migration was already high before the turn of the century. As migration spread across the island during the 1900s, other areas caught-up rate-wise.[11]

Second, I underline the relevance of *ex-ante* migration networks. This concept has received little attention in the literature - see Battiston (2018) for an exception. Ex-ante networks are simply defined as networks that might arise before passengers set foot on Ellis Island. Within the context of an underdeveloped transportation system, I find that migrants that were located further away from ports had an intrinsic incentive to travel with their fellow citizens. Subsequently, migration networks in the United States should be thought as the continuation of a process that already started before the actual departure of migrants. This is not to say that migrants had no interactions at destination or during the transoceanic voyage; rather, it is likely that *ex-ante* networks interacted with other forms of networks, influencing location and occupational choices.

Third, the paper highlights how cyclical migration was quite common, accounting for up to 20 per cent of the flow from certain municipalities. In order to grasp the magnitude of the event I count how many times a given individual appears in the archive. I opt for a *very* conservative matching approach with no buffers: two records pertain to the same individual if and only if all the information utilized to identify them did not change over time – name, surname, municipality of origin, year of birth, and marital status. My estimates are therefore a lower bound.[12]

The second chapter, co-authored with Prof Rowena Gray and Prof Gaia Narciso, explicitly links Italian international migration to the parallel integration of agricultural commodity markets across the Atlantic Ocean. In order to investigate the elasticity of out-migration to price shocks, we exploit exogenous international price fluctuations across agricultural commodities and province-wise agricultural commodity mix heterogeneity. We construct

[11] Further research should test whether convergence is inevitable or there are certain factors that trigger it.

[12] Unfortunately, I do not know whether passengers eventually stayed in the United States or opted to return home.

an index of Global Price Exposure (GPE) that could be thought as a Bartik, shift-share, measure.[13] Our study highlights how international out-migration from Italian provinces positively responded to higher agricultural commodity prices and suggests that negative price shocks tightened liquidity constraints thereby reducing migration out-flows.

The third and last chapter, co-authored with Prof Yannay Spitzer and Prof Ariell Zimran, evaluates the impact of a severe natural disaster on municipality-level international migration from Calabria and Sicily. These areas were hit by one of the deadliest earthquakes in modern European history on December 28th 1908 – it claimed more than 100,000 lives and destroyed the two major commercial hubs of the area as well as several neighboring villages. The setting is very peculiar for two sets of reasons: on one hand, out-migration from both Calabria and Sicily had already morphed into a mass phenomenon; on the other, there was a polarization of destination choices toward the United States – which, as already pointed out, implemented open-border policies. These conditions allow to investigate the international migration response to a severe natural disaster within a framework that could inform modern debates on refugee crises.

According to our DD estimations and event studies, the effect of the earthquake is surprisingly limited and differs in Calabria and Sicily. In the former region, there is evidence of a short-lived, statistically significant, negative effect on migration rates followed by a rapid recovery. As far as Sicily is concerned, we fail to reject the null hypothesis that migration rates were not affected by the earthquake. We explore two possible mechanisms. First, we study the effect of networks – defined as the cumulative stock at destination before 1904 – and find no statistically significant evidence of them having a major role. We interpret this result arguing that most individuals had already some sort of connections in the US by the time the earthquake took place, suggesting that migration rates became less sensitive

[13]Indeed, it is a weighted average of price changes across agricultural commodities that assigns more weight to those commodities that capture a larger share of initial land in each province. Despite every geographical unit is exposed to the same price vector, differences in commodity mix determine a differential out-migration response.

to networks over time. Second, we exploit the fact that Calabria was hit by two successive earthquakes in 1905 and 1907 while Sicily did not suffer from either. In other words, we test for a damage-accumulation hypothesis and find that the migration response to the 1908 earthquake was more negative whenever the municipality was hit by a natural disaster in the recent past.

Overall, the arguments developed by this project focus on what the literature would define as *push* rather than *pull* variables. The accent is put on how migration choices varied across Italian geographical areas while the country as a whole underwent the first globalization era of both labor and commodity markets. While significantly adding to the literature, these essays are by no means the final word on the subject. Rather, they should be considered as a stepping stone toward a more comprehensive treatment of determinants. Future research should examine international migration flows directed toward other major destinations separately – e.g. Argentina and Brazil – and shed light on within-province variation in terms of timing and preferences for certain countries, as well as self-selection. Similarly, this project is silent about the economic consequences of mass migration on both origin and destination countries: a comprehensive temporal analysis of causes cannot abstract from it as changing economic conditions might also modify incentives in a feedback loop.

I conclude this brief introduction drawing on Gallo (2012). From a historical perspective, it is useful to underline how Italian international migration did not materialize unexpectedly: Italy – as much as other European countries – had a pre-unification tradition of short-lived, temporary movements within confined geographical areas that lasted for weeks or months at the very maximum. By their very own nature, these restless drifts were not measurable - there was not a technical change of residence and laborers always returned - but represented a common income-smoothing strategy for as many as 200 thousands individuals who moved across the peninsula each year.[14] Farmers from the Alps descended toward the Po plain during

[14]Note that I am not referring to internal migration thought as a temporary or permanent resettlement over a longer time horizon.

rice harvest seasons; so did Central Italian agricultural laborers who reached Aplulia and Latium to harvest wheat every coming year; likewise, individuals from Calabria streamed to Sicily during the harvest seasons of citrus. It seems to me that there is a continuum between these flows and the international migration of the 20th century: technological progress, wage differentials and employment opportunities on the other side of the Ocean only stretched the radius and temporal dimension of these purely economic, perennial, strategies.

In essence, this project revolves around the belief that knowing *what* determined international out-migration in the past is crucial to understand *how* policies could be shaped today or, at the very least, *how* this issue could be framed.

References

Bandiera, Oriana, Imran Rasul and Martina Viarengo. 2013. "The making of modern America: Migratory flows in the age of mass migration." *Journal of Development Economics* 102:23–47.

Battiston, Diego. 2018. "The Persistent Effects of Brief Interactions: Evidence from Immigrant Ships." *Mimeo, London School of Economics* .

Bevilacqua, Piero. 2001. *Storia dell'emigrazione italiana*. Donzelli Editore.

Collins, William J. and Ariell Zimran. 2018. "The Economic Assimilation of Refugees: Evidence from Irish Famine Immigrants to the United States." *Manuscript, Vanderbilt University* .

Faini, R and A Venturini. 1994. Italian emigration in the pre-war period. In *Migration and the international labor market, 1850-1939*, ed. Timothy J Hatton and Jeffrey G Williamson. Taylor and Francis pp. 72–90.

Fauri, Francesca. 2015. *Storia economica delle migrazioni italiane*. Il mulino.

Ferenczi, I and Walter F Willcox. 1929. "International Migrations." *NBER* .

Foerster, Robert Franz. 1919. *The Italian emigration of our times*. Vol. 20 Harvard University Press.

Gallo, Stefano. 2012. *Senza attraversare le frontiere: le migrazioni interne dall'unità a oggi*. Gius. Laterza & Figli Spa.

Ginzburg, Carlo. 1978. "Spie, Radici di un paradigma scientifico." *Rivista di storia contemporanea* 7(1):1.

Goldin, Claudia. 1994. The political economy of immigration restriction in the United States, 1890 to 1921. In *The regulated economy: A historical approach to political economy*. University of Chicago Press pp. 223–258.

Gomellini, Matteo and Cormac O'Grada. 2013. *The Oxford Handbook of the Italian Economy Since Unification*. Oxford University Press, New York chapter Migrations.

Gould, John D. 1979. "European inter-continental emigration 1815-1914: patterns and causes." *Journal of European Economic History* 8(3):593.

Hatton, Timothy J and Jeffrey G Williamson. 1998. *The age of mass migration: Causes and economic impact*. Oxford University Press on Demand.

O Grada, Cormac and Kevin H O'Rourke. 1997. "Migration as disaster relief: Lessons from the Great Irish Famine." *European Review of Economic History* 1(1):3–25.

Sánchez-Alonso, Blanca. 2013. "Making sense of immigration policy: A rgentina, 1870–1930." *The Economic History Review* 66(2):601–627.

Sori, Ercole. 1979. *L'emigrazione italiana dall'Unità alla seconda guerra mondiale*. Vol. 188 Il mulino.

Spitzer, Yannay. 2016. "Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914." *Manuscript* The Hebrew University of Jerusalem.

Spitzer, Yannay and Ariell Zimran. 2018. "Migrant Self-Selection: Anthropometric Evidence from the Mass Migration of Italians to the United States, 1907-1925." *Journal of Development Economics* .

Chapter 1

Diffusion, Networks and Cycles

US-bound Italian migration (1892-1912)

Gaspare Tortorici [*]

March 2019

Abstract

This paper investigates the municipal dimension of Italian mass migration to the United States at the turn of the 20th century (1892 - 1912). The analysis revolves around Ellis Island administrative records and comprises the universe of Italian immigrants entering the United States through the port of New York - a rough total of 3.5 million observations. The contribution is threefold: it disaggregates US-bound out-migration from Italy in its municipal dimension; it sheds light on the concept of ex-ante migration networks, evidencing their role within migration strategies; it engages with the emerging strand of literature on repeat migration, highlighting its relevance in the past. As a whole, the paper further characterizes the age of mass migration as a globalized era and documents a relevant degree of mobility that contrast with traditional accounts of transoceanic migrations.

Keywords: *Age of mass migration, migration networks, temporary migration*
JEL Codes: F22, J61, O15

[*]Department of Economics, Trinity College Dublin (tortorig@tcd.ie). I am grateful to Prof Gaia Narciso for her guidance and support. For very helpful comments and discussions, I thank Giacomo Gabbuti, Andrea Guariso, Martina Kirchberger, Ronan Lyons, Bruno Morando, Cormac Ò Gràda, Hillel Rapoport, Davide Romelli, Marvin Seusse, Yannay Spitzer, Ariell Zimran, and seminar speakers at Trinity College Dublin and Paris School of Economics. I acknowledge funding from the Irish Research Council.

1.1 Introduction

Pre-WWI Italian overseas out-migration represents one of the largest instances of voluntary population movements in history: more than 10 million individuals left the Kingdom of Italy between the early 1870s and the late 1910s either temporarily or permanently (ISTAT, 2011).

Such flows are framed within the broader context of the age of mass migration,[1] an era that largely overlaps with the first globalization. Favored by technological improvements in shipping and virtually no restrictions to international migration, both commodity and labor markets across the Atlantic Ocean displayed an unprecedented degree of interdependence. Indeed, major destination countries enforced permissive immigration policies until the outbreak of WWI.[2]

Despite having affected the entire country, Italian international out-migration shows an extraordinary degree of heterogeneity across several dimensions: timing, rates and destination choices largely varied province-wise (Fauri, 2015). Between the 1870s and the 1890s, the lion's share of Italian out-migration arose from northern regions (Veneto, Friuli Venezia-Giulia, Piedmont and Lombardy) and was directed toward Europe (Austro-Hungarian Empire and France mainly) and Latin America (Argentina, Brazil and Uruguay in this order). At the turn of the 20th century the picture reverted and overseas out-migration from southern Italy boomed, displaying a marked polarization of preferences for the United States.[3] Spitzer and Zimran (2018) argue that - while negative on aggregate - the degree of local self-selection into migration differed, with US-bound passengers from the south being more positively drawn

[1] Starting from the 1840s, all major European countries - with the notable exception of France - undergone processes of mass out-migration. These flows were mostly absorbed by the United States and, in the case of Portugal and Spain, Latin America. See Baines (2004) for a concise, yet comprehensive, account.

[2] The United States introduced a series of progressively stricter restrictions (1917, 1921 and 1924 Immigration Acts) with the intent to control in-bound flows in terms of quantity, quality and origin. See Goldin (1994) for a comprehensive treatment of the political motives that brought this change about. As far as Asian immigration is concerned, the US enforced the Chinese Exclusion Act in 1888. Argentina adopted a similar approach while Brazil subsidized *colono* contracts until the practice was discouraged by Italian law in 1902. See Sánchez-Alonso (2013) for an overview of immigration policies toward major Latin American countries.

[3] Southern and eastern European international out-migration has often been deemed as late - in the sense that it did not develop into a mass event before the 1900s even though both traditional push and pull factors were already apparent, i.e. wage differentials and dire economic conditions for agricultural laborers (Hatton and Williamson, 2005).

from their origin populations.

Despite having been hampered by data availability, the idea that Italian international migration should be investigated through a magnifying glass is not new in the literature: early contributions already questioned the reliability of national-level migration narratives, arguing that they might be subject to aggregation bias (Foerster, 1919); more recent studies echoed this insight and developed both regional and provincial analyses.[4] For example, Chapter 2 provides a provincial analysis of out-migration determinants that exploits exogenous fluctuations of international agricultural prices (O'Rourke, 1997).[5]

I use a recently digitized large demographic collection, Ellis Island administrative records, to reconstruct Italian US-bound municipal emigration flows from 1892 to 1912. The original data contain information on virtually all passengers that entered the United States through the port of New York, then the major immigration hub in the country. After having identified Italian passengers according to their ethnicity, I focus on those who arrived in the US before the outbreak of WWI - an approximate total of 3.5 million individual-level observations. Although the characterizing feature of this archive is the availability of each passenger's last residence, spelling mistakes are extremely common. In order to overcome this issue, I used a combination of automatic and manual matching that required a significant effort - detailed in Section 1.2. The resulting data set contains information on US-bound absolute migration figures from 5,923 municipalities, 110 provinces and 20 regions.[6]

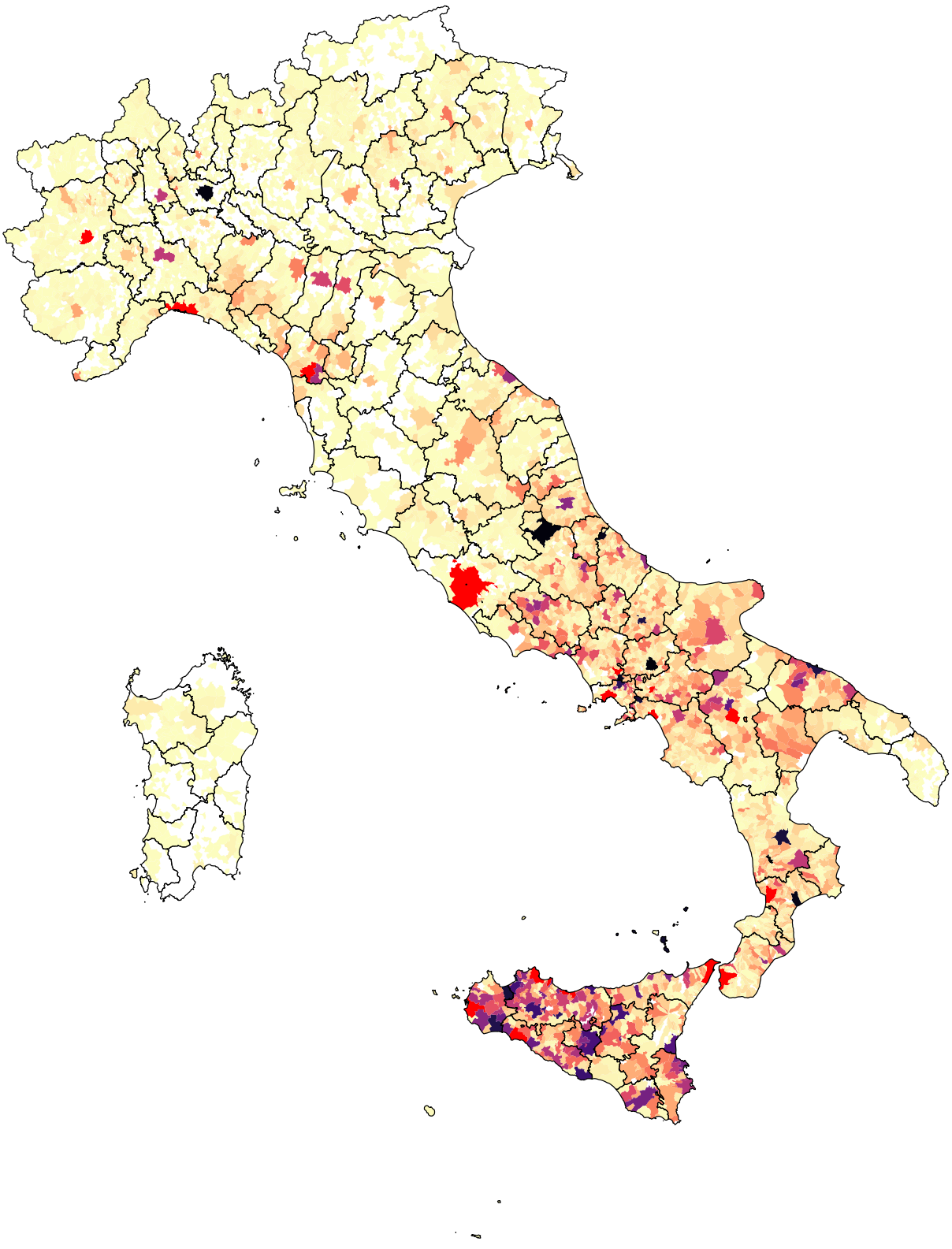
Figure 1.1[7] - in the following page - displays the major contribution of the data set and highlights the extraordinary degree of variability in US-bound Italian out-migration. Figures

[4]See also, among others, Faini and Venturini (1994) or Ardeni and Gentili (2014). This approach is likely to be crucial in other contexts too. For instance, Ò Gràda and O'Rourke (1997) shows that Irish post-famine out-migration rates varied substantially across counties, with poorer ones having a worse death to out-migration ratio because of binding liquidity constraints.

[5]See Gomellini and Ò Gràda (2013) for a regional analysis of out-migration determinants and Hatton and Williamson (1998) for a provincial study based on 1901 and 1911 censuses.

[6]Municipalities, provinces and regions are defined using current borders. Municipal data will be made available through my personal website after publication. Provincial and regional absolute migration figures from 1901 can already be found in tables A1.1 and A1.2.

[7]Footnotes have been omitted in the interest of readability. I report them here. *Source*: Ellis Island administrative records. *Notes*: Red polygons denote aggregate migration greater than 6,000 individuals.


Thousands of passengers
1
2
3
4
5

A1.1 and A1.2 show, respectively, regional and provincial US-bound aggregate numbers.

It is interesting to note that not all Italian municipalities were affected by the phenomenon. Despite attrition might represent an issue, only 5,923 over 8,104 municipalities sent at least a migrant to the US through Ellis Island. Why did Italian out-migration to the United States exhibit such degree of heterogeneity across municipalities? At a glance, the map seems to reveal a correlation between altitude and out-migration: indeed, rates appear to be comparatively higher along the Apennines than in surrounding, hilly or coastal, areas. For example, there is anecdotal evidence suggesting that out-migration from Tuscany followed topographic contour lines (Commissariato dell'emigrazione, 1902). The map also shows that certain provinces of the south – e.g. Lecce and Brindisi in Apulia – did not experience a significant US-bound out-migration flow while being exposed to the same economic environment that produced a mass movement of people in neighboring areas.

In this paper, I provide descriptive evidence of a diffusion process based on path-dependency, ex-ante migration networks and repeat migration flows. I find that migrants located further away from ports tended to travel more with their fellow-citizens. While traveling together might have sheer economic advantages,[8] the fact that propensity increased with distance from ports might explain why several inland municipalities (e.g. in Abruzzo or Calabria) exhibited a distinct polarization of destination choices. Similarly, Karadja and Prawitz (forth.) show that during the initial wave of Swedish US-bound emigration, areas located further away from ports displayed lower out-migration because they faced higher transportation costs.

My analysis also shows that the ratio between repeat and regular migration is comparatively higher where out-migration was more widespread in the first place. Given that the vast majority of these municipalities was located in poorer areas, this result suggests that repeat migrants were able to overcome liquidity constraints through either a successful stay in the US or a multi-stage migration strategy. *Golondrinas* who arrived in the US for the first time were on average about 24.3 years old, 9 months younger than regular migrants. However,

[8] According to the *Archivio Centrale dello Stato*, groups including more than 10 people benefited from a discount when using the national railway system.

both repeat and regular migrants exhibited similar time preferences because they both tended to leave Italy during the same months and years.

More generally, this paper engages with the flourishing literature on the age of mass migration and puts forward a new perspective on the Italian case, highlighting less known elements and stressing the role of international mobility.

Several authors have developed other country-specific analyses and examined this phenomenon from different standpoints. In the context of Norway, for instance, Abramitzky, Boustan and Eriksson (2012) find that wealth and the propensity to migrate were negatively associated; similarly, self-selection from urban areas was mainly negative and returns to migration lower than previously thought (Abramitzky, Boustan and Eriksson, 2013). In a recent contribution, Eriksson (2018) highlights how Norwegians living in larger enclaves had lower occupational earnings while being more likely to be employed in the primary sector.

Spitzer (2015, 2016) analyzes Jewish migration from the Pale of Settlement at the local level, highlighting a process of spatial diffusion. Initially clustered in confined areas, migration spread and accelerated thanks to networks. While pogroms did contribute to increase US-bound out-migration rates, they did not determine the initial migration wave from the Pale of Settlement. In its initial stages, out-migration was therefore largely interpretable as a continuation of previous trends attributable to economic factors.

The rest of the paper is divided as follows: section 1.2 describes the matching process and the relation between Ellis Island data and other sources commonly used in the context of Italian international mass migration before WWI; section 1.3 provides an overview of the main descriptives; section 1.4 describes migration networks ex-ante; section 1.5 is centered on repeat migration flows while section 1.6 offers some conclusive remarks.

1.2 Data

Following the Steerage Act (1819), first officials of vessels arriving in the United States were legally required to produce a list of passengers on board - and held responsible for

any individual entering the country illegally.[9] Historically, such obligations allowed for the constant recording of migration flows through US ports. This data set traces original manuscripts and comprises the universe of Italian immigrants setting foot on Ellis Island between 1892 and 1912, a rough total 3.5 million records.[10] Individuals are indexed according to their name, surname, gender, date and port of departure, vessel denomination, marital status, and the last municipality in which they permanently resided before embarking on the voyage toward Ellis Island.[11]

Despite having been utilized in recent analyses on the age of mass migration[12] the municipal dimension of Ellis Island administrative records has remained untapped. Indeed, there are several layers of errors that prevent researchers from immediately and univocally identifying these municipalities. On one hand, the recent digitization of ship manifests inevitably entangled a degree of approximation whenever hand-written documents were illegible, resulting in omissions and improper compilations. On the other, linguistic barriers between migrants and port officials at Ellis Island led to all sorts of inaccuracies despite US authorities hired Italian personnel following the intensification of Italian immigration. The fact that an overwhelming majority of passengers were illiterate and that passports were only made compulsory after 1901 contributed to decrease record precision. Extra, missing, and misplaced characters are therefore extremely common. I partially overcome such limitations through an *ad hoc* string-matching script.

1.2.1 Matching

Large demographic collections lie at the core of the recent developments in the economic history literature. However, manipulating *big data* requires a significant reconstruction effort

[9] Figure A1.3 in the appendix shows how actual Ellis Island manifests looked like.

[10] Data is available until 1924. I accessed the archive at http://www.jewishgen.org, through a purposely designed Python script that relies on the interface designed and provided by Steve Morse, creator of the Intel 8086 microprocessor chip.

[11] I cannot account for immigrants entering the United States through other ports - e.g. Philadelphia, Baltimore, Boston, New Orleans. However, by the 1890s more than four fifths of nation-wide immigrants transited through Ellis Island. Previously a depot, Ellis Island began serving as the main Federal Immigration center on January 1st 1892, after the closure of Castle Garden.

[12] See, for example, Bandiera, Rasul and Viarengo (2013).

as volume and *fuzziness* are often directly proportional to each other. Several authors have therefore developed different methodologies to deal with such complexity. A particularly interesting case revolves around the idea of *matching* - which is generally understood as the process of linking individuals across different data sets.[13] Having substantial implications on sample representativeness and inference, record linking is non-neutral: match, false match (Type I error) and missed match (Type II error) rates are largely methodology-driven. In a pivotal paper, Bailey et al. (2017) argue that unsupervised matching generates, on average, a Type I error rate of 32 %. In other terms, the authors suggest that match rates should be accompanied by match quality as Type I errors bias sample features even when records are hand-linked.

In order to minimize this issue, I designed an *ad hoc* R script that adopts a cautious approach and is aimed at resolving case-specific issues ex-ante through a combination of automated tools and human supervision.[14] For clarity and completeness, note that, while drawing from the current literature, I explicitly define matching as *the process of comparing fuzzy strings - municipality denominations - to a pre-defined, correctly spelled, list.*
In particular, I first re-create an encompassing list of Italian municipalities that includes all their historical denominations and specify what province and region they currently belong to. This is a crucial step because administrative borders and denominations shifted significantly across time.[15] I then expand the list including a modified set of denominations that accounts for abbreviations, prepositions, apostrophes etc. For instance, Sant'Agata de'Goti (Benevento) would be also associated, ex-ante, with Sant Agata de Goti, S Agata de Goti, Santa Agata de Goti, Sant Agata Goti etc. I therefore group last residences by spelling and contrast each of them to the above-mentioned list.[16]

[13] See, among others, Aizer et al. (2016); Collins and Wanamaker (2017); Eli, Salisbury and Shertzer (2016). See also Abramitzky, Mill and Perez (2018) for a novel, fully-automated, matching approach.

[14] Figure A1.4 graphically describes the procedure I employed.

[15] Raw data on historical administrative changes has been taken and re-elaborated from http://www.elesh.it/storiacomuni/documentazione.asp.

[16] Intuitively, the same spelling mistake could appear more than once; grouping last residences allows to quicken the process as it decreases the overall number of comparisons to be made - which increases exponentially in relation to sample size due to the dimensional problem.

The result of this process is a combination of perfect and imperfect matches. In the first case, I simply extract the data point; in the latter, the script returns a set of suggestions based on a minimum similarity threshold. After this step, I manually evaluate suggestions on a case by case basis. In essence, the script computes an index of string similarity. While I experimented with both Machine Learning techniques (e.g. Random Forest Classifiers)[17] and different measures of string-distance (Demerau-Levenshtein, Jaro-Winkler, Hamming), I eventually opted to employ the simplest definition available, Levenshtein's: the distance between two strings is given by the minimum number of character insertions, deletions or substitutions needed to obtain one from the other. This choice is motivated on the premise that, as explained in the following paragraphs, Italian spelling might present peculiar traits.

Excluding missing data points (unevenly clustered before 1901), such procedure allows to retrieve about 55 per cent of the remaining available data, a grand total of about 1.6 million observations. The approach I adopted should minimize Type I errors to a minimum.

Other authors have attempted to geo-locate the last residences of the Italian passengers listed in Ellis Island records using slightly different procedures. The methodologically closer paper is Spitzer and Zimran (2018).[18] Their approach is elaborate and convincing. However, there are two major differences between their methodology and the one described in this chapter. First, the bulk of matches in Spitzer and Zimran (2018)'s paper is based on Google Maps algorithms. Indeed, the authors only try to match last residences with a pre-compiled list of municipalities whenever Google's algorithm did not produce a perfect match or suggestions were inappropriate. Second, they used historical rather than current borders: this is a major advantage because it allows them to reconstruct migration rates from Italian municipalities. Indeed, ISTAT municipal population data are only available for historical administrative units. Note that Chapter 3 is based on this very data set, kindly provided by the authors to conduct the analysis of earthquake-induced migration effects at the municipal level.

Even if I am not able to compare the two data sets in terms of municipal out-migration

[17] See Feigenbaum (2016) for an approach based on Machine Learning.
[18] See also Battiston (2018).

because of administrative changes and conflicting border definitions, I check whether there are any significant differences at the province-year level.[19] Results are comforting: indeed, there is an almost perfect spatio-temporal correlation, despite the total number of matches is higher in Spitzer and Zimran (2018) - a grand total of about 2.2 million observations. This might be due to the fact that my approach is significantly more conservative and stresses the desirably of avoiding false matches. Indeed, alongside spelling inaccuracy, there are several concerns specifically tied to Italian spelling and municipal denominations that might be masked by automated approaches grounded on Google APIs. In particular:

Saint-based toponyms The current denomination of more than a eighth Italian municipalities consists in the name of a Christian saint - e.g. Santa Maria, Santo Stefano, San Giovanni etc. - followed by a specific appellation - e.g. San Giovanni Gemini, San Giovanni la Punta etc. Unfortunately, this additional piece of information is missing more often than not. Without any further details it is therefore impossible to unambiguously singling these municipalities out as they might belong to different provinces and regions.

Suffixed/Prefixed toponyms Several Italian municipalities share part of their denomination - be it a noun, a prefix or a suffix - with others. For instance, there are 8 different Francavilla scattered across the peninsula: Francavilla al mare (Chieti); Francavilla Angitola (Vibo Valentia); Francavilla Bisio, (Alessandria); Francavilla d'Ete (Fermo) etc.

Abbreviations and prepositions Santo or San (Saint in English) might have been abbreviated into S., St., Sant. etc. Similarly, Inferiore/Superiore (Lower/Upper in English) often become Inf/Sup. Several municipalities incorporate a preposition/apostrophe into their denomination, e.g. Castelnuovo del Garda (Verona), Magliano de' Marsi (L'Aquila).

Perfectly homonym toponyms This issue poses similar problems even though municipalities having exactly the same denomination are not common.

[19] In order to do this, I re-computed provincial absolute migration figures from Spitzer and Zimran (2018)'s version of the data.

2/3-letter toponyms Whenever the denomination of a municipality consists of 2/3 letters it is hard to tell whether the record refers to the actual municipality, hides an omission or corresponds to an abbreviation - e.g. Alì and Alì Terme are distinct administrative units.

Italian toponyms in the United States There are about forty US cities named after Italian localities in the US. The list includes cities such as Palermo, Salerno and Naples, characterized by high out-migration rates. In the absence of any further details, it is not clear what the record refers to.

Historical changes and shifting borders Several municipalities have undergone processes of absorption, aggregation and dissolution over time; so did provincial and regional borders. Passengers might therefore have declared that their last residence corresponded to a town that does not exist anymore - in the same administrative sense. For instance, Nicastro, a locality characterized by one of the highest migration rates in the province of Catanzaro (Calabria), was merged with Sambiase and Sant'Eufemia Lamezia in 1968, giving birth to Lamezia Terme.

Similar-sounding toponyms Whenever there is a slight spelling mistake, string-distance-based measures might return exactly the same value. E.g. Cassano/Bassano; Rossano/Fossano etc.

1.2.2 Relation to other sources

There were two Italian governmental bodies responsible for the collection of emigration data during this temporal horizon: the *Dipartimento Generale della Statistica* (henceforth, DGS), active since 1876, and the *Commissariato Generale dell'Emigrazione* (hereafter, CGE), established in 1901 after a comprehensive emigration-related reform (Legge n. 23, 31/12/1901). Despite having the same objective, both agencies developed arbitrary definitions of emigrants that reflected themselves in partially discordant series. Discrepancies were further amplified by the lack of standardized collection techniques - see Marucco (2001) for a debate on competing

standpoints. Indeed, DGS estimations essentially relied on the number of passports issued by Italian municipalities. Intuitively, not all of those who obtained this document eventually departed while others might have used the same passport more than once during its period of validity. On the other hand, CGE series are extrapolated from official passengers' lists as each transoceanic company had to report the exact number of individuals on board for fiscal purposes - a proportional tax was levied. This is, in substance, the same mechanism utilized to create the Ellis Island data set, except for the fact that individuals appearing on ship manifests were counted ex-ante.[20] However, aside from offering comparatively more accurate estimates, the major advantage of using Ellis Island records is the extraordinary level of spatial disaggregation: while the available Italian official data had been aggregated province-wise, Ellis Island records preserve municipality-level information.

All the available statistics were merged into a single publication, the *Annuario statistico dell'emigrazione italiana dal 1876 al 1925* (ISTAT, 1926), shortly after the institution of ISTAT. While hardly acknowledged, part of the volume had been reported in the most comprehensive historical reference about international migration to the United States, Ferenczi and Willcox (1929), determining a concurrence of aggregate official sources on both sides of the Atlantic.

Figure 1.2 - in the following page - plots DGS and CGE series against Ellis Island data, evidencing similar trends but numerous partial inconsistencies. Despite only capturing Italian emigrants entering the United States through New York, Ellis Island data points are often higher, suggesting an actual under-reporting issue in CGE aggregates. This is not trivial to ascribe - especially in light of return migration being even higher than previously thought (Bandiera, Rasul and Viarengo, 2013).[21]

Note that both Ellis Island and official pre-1901 figures are likely to be unreliable because migrants were not legally bound to obtain a passport before embarking on transoceanic

[20] Collection techniques were modified again in 1916: passports were issued together with two coupons, each of which was retained by local authorities any time the migrant departed from Italy and traveled back.

[21] Given that the US implemented an open-border policy throughout the period, nationality misreporting is assumed to be negligible as there was little incentive to do so - at least for Italians before the first restrictive Immigration Act was enforced in 1917.

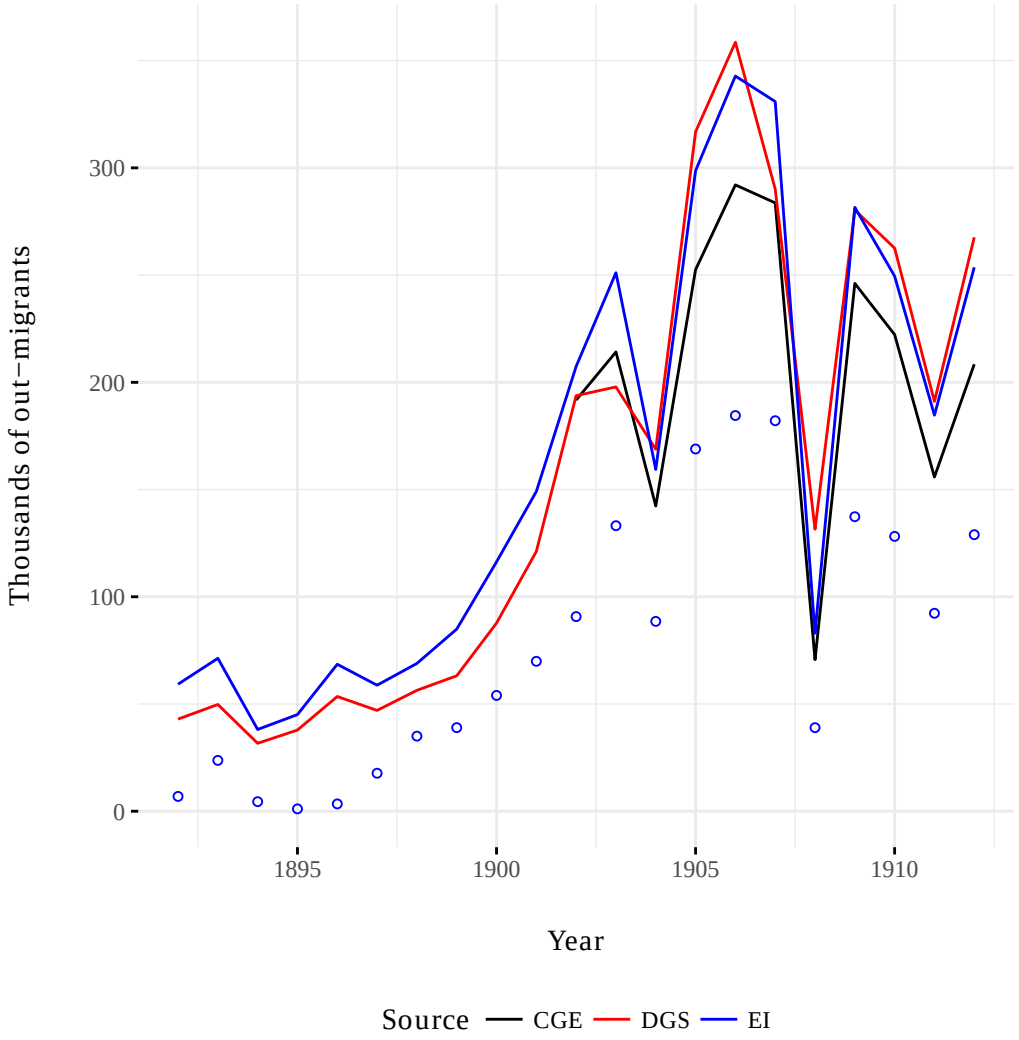


Figure 1.2: Data series in comparison.
Source: Ellis Island administrative records and *Annuario Statistico dell'emigrazione italiana*.
Notes: Blue circles denote the Ellis Island subset I am working with.

vessels (Keeling, 2013). Throughout the 1890s, American authorities either struggled or paid less attention to properly compile passengers lists, often omitting or transcribing residences according to macro-areas, e.g. Italy, northern/southern Italy. As a consequence, the largest amount of missing information in the data set relates to these years. The approximate nature of pre-1901 series gives rise to the main limitation of both the data set and the analysis below.

1.3 Overview

Figure 1.3 - in the following page - shows regional emigration trends toward Ellis Island, highlighting how out-migration flows co-moved, while having different magnitudes.

Figure 1.4[22] - at page 16 - maps the annual evolution of municipal migration from Sicily to the United States and visually displays the progressive diffusion of the phenomenon from 1901. Initially, migration was clustered in the province of Palermo, along the Western coast of the island (notably Sciacca), and in correspondence of the other major ports (Messina and Catania). This tendency remained relatively constant toward the late 1890s, before emigration spread across the island at the turn of the 20th century, morphing into a semi-universal event. There are also provincial patterns emerging from the map: for instance, out-migration from the areas orbiting around Syracuse rapidly increased and caught-up both in absolute and relative terms. However, when emigration from Sicily peaked in 1907, municipalities with a well-established recent emigration past were still numerically more represented, pointing toward path-dependency and persistence.[23] According to a suggestive hypothesis formulated by Gould (1979), information about potential destinations in the US might have originally entered the island through major ports and coastal cities[24] simply because of geography; it then spread toward inner areas, accumulated, and progressively expanded migration networks, easing logistic and financial constraints. Naval companies were granted the right to operate on transoceanic routes from Palermo and Messina after, respectively, 1901 and 1904.[25]

[22] Footnotes have been omitted in the interest of readability. I report them here. *Source*: Ellis Island administrative records. *Notes*: Red polygons denote aggregate migration greater than 1,000 individuals. Darker polygons are associated with higher migration.

[23] There are other examples across Europe. For example, using data on German migrants from the 1830s to the 1850s, Wegge (1998) shows that migration from Hesse-Kassel was path-dependent.

[24] Emigration from southern Italy might have followed trade routes: indeed, commercial ties between Sicily and the US had been blooming during the first half of the 19th century. Most of the northern American demand for citrus had been met by Sicilian suppliers, unevenly clustered in the province of Palermo and Messina (Dimico, Isopi and Olsson, 2017).

[25] Beforehand, Sicilian passengers had to reach Naples first. Port proximity might have had a catalyst role in favoring and fostering emigration. Palermo remained a largely islander port, never becoming a first-choice for migrants hailing from Calabria.

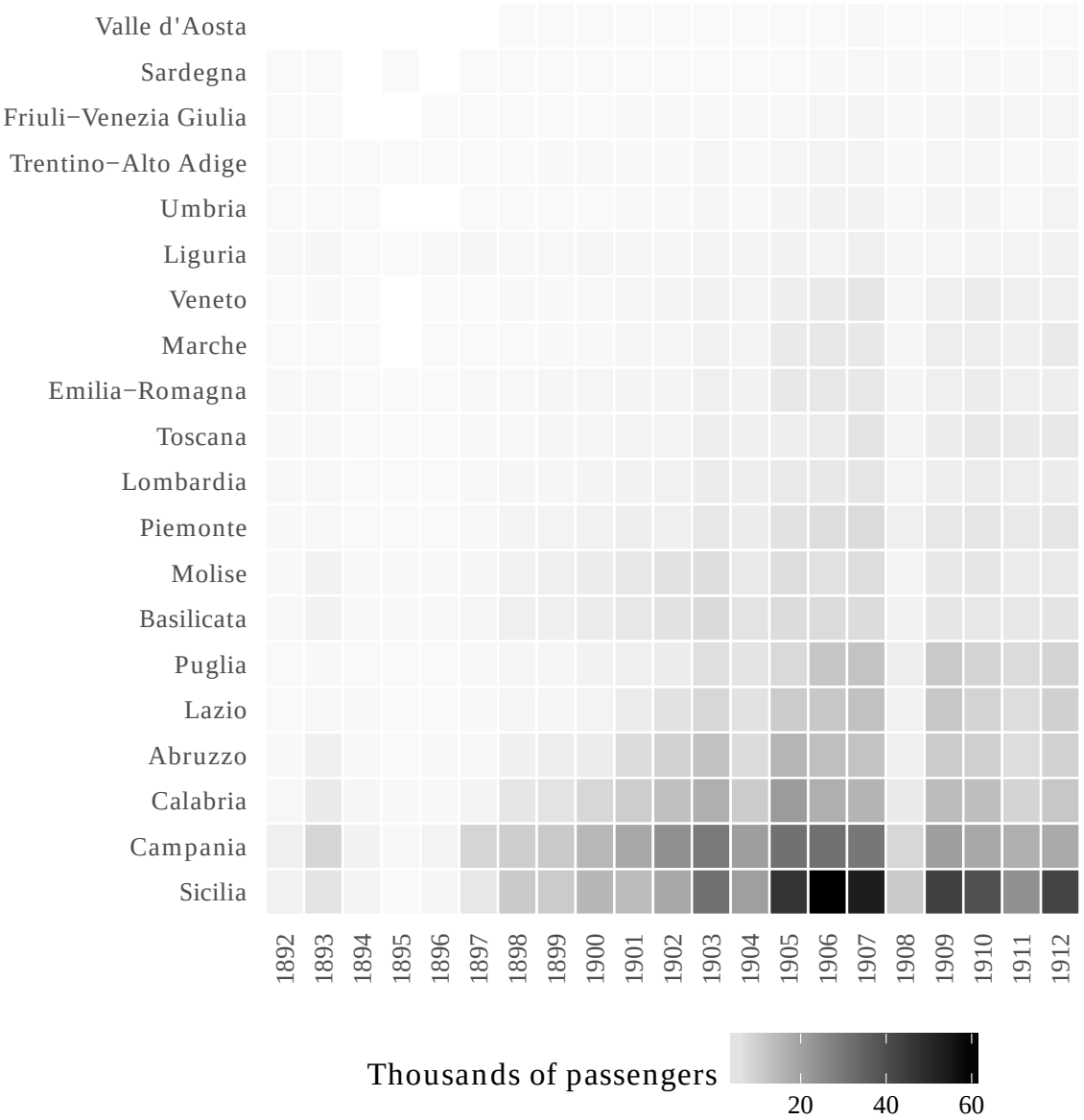
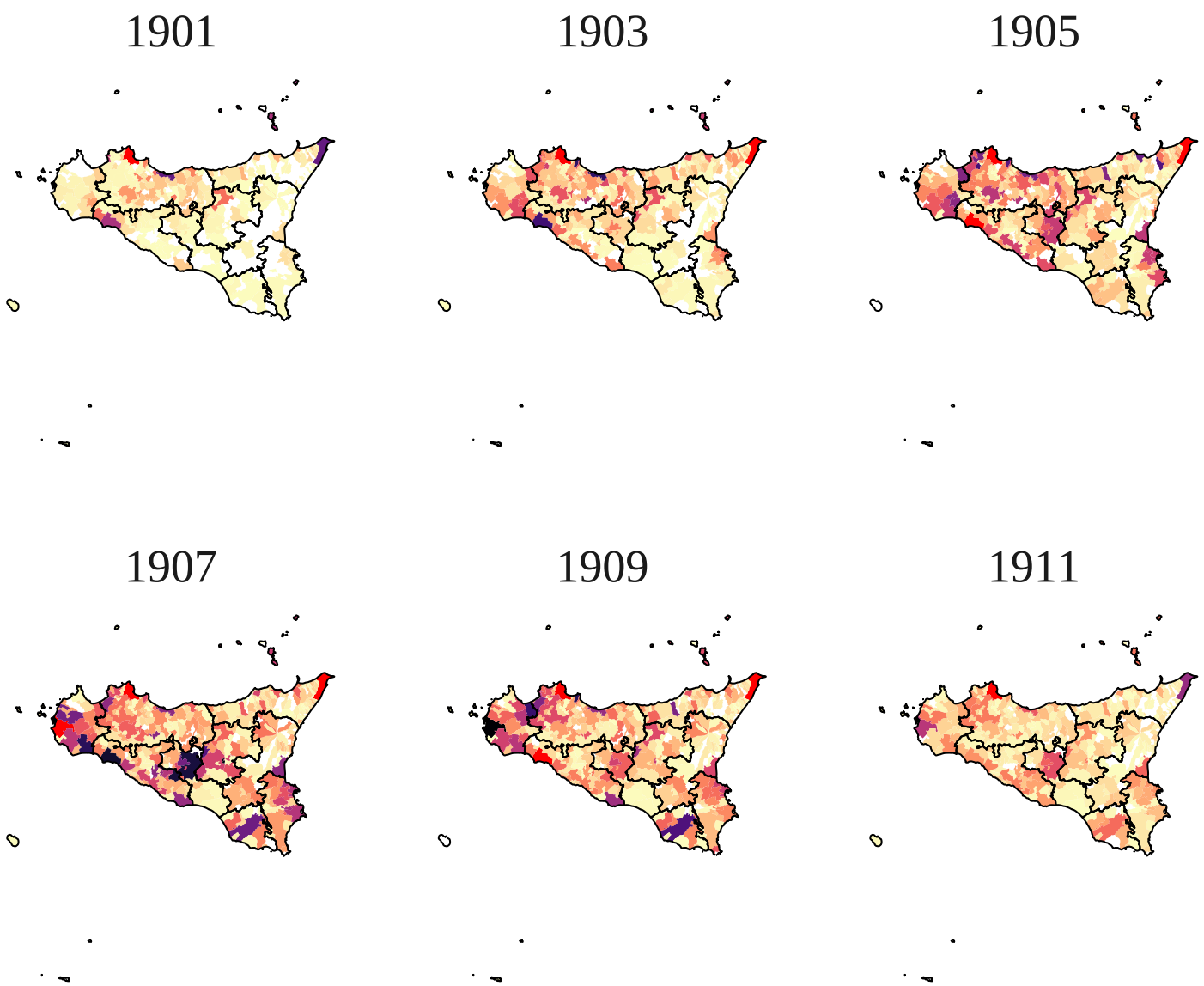


Figure 1.3: Regional emigration toward the United States.
Source: Ellis Island administrative records.
Notes: There are 20 regions; current borders.


1901
1903
1905
1907
1909
1911
Thousands of passengers
0.25
0.50
0.75

MacDonald and MacDonald (1964) argue that pre-WWI migration patterns are particularly suited for the study of chain migration because Italian authorities could not actively promote out-migration and migration flows developed through networks. In the case of Sicily and southern Italy, the authors maintain that the rise of migration chains could be imputed to two, partially overlapping, reasons. Before the 1900s, migration in the US was favored and sponsored by *padroni*, influential businessmen often linked to the local mafia. These individuals recruited migrants and channeled them into sectors heavily reliant on migrant labor, e.g. construction. Such flows were not self-perpetuating because migrants tended to become more autonomous once they familiarized with the US labor market. Second, serial migration occurred through district-or-municipality-level networks that allowed the establishment of migrants in certain niches and/or segments of the labor market. Note that Briggs (1972) stresses the idea that migrants hailing from a certain municipality could still opt to relocate away from their town-based communities, despite having initially benefited from them: indeed, the author documents several examples of chain migration from Termini Imerese - a coastal town near Palermo - to Buffalo, Rochester, Utica and Kansas City. On the other hand, Yans-McLaughlin (1977) interestingly argues that the study of chain migration is somewhat biased because it does not account for unsuccessful experiences and returnees.

Figure 1.5 - in the following page - conveniently displays the aggregate age, marital, and gender composition of Italian emigration flows to Ellis Island. After having controlled for gender, the pyramid is symmetric, meaning that there were similar shares of married and unmarried individuals; it is not balanced though: in line with previous examinations that characterized transoceanic migration as a predominantly young-men-driven event, men were over-represented with respect to women. However, when excluding children ($<$ 15 years old), married men cumulatively outnumber singles, suggesting more complex dynamics. There is some degree of variability across regions: women represented a sizable share of the outflow from Sicily, Liguria, and Piemonte - over 30 per cent - while only adding to 10 per cent in the case of Lazio. Anecdotal evidence suggests that gender-specific emigration had been

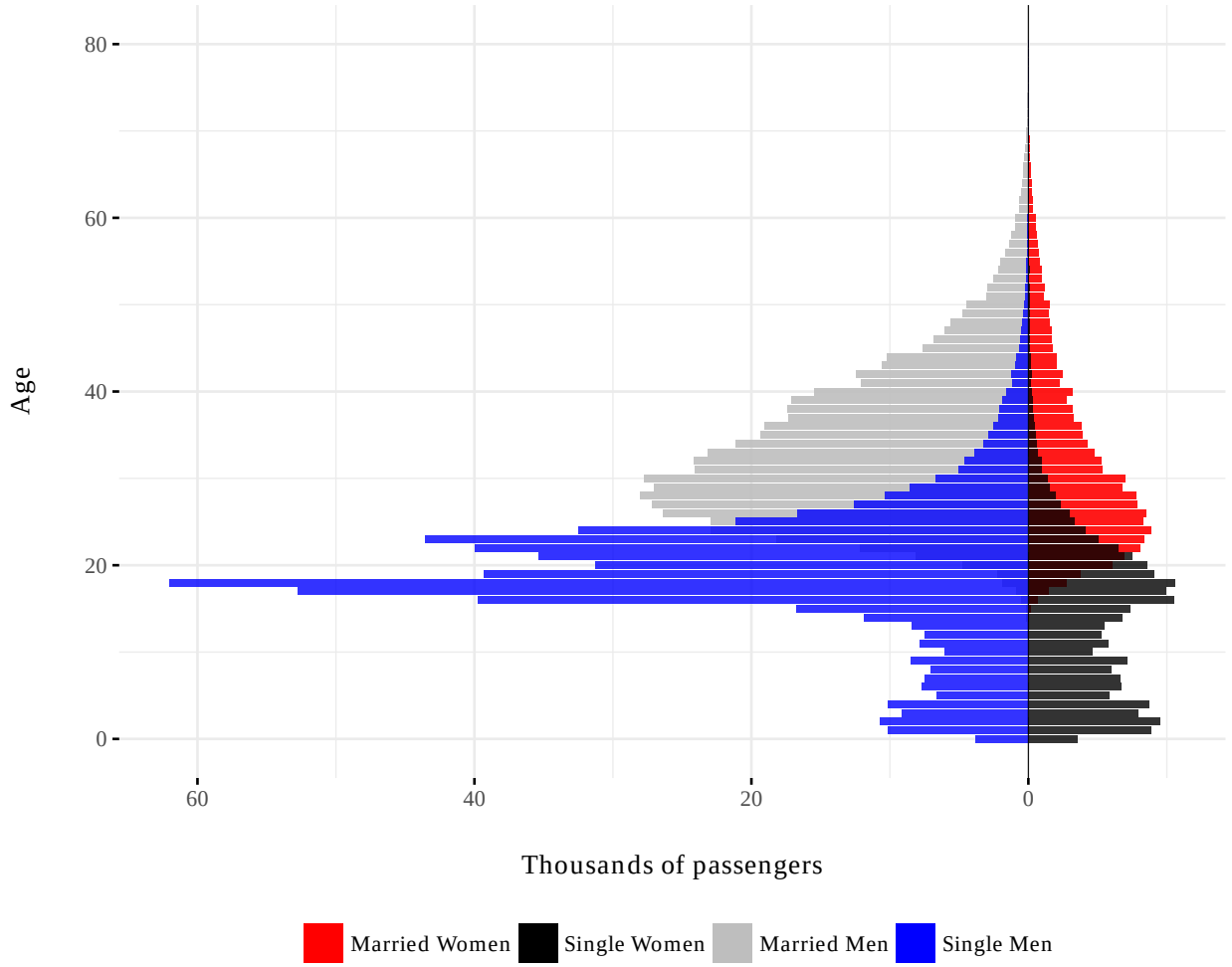


Figure 1.5: Age (in years), gender, and marital status of Italian passengers.
Source: Ellis Island administrative records.
Notes: Individuals have been omitted if older than 80.

relatively common, with household strategies often being determined by job opportunities abroad. For instance, nannies from Calabria flowed into the Suez area following the opening of the construction site in the mid-1850s; similarly, women textile laborers were employed in New York throughout the 1890s and 1900s (Corti and Sanfilippo, 2012).

Figure 1.6 - in the following page - concisely summarizes monthly arrivals, highlighting a clear underlying trend. The intra-year cyclicality reveals an uneven distribution of arrivals across months and seasons. While emigrants generally opted to leave Italy during springtime, figures are lower during the Summer, with a slight increase in Autumn. Winter departures are less frequent. The timing of migration might have been influenced by agricultural rhythms -

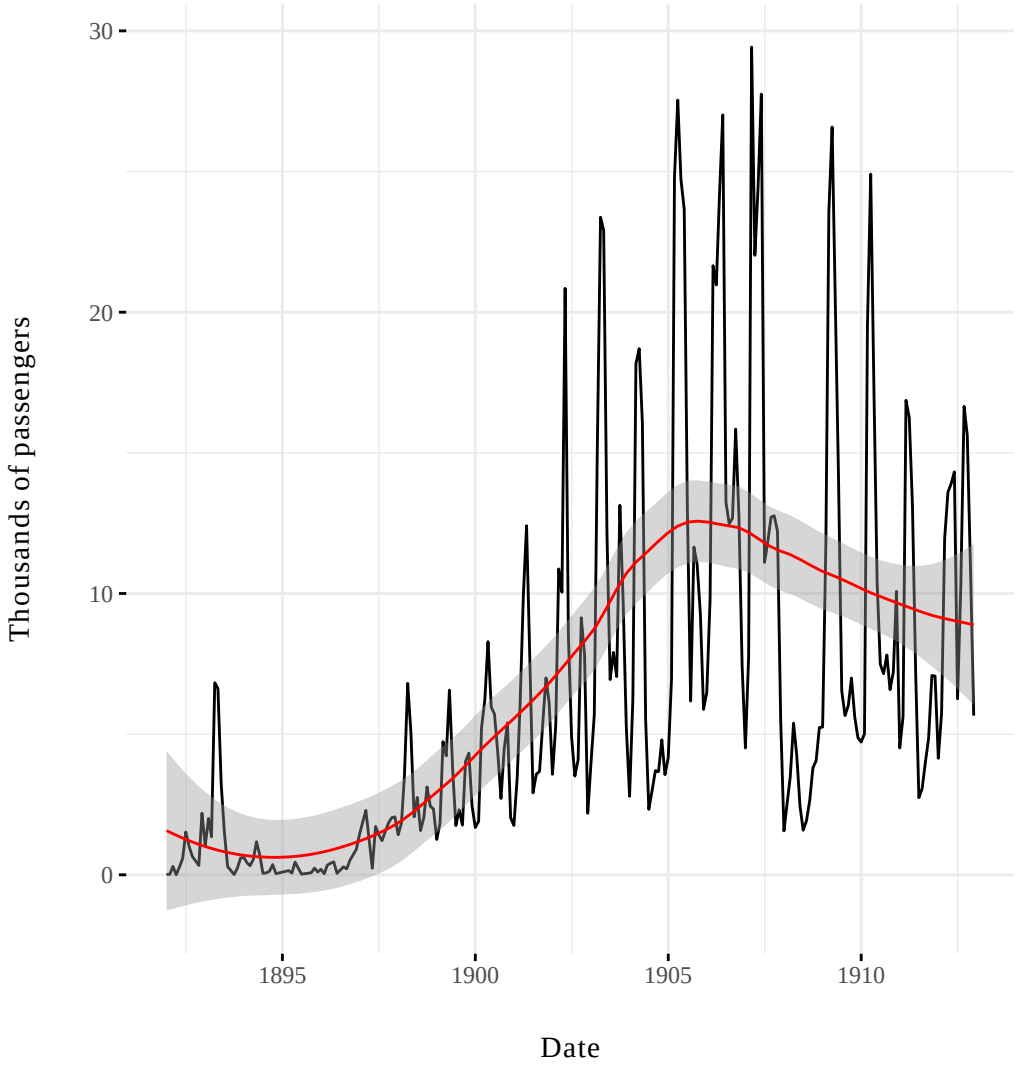


Figure 1.6: Monthly arrivals of Italian passengers.
Source: Ellis Island administrative records.
Notes: The red line fits a polynomial trend line.

as well as weather conditions. For instance, olive and grape harvests in Autumn might have increased seasonal employment opportunities - especially in the *mezzogiorno*. Citrus harvests from November to January attracted seasonal workers from neighboring regions to both Sicily and Calabria. Similarly, intra-annual fluctuations of fare prices were common, with travel expenses being constantly lower in the last quarter of each year (ISTAT, 1926). Finally, the notoriously inadequate conditions of Italian ports were likely to discourage emigrants from embarking on transoceanic vessels during wintertime.

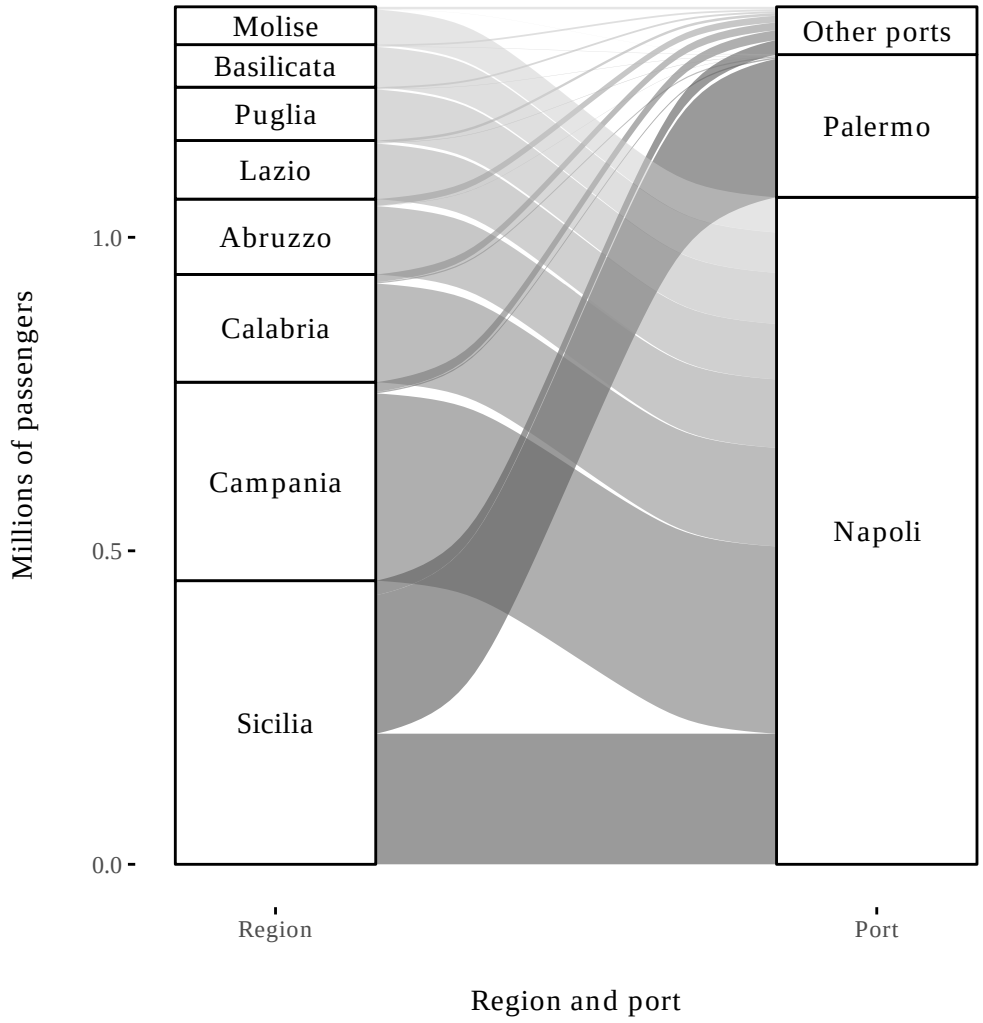


Figure 1.7: Port relevance per region.
Source: Ellis Island administrative records.
Notes: This graph only considers southern Italian regions.

The alluvial diagram depicted in figure 1.7 evidences the role of Naples and Palermo as the major migration hubs for, respectively, southern Italy and Sicily. The case of northern regions is more nuanced and heterogeneous. Indeed, departures from Naples were more frequent than those from Genoa. There is surprising evidence of non-negligible shares of northern Italians departing from other European hubs: Le Havre, Marseilles, Trieste (then part of the Austro-Hungarian empire), Hamburg, Bremen, Southampton, and Bolulogne-sur-Mer. In particular, departures from Le Havre outnumbered those from Genoa.[26] Rinaldetti (2013), descriptively

[26]Note that after 1901, would-be migrants leaving from Italian ports were legally required to hold a passport

documents these dynamics at the micro level, focusing on Fossato di Vico, a Medieval town in the province of Perugia: most of the emigrants - interestingly women - reached Le Havre before setting sail to Ellis Island, despite Naples and Genoa were incomparably closer. As already mentioned, migrants traveling on the European railroad system were eligible for a discount, proportional to their group size. The significance of ex-ante migration networks becomes therefore apparent.

1.4 Migration networks ex-ante

Migration networks could be thought as the universe of migrants hailing from a specific area in a given destination. Intuitively, diasporas might significantly lower the psychological, logistic, and financial costs involved in national and transnational migrations through different channels. On one hand, well-established local communities might facilitate the integration of their fellow-citizens into foreign labor markets; on the other, they might transfer valuable information about customs, best-practices, lodgings etc. therefore favoring integration as a whole (Abramitzky and Boustan, 2017). Munshi (2014) reports that, during the age of mass migration, recent migrants who were not familiar with the US labor market naturally relied on ethnic networks to access segments of the economy characterized by a relevant degree of flexibility.[27]

However, much of the quantitative literature is silent about possible spillover effects triggered by both returnees and repeat migrants. Moreover, there is an inclination toward host-country-focused, ex-post, investigations that neglect the sub-regional traits of migration flows and the crucial importance of networks before departure, ex-ante.

The very idea of network is quite sensitive to the geographical unit of reference, both in terms of origin and destination areas. As already mentioned, analyzing Italian migration networks at the aggregate level might introduce a severe bias: indeed, there is abundant

and this was not the case abroad. Ships often called in various ports around Europe before committing to the transoceanic voyage.

[27]Interestingly, the author also documents the presence of caste-based networks in India and similar mechanisms in Mumbai in the late 19th century.

descriptive evidence of urban clustering based on migrants' towns rather than countries of origin (MacDonald and MacDonald, 1964), meaning that the location choices of newcomers were heavily influenced by the geographical distribution of those who preceded them. Reaping the benefits of family-and-friends effects in the United States did not equate to being Italian for having the same nationality was arguably not a sufficient condition. The migration decisions of migrants hailing from different regions should therefore be, to a certain extent, considered independent.[28] Vecoli (1983) opens his paper with a very significant quote taken from (Thistlethwaite, 1960, p. 5): *"It is not useful to think of immigrants as an undifferentiated, mass movement of 'peasants' or indeed 'artisans' thronging towards immigrant ports from vaguely conceived countries of origin like Italy, Germany, or even Poland or Ireland. Seen through a magnifying glass, this undifferentiated mass surface breaks down into a honeycomb of innumerable particular cells, districts, villages, towns, each with an individual reaction or lack of it to the pull of migration"*. More recent contributions show that both weak and strong ties are more likely to arise if migrants hail from localized areas in origin countries (Comola and Mendola, 2015) and referral-based networks help reducing information asymmetries in the labor market (Dustmann et al., 2015)

There are several examples of networks based on a more precise definition of origin. According to the Commissariato dell'emigrazione (1902), Italians from different regions tended to cluster in different areas: for instance, the Italian population of Harlem (New York) was mainly composed of immigrants hailing from Calabria, Campania and Basilicata. Vecoli (1983) highlights how the spatial distribution of Italian migrants in Chicago gave rise to sixteen *Little Italies* scattered across the city rather than a unique Italian neighborhood. These areas were populated by migrants who came from different areas in Italy, settled in the US at different times and specialized in numerous niches - e.g. migrants from Lucca settled in the North Side and specialized in ceramics while those from Genoa populated central areas

[28] Interestingly, US authorities did differentiate between northern and southern Italians in the Dillingham report.

and focused on food services.[29]

There is reason to believe that Italians were not special. Indeed, Anbinder and McCaffrey (2015) document that Irish migrants from different counties exhibited specific location preferences in the US: post-famine migrants from Wexford were more represented in Savannah; migrants from Sligo and Kerry clustered in New York's Five Points neighborhood; the Irish population of Cleveland was chiefly drawn from county Mayo. Similarly, Bergquist (1984) highlights that community-based, German-speaking, enclaves were present in several US cities.

The attraction of migration networks in the United States - hence the pace of immigration - might have been influenced by the dimension, tightness and vintage of local-level diasporas: the relationship between networks and the propensity to migrate could therefore be non-linear, as recent contributions have highlighted for both income and wealth (Bazzi, 2017). In other words, networks might exert little attraction below a certain threshold, discourage migration in case of congestion while remaining functioning otherwise. On the other hand, considering where *exactly* migrants are from helps disentangling the dynamics of migration: individuals hailing from large cities might disperse across destinations more than those coming from villages as ex-ante personal connections are more likely to arise in the latter.[30]

Despite not containing information on migrants' residences in the United States, Ellis Island records do convey granular details about migration networks ex-ante. This topic is of interest as networks might have played a crucial role both before departure and during the actual voyage (Battiston, 2018). Indeed, within the context of an underdeveloped railroad and transportation system, reaching departure ports from inland localities was a journey in its own right. Given that there were only 4 major ports - 2 of which in Sicily - that directly liked Italy to Ellis Island, migrants from Abruzzo had to reach Naples - the closest option - first. This holds true for would-be migrants from the Alps willing to embark at the

[29]It is interesting to notice that such parochialism is reflected in the names of churches: *la Chiesa dei Genovesi*, *la Chiesa di Pompei*, *la Chiesa dei Siciliani* etc.

[30]Provincial and regional networks might also have a key role as interactions.

ports of Genoa, Trieste, or Le Havre. Even when considering areas closer to minor ports, it seems plausible that emigrants could have opted to leave in groups during their coastal navigation transfers to major gateways. After 1901, emigrants flowing to Genoa, Naples, and Palermo were legally granted the right to be accommodated in authorized inns at the expense of transoceanic naval companies. Molinari (2001) highlights how these obligations largely remained unfulfilled and how emigrants were frequently exposed to frauds, speculation and health risks: for instance, Naples had been plagued by cholera from 1884 and neighborhoods surrounding the port were particularly affected (Snowden, 2002).

In order to get a sense of the phenomenon, I restrict the analysis to the port of Naples. Throughout the 1900s, departures from there were very frequent: in 1907, there were as many as 240 departures, an average of 20 a month.[31] If ex-ante migration networks did matter, migrants from a given municipality should systematically cluster on a vessel rather than dispersing across them. This effect should also decrease over time and with proximity to ports, thanks to knowledge diffusion and spillovers.

I compute the sum, v_t, of transoceanic vessels that left from Naples during any given month of the entire time frame T (1892-1912); I calculate how frequently a vessel with migrants from municipality i on board docked at Ellis Island in any given period - $v_{i,t}$. Eventually, I simply take the ratio, averaging across municipalities to obtain μ_i.

$$\mu_i = \frac{\sum_{t=1}^{T}(\frac{v_{i,t}}{v_t})}{T} \qquad (1.1)$$

where $T = 252$ as there are 12 months times 21 years periods. In terms of interpretation, the higher the index, the more passengers from certain municipalities tended to travel together. I define also μ_y, as the cumulative yearly average across municipalities.

Figure 1.8 - in the following page - visually summarizes the the relationship between μ_i and the distance from Naples: the further emigrants were from the port, the more they tended to

[31] The frequency of departures assures that the analysis does not merely mirror departure discontinuities set by naval companies.

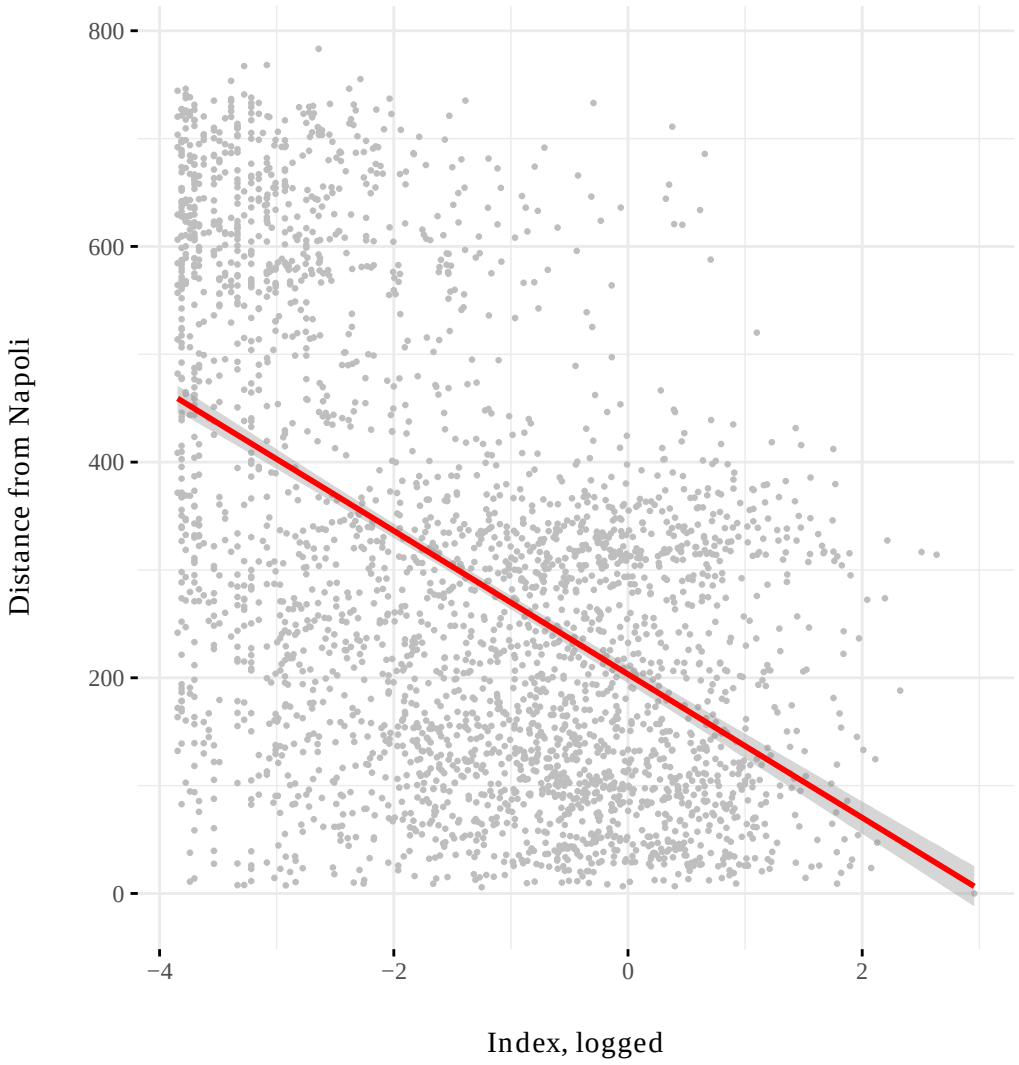


Figure 1.8: Migration networks ex-ante and distance from Naples (in Km); R = - 0.51. The index is logged.
Source: Ellis Island administrative records.
Notes: The index has been constructed according to equation (1.1).

travel with their fellow citizens and cluster on single vessels, rather than dispersing across them.[32] Interestingly, the relationship between the distance from Naples and the overall tendency in traveling together is negative but does not reduce to naught while getting closer to the port: locations that are in the proximity of Naples do exhibit a similar tendency, only less evident. In other words, traveling together might have had intrinsic benefits,

[32]This result could be partially read as a reflection of absolute migration flows being higher from larger cities. However, given the passenger-wise capacity and tonnage of transoceanic vessels, it would not be clear why these tendencies only manifests themselves in villages and peripheral municipalities.

regardless of the distance to the migration hub. Drawing upon this idea, location choices and migration networks in the US might be partly explained through the analysis of ex-ante patterns. Migrants arriving in the US did interact with one another during the voyage but it is important to highlight that networks at destination were likely to reflect networks ex-ante because several migrants were already part of them.

Moreover, figure A1.5 shows that, on aggregate, μ_y is evidently decreasing, suggesting that the diffusion of information eased migration. Interestingly, the index spikes in 1908, when migration from Italy declined as a response to the Panic of 1907 in the US: harder economic conditions made migrants more likely to travel together. Note that the result is unlikely to be driven by lowered transoceanic fares: as argued by Keeling (1999), fare prices did not steadily decreased from the 1880s - possibly as a result of cartels between naval companies.

These results point toward a general dynamic: during the initial stages of migration, migrants would rather leave with their fellow-citizens, independently of how frequent transoceanic crossing was; as information spilled over and paths became well-trodden, migrants departed more frequently, distributing themselves on more than one vessel - with the pace of emigration being accelerated by migration networks in the US. Coupled with evidence of pre-1901 emigration being less intense, these results descriptively validate the idea that pioneers had a significant role in tracing routes that other migrants would successively follow *en masse* (Gould 1979; Moretti 1999).

1.5 Repeat migration

Recent contributions stress the significance of return migration from the US, suggesting that about 70 per cent of migrants opted to temporarily remain abroad before eventually relocating to Italy over the period 1900 - 1910 (Bandiera, Rasul and Viarengo, 2013).[33] Although still open to debate, these findings challenge the traditional idea that transatlantic migration was an irreversible choice.

[33]Figures are even higher for the following decade, exceeding 80 per cent. Note how this phenomenon had already been noticed before 1900 by contemporary observers. See, among others, Schuyler (1889).

Temporary and repeat migrations hardly were new phenomena. There is abundant anecdotal evidence suggesting that short-distance movements within Italy were quite frequent even before the 1860s and were used as an income-smoothing strategy. A handful of authors argue that the age of mass migration uncovered historically established, temporary, approaches, stretching routes across continents. Indeed, numerous qualitative accounts of Italian emigration to Latin America have drawn attention to the presence of repeated, intra-year, streams of individuals taking advantage of inverse seasonality. *Glondrinas* (birds of passage), were mainly employed in the agricultural sector and could participate in two harvests per year.[34] However, Italian emigration to the United States had a mainly rural-urban nature and, if present, repeat movements might have had different motivations, objectives, and features.

These considerations inevitably contribute to a subtler characterization of transoceanic migration. Indeed, most studies implicitly suggest that return migration is a well-defined, unique, event - see Ward (2016) for an exception. In other words, cyclicality and multiple individual crossings are neglected. Moreover, the temporal horizon of repeat migration is not neutral: spells might either be seasonal or span across several years with individuals re-entering the United States multiple times. According to Dustmann and Görlach (2016), migration strategies include a temporary dimension that might modify the socioeconomic behavior of migrants and their households, therefore impacting differently on host economies.

While Ellis Island manifests do report whether migrants were re-entering the United States, the information has not been systematically digitized. I indirectly identify individuals who arrived in New York more than once through a simple numerical exercise that tallies the frequency of individual registrations. Data are initially subset considering only migrants appearing twice or more in the Ellis Island data set, i.e. having identical surname, name, marital status, year of birth and residence.[35] I assume that two entries refer to the same

[34]The abolition of slavery in Brazil (1888) boosted the labor supply of agricultural jobs in *fazendas*. The Argentinian government encouraged new settlements and economically supported those willing to farm rural areas (Bevilacqua, 2001).

[35]Identical records on a given ship are obviously only counted once. Similarly, if the same individual

individual only if both marital status and last residence did not change over time. In other words, US returnees that changed residence, got married, divorced, or lost their spouse between successive arrivals are excluded. Additionally, despite being the main gateway to the United States, Ellis Island was not the only immigration port. Consequently, US returnees re-entering the United States from Boston or Philadelphia only appear once. Given the *very* conservative nature of these assumptions, repeat migration figures represent a lower bound estimate. Moreover, the fact that spelling mistakes are common and often prevent correct identifications, further biases estimates downwards.[36]

The resulting sub-sample of repeat migrants consists of about 70,000 records, for a total of 30,837 migrants - 80 % of which were male. Table 1.1 - at page 31 - summarizes the main findings, grouped region-wise. Overall, 2,142 municipalities were affected by repeat migration, 36.1 % of all 5,923 migration municipalities - see columns (4) and (5), table 1.1. The average number of arrivals per migrant is 2.07, suggesting that the vast majority of repeat migrants only entered the US twice - see column (1), table 1.1. There is little variation across regions and provinces. Indeed, grouping repeat migrants according to their province of origin reveals that, on average, the highest number of successive arrivals was 2.12, from Palermo.

The geographical distribution of the individual observations is not homogeneous: indeed, Sicily accounts for 42 % of all observations, Campania for 17 % and Calabria for about 11 per cent. The rest of southern Italian regions are significantly more represented than any northern areas - see column (2), table 1.1.

Repeat migrants were on average 24.3 years old when they first arrived in the US and 27.05 the last time they re-entered the country. They were therefore about 9 months younger than the average migrant - who was 25.05 years old - when they first reached Ellis Island. While there is some interesting regional variation in these figures, no clear-cut pattern emerges in terms of north/south divide - see columns (6) and (7), table 1.1.

appears twice within a month - an approximate estimate of transatlantic crossing - I infer that records must pertain to different migrants.

[36] An example above all: when re-entering the United States through Ellis Island, *Fiorello La Guardia*, the Italian-Jewish mayor of New York city, had his name misspelled - *Forello* in lieu of *Fiorello*.

There is an almost perfect correlation in terms of time preferences between repeat and regular migrants. Indeed, all passengers traveling to Ellis Island opted to leave during Spring and Autumn. This piece of information differentiates US-bound *golondrinas* from their South American counterparts because they evidently did not engage in successive migration to exploit inverse seasonality and be employed in the agricultural sector twice a year.

I also compute a municipal-level measure that conveys information on the relative weight of repeat migration.[37] It equates to the ratio between repeat and non-repeat migration flows from any given municipality, averaged across the entire time frame - see column (3), table 1.1. Given the high level of disaggregation, results are best shown through a map, figure 1.9 - at page 30. Figures vary widely across the geographical units of reference and clearly highlight how repeat migration was an eminently southern phenomenon. Sicily is the region that displays the highest degree of dynamism, with an average ratio between repeat and circular migration above 5 per cent. The phenomenon is very heterogeneous even within this region. Indeed, while repeat migration was not common in certain migration municipalities, others - like coastal towns and Northwestern areas - exhibit rates ranging from than 10 to 20 per cent. Repeat migration seems to be correlated with a well-established migration past, i.e. it is more evident in areas that were more affected by out-migration toward the United States in the first place. Another good example might be a specific area of Tuscany: repeat migration is more relevant in the provinces of Lucca and Massa Carrara, those that initially sent more migrants to the United States.

I also add a temporal dimension to the analysis by computing the aggregate repeat migration ratio *ad annum*. Figure 1.12 - at page 36 - contrasts repeat and non-repeat temporal trends, suggesting that individuals willing to enter and re-enter the United States faced similar incentives.

[37]Note that, if different records refer to the same individual, the entire emigration flow should then be adjusted downwards. In other words, official figures are usually inflated because the number of people entering the US does not correspond to the number of records.

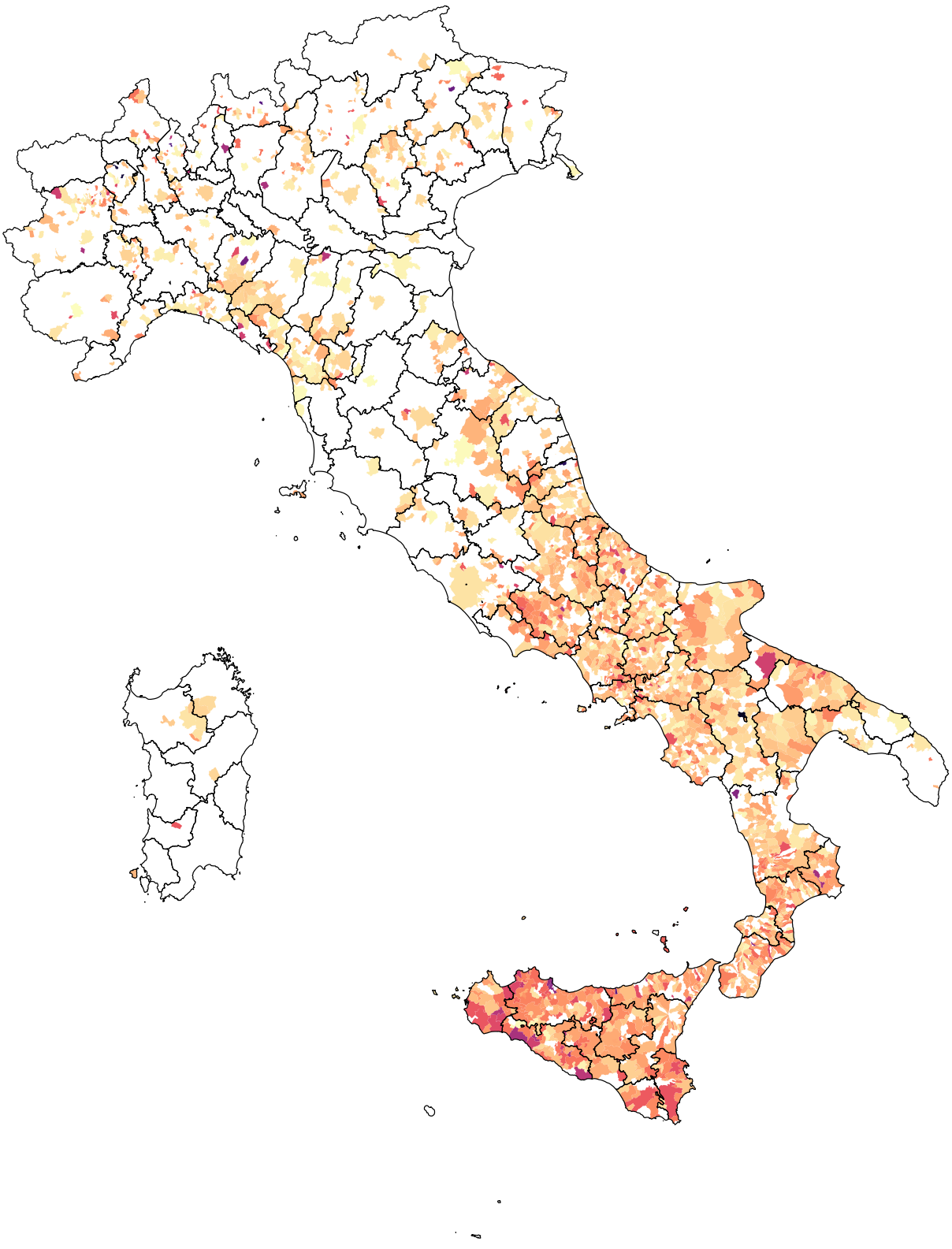

5
10
15
20
25

Table 1.1: Golondrinas

Code	Reg	(1) Avg. n. of arrivals	(2) Repreat migrants	(3) Share over total	(4) Repeat municipality	(5) Migration municipality	(6) Avg. age first arrival	(7) Avg. age last arrival
1	Abruzzo	2.04	1901	3.75	203	297	26.09	29.72
2	Basilicata	2.03	849	2.98	89	124	24.96	27.72
3	Calabria	2.05	3331	4.10	276	401	26.68	30.17
4	Campania	2.06	5416	3.74	388	536	25.14	28.16
5	Emilia-Romagna	2.07	216	3.10	54	247	26.52	30.10
6	Friuli-Venezia Giulia	2.00	42	3.59	21	150	30.21	34.12
7	Lazio	2.06	1894	4.41	134	318	26.25	29.51
8	Liguria	2.01	124	3.92	32	187	20.58	23.65
9	Lombardia	2.06	253	4.34	99	893	27.34	30.62
10	Marche	2.03	337	3.87	61	188	26.62	29.99
11	Molise	2.03	894	3.41	97	134	25.68	29.26
12	Piemonte	2.02	345	4.07	123	899	25.47	29.21
13	Puglia	2.05	1551	4.08	103	192	24.93	27.47
14	Sardegna	2.00	14	3.22	12	138	26.50	27.93
15	Sicilia	2.11	13111	5.69	294	368	22.57	24.49
16	Toscana	2.03	227	2.57	51	192	26.11	29.68
17	Trentino-Alto Adige	2.00	46	3.94	26	213	23.93	26.80
18	Umbria	2.05	125	3.35	20	80	26.60	29.76
19	Valle d'Aosta					46		
20	Veneto	2.02	161	3.13	67	397	25.96	29.49

Notes: This table has been produced using current administrative borders. There are 20 regions. Column (1) shows the average number of arrivals in Ellis Island of Italian passengers coming from different regions; column (2) reports the number of repeat migrants; column (3) reports the share of repeat over regular migrants; column (4) documents whether or not the municipality sent any repeat migrants while column (5) points out if the municipality sent any migrants at all; column (6) and (7) inform about the average age of repeat migrants when they first/last arrived in the US.

However, it is interesting to note that repeat migration has been more intense where, in principle, incentives to emigrate were higher, i.e. in poorer, backward areas. How is it possible? There are three, overlapping, threads of reasoning. First, it might be the case that ex-ante and ex-post migration networks considerably lowered financial, logistic and information costs, allowing a substantial ease of movement. Second, repeat migrants did not fundamentally differ at the time of their initial departure - i.e. their degree of self-selection was not different. However, they were positively self-selected afterwards for they were able to finance continuous transoceanic crossings, arguably as a consequence of positive labor outcomes and successful experiences.[38] In other words, stays in the United States were well worth the transoceanic voyage. Moreover, repeat migrants might have re-entered the United States with initially left-behind families or just paid them a visit - but unfortunately I cannot explicitly test this hypothesis as I am not able to univocally define households. Third, using anthropometric evidence Spitzer and Zimran (2018) argue that Italian migrants hailing from southern provinces were more positively self-selected then those from the north. It might be the case that repeat migrants differed from once-off migrants in terms of how financially constrained they were to begin with. It would follow that repeat migrants really were *golondrinas*, in the sense that their continuous movement represented a well-defined, multiple-stage, migration strategy.

1.5.1 Returnees to the United States

The previous subsection relies on the assumption that repeat migrants' last residences did not change over time precisely because migration was conceived as temporary. In other words, when questioned by US authorities, repeat migrants declared to be hailing from a given municipality in Italy even though they had previously been admitted to the US.[39]

[38] I cannot exclude that repeat migrants were unsuccessful individuals who opted to give the US another chance after an initially disastrous experience. According to Ward (2017), there is a mismatch between actual returnees and migrants who initially intended to do so; the difference originated from bleaker-than-expected economic outcomes, especially for southern and eastern Europeans. Migration experiences partially improved after the introduction of immigration quotas in 1921 and 1924.

[39] As already mentioned, Ellis Island administrative records do contain dichotomous - Yes/No - information about immigrants having set foot on US soil before. The last residences of individuals re-entering the US are

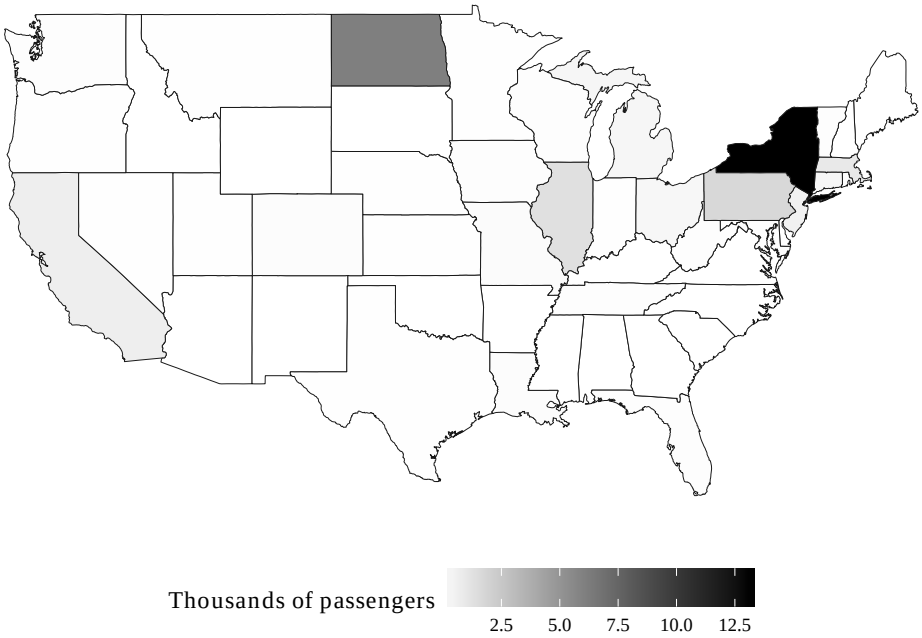


Figure 1.10: Destinations of Italian return passengers to the United States.
Source: Ellis Island administrative records.

Nevertheless, the data set also contains about 56,000 ethnically Italian individuals who declared to reside in the United States at the time of their registration on Ellis Island.[40] While I cannot definitively differentiate this form of repeat migration from the one described

often in Italy.

[40]The methodology I utilized to identify these individuals is similar to the one I presented in section 1.2. I initially constructed a list of US localities that comprises all city councils, counties, states, and state abbreviations contained in the US census of 1910; I then contrasted it with Ellis Island data and isolated perfect matches. Imperfect matches were polished manually. Note that the figures might be much higher than the one I present here. Indeed, there is a plethora of US localities having the same toponyms e.g. Washington, Springfield, Franklin, Greenville, and/or bearing the denomination of municipalities scattered across Europe and Canada, e.g. Turin, Rome, Venice, London, Paris, Liverpool, Amsterdam, York, etc. I only retain these observations when further information is provided, e.g. Oxford (Alabama), Oxford (Arkansas).

in the previous section, declaring to reside in the US when re-entering the country suggests a somehow higher degree of integration and/or the will to settle there. Unfortunately, it is not possible to trace back these individuals and obtain their last residences in Italy. On the contrary, figure 1.10 - in the previous page - maps where these repeat immigrants were traveling to. There seems to be an overlap between repeat migrants' residences and areas where Italian immigration had a well-established and numerically significant tradition. Indeed, most of these migrants resided on the Atlantic Coast - in order, New York (about a fourth of the total), Pennsylvania, Massachusetts, New Jersey, and Connecticut - and in the Great Lakes region - Illinois and Michigan. There also was a sizable presence of Italian migrants in California.

Timing-wise, this type of return migration only gained momentum after 1904 and peaked in 1909 - after having dipped in 1908 like other forms of Italian migration to the United States - as shown in figure 1.12, at page 36.

1.5.2 Italian migration from other countries

I also highlight the fact that among the universe of ethnically Italian migrants setting foot on Ellis Island, there were about 30,000 individuals - again, a lower bound estimate - whose last residence was located in any other country aside from Italy and the United States. While being numerically less relevant, the presence of these flows points toward an extraordinary degree of mobility. Figure 1.11 - in the following page - highlights how the last residences of these migrants were scattered all across the globe. There seems to be a correlation between the size of the Italian diaspora in each of these countries and the amount of migrants arriving in the United States from there (Bevilacqua, 2001). In particular, about a third of them hailed from France, notably from Paris and Marseilles - two traditional Italian migration hubs. Italian emigrants might have tried to settle in other destinations - those where migration networks were already developed and migration had a long-standing tradition - and then opted to relocate to the Unites States, with nothing guaranteeing that their journey effectively

ended there. This multiple-stage migration process might have been favored by the progressive diffusion of information. The timing of such arrivals is correlated with other types of migration flows - see figure 1.12, at page 36.

Figure 1.11: Last residence of ethnically Italian passengers (excluding, Italy and the United States). *Source*: Ellis Island administrative records.

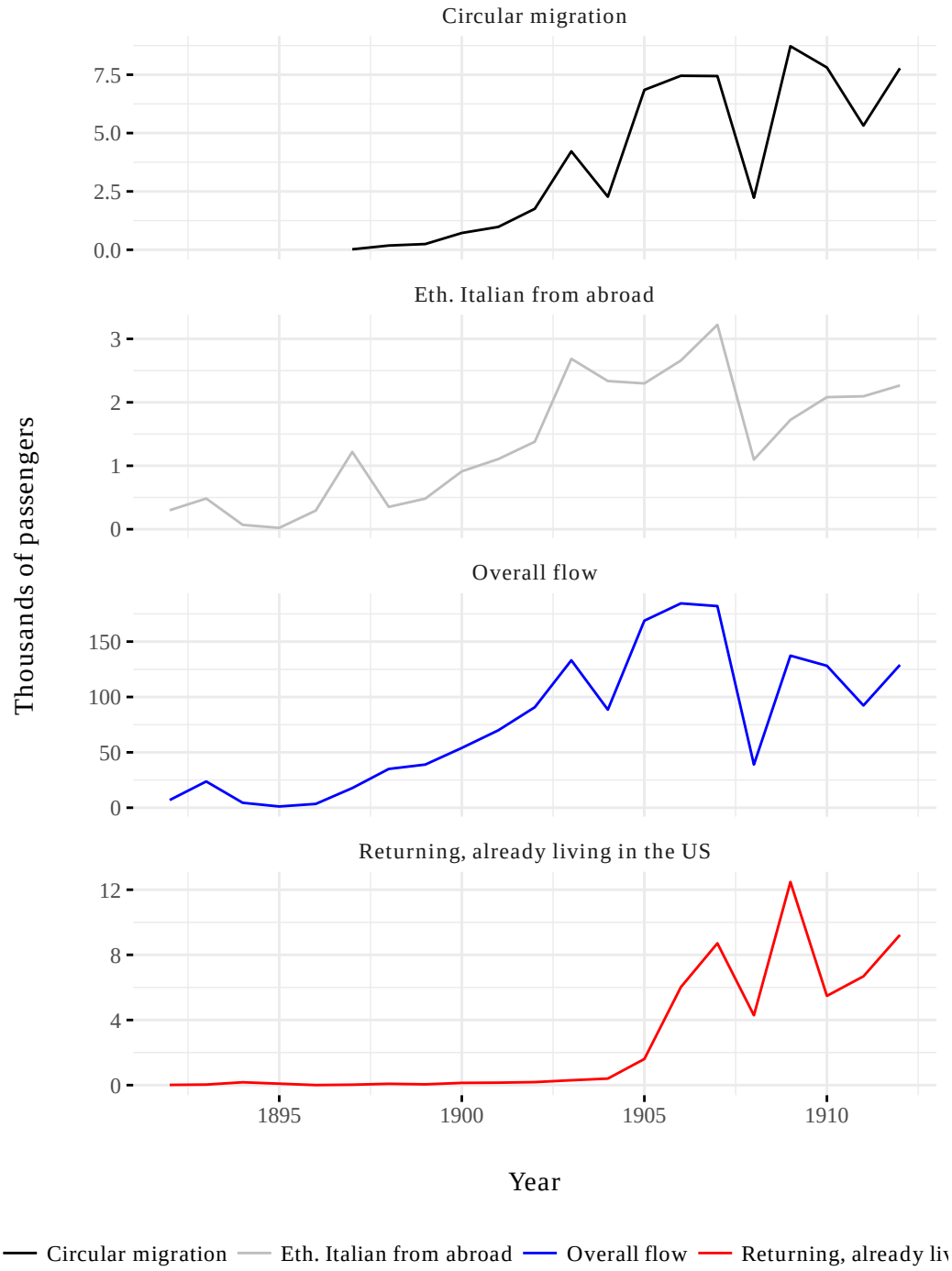


Figure 1.12: Time trends of Italian emigration towards the United States.
Source: Ellis Island administrative records.

1.6 Conclusion

"*[...] the ecological model [of migration] does not adequately explain the peculiarities of the spatial distribution of the Italian neighborhoods [in Chicago]. Land use patterns, industrial location, labor market, and housing market did not in themselves determine where the Italians finally lighted*" (Vecoli, 1983, p. 16). Drawing on this insight, this article makes a contribution to the literature on migration, with a particular focus on the pre-WWI period. It shows how a comprehensive account of migration flows cannot abstract from the spatial heterogeneity of migration itself. It also stresses the fact that migration is a *dynamic* process, one in which people might come and go, displaying a repeat attitude that is often related to what others in their networks are doing. I think that this is a relevant concept that is too often hidden in the literature. Indeed, econometric analyses of the consequences of migration on both destination and origin countries rarely account for the possibility that migrants do not stay or move multiple times - e.g. when considering their impact on wages. Overall, this analysis allows to deepen the debate on the age of mass migration in unprecedented manners, but it is by no means a *finale*. Similar efforts about Latin American and European countries are both desirable and essential to obtain a wider picture. Finally, this paper challenges the view of emigration being a single and immutable choice and contributes to characterizing the turn of the 20th century as the first truly globalized era.

References

Abramitzky, Ran and Leah Boustan. 2017. "Immigration in American economic history." *Journal of Economic Literature* 55(4):1311–45.

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2012. "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." *The American economic review* 102(5):1832–1856.

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2013. "Have the poor always been less likely to migrate? Evidence from inheritance practices during the Age of Mass Migration." *Journal of Development Economics* 102:2–14.

Abramitzky, Ran, Roy Mill and Santiago Perez. 2018. Linking Individuals Across Historical Sources: a Fully Automated Approach. Technical Report No. 24324 NBER.

Aizer, Anna, Shari Eli, Joseph Ferrie and Adriana Lleras-Muney. 2016. "The long-run impact of cash transfers to poor families." *The American Economic Review* 106(4):935–71.

Anbinder, Tyler and Hope McCaffrey. 2015. "Which Irish men and women immigrated to the United States during the Great Famine migration of 1846–54?" *Irish Historical Studies* 39(156):620–642.

Ardeni, Pier Giorgio and Andrea Gentili. 2014. "Revisiting Italian emigration before the Great War: a test of the standard economic model." *European Review of Economic History* 18(4):452–471.

Bailey, Martha, Connor Cole, Morgan Henderson and Catherine Massey. 2017. How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth. Technical Report No. 24019 NBER.

Baines, Dudley. 2004. Population, migration and regional development, 1870-1939. In *The Economic History of Britain Since 1700*, ed. Roderick Floud and Deirdre N. McCloskey. Vol. 1700 Cambridge University Press pp. 25–55.

Bandiera, Oriana, Imran Rasul and Martina Viarengo. 2013. "The making of modern America: Migratory flows in the age of mass migration." *Journal of Development Economics* 102:23–47.

Battiston, Diego. 2018. "The Persistent Effects of Brief Interactions: Evidence from Immigrant Ships." Unpublished manuscript, London School of Economics.

Bazzi, Samuel. 2017. "Wealth heterogeneity and the income elasticity of migration." *American Economic Journal: Applied Economics* 9(2):219–255.

Bergquist, James M. 1984. "German Communities in American Cities: An Interpretation of the Nineteenth-Century Experience." *Journal of American Ethnic History* 4(1):9–30.

Bevilacqua, Piero. 2001. *Storia dell'emigrazione italiana*. Donzelli Editore.

Briggs, John W. 1972. *Italians in Italy and America: a study of change within continuity for immigrants to three American cities, 1890-1930*. University of Minnesota.

Collins, William J and Marianne H Wanamaker. 2017. Up from Slavery? African American Intergenerational Economic Mobility Since 1880. Technical Report No. 23395 NBER.

Commissariato dell'emigrazione. 1902. Bollettino dell'emigrazione. Technical report Ministero degli affari esteri.

Comola, Margherita and Mariapia Mendola. 2015. "Formation of migrant networks." *The Scandinavian Journal of Economics* 117(2):592–618.

Corti, Paola and Matteo Sanfilippo. 2012. *L'Italia e le migrazioni*. Laterza.

Dimico, Arcangelo, Alessia Isopi and Ola Olsson. 2017. "Origins of the sicilian mafia: The market for lemons." *The Journal of Economic History* 77(4):1083–1115.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg and Herbert Brücker. 2015. "Referral-based job search networks." *The Review of Economic Studies* 83(2):514–546.

Dustmann, Christian and Joseph-Simon Görlach. 2016. "The economics of temporary migrations." *Journal of Economic Literature* 54(1):98–136.

Eli, Shari, Laura Salisbury and Allison Shertzer. 2016. Migration in Response to Civil Conflict: Evidence from the Border of the American Civil War. Technical Report No. 22591 NBER.

Eriksson, Katherine. 2018. Ethnic enclaves and immigrant outcomes: Norwegian immigrants during the Age of Mass Migration. Technical Report No. 24763 NBER Working Paper.

Faini, R and A Venturini. 1994. Italian emigration in the pre-war period. In *Migration and the international labor market, 1850-1939*, ed. Timothy J Hatton and Jeffrey G Williamson. Taylor and Francis pp. 72–90.

Fauri, Francesca. 2015. *Storia economica delle migrazioni italiane*. Il mulino.

Feigenbaum, James J. 2016. "A Machine Learning Approach to Census Record Linking." Unpublished manuscript, Harvard University, Department of Economics.

Ferenczi, I and Walter F Willcox. 1929. "International Migrations." *NBER* .

Foerster, Robert Franz. 1919. *The Italian emigration of our times*. Vol. 20 Harvard University Press.

Goldin, Claudia. 1994. The political economy of immigration restriction in the United States, 1890 to 1921. In *The regulated economy: A historical approach to political economy*. University of Chicago Press pp. 223–258.

Gomellini, Matteo and Cormac Ò Gràda. 2013. *The Oxford Handbook of the Italian Economy Since Unification*. Oxford University Press, New York chapter Migrations.

Gould, John D. 1979. "European inter-continental emigration 1815-1914: patterns and causes." *Journal of European Economic History* 8(3):593.

Hatton, Timothy J and Jeffrey G Williamson. 1998. *The age of mass migration: Causes and economic impact*. Oxford University Press on Demand.

Hatton, Timothy J and Jeffrey G Williamson. 2005. *Global migration and the world economy: Two centuries of policy and performance*. MIT press Cambridge, MA.

ISTAT. 1926. *Annuario statistico dell'emigrazione italiana*. Commissariato generale dell'emigrazione.

ISTAT. 2011. *L'Italia in 150 anni: sommario di statistiche storiche, 1861-2010*. ISTAT.

Karadja, Mounir and Erik Prawitz. forth. "Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States." *Journal of Political Economy* .

Keeling, Drew. 1999. "The transportation revolution and transatlantic migration, 1850-1914." *Research in Economic History* 19:39–74.

Keeling, Drew. 2013. "Oceanic travel conditions and American immigration, 1890-1914." *MPRA* (47850).

MacDonald, John S and Leatrice D MacDonald. 1964. "Chain migration ethnic neighborhood formation and social networks." *The Milbank Memorial Fund Quarterly* 42(1):82–97.

Marucco, Dora. 2001. Le statistiche dell'emigrazione italiana. In *Storia dell'emigrazione italiana*, ed. Piero Bevilacqua. Vol. 1 Donzelli Editore.

Molinari, Augusta. 2001. Porti, trasporti, compagnie. In *Storia dell'emigrazione italiana*, ed. Piero Bevilacqua. Vol. 1 Donzelli Editore.

Moretti, Enrico. 1999. "Social networks and migrations: Italy 1876-1913." *International Migration Review* pp. 640–657.

Munshi, Kaivan. 2014. "Community networks and migration." *The Oxford Handbook of the Economics of Networks* pp. 630–648.

Ò Gràda, Cormac and Kevin H O'Rourke. 1997. "Migration as disaster relief: Lessons from the Great Irish Famine." *European Review of Economic History* 1(1):3–25.

O'Rourke, Kevin H. 1997. "The European grain invasion, 1870-1913." *The Journal of Economic History* 57(04):775–801.

Rinaldetti, Thierry. 2013. "Italian Women, Mobility, and Labor Migrations The Case of Fossato di Vico in the Umbrian Apennines (1900-1914)." *Italian Americana* pp. 63–82.

Sánchez-Alonso, Blanca. 2013. "Making sense of immigration policy: A rgentina, 1870–1930." *The Economic History Review* 66(2):601–627.

Schuyler, Eugene. 1889. "Italian immigration into the United States." *Political Science Quarterly* 4(3):480–495.

Snowden, Frank M. 2002. *Naples in the Time of Cholera, 1884-1911*. Cambridge University Press.

Spitzer, Yannay. 2015. "The Dynamics of Mass Migration: Estimating the Effect of Income Differences on Migration in a Dynamic Model of Discrete Choice with Diffusion." Unpublished manuscript, Brown University, Deptartment of Economics.

Spitzer, Yannay. 2016. "Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914." Unpublished manuscript, Norwestern University, Department of Economics.

Spitzer, Yannay and Ariell 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.

Thistlethwaite, Frank. 1960. "Migration from Europe overseas in the nineteenth and twentieth centuries." *Population Movements in Modern European History* pp. 73–92.

Vecoli, Rudolph J. 1983. "The Formation of Chicago's" Little Italies"." *Journal of American Ethnic History* 2(2):5–20.

Ward, Zachary. 2016. "There and back (and back) again: Repeat migration to the United States 1897–1936." Unpublished manuscript, The Australian National University, Department of Economics.

Ward, Zachary. 2017. "Birds of passage: Return migration, self-selection and immigration quotas." *Explorations in Economic History* 64:37–52.

Wegge, Simone A. 1998. "Chain migration and information networks: evidence from nineteenth-century Hesse-Cassel." *The Journal of Economic History* 58(04):957–986.

Yans-McLaughlin, Virginia. 1977. *Family and community: Italian immigrants in Buffalo, 1880-1930*. University of Illinois Press.

Appendix 1

Figures and Tables

Figure A1.1: Regional US-bound emigration in abolute terms.
Source: Ellis Island administrative records.
Notes: This map has been produced using current administrative borders. There are 20 regions.

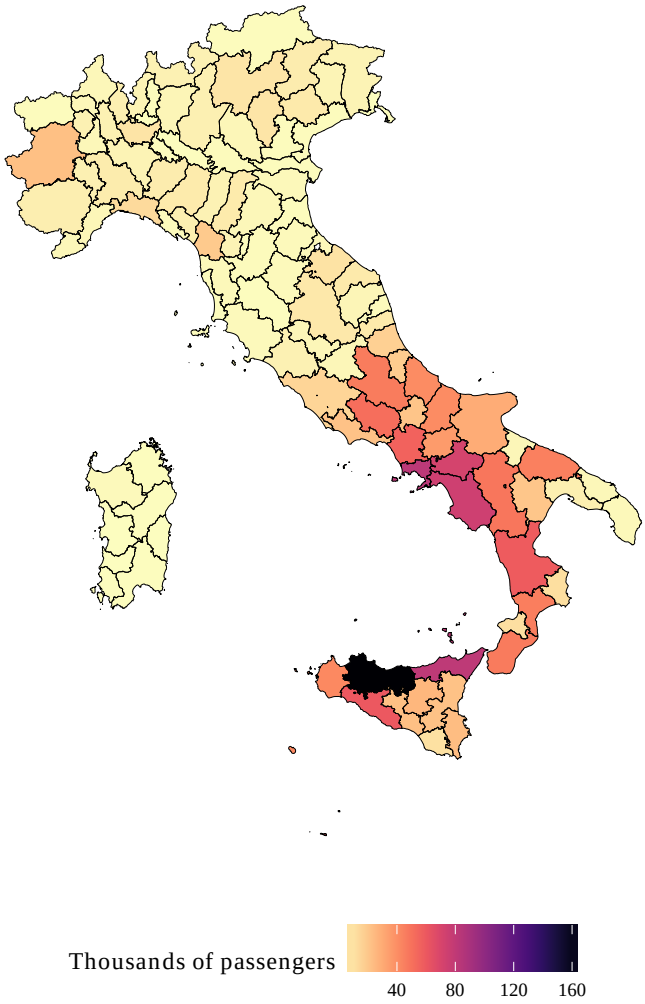


Figure A1.2: Provincial US-bound emigration in abolute terms.
Source: Ellis Island administrative records.
Notes: This map has been produced using current administrative borders. There are 110 provinces.

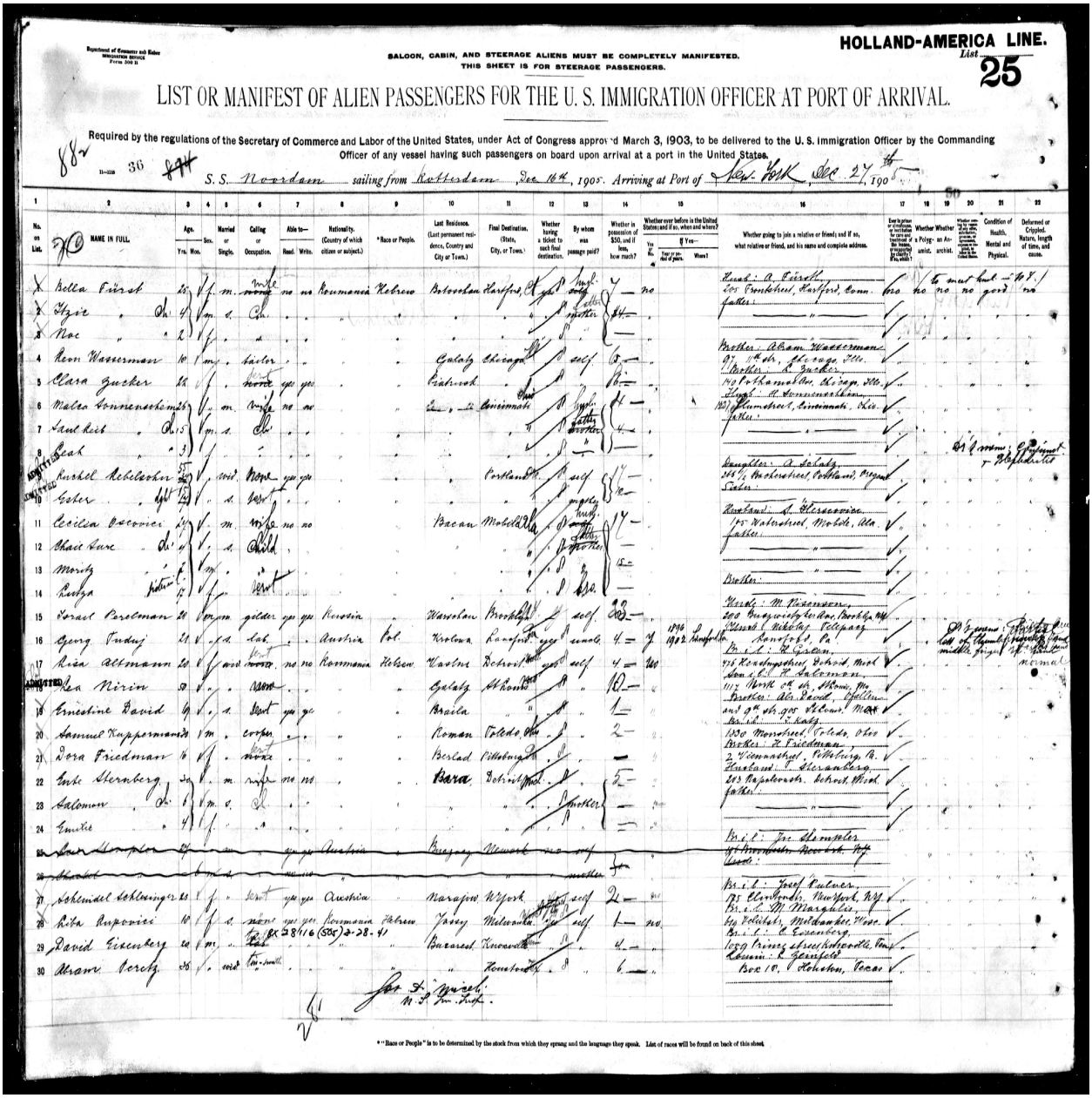


Figure A1.3: Original ship manifest.
Source: The Statue of Liberty - Ellis Island Foundation.

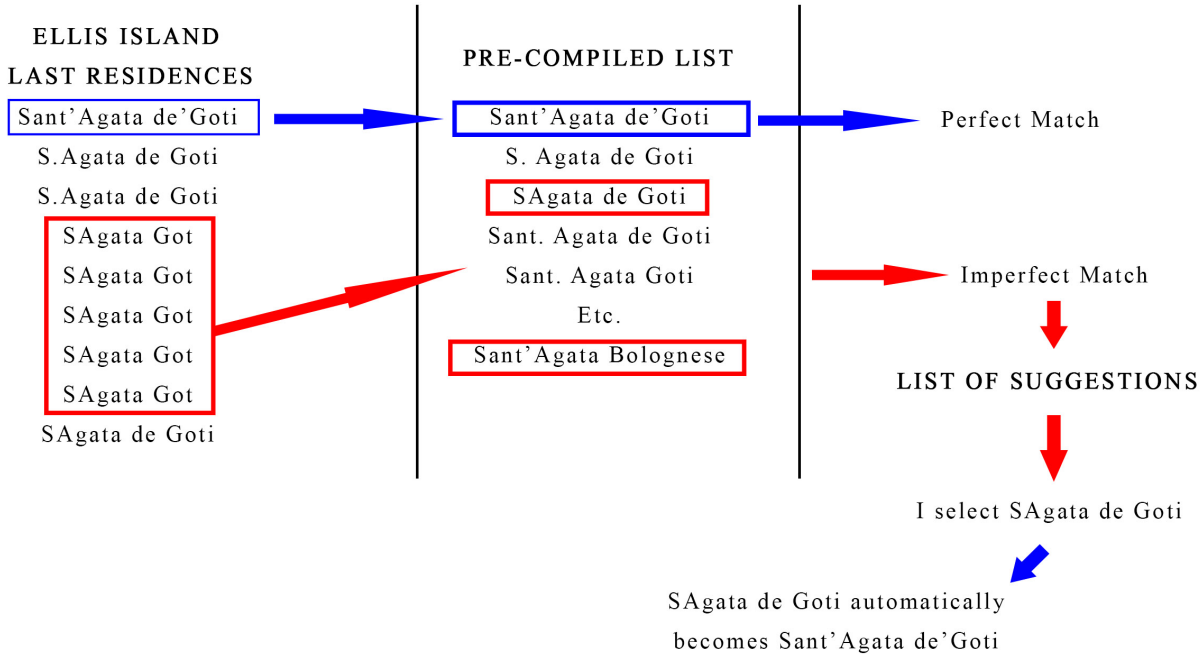


Figure A1.4: Matching scheme.
Notes: This scheme summarizes the matching process I used to construct the Ellis Island data. Ellis Island last residences are grouped by spelling mistake and confronted to a pre-compiled list that includes variants of each municipality's correct spelling. I extract perfect matches and choose the most appropriate option whenever the script returns suggestions.

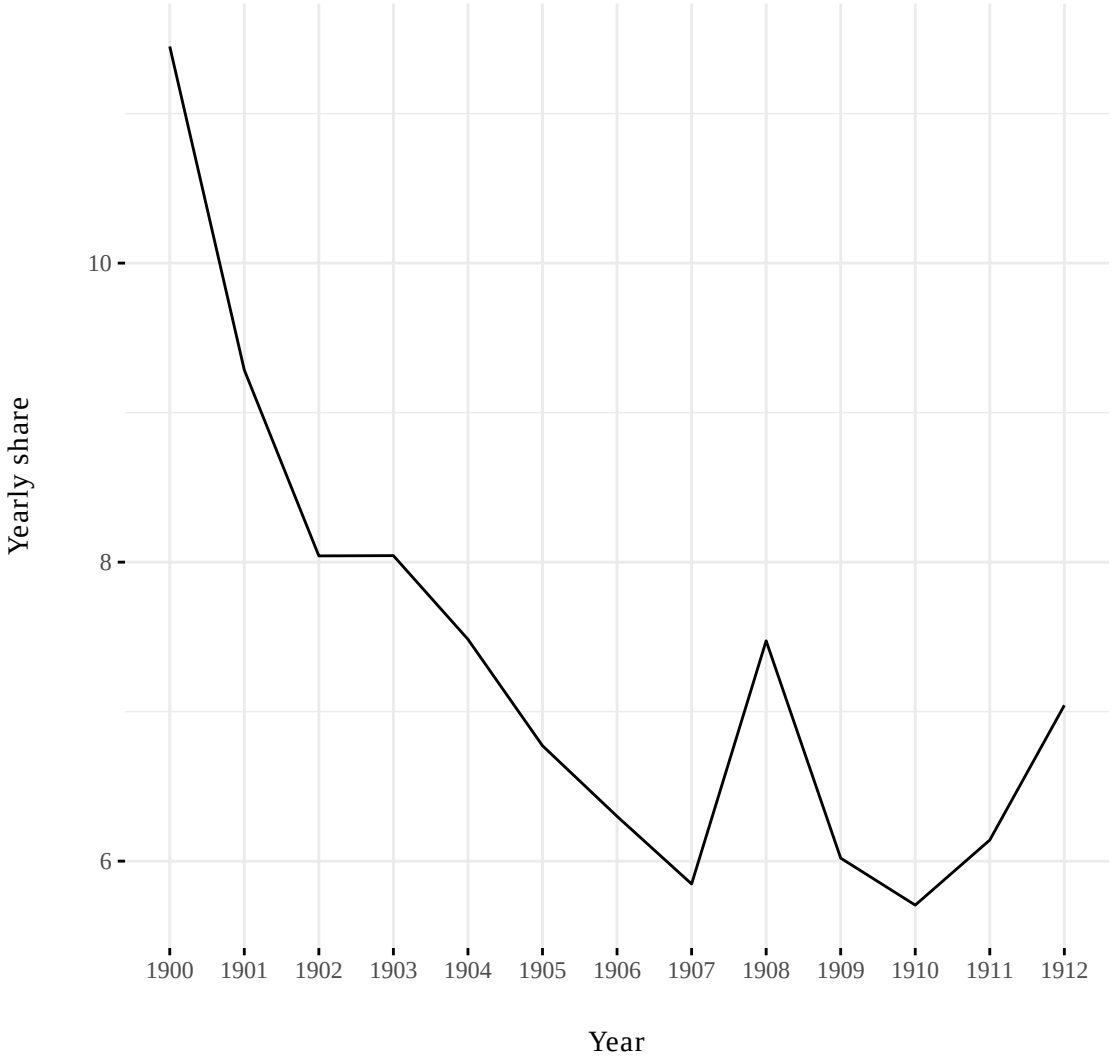


Figure A1.5: Yearly share of migrants travelling together.
Source: Ellis Island administrative records.
Notes: This plot is based on a variant of equation (1.1) and depicts the measure μ_y. It shows how the importance of ex-ante networks declined over time, proportianlly to the spread of information.

Table A1.1: Regional US-bound out-migration since 1901, absolute figures

Code	Region	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912
1	Abruzzo	6410	8616	12462	6368	15046	12805	11891	2169	10145	9201	6141	8822
2	Basilicata	3912	4938	6866	4841	6257	6581	6255	1556	4046	3806	3803	4507
3	Calabria	9551	12683	16347	9889	21118	16482	15277	3336	13667	13018	8179	10912
4	Campania	18107	23735	29039	20341	31380	31450	29759	7411	20483	17998	16390	17622
5	Emilia-Romagna	1145	1164	2066	1617	3689	3823	3712	1019	2093	2849	2137	2323
6	Friuli-Venezia G.	173	161	385	514	493	813	1221	221	685	985	585	780
7	Lazio	2889	4946	7339	4961	10177	11029	12273	1512	10995	8278	6154	8929
8	Liguria	549	698	1332	1345	1432	1332	2099	727	883	1272	1364	1565
9	Lombardia	1569	1697	2915	2293	3446	3768	4408	1363	2350	3031	2528	2926
10	Marche	918	923	1830	1272	3204	3960	3675	603	2684	2665	1953	3211
11	Molise	3825	5055	5828	3435	6088	5247	6145	1008	3702	4182	2990	3540
12	Piemonte	2351	2139	3876	2790	5046	5897	6581	2095	3854	4305	3455	4267
13	Puglia	2070	2837	5653	4657	7080	11149	11943	2647	10513	8256	6492	8101
14	Sardegna	31	45	68	50	46	120	489	126	208	403	291	520
15	Sicilia	13709	18158	31497	20187	47445	60918	54360	10419	43923	39402	23751	43079
16	Toscana	1229	1305	2480	1854	2432	3090	4797	1294	2775	3635	3188	3626
17	Trentino-Alto Adige	273	307	893	472	966	1265	1387	260	930	777	540	717
18	Umbria	283	277	649	388	1113	1476	1429	325	1183	1037	460	1198
19	Valle d'Aosta	9	27	42	24	38	72	131	42	45	86	62	51
20	Veneto	911	1047	1581	1233	2421	3246	4257	835	2175	2981	1850	2264

Notes: This table has been produced using current administrative borders. There are 20 regions.

Table A1.2: Provincial US-bound out-migration since 1901, absolute figures.

Code	Province	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912
1	Agrigento	1508	1941	3442	2676	5498	7844	8286	1076	5238	5190	3439	6939
2	Alessandria	128	213	520	304	612	807	835	206	479	626	443	544
3	Ancona	266	299	467	328	917	1172	830	154	745	579	501	629
4	Aosta	9	27	42	24	38	72	131	42	45	86	62	51
5	Arezzo	77	63	34	40	96	143	284	23	119	131	102	135
6	Ascoli Piceno	380	305	585	169	1013	1079	889	78	592	560	359	813
7	Asti	162	143	267	153	366	488	769	146	501	547	484	478
8	Avellino	4787	5943	7073	4262	6564	6103	6335	1133	4499	4390	3920	4022
9	Bari	550	905	2788	2896	3639	6129	5914	1733	6388	4848	3579	4197
10	BAT	48	62	94	104	318	461	341	86	289	312	271	276
11	Belluno	309	354	539	368	640	550	830	183	485	629	396	349
12	Benevento	2170	3138	3676	2148	4067	2974	3045	433	2700	2423	1801	1948
13	Bergamo	111	161	162	151	302	331	413	97	277	377	155	253
14	Biella	275	240	359	223	375	340	431	159	269	395	316	347
15	Bologna	54	62	200	195	202	141	245	113	155	155	154	159
16	Bolzano	67	36	80	35	94	66	91	16	41	76	56	48
17	Brescia	241	194	458	272	450	638	588	148	186	438	332	302
18	Brindisi	35	44	17	25	86	251	400	33	309	265	316	165
19	Cagliari	14	2	35	16	13	23	46	20	16	28	21	19
20	Caltanissetta	451	830	1374	867	2160	3088	3383	625	1837	1487	1472	3010
21	Campobasso	2500	3129	3766	2544	3886	3544	4391	707	2519	2848	2033	2512
22	Carbonia-Iglesias	1	9	3	6	5	19	24		8	10	71	45
23	Caserta	3238	4502	5126	3703	6277	5576	5338	1363	4368	3103	2419	3151
24	Catania	227	244	663	659	1996	3210	3592	640	2745	2047	1549	3247
25	Catanzaro	2970	3580	3839	2727	5792	4192	3621	853	3234	3541	2057	2543
26	Chieti	2093	3423	4745	2124	5245	3845	3671	800	3031	2483	2275	3059
27	Como	151	171	315	192	370	358	449	149	257	206	217	289
28	Cosenza	3562	4538	5199	3167	5933	4986	4659	1371	4274	4367	3167	3464
29	Cremona	5	18	30	20	24	68	98	31	43	40	42	62
30	Crotone	522	881	1368	410	1520	1249	1121	168	830	836	600	819
31	Cuneo	229	232	352	291	486	603	802	258	455	503	345	450
32	Enna	541	675	1341	891	2403	3690	4204	771	2669	2168	1544	2935
33	Fermo	18	20	121	68	91	95	115	33	84	139	84	158
34	Ferrara	130	60	62	80	137	110	239	45	67	70	94	73
35	Firenze	38	21	39	88	63	94	167	80	101	64	62	122
36	Foggia	1245	1627	2417	1410	2659	3601	4260	628	2765	1862	1730	2795
37	Forlì-Cesena	30	18	31	36	127	138	147	30	92	165	139	196
38	Frosinone	1807	3297	4112	2341	6393	6043	5916	584	5768	4648	3017	5128
39	Genova	436	420	919	1016	1014	922	1362	543	587	780	916	1012
40	Gorizia		1	5	20	21	13	21		13	9	14	17

Notes: This table has been produced using current administrative borders. There are 110 provinces.

Table A1.2: Provincial US-bound out-migration since 1901, absolute figures, ctd.

Code	Province	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912
41	Grosseto	3	13	27	13	129	186	192	19	120	219	71	168
42	Imperia	27	126	98	66	160	143	178	55	114	161	87	90
43	Isernia	1325	1926	2062	891	2202	1703	1754	301	1183	1334	957	1028
44	L'Aquila	2842	3288	4226	2431	5656	4443	4330	853	3920	4120	2101	2853
45	La Spezia	19	23	108	65	49	42	189	56	79	108	128	151
46	Latina	852	1312	2117	1241	2043	2370	2520	411	1957	1412	1505	1744
47	Lecce	68	65	66	87	77	181	223	87	148	178	133	127
48	Lecco	28	46	62	67	48	90	122	34	59	69	60	61
49	Livorno	26	14	50	35	36	54	42	15	62	26	62	65
50	Lodi	6	2	20	9	23	11	41	20	6	15	15	14
51	Lucca	736	847	1428	1175	1209	1374	2374	828	1461	1873	1968	2004
52	Macerata	114	76	232	152	331	428	295	93	268	262	201	424
53	Mantova	21	13	44	46	163	176	139	43	50	59	60	67
54	Massa-Carrara	180	95	388	256	383	486	880	118	514	500	373	426
55	Matera	1101	1526	2087	899	1917	2111	1750	469	1382	1252	1181	1538
56	Medio campidano		3		1	1	8	4	2	12	39	4	9
57	Messina	3086	3647	6984	4642	10814	11103	8308	1742	7927	6303	3596	5433
58	Milano	483	439	767	562	797	897	949	393	672	697	561	744
59	Modena	248	291	554	284	670	1091	757	355	476	588	395	374
60	Monza	1	2	15	6	14	12	16	5	7	12	7	12
61	Napoli	3573	4583	5858	5209	7058	9247	8450	2571	4949	4134	4224	4277
62	Novara	193	133	315	239	462	597	642	89	468	314	264	272
63	Nuoro	1	7	2	4	3	2	92	2	15	47	24	30
64	Ogliastra	1					3	1		3		2	3
65	Olbia-Tempio		10	2	7	2	11	38	3	9	58	8	164
66	Oristano	6	3	5	1	2	12	11		9	9	9	10
67	Padova	8	30	39	44	55	51	142	26	38	40	51	62
68	Palermo	7036	9503	14502	8397	17889	20480	15106	3850	13618	12189	7451	11764
69	Parma	216	273	434	277	688	756	752	214	417	728	678	605
70	Pavia	166	154	227	286	372	416	486	143	217	267	313	332
71	Perugia	271	239	569	352	916	1097	861	292	941	731	340	791
72	Pesaro e Urbino	140	223	425	555	852	1186	1546	245	995	1125	808	1187
73	Pescara	896	1321	1924	847	1959	1702	1883	247	1675	1324	1005	1440
74	Piacenza	290	290	440	490	702	734	705	85	344	511	424	362
75	Pisa	36	42	135	96	132	234	273	53	95	235	169	171
76	Pistoia	84	158	331	104	303	434	484	132	239	496	332	462
77	Pordenone	75	79	140	88	170	337	547	107	305	429	358	371
78	Potenza	2811	3412	4779	3942	4340	4470	4505	1087	2664	2554	2622	2969
79	Prato	38	27	19	22	42	61	39	11	31	65	36	44
80	Ragusa	35	54	132	109	638	1275	1482	225	1674	1821	983	1797

Notes: This table has been produced using current administrative borders. There are 110 provinces.

Table A1.2: Provincial US-bound out-migration since 1901, absolute figures, ctd.

Code	Province	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912
81	Ravenna	27	11	26	23	43	48	109	26	65	48	28	27
82	Reggio Calabria	1875	2986	4763	2990	6386	5096	4949	826	4410	3105	1863	3190
83	Reggio Emilia	115	141	239	210	986	549	427	127	330	392	137	236
84	Rieti	63	16	122	41	185	302	407	39	353	405	309	325
85	Rimini	35	18	80	22	134	256	331	24	147	192	88	291
86	Roma	148	291	733	1302	1326	1665	2500	437	2208	1188	1016	1184
87	Rovigo	7	20	20	27	55	92	141	62	45	65	23	51
88	Salerno	4339	5569	7306	5019	7414	7550	6591	1911	3967	3948	4026	4224
89	Sassari	8	11	21	15	20	42	273	99	136	212	152	240
90	Savona	67	129	207	198	209	225	370	73	103	223	233	312
91	Siena	11	25	29	25	39	24	62	15	33	26	13	29
92	Siracusa	156	230	733	487	2070	3641	4267	738	3035	2772	1622	3111
93	Sondrio	140	216	344	278	363	395	616	134	308	408	393	339
94	Taranto	124	134	271	135	301	526	805	80	614	791	463	541
95	Teramo	579	584	1567	966	2186	2815	2007	269	1519	1274	760	1470
96	Terni	12	38	80	36	197	379	568	33	242	306	120	407
97	Torino	1098	949	1776	1310	2354	2654	2448	1031	1347	1653	1310	1749
98	Trapani	669	1034	2326	1459	3977	6587	5732	752	5180	5425	2095	4843
99	Trento	206	271	813	437	872	1199	1296	244	889	701	484	669
100	Treviso	125	191	260	322	426	865	1386	250	544	749	658	881
101	Trieste	1	3	28	27	40	62	40	12	33	32	16	43
102	Udine	97	78	212	379	262	401	613	102	334	515	197	349
103	Varese	216	281	471	404	520	376	491	166	268	443	373	451
104	Venezia	46	14	30	40	51	78	90	36	28	44	48	44
105	Verbano-C.-O.	170	144	160	174	204	175	314	96	165	129	159	185
106	Vercelli	96	85	127	96	187	233	340	110	170	138	134	242
107	Verona	118	125	248	181	291	332	278	100	157	257	220	201
108	Vibo Valentia	622	698	1178	595	1487	959	927	118	919	1169	492	896
109	Vicenza	298	313	445	251	903	1278	1390	178	878	1197	454	676
110	Viterbo	19	30	255	36	230	649	930	41	709	625	307	548

Notes: This table has been produced using current administrative borders. There are 110 provinces.

Chapter 2

Globalization, Agricultural Markets and Mass Migration

Italy, 1881-1912

Rowena Gray,[*] *Gaia Narciso,*[‡] *Gaspare Tortorici*[¶]

March 2019

Abstract

Despite the significant attention paid to the current consequences of globalization for migration behavior, there are only a few historical accounts about the effect of commodity market integration at a disaggregated level. We set our paper within the context of the first globalization era, when migration flows were largely unregulated, and highlight how exogenous shocks in agricultural commodity prices influenced international migration flows from Italian provinces between 1881 and 1912. To do this, we construct an index of global price exposure based on initial provincial agricultural production structures. Our analysis quantifies the contribution of globalization-induced agricultural price shocks to migration decisions, alongside more traditional explanatory factors such as migrant networks and landholding systems. We find evidence that agricultural price shocks are positively related to the propensity to migrate, as migration tended to increase in proportion with agricultural commodity prices. This result suggests that liquidity constraints were binding until agricultural incomes reached a certain threshold. These findings can inform our understanding of present-day migration responses in developing countries, in the face of even more rapid globalization but higher barriers to legal migration.

Keywords: *Age of mass migration; determinants of migration; agricultural shocks*
JEL Codes: N93; N13; F22; O15

[*]University of California (Merced), [‡]Trinity College Dublin, [¶]Trinity College Dublin.
The authors wish to thank Kris Mitchener, Gabriele Cappelli, Ruben Durante, Cormac Ó Gráda, Andrea Guariso, Tim Hatton, Martina Kirchberger, Ronan Lyons, Bruno Morando, Hillel Rapoport, Davide Romelli, Yannay Spitzer, Greg Wright, Ariell Zimran, and seminar participants at Trinity College Dublin, the Paris School of Economics and the "Migration and the Labour Market" workshop. We also thank Isaac Dempsey for providing excellent research assistance. Tortorici acknowledges funding from the Irish Research Council.

2.1 Introduction

The consequences of globalization for migration have been recently explored in the context of developing countries. For example, Dix-Carneiro and Kovak (2017) find that Brazilian trade liberalization had little effect on migration, while Majlesi and Narciso (2018) provide evidence that Chinese import competition reduced Mexican migration to the US. Our paper takes a historical perspective and investigates the effect of globalization-induced agricultural price shocks on international out-migration from Italian provinces over the period 1881 to 1912. This period was largely free of binding legal barriers to migration, allowing us to focus on the theoretical tension between two opposite mechanisms through which price shocks might affect migration. First, negative agricultural price changes may reduce agricultural incomes and create a labor surplus that may not be absorbed by an infant industrial sector, thus incentivizing migration. Second, negative agricultural price changes may reduce agricultural incomes and limit the ability to cover the up-front cost involved in any migration decision. Our analysis provides evidence that agricultural price shocks were positively related to the propensity to migrate and suggests that liquidity constraints were binding. We thus provide new evidence on this issue, a key component of the puzzle identified by Hatton and Williamson (2005): during the initial stages of the first globalization era, sending regions were all in northwestern Europe despite wage differentials with the New World were higher in southern countries; on the other hand, migration rates typically increased proportionally to economic growth, suggesting that liquidity constraints must have been large enough to counteract rising incomes.[1] Using annual migration flows and aggregate GDP data, Covarrubias, Lafortune and Tessada (2015) find that liquidity constraints were indeed binding before WWI. We further extend the debate by presenting evidence based on finer data and a more plausibly exogenous shock to income.

[1] Hatton (2010) highlighted the difference between countries like Ireland, where narrowing wage differentials reduced emigration, and Italy, where convergence was lower. Using modern data, Belot and Hatton (2012) show that any study of migrant selection must take into account the fact that those with the least skill typically cannot afford to migrate.

In order to estimate whether agricultural price shocks foster or restrain migration within the above-outlined framework, we construct an index of international price exposure – a proxy for commodity market integration during the first globalization era[2] – based on the agricultural production structure of each province in Italy. The intuition behind this measure is straightforward: international agricultural price shocks affect each geographical unit of analysis differently, proportionally to the share of land allocated to each crop. Our analysis suggests that migration increased proportionally to agricultural prices. For example, looking at a period of rising prices, from 1893 to 1900, we estimate that the resulting boost to agricultural incomes can explain up to one half of the rise in the total migration rate at that time. We argue that higher agricultural commodity prices translated into higher income, thus relaxing liquidity constraints and allowing would-be migrants to afford passage abroad.[3]

We also investigate the role of income uncertainty – proxied by international agricultural commodity price fluctuations – and show how increased uncertainty tended to increase migration rates. Overall, our results suggest that Italians were more sensitive to changes in the level of income than its volatility, the effect of the latter being about one third as large as of the former.[4]

We focus on Italy during the years 1881 to 1912 for three reasons. First, Italian pre-WWI migration represents one of the most significant voluntary, mainly economic-driven, population movements in history: more than 11 million Italian migrants left Italy from 1870 to WWI (ISTAT, 2011), with only Russian Jews being more likely to migrate.[5] Second, Italian international out-migration presents an extraordinary level of – largely unexplained – spatial

[2]Throughout the paper we use the term globalization – or first globalization era – as shorthand for the general process of labor movements and commodity market integration across the Atlantic Ocean.

[3]Symmetrically, the phenomenon reversed in the face of sharp price plunges. Harvey et al. (2017) document a positive linkage between commodity prices and short-run incomes.

[4]In theory, we might expect that volatility would have affected seasonal more than gross migration. However, our data do not allow to make a distinction between temporary and permanent migrants. Therefore, we cannot test this hypothesis. Hatton (1995)'s model of migration explores the role of employment uncertainty as an emigration determinant. Moreover, data for UK outflows suggest that volatility explains short-run fluctuations more than long-run trends.

[5]Spitzer (2016) documents that, as of 1897, one third of the entire Russian Jewish population had emigrated.

and temporal heterogeneity across provinces, in terms of destinations and magnitudes (Fauri, 2015). Finally, the consequences of agricultural commodity market integration on migration are more likely to materialize in developing economies largely dependent on the primary sector. Indeed, about two thirds of the Italian labor force was employed in agriculture at the beginning of the 20th century (Felice, 2017). Moreover, Italian wages were significantly lower than the European average and roughly equivalent to a fifth of those in the US (Williamson, 1995). Anecdotally, a series of government inquiries (Franchetti and Sonnino 1877; Jacini 1886; Faina 1909) offered a dismal depiction of agricultural laborers across the peninsula. Often illiterate and exposed to malarial illnesses, most farmers lived in near-subsistence conditions.

The time-frame of our study is of particular interest because it is part of a broader age of convergence started in the 1850s and ended with WWI. O'Rourke and Williamson's *Globalization and History* attributes this global phenomenon to the "*open economy forces of trade and mass migration*" (O'Rourke and Williamson, 1999, p. 5). The era was also characterized by significant commodity price swings, somewhat masked by an overall downward trend for most commodities, as Asian silk, Indian rice and Russian and American cereals flooded European markets for the first time (O'Rourke, 1997). Such swings created what Díaz Alejandro (1984) has called the "*commodity lottery*", whereby price movements generated winners and losers as globalization forces increased exposure to shocks in the international market and often made terms of trade less favorable for peripheral and developing countries.

Our paper contributes to the literature on migration at large, with a specific focus on agricultural commodity market integration as a key determinant. Indeed, we construct emigration rates from Italian provinces and examine the role of agricultural price shocks as a driver of migration. According to O'Rourke (1991), the decline of agricultural labor demand in rural Ireland should be imputed to the Great Famine rather than international commodity price shocks. Indeed, he argues that the structure of the Irish agriculture drastically changed following a shift away from tillage and potato due to the blight. We provide indirect evidence

that supports this line of reasoning: while out-migration from Italy morphed into a mass phenomenon, the agricultural production structure of Italian provinces was not overthrown by price plunges - which partially recovered after 1894.

Dis-aggregation is a crucial element of our analysis: given the substantial cross-province heterogeneity of migration, the focus on national data may hide local factors driving differential rates.[6]

A handful of papers focus on the role of income shocks in historical migration decisions. Spitzer (2015) analyses the impact of temporary income shocks – proxied by changes in agricultural yields – on Jewish migration from Russia. He shows that economic shocks might explain the timing of migration but underlying forces such as wage differentials and networks remained key driving factors. Abramitzky, Boustan and Eriksson (2012, 2013) focus on Norwegian migration to the US and exploit variation in expected inheritance based on birth order to show that wealth was negatively correlated with migration, arguing that self-selection was negative from urban areas. Persaud (2017) focuses on income volatility - proxied by fluctuations in rice prices - as a push factor for Indian indentured servant migration.

We find further evidence that Italian would-be migrants were liquidity constrained during our period of reference. Coan (1997) presents interviews with several Italian-Americans who entered the US through Ellis Island. *Mario Vina* reports not having being able to migrate immediately after his father as remittances were used to hospitalize a sick relative. He describes how "*in 1909, the year we came, my father got smart. He sent no cash, just the paid tickets*" (Coan, 1997, p. 38). Others vividly describe the role of networks in facilitating migration: *Peter Mossini* was only able to leave Sicily in 1921 despite his father had worked in the US between 1901 and 1912; indeed, he only sent enough money to cover the passage for one of his children at a time, forcing Peter to wait until his siblings earned enough to pay for his transoceanic voyage (Coan, 1997, p. 45).

[6]Echoing Moretti (1999), recent contributions - such as Ardeni and Gentili (2014) - have stressed the desirability of sub-national examinations. This is evident in Spitzer and Zimran (2018): the authors present anthropometric evidence that, while negative at the national level, the degree of self-selection of US-bound Italian passengers varied across provinces, being more positive in southern areas.

There is anecdotal evidence that during the early stages of Italian migration many small landowners and day laborers would borrow to purchase their tickets (Bevilacqua, 2001). Given that credit markets were underdeveloped, moneylenders were likely to be the main source of loans. On the other hand - as already argued in Chapter 1 - repeat migration was a very common phenomenon, especially from municipalities located in comparatively poorer provinces. The liquidity constraint argument developed in this chapter might therefore appear to be somehow counter-intuitive, i.e. how can migrants finance multiple journeys if liquidity constraints were binding? Both phenomena - liquidity constraints and repeat migration - could have coexisted if income never reached the threshold beyond which it constitutes a restraining force to migration. In other words, income growth represented a push factor throughout the period and its importance declined over time as would-be migrants became connected to networks. Within this framework, migration is assumed to be a multi-staged investment, evidently well-worth the up-front cost involved in it. Once a migrant could fiance its journey for the first time, successive returns and chain migration were made possible by good employment/salary opportunities abroad.

2.2 Data

Our data set is the result of a significant digitization effort. We combine annual information on international emigration and population at the province level along with agricultural production structures, international agricultural commodity prices and landholding systems.

2.2.1 Emigration Data

Given that provincial emigration data are only available in absolute terms, the reconstruction of annual migration rates required the digitization of population series – to be found in yearly publications made available by ISTAT (Direzione Generale della Statistica, ad annum).[7]

[7]Inter-census population data are available on a yearly basis only after 1896. We conduct a robustness check constructing out-migration rates using population figures taken from the 1881 Census. Results are shown in table A2.12 and support the main findings discussed below.

Figures are taken from the *Annuario statistico dell'emigrazione italiana dal 1876 al 1925* (ISTAT, 1926), a comprehensive volume that contains the universe of emigration statistics produced by the two main governmental migration agencies operating during our reference period: the *Direzione Generale della Statistica* (henceforth, DGS) and the *Commissariato Generale dell'Emigrazione* (henceforth, CGE), active since 1901.[8] We improve on the existing literature by employing a longer time series, 1881-1912, at a significantly higher level of disaggregation, i.e. the 69 Italian provinces rather than the 16 regions. The most closely related studies are Gomellini and Ò Gràda (2013) – who carry out a regional analysis over a similar time frame – and Hatton and Williamson (1998) – who employed provincial data for only two census years, 1901 and 1911.

As discussed in Chapter 1, DGS series are available since 1876 – unlike CGE figures, which are only available from 1902 onward. We therefore opted to focus on those. According to Italian law, would-be migrants had to undergo a two-step procedure to leave the country. First, they had to apply for a *nulla osta*, a document certifying their eligibility for a passport – which was then issued by local authorities. Second, they had to collect the actual passport – which was then valid for 3 years. DGS series are based on the number of *nulla osta* issued by each municipality, aggregated to the province level.

These data exhibit two caveats. First, would-be migrants could decide not to leave and, in rare cases, local authorities could reject passport applications. Conversely, several destination countries allowed European migrants to enter without passports, decreasing incentives to follow standard legal procedures in sending countries (Keeling, 2013).[9] The DGS series thus measures emigration with error, although there is no reason to expect any systematical variation across provinces. Second, the DGS series only differentiates between European and overseas emigration at large, masking destination-specific patterns as well as any gender, age,

[8]The two bodies operated autonomously and developed different, often contradictory, definitions of migrants. They also used different data collection techniques. Such inconsistencies help explain why, despite similar overall trends, there might be discrepancies in absolute numbers. See section 1.2 in Chapter 1 and Bevilacqua (2001) for an in-depth treatment of the matter.

[9]Italian migrants could also depart from other European ports, despite they were located further away from origin municipalities. In this case, passports were not required to embark on transoceanic vessels.

and occupational dimensions. The occupational decomposition of flows is available at the national level and confirms that about 70% of Italian passport applicants were employed in the agricultural sector. In light of this piece of evidence, our analysis is likely to capture the main driving forces of provincial out-migration.

Our paper shares a limitation with the existing literature as it is not possible to account either for return or for seasonal migration. Indeed, recent contributions such as Bandiera, Rasul and Viarengo (2013)[10] have highlighted how considerable these phenomena were during the era of mass migration. Since we are interested in investigating how a changing economic environment influences migration at large, such constraints should not invalidate our analysis despite inevitably inflating the measure of migration stocks at destination.

In order to capture US-bound migration, we complement provincial figures with Ellis Island administrative records.[11] Indeed, we make use of the data set resulting from the process described in Chapter 1. In particular, we employ yearly data at the province level from 1892 to 1912. The cross-column differences in sample sizes in the tables below are mainly attributable to the fact that US-bound migration figures are only available since 1892 - rather than 1876. Given that we could not match all migrants to their last residence of origin, this series represents a lower bound estimate of the actual US-bound flow due to attrition.[12]

The emigration data presented in figure 2.1 - at page 60 - show a significant degree of variability across provinces, regardless of destination choices. Emigration rates approach zero for some southern provinces in the early 1880s, while *maxima* are driven by the provinces of Udine and Belluno, in the north-east. These areas experienced intensive emigration toward European countries – even compared with the later emigration rates of southern provinces toward the US.[13] Table 2.1 - in the following page - presents summary statistics for the

[10]The authors argue that return migration flows during this period were much higher than previously thought.

[11]See also Bandiera, Rasul and Viarengo (2013), Spitzer and Zimran (2018), and Ward (2017).

[12]For completeness, note that information on last residence is particularly coarse before 1901. Our analysis might therefore underestimate emigration to Ellis Island throughout the 1890s.

[13]Before 1900, Europe and South America were popular destinations; after 1900, the US became increasingly dominant – despite a sizable stream of Italian migrants still opting for Europe.

Table 2.1: Summary Statistics

Variables	Minimum	Maximum	Mean	SD
GPE (*log*)	2.18	2.31	2.26	0.02
GPE, 5-year volatility (*log*)	1.65	2.04	1.91	0.06
Global emig. rate per 1000	0.00	143.75	14.85	16.32
Global emig. rate per 1000 (*log*)	-6.46	4.97	1.98	1.58
Global migration network	0.00	1228871.00	62131.30	106544.89
Transoceanic emig. rate per 1000	0.00	58.68	7.99	10.02
Transoceanic emig. rate per 1000 (*log*)	-6.35	4.07	1.15	1.68
Transoceanic migration network	0.00	324715.00	33373.53	50096.87
US emig. rate per 1000	0.00	26.12	2.74	4.27
US emig. rate per 1000 (*log*)	-6.72	3.26	-0.41	2.03
US migration network	0.00	155449.00	4768.44	13511.42
European emig. rate per 1000	0.00	137.38	6.95	13.63
European emig. rate per 1000 (*log*)	-6.46	4.92	0.63	1.93
European migration network	0.00	1106798.00	29094.44	87759.11
Population (1881 census)	114295.00	1114991.00	412458.67	222663.86
Yearly population	114295.00	1743723.00	466499.81	269462.58
Share of non-landowners	41.11	97.32	84.43	10.25
Provincial industrialization index	0.43	2.26	0.91	0.34
Year			1881	1912

Notes: The GPE index and the 5-year volatility measure are defined, respectively, as in equations (2.1) and (2.2). Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. Provincial industrialization in based on Ciccarelli and Fenoaltea (2012). Migration data for overall, transoceanic and European migration has been digitized from the *Annuario statistico dell'emigrazione italiana dal 1876 al 1925* (1926). US-bound migration figures are based on Ellis Island data.

variables used in the analysis. Migrant networks are defined as the one-year lagged sum of absolute migration.

2.2.2 Global Price Exposure Index and other Variables

In order to investigate the impact of agricultural commodity price shocks on international out-migration from Italian provinces, we construct a Global Price Exposure (GPE) index that captures provincial exposure to global price movements across relevant agricultural commodities produced in Italy during the sample period. We construct this measure annually for the period 1881 to 1912.

We proxy international commodity prices with import prices net of tariffs and bilateral

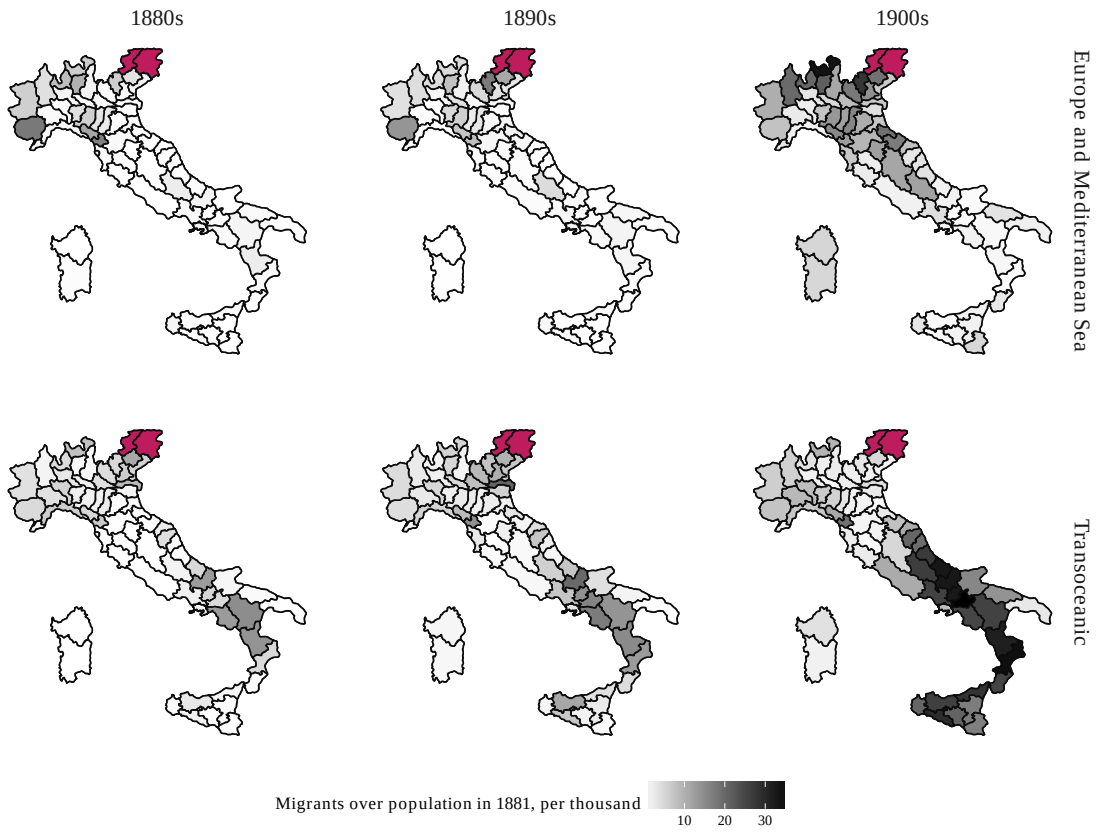


Figure 2.1: Provincial emigration rates per decade.
Source: Annuario Statistico dell'emigrazione italiana dal 1976 al 1925 (1926).
Notes: These maps were created without including the provinces of Udine and Belluno - highlighted in dark red - because emigration rates from both areas were high enough to mask the actual cross-province variation elsewhere. We included both Udine and Belluno in the analysis. Estimation results are robust to their exclusion.

agreements, digitized from the *Annuario Statistico Italiano* (henceforth, ASI).[14] This choice is motivated on the basis of endogeneity concerns as commodity prices are required to be exogenous with respect to the economic activity of each province. Indeed, given that pre-WWI Italy can be regarded as a developing country, it was not a price maker in international grain markets.

Furthermore, we note that any changes in trade policies would be absorbed by the year fixed effects we employ in all specifications, as they are enforced nationally. Figure 2.2 - in the following page - shows that the production structure of Italian provinces was quite

[14]The price series covers the period from 1881 to 1912. We check whether series are non-stationary using the Augmented Dickey-Fuller test.

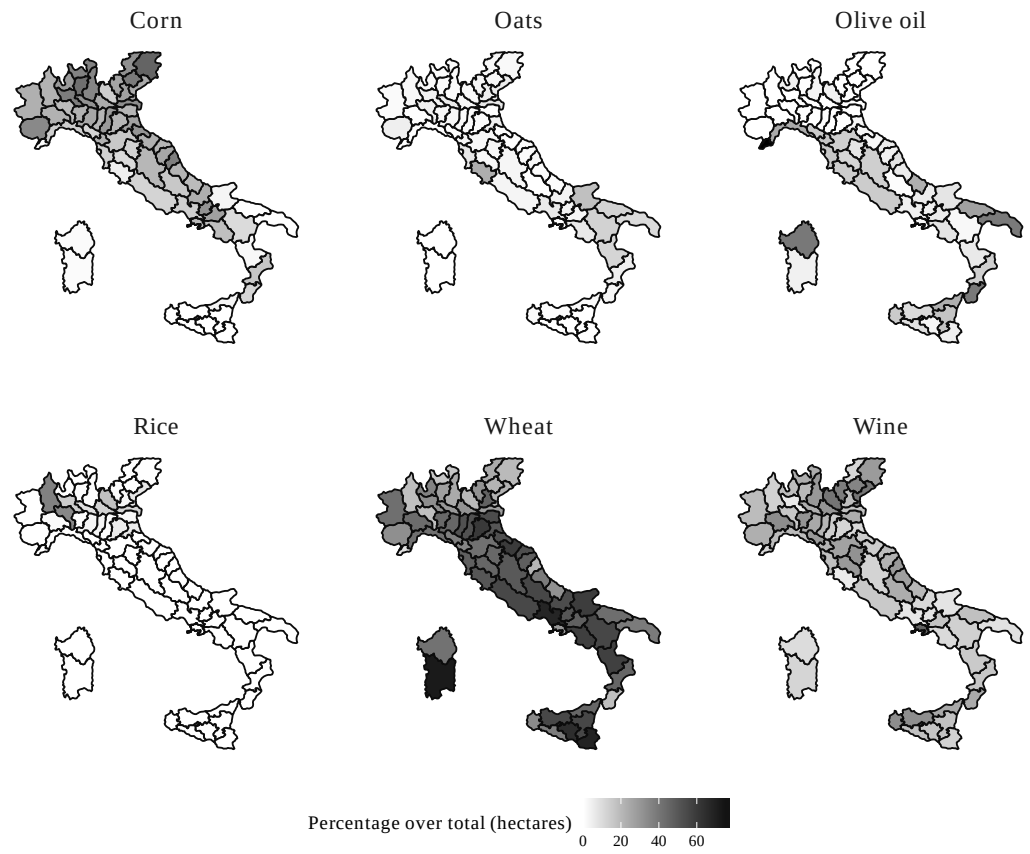


Figure 2.2: Provincial crop distribution in terms of acreage.
Source: Annuario Statistico Italiano (1886).
Notes: These maps are based on the share of land allocated to each of the major crops farmed in Italy (wheat, corn, rice, oats, wine and olive oil), computed as an average over the period 1876-1881.

diverse – i.e. different crops were produced in different areas – and fluctuations in, say, wheat prices might not affect provinces equally. In order to account for such heterogeneity, we weigh price movements using purposely-digitized information about crop allocations in terms of acreage, averaged over the period 1876-1881.[15] Our measure incorporates all major agricultural commodities produced in Italy: wheat, corn, oats, wine, olive oil and rice. Data have been digitized from the ASI (1886).[16]

Figure 2.2 highlights that wheat was the most farmed crop across the peninsula. Given

[15] As an alternative specification, we construct a GPE index based on crop shares averaged over the period 1879-1883. Table A2.2 in the appendix presents the results using this alternate GPE definition.

[16] To the best of our knowledge, there is no systematic recording of acreage data at the province level after 1894. Data are available as 1876-1881 averages (ASI, 1887-1888) or yearly measures for 1891 and 1894 (ASI, 1892, 1895).

that our GPE index is essentially a weighted average, a significant part of the exogenous variation we are exploiting in order to identify the effect of international price shocks on provincial out-migration rates is tied to wheat prices. While producing wheat, Italy was also a net importer of it. Corn represented a semi-substitute for wheat, especially in the north-west, while rice was mainly farmed along the Po river and in certain provinces of Piedmont. While not prevalent, the production of olive oil and wine was common across the country.

In constructing our GPE, we employ acreage rather than output because many authors have expressed concerns about the reliability of historical Italian data on provincial crop production.[17] Following the recent literature, we keep the share of land assigned to each crop fixed to the beginning of our time frame in order to avoid potentially endogenous re-allocation issues.[18] While provincial crop mixes might have evolved over time, changes were not significant enough to determine a radical shift in crop prevalence within each unit of analysis - see Figure A2.1 and A2.2. Our Global Price Exposure (GPE) index is constructed as follows:

$$GPE_{i,t} = \sum_{j=1}^{J} \gamma_{i,j,76-81} \times P_{j,t} \qquad (2.1)$$

where $\gamma_{i,j,76-81}$ represents the 1876-1881 average share of crop j farmed in province i, and $P_{j,t}$ is the international price level of crop j at time t. As already hinted at, the use of pre-period crop shares, $\gamma_{i,j,76-81}$, mitigates concerns about the endogenous response of local crop distribution following price changes. Given that the GPE index can be thought as a weighted average, it is in line with the recent measures of local labor market exposure to import competition developed by Topalova (2007) and Autor, Dorn and Hanson (2013); while reflecting provincial exposure to agricultural price changes over time, it is conceptually related to the the literature that examines the competition-induced consequences of China's

[17] See, for example, Fenoaltea (2011) who argues that the agricultural crisis of the 1880s might actually be attributed to misreporting. This is also the reason why we avoid computing yields.

[18] Note that we use average shares as yearly measures might be influenced by crop rotations. Our approach represents a contribution to the literature as most studies only account for the main crop, as in Bazzi (2017) and Persaud (2017).

accession to the WTO on international labor markets.[19] It can be also compared to the composite commodity price indices employed in the macroeconomics literature.[20]

We argue that international prices are driven by global demand and supply forces and are independent of changes in these factors at the province level. We expect that changes in supply costs dominated during these years, given the advances in transportation technology.

Figure 2.3 - in the following page - displays the clear downward trajectory of international real prices (in 1912 *lire*) of the six major agricultural commodities farmed in Italy up to the mid-1890s. Prices of the dominant crop, wheat, reached their *nadir* in 1894, at about half their 1879 level. Prices partially recovered thereafter but exhibited evident year-to-year fluctuations. Italian wheat exports dropped to zero while imports followed an upward trend.

We argue that these price movements were perceived by Italian farmers as a result of agricultural commodity market integration. One concern might be that declining/raising international crop prices could have an ambiguous *a priori* effect, translating themselves into both lower/higher incomes and domestic consumption prices. However, such debate is interwoven with a discussion on tariffs and trade policies. As detailed in section 2.3, grain protection varied over time and increased as international prices fell – partly offsetting the possible benefits of cheaper imports, a positive supply shock. We use international prices in order to avoid endogeneity issues. In other words, we wish to use only the part of prices that is determined by global demand and supply forces - i.e. that is not influenced by Italian production. On the other hand, there is evidence of rising living standards and consumption from the 1880s onward (Fenoaltea, 2011). This is not in contrast with our central hypothesis because it relies on the idea that out-migration increased following the weakening of liquidity constraints, not because of increased hardship.[21]

[19] See, for example, Pierce and Schott (2016) and Acemoglu et al. (2016).

[20] See, for example, Harvey et al. (2017). The authors improved on the well-established Grilli-Yang series by looking at a wider range of commodities, adding oil and metals to agricultural staples.

[21] In order to investigate this issue further, we excluded provinces where wheat or corn ware a comparatively small fraction of production, i.e. less than 20 per cent of the total acreage. In other words, we want to analyze how price fluctuations worked where the cost of living channel was prevalent. The results are shown in table A2.15 and are comparable to those of table 2.4, although there is some loss of precision due to a smaller sample size.

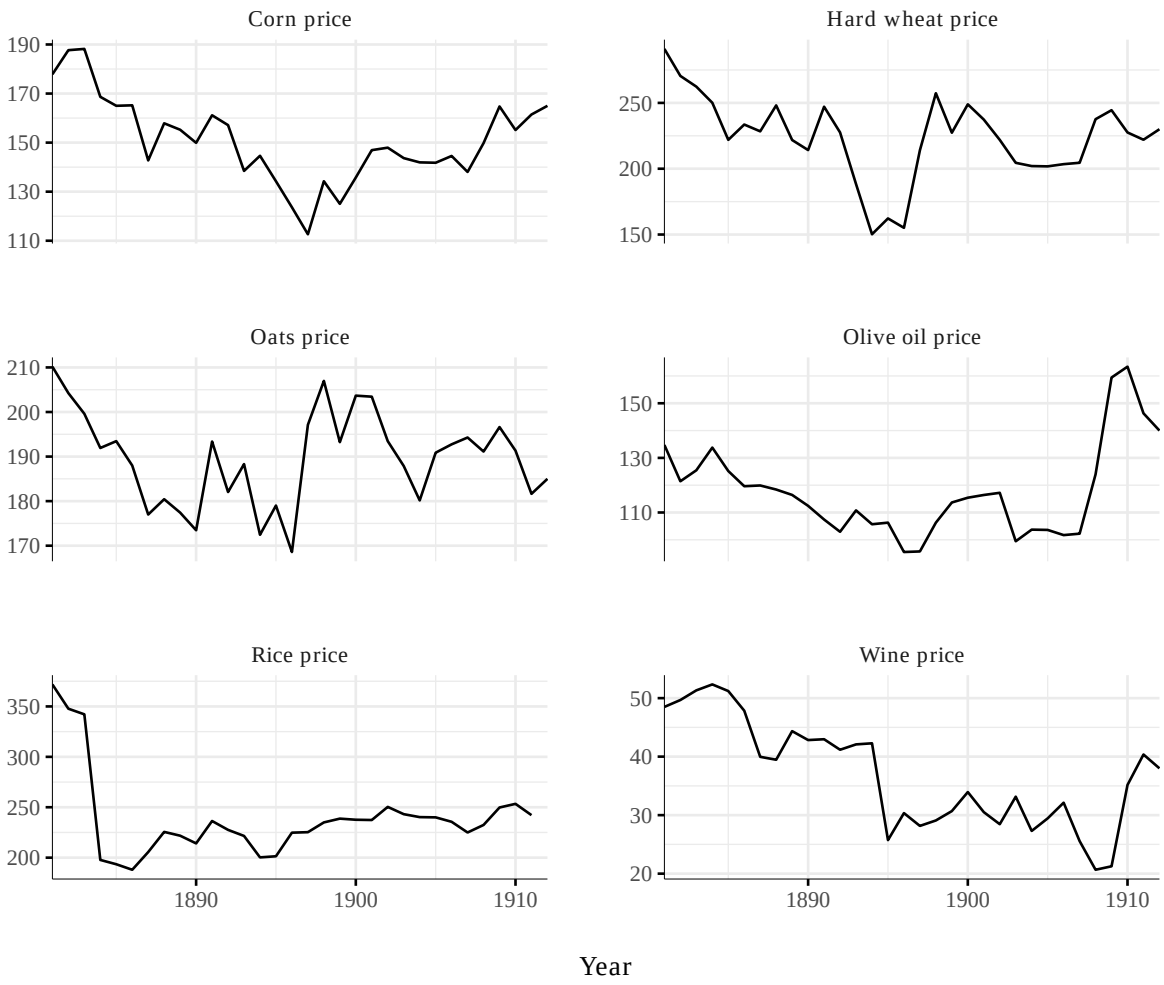


Figure 2.3: Real international crop prices (1912 £).
Source: Annuario Statistico Italiano (*ad annum*).
Notes: These graphs trace the price trajectories of the major crops farmed in Italy (wheat, corn, rice, oats, wine and olive oil) from 1881 to 1912.

Our measure is admittedly broad and might capture variations in global supply and demand that could have driven agricultural price changes: we therefore interpret our GPE index as a proxy for general commodity market trends. In addition to the GPE measure, we also construct a measure of price volatility in line with Persaud (2017). Analyzing the determinants of Indian indentured servitude, he argues that volatility in the price of rice, the dominant crop, could be thought as a proxy for income uncertainty, a potential push-factor driving migration. More formally, both unfavorable and uncertain prices for agricultural goods could reduce the demand for farm labor by lowering the expected marginal value

product of workers, thus reducing their wage. In addition, low and variable prices will reduce marginal profits for small landholders who may be forced to sell or abandon their plots, or encourage temporary migration to finance buying more land to compensate for falling prices (Cinel, 1991, 166). Both of these channels will increase the incentives for affected workers and landholders to migrate.[22]

For each crop j and time period t, price p volatility is computed over 5-year windows; it corresponds to the standard deviation of logged inter-temporal international price ratios. The GPE volatility index is then constructed similarly to the GPE presented in equation (2.1), using average acreage shares at the provincial level interacted with this measure of volatility in international agricultural commodity prices for each crop:[23]

$$volatility_{i,t} = sd[ln(\frac{p_{j,t-h+1}}{p_{j,t-h}})] \quad with \quad h = 1...5 \qquad (2.2)$$

Our extended specifications control for factors such as migrant stocks and landholding structures - which have been found important in other studies (Hatton and Williamson, 1998). We include migrant stocks at destinations as they might significantly lower the psychological and financial costs involved in transnational migration through different channels, thus alleviating liquidity constraints. For example, recent migrants could get information on job vacancies, local customs and lodgings within their communities.

We digitized information about landownership from the 1881 Italian Census. The share of non-landowners employed in the agricultural sector varies across provinces, with lower values to be found in northwestern Italy. Overall, the mean is quite high, with about 85 per cent of those employed in the primary sector not owning land - see table 2.1.

Finally, we include a measure of industrialization, as better employment opportunities

[22] Barkan (2013) documents how small-holders sold their plots to fund their migration, thus escaping liquidity constraints.

[23] We also compute volatility by varying the time window, from 3 up to 7 years. The results obtained using these alternative volatility measures are consistent with our findings and are presented in tables A2.5, A2.6, A2.7, A2.8. The trade-off in deciding on the appropriate time-window is between sample size reduction and having a sensible measure (volatility over 2 years would not be very informative).

in the secondary sector might represent an alternative to migration for would-be migrants employed in agriculture. We measure provincial industrialization using the recent estimates by Ciccarelli and Fenoaltea (2013). Industrialization varied significantly across Italian provinces with the *maximum* being Milan and the *minimum* Sassari, the northernmost province of Sardinia at that time. Overall, aside from Naples, the south was on average less industrialized than the north.

2.3 Econometric specification

Our estimation relies on the idea that international price fluctuations transfer to local economies and affect international out-migration through consequential income shocks. For this to hold, two conditions must be satisfied. First, agricultural commodity prices should be positively related to agricultural incomes. Second, local – in this case provincial – and global agricultural markets should exhibit a significant degree of integration. Both assumptions are supported in the historical literature.

Italy remained a fairly open economy and progressively integrated into the global market throughout the entire time span of reference. However, rapidly increasing competition on traditionally land-intensive crops, e.g. wheat and corn, alarmed land-owners who envisaged reduced domestic market shares. Even though protectionist measures remained mild until the early-1880s, Italy levied a series of successive tariffs on wheat from the mid-1880s – with the tariff corresponding to 40 per cent, *ad valorem*, by 1913. In general, overall agricultural protection remained relatively low, with the exception of a short spell in the 1890s. Italian agricultural trade policies did not define it as an unusually protectionist economy in the European context. Indeed, there is evidence that tariffs did not generate major compositional shifts in agricultural mixes, while possibly increasing wheat production (Federico and O'Rourke, 2000). Overall, protectionism did not prevent international prices from falling across Europe as technological change and market integration acted in the opposite direction. This point is supported by evidence of wheat prices co-moving within

Italian major markets and across Europe.[24] As explained in the previous section, our GPE index is based on import prices net of tariffs and bilateral agreements, weighted using crop shares in terms of acreage. The main econometric specification is as follows:

$$log(migration\_rate) = \beta log(GPE_{i,t}) + X_{i,t}^{'}\gamma + \delta_i + \vartheta_t + \varepsilon_{i,t} \qquad (2.3)$$

where $log(migration\_rate)$ is the out-migration rate from province i at time t and, β, the key coefficient of interest, captures the effect of agricultural commodity market integration through international price shocks on provincial out-migration rates. Note that both migration rates and the GPE index are logged, so β can be interpreted as an elasticity.

We run three separate sets of regressions, taking into account different macro-area destinations: Europe; Mediterranean and overseas destinations at large (Canada, US, Latin America and Oceania) and the United States.[25] Equation (2.3) includes time-varying characteristics at the province level ($X_{i,t}^{'}$), namely provincial industrialization, land ownership and migration networks. The specification also includes province and time fixed effects (δ_i and ϑ_t). Standard errors are clustered at the province level. Further, we interact the GPE index with the provincial share of non-landowners[26] in 1881 because globalization might have had differential effects depending on what landholding structure prevailed. Additionally, we estimate specifications that include an interaction between our GPE and a dummy variable for southern provinces to account for the diverse economies of the two macro-areas.

2.4 Results

How did globalization – thought of as agricultural commodity market integration – influence international emigration choices from Italian provinces between 1881 and 1912?

Table 2.2 - in the following page - presents baseline estimates of the impact of the GPE

[24] See figure A2.3 in the appendix for additional details.

[25] We conducted a similar analysis using internal migration rates - unfortunately only available after 1906 - and found no statistically significant results - see table A2.13. This is not surprising because international migration flows dwarfed internal migration.

[26] This class includes sharecroppers, day laborers, salaried laborers as well as renters.

Table 2.2: Globalization and provincial emigration rates

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	1.3324*	1.7217**	0.6433***	0.7187
	[0.668]	[0.854]	[0.238]	[0.613]
Observations	1,442	1,432	1,115	1,423
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.543	0.435	0.659	0.535

Notes: This table reports the estimates of equation (2.3) for all possible destinations. It includes our GPE measure - as defined in equation (2.1) - without any additional controls. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

index on provincial emigration rates toward different international macro-areas. Columns 1-4 pertain to, respectively, total overseas; transatlantic; US (Ellis Island); and European out-migration rates. We find that the GPE index has a statistically significant and positive effect on emigration rates, with transoceanic migration being the most responsive.[27] The log-log regression results show that the effect of a 1% increase in GPE ranges from a 0.64% increase in the Ellis Island out-migration to a 1.72% increase in transoceanic migration. Given that our international agricultural price index is essentially a weighted average of international prices – with weights reflecting initial provincial crop intensities – we interpret this coefficient as indicating that emigration increased in periods when international prices were rising. The rationale behind this mechanism relies on the idea that agricultural incomes – especially those of small land-holders – increase proportionally to agricultural commodity prices. Conversely, exogenous price plunges are expected to have a negative effect on agricultural incomes and prevent would-be migrants from migrating as liquidity constraints bind. Our finding is in line with Bazzi (2017), who explores the relationship between income and propensity to

[27]Results are robust to lagging the GPE measure for one year. Table A2.4 shows that coefficients increase in size, indicating that the response to price shock was not immediate.

migrate in contemporary Indonesia and finds an inverted U-shaped relation between wealth and migration. Throughout our sample period, Italy remained a slowly-developing country where liquidity constraints prevailed despite negative price shocks made migration more appealing. This result is supported by Faini and Venturini (1994) who argue that Italy was indeed caught in a poverty trap that curbed out-migration. Spitzer and Zimran (2018) also suggested that migrants from poorer Italian provinces and those who were able to finance their own trips were more positively selected, consistent with the presence of significant liquidity constraints.[28] Evidence of liquidity constraints have further been highlighted by Covarrubias, Lafortune and Tessada (2015), who showed that the average skill level of US-bound migrants decreased proportionally to GDP growth across 39 different countries between 1899 and 1932.

Our results are in line with the theoretical model designed by Ciccarelli et al. (2018): after having included liquidity constraints, areas with initially low wages could respond to increasing real incomes by migrating. The authors believe that this effect could be magnified in areas with more concentrated landholding or monopsonistic labor markets. Finally, Spain and Italy are often analyzed together as development and migration laggards. For example, Sánchez-Alonso (2000) found that agricultural wages and migration rates were positively correlated at the province level for pre-1914 Spain.

Table 2.3 - in the following page - expands our baseline specification and includes a measure of GPE volatility over a 5-year window. We add this independent variable because would-be migrants might have been influenced by price – and therefore income – uncertainty, over and above their response to changing income trends. Indeed, we find evidence that this was the case. The coefficient is statistically significant both in columns 1 and 2 – for aggregate and transoceanic emigration rates – while there seems to be no influence on emigration toward the US (column 3) and to other European countries (column 4).

Overall, although uncertainty might have driven emigration – which could then be thought of as a coping strategy – our results do not offer a clear insight on how price volatility might

[28] Wegge (1998), looking at the mid-nineteenth century, found that liquidity constraints prevented the poorest laborers from German villages from migrating, and led to the middle class dominating migrant ranks.

Table 2.3: Globalization, 5-year volatility and provincial emigration rates

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	1.6450**	1.9399**	0.6889***	0.8597
	[0.732]	[0.872]	[0.222]	[0.708]
5-year price volatility	0.6570**	0.6279*	-0.1508	-0.1353
	[0.258]	[0.326]	[0.235]	[0.318]
Observations	1,375	1,369	1,115	1,356
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.538	0.374	0.660	0.544

Notes: This table reports the estimates of equation (2.3) for all possible destinations. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

affect individuals differently across the income distribution. For instance, Persaud (2017) presents evidence that rice price volatility – the major crop in India, where the study is set – was a key driver of indentured emigration especially for lower status individuals. A possible interpretation for the fact that our volatility coefficient is not significant across all specifications could be seasonal migration – which we cannot capture – toward both North and South America.

Table 2.4 - in the following page - presents our main specification.[29] It includes a full set of controls, i.e. provincial industrialization, migrant stocks at destination and landholding structures.[30] Industrialization is not significant except for migration toward the US, where it has the anticipated negative sign, indicating that more industrially developed provinces were able to absorb the surplus of agricultural labor, thus decreasing propensity to migrate.

[29]Given that sample size differs across destinations, we run bridge regressions in order to make sure that results are not sensitive to the time-frame of reference. Findings and consistent and shown in table A2.14.

[30]Table A2.11 in the appendix presents the estimation results including welfare ratios from Federico, Nuvolari and Vasta (forthcoming). Note that we do not control for wages in our main specification, for two reasons. First, the provincial series only began in 1905 and second, wages are likely to be endogenous to agricultural or industrial business cycles.

Table 2.4: Full specification

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	3.6225**	2.2614*	1.1432	5.8793***
	[1.409]	[1.222]	[1.048]	[1.667]
5-year price volatility	0.4646***	0.4061**	-0.0707	-0.0129
	[0.171]	[0.183]	[0.166]	[0.246]
Provincial industrialization	-0.0208	-0.5223	-1.2911***	0.2606
	[0.456]	[0.531]	[0.228]	[0.505]
GPE x Share of non-landhowners	-0.0404**	-0.0237	-0.0085	-0.0665***
	[0.016]	[0.015]	[0.012]	[0.019]
Global migrant network	0.9342***			
	[0.057]			
Transiceanic migrant network		0.8653***		
		[0.044]		
EI (US) migrant network			0.2377***	
			[0.048]	
European migrant network				0.9898***
				[0.057]
Observations	1,374	1,368	1,051	1,355
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.761	0.639	0.721	0.751

Notes: This table reports the estimates of equation (2.3) for all possible destinations. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

These coefficients are in line with the fact that migration shifted southward at the turn of the 20th century. Indeed, southern provinces were comparatively less industrialized than northern ones and the secondary sector could not absorb agricultural labor surplus to any significant extent. As expected, migrant networks – despite not discounting for return and repeat migration – have a crucial role in explaining provincial emigration rates, as existing migrant stocks at destination facilitated chain migration through financial and information channels, e.g. remittances (McKenzie and Rapoport 2010; and Gomellini and Ò Gràda 2013).

This result is consistently statistically significant at the 1 per cent level and similarly sized across all destinations we consider. Moretti (1999) described how these networks allowed potential migrants to overcome, to some extent, the significant liquidity constraints they faced.

Table 2.4 also explores the role of land-tenure and includes an interaction term between our GPE index and the share of non-landowners in each province, as measured in 1881. Albeit small, the estimated coefficient is negative, indicating that small owners' incomes might have been more responsive to price fluctuations. On the other hand, landless laborers might have benefited from food price reductions, being able to count on a higher income. Another possible mechanism might be that higher agricultural prices increased incentives to invest in agriculture: given that temporary and seasonal migration had remained quite high throughout this period, small plot-owners' decision to migrate may have been motivated by a desire to acquire extra capital. According to Hatton and Williamson (1998) differentials in north/south migration patterns might be due to the relative abundance of non-agricultural labor opportunities in relatively more industrialized northern areas.

Table 2.5 - in the following page - presents the results of a specification that includes an interaction term between our GPE index and a dummy variable for southern provinces: if there is a non-linear relationship between income and migration, we might expect to see a differential effect in these areas. We find that there is a significant difference when comparing transoceanic emigration rates at large (column 2) to emigration rates to the US (column 3). The effect of globalization on transoceanic emigration rates is lower in southern provinces, but we find the opposite result when we look at emigration rates toward Ellis Island. Indeed, the estimated effect of the GPE index doubles in southern provinces in the latter case. This finding is consistent with the overall migration patterns discussed in section 2.2: emigration centers gradually shifted from northern to southern provinces while preferences converged on the US. This result also suggests that liquidity constraints on migration bound more tightly in southern provinces, which were less developed and had smaller migrant stocks at

Table 2.5: Full specification - North/south dynamics

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	3.6578***	2.2740**	1.5377**	5.8641***
	[1.274]	[1.133]	[0.656]	[1.690]
5-year price volatility	0.4300**	0.3788**	-0.0660	0.0058
	[0.172]	[0.188]	[0.159]	[0.247]
Provincial Industrialization	-0.0635	-0.5502	-0.9885***	0.2837
	[0.464]	[0.534]	[0.204]	[0.507]
GPE x Share of non-landhowners	-0.0368**	-0.0204	-0.0207**	-0.0683***
	[0.015]	[0.015]	[0.008]	[0.019]
GPE x South	-0.7971*	-0.6805*	1.5745***	0.4164
	[0.416]	[0.391]	[0.235]	[0.477]
Global migrant network	0.9348***			
	[0.056]			
Transoceanic migrant network		0.8691***		
		[0.043]		
EI (US) migrant network			0.2385***	
			[0.044]	
European migrant network				0.9911***
				[0.057]
Observations	1,374	1,368	1,051	1,355
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.762	0.640	0.739	0.751

Notes: This table reports the estimates of equation (2.3) for all possible destinations. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. The south dummy excludes central Italian provinces. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

destination throughout the 1880s and 1890s.

In terms of magnitude, we focus on the coefficients on global and transoceanic migration. These are the most stable coefficients across our specifications. This is not surprising because Ellis Island migration is imperfectly measured and European migration might suffer from our inability to account for return migration. The coefficients on global emigration range from 1.33%-3.66% across tables 2.2-2.5, suggesting an elastic response of migration to all

destinations as a result of agricultural commodity price increases. Considering the sub-period 1893-1900, we see that the GPE index increased by almost a standard deviation, 18.26%. This implies a 24.29-66.83% increase in global migration rates. The actual surge in global migration across provinces over that 8-year span was 138.6%, so it appears that commodity price trends can explain almost 50% of it.

2.5 Robustness checks

We conduct a series of robustness checks and present them in the appendix.

First, we construct a revised GPE index that uses output rather than acreage shares.[31] This approach is rooted in the idea that international crop prices might conceal relevant information about harvests in Italian provinces, therefore disregarding the price implications of particularly favorable or unfavorable years. In other words, there might be an inverse relationship between local yields and prices. Moreover, using output data mitigates concerns about inter-cropping, particularly widespread in certain areas. Results, shown in table A2.1 are consistent with the main findings of the paper.

As explained in section 2.2, we calculate the GPE index on the basis of provincial crop shares averaged over the period 1876-1881. In order to check whether results are sensitive to different measures of acreage data, we use 1879-1883 average shares. Table A2.2 highlights that this is not the case.

In the analysis presented so far, we use international price series rather than national agricultural commodity prices to preserve the exogeneity assumption. Indeed, one may be concerned that there may be factors jointly impacting both domestic prices and out-migration rates, such as a bad harvest or harsh weather conditions. As a robustness check, we replicate the analysis using national, rather than international, agricultural commodity price series taken from the *Sommario di Statistiche Storiche 1861-2010*. The results, presented in table A2.3, support our main findings. This is indeed not surprising. Figure 2.4 - in the following

[31] The output data were taken from the same source we used for acreage - ASI (1886) - and represent an average computed over the period 1876-1881.

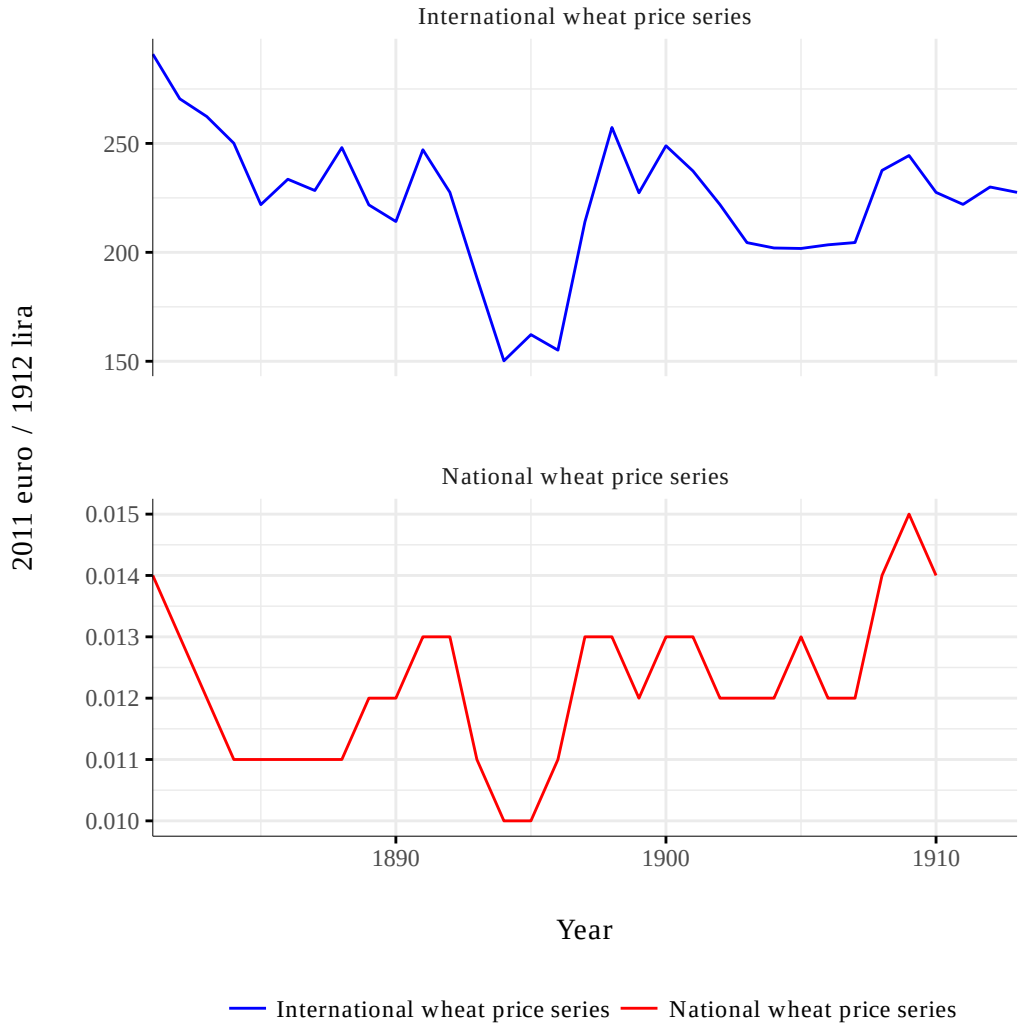


Figure 2.4: National and international wheat price series (2011 €, 2012 £).
Source: Annuario Statistico Italiano (*ad annum*) and *Sommario di Statistiche Storiche 1861-2010* (2011).
Notes: These graphs trace the trajectories of national and international wheat prices from 1881 to 1912.

page - highlights a striking degree of correlation between national and international wheat price series, suggesting that markets were very integrated and price trends mainly driven by global technology shocks. Furthermore, figure A2.3 shows that local (major-city-wise) wheat price series – purposely digitized from the *Movimento dei prezzi di alcuni generi alimentari dal 1862 al 1885* (Direzione Generale di Statistica, 1886) – co-move and don't exhibit much variation in terms of levels since the early 1860s, suggesting that there was considerable integration across internal markets too.

Next, we measure volatility over different time windows - from 3 to 7 years. The results of this exercise are largely unchanged and reported in tables A2.5, A2.6, A2.7 and A2.8.

The core of our interpretation relies on a poverty-trap-like argument: would-be migrants were not able to finance their passage because income levels were not sufficiently high. However, as discussed in Chapter 1, return and repeat migration were both common phenomena, especially after 1901. Are these two pieces of apparently discordant evidence reconcilable? It might be the case that both types of migration co-existed while migration underwent a process of spatial diffusion based on local networks. This reasoning would imply that the role of liquidity constraints decreased over time, as networks in destination countries accrued. In order to test this hypothesis, we split the sample into two periods, pre-1894 and post-1894. This allows us to capture two important aspects: first, the effect of our GPE index when agricultural commodity prices were generally trending downwards (pre-1894) or upwards (post-1894); second, the fact that migration networks progressively expanded. Results are presented in tables A2.9 and A2.10. The estimated coefficient of the GPE index is positive and statistically significant in both cases and its magnitude is larger in the former period, suggesting that incentives did play an important role and liquidity constraints were stronger earlier on. A price increase when prices are generally spiraling downwards matters more than a price increase when prices have already recovered.

As an additional robustness check, we run a specification including the welfare ratios generously provided by Federico, Nuvolari and Vasta (forthcoming), unfortunately available only from 1905. Results are presented in table A2.11: we do not find any statistically significant effect of welfare ratios on out-migration rates, possibly because the series is relatively short and migration networks were already fully functioning, thus decreasing the sensitivity of migration to wage differentials – which were in any case gradually falling.

Finally, we try an alternative way of computing migration rates. As already mentioned, pre-1896 population data are not available for all years, which means that some observations are dropped in the early part of the sample. We construct out-migration rates using provincial

population figures taken from the 1881 Census only. Table A2.12 presents the results of this alternative measure. They are in line with the previous findings.

2.6 Conlcusion

Analyses of Italian international migration flows at the sub-national level have been relatively limited in the historical literature on the age of mass migration. We fill this gap by exploiting the extraordinary heterogeneity across Italian provinces in terms of out-migration rates and agricultural production structures.

This paper revolves around exogenous shocks to international agricultural commodity prices – and therefore agricultural incomes – that occurred in Italy as a result of the first wave of globalization. We use beginning-of-period crop mixes to assign differential treatment intensity to each province and explore the effect of price fluctuations and uncertainty on migration flows. We find that higher prices translated into higher agricultural incomes and therefore higher migration rates; this mechanism suggests that would-be migrants faced binding liquidity constraints until their income reached a given threshold, despite sharp negative price shocks increased incentives to relocate.

This article represents an opportunity to compare the current globalization process with its first historical counterpart, drawing lessons for today's developing countries by analyzing then-developing Italy. Indeed, as transport and communication technology continues to advance and more countries join international trade agreements, the consequences of market integration will likely become increasingly salient. Today, many developing countries are vulnerable to commodity price shocks and exhibit even greater wage differentials with potential destinations. However, stringent immigration laws prevent most would-be migrants from embarking on international migration. Our study provides suggestive evidence of what might occur should developing countries gain more access to migration, during their ongoing process of globalization.

References

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2012. "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." *The American economic review* 102(5):1832–1856.

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2013. "Have the poor always been less likely to migrate? Evidence from inheritance practices during the Age of Mass Migration." *Journal of Development Economics* 102:2–14.

Acemoglu, Daron, David Autor, David Dorn, Gordon H Hanson and Brendan Price. 2016. "Import competition and the great US employment sag of the 2000s." *Journal of Labor Economics* 34(S1):S141–S198.

Ardeni, Pier Giorgio and Andrea Gentili. 2014. "Revisiting Italian emigration before the Great War: a test of the standard economic model." *European Review of Economic History* 18(4):452–471.

Autor, David, David Dorn and Gordon H Hanson. 2013. "The China syndrome: Local labor market effects of import competition in the US." *The American Economic Review* 103(6):2121–68.

Bandiera, Oriana, Imran Rasul and Martina Viarengo. 2013. "The making of modern America: Migratory flows in the age of mass migration." *Journal of Development Economics* 102:23–47.

Barkan, Elliott Robert. 2013. *Immigrants in American history: Arrival, adaptation, and integration*. Vol. 1 ABC-CLIO.

Bazzi, Samuel. 2017. "Wealth heterogeneity and the income elasticity of migration." *American Economic Journal: Applied Economics* 9(2):219–255.

Belot, Michèle and Timothy J Hatton. 2012. "Immigrant Selection in the OECD." *The Scandinavian Journal of Economics* 114(4):1105–1128.

Bevilacqua, Piero. 2001. *Storia dell'emigrazione italiana*. Donzelli Editore.

Ciccarelli, Carlo, Alberto Dalmazzo, Daniela Vuri et al. 2018. "Home Sweet Home: the Effect of Sugar Protectionism on Emigration in Italy, 1876-1913." *CEIS Tor Vergata Research Paper Series* 16(No. 437).

Ciccarelli, Carlo and Stefano Fenoaltea. 2013. "Through the magnifying glass: provincial aspects of industrial growth in post-Unification Italy." *The Economic History Review* 66(1):57–85.

Cinel, Dino. 1991. *The national integration of Italian return migration, 1870-1929*. Cambridge University Press.

Coan, Peter M. 1997. *Ellis Island interviews: In their own words*. Barnes & Noble Publishing.

Covarrubias, Matías, Jeanne Lafortune and José Tessada. 2015. "Who Comes and Why? Determinants of Immigrants Skill Level in the Early XXth Century US." *Journal of Demographic Economics* 81(1):115–155.

Díaz Alejandro, C F. 1984. Latin America in the 1930s. In *Latin America in the 1930s*. Springer pp. 17–49.

Direzione Generale della Statistica. ad annum. *Annuario Statistico Italiano*. ISTAT.

Direzione Generale di Statistica. 1886. *Movimento dei prezzi di alcuni generi alimentari dal 1862 al 1885*. ISTAT.

Dix-Carneiro, Rafael and Brian K Kovak. 2017. "Trade liberalization and regional dynamics." *American Economic Review* 107(10):2908–46.

Faina, Eugenio. 1909. *Inchiesta Parlamentare sulle Condizioni dei Contadini nelle Provincie Meridionali e della Sicilia*. Vol. I, II, IV Parlamento Italiano.

Faini, R and A Venturini. 1994. Italian emigration in the pre-war period. In *Migration and the international labor market, 1850-1939*, ed. Timothy J Hatton and Jeffrey G Williamson. Taylor and Francis pp. 72–90.

Fauri, Francesca. 2015. *Storia economica delle migrazioni italiane*. Il mulino.

Federico, Giovanni, Alessandro Nuvolari and Michelangelo Vasta. forthcoming. "The Origins of the Italian Regional Divide: Evidence from Real Wages, 1861-1913." *he Journal of Economic History* (No. 748).

Federico, Giovanni and Kevin H. O'Rourke. 2000. *The Mediterranean Response to Globalisation before 1950*. Routledge, London chapter Much Ado about Nothing? Italian Trade Policy in the Late 19th Century, pp. 269–96.

Felice, Emanuele. 2017. "The roots of a dual equilibrium: GDP, productivity and structural change in the Italian regions in the long-run (1871-2011)." *Economic History Working Papers* (40).

Fenoaltea, Stefano. 2011. *The reinterpretation of Italian economic history: From unification to the Great War*. Cambridge University Press.

Franchetti, Leopoldo and Sidney Sonnino. 1877. *La Sicilia nel 1876*. G. Barbera.

Gomellini, Matteo and Cormac Ò Gràda. 2013. *The Oxford Handbook of the Italian Economy Since Unification*. Oxford University Press, New York chapter Migrations.

Harvey, David I, Neil M Kellard, Jakob B Madsen and Mark E Wohar. 2017. "Long-run commodity prices, economic growth, and interest rates: 17th century to the present day." *World Development* 89:57–70.

Hatton, Timothy J. 1995. "A model of UK emigration, 1870-1913." *The Review of Economics and Statistics* pp. 407–415.

Hatton, Timothy J. 2010. "The cliometrics of international migration: a survey." *Journal of Economic Surveys* 24(5):941–969.

Hatton, Timothy J and Jeffrey G Williamson. 1998. *The age of mass migration: Causes and economic impact*. Oxford University Press on Demand.

Hatton, Timothy J and Jeffrey G Williamson. 2005. *Global migration and the world economy: Two centuries of policy and performance*. MIT press Cambridge, MA.

ISTAT. 1926. *Annuario statistico dell'emigrazione italiana*. Commissariato generale dell'emigrazione.

ISTAT. 2011. *L'Italia in 150 anni: sommario di statistiche storiche, 1861-2010*. ISTAT.

Jacini, Stefano. 1886. "Relazione finale sui risultati dell'inchiesta." *Atti della Giunta per la Inchiesta Agraria sulle condizioni della classe agricola* .

Keeling, Drew. 2013. "Oceanic travel conditions and American immigration, 1890-1914." *MPRA* (47850).

Majlesi, Kaveh and Gaia Narciso. 2018. "International import competition and the decision to migrate: Evidence from Mexico." *Journal of Development Economics* 132(5):75–87.

McKenzie, David and Hillel Rapoport. 2010. "Self-selection patterns in Mexico-US migration: the role of migration networks." *The Review of Economics and Statistics* 92(4):811–821.

Moretti, Enrico. 1999. "Social networks and migrations: Italy 1876-1913." *International Migration Review* pp. 640–657.

O'Rourke, Kevin. 1991. "Did the great Irish famine matter?" *The Journal of Economic History* 51(1):1–22.

O'Rourke, Kevin H. 1997. "The European grain invasion, 1870-1913." *The Journal of Economic History* 57(04):775–801.

O'Rourke, Kevin H and Jeffrey G Williamson. 1999. *Globalization and history: the evolution of a nineteenth-century Atlantic economy*. Mit Press.

Persaud, Alexander. 2017. "Risk mitigation and selection under forward contracts: 19th-century Indian indentureship." Unpublished manuscript, Department of Economics, University of Michigan.

Pierce, Justin R and Peter K Schott. 2016. "The surprisingly swift decline of US manufacturing employment." *The American Economic Review* 106(7):1632–1662.

Sánchez-Alonso, Blanca. 2000. *Those who Left and Those who Stayed Behind: Explaining Emigration from the Regions of Spain, 1880-1914*. Cambridge University Press.

Spitzer, Yannay. 2015. "The Dynamics of Mass Migration: Estimating the Effect of Income Differences on Migration in a Dynamic Model of Discrete Choice with Diffusion." Unpublished manuscript, Brown University, Deptartment of Economics.

Spitzer, Yannay. 2016. "Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914." Unpublished manuscript, Norwestern University, Department of Economics.

Spitzer, Yannay and Ariell 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.

Topalova, Petia. 2007. Trade liberalization, poverty and inequality: Evidence from Indian districts. In *Globalization and poverty*. University of Chicago Press pp. 291–336.

Ward, Zachary. 2017. "Birds of passage: Return migration, self-selection and immigration quotas." *Explorations in Economic History* 64:37–52.

Wegge, Simone A. 1998. "Chain migration and information networks: evidence from nineteenth-century Hesse-Cassel." *The Journal of Economic History* 58(04):957–986.

Williamson, Jeffrey G. 1995. "The evolution of global labor markets since 1830: background evidence and hypotheses." *Explorations in Economic History* 32(2):141–196.

Appendix 2

Figures and Tables

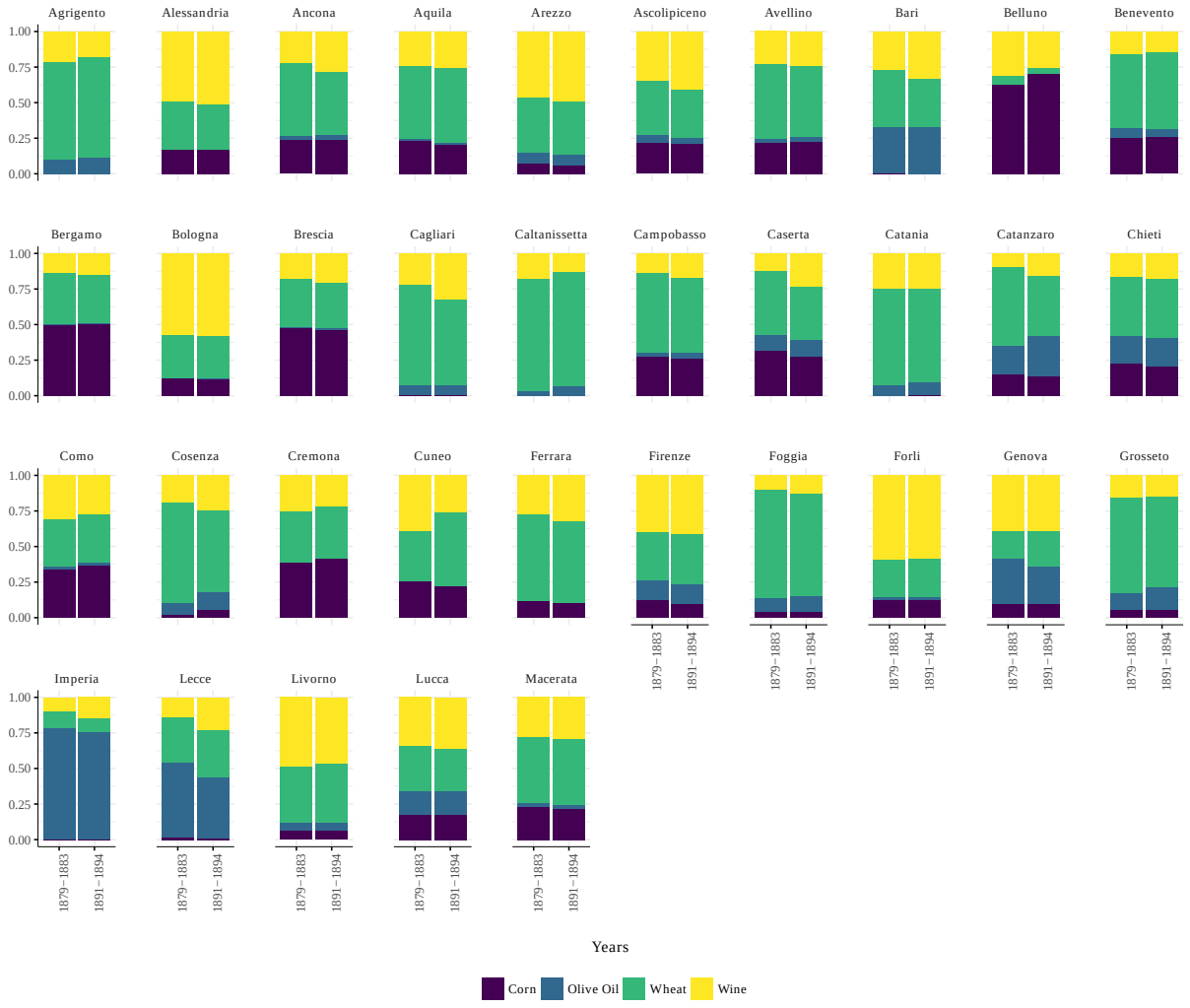


Figure A2.1: Crop reallocation, panel (a).
Source: *Annuario Statistico Italiano* (1886, 1892, 1895).
Notes: The comparison is based on 1876-1881 and 1891-1894 averages across all provinces. In order to preserve readability data only include wheat, corn, wine and olive oil.

Figure A2.2: Crop reallocation, panel (b).
Source: *Annuario Statistico Italiano* (1886, 1892, 1895).
Notes: The comparison is based on 1876-1881 and 1891-1894 averages across all provinces. In order to preserve readability data only include wheat, corn, wine and olive oil.

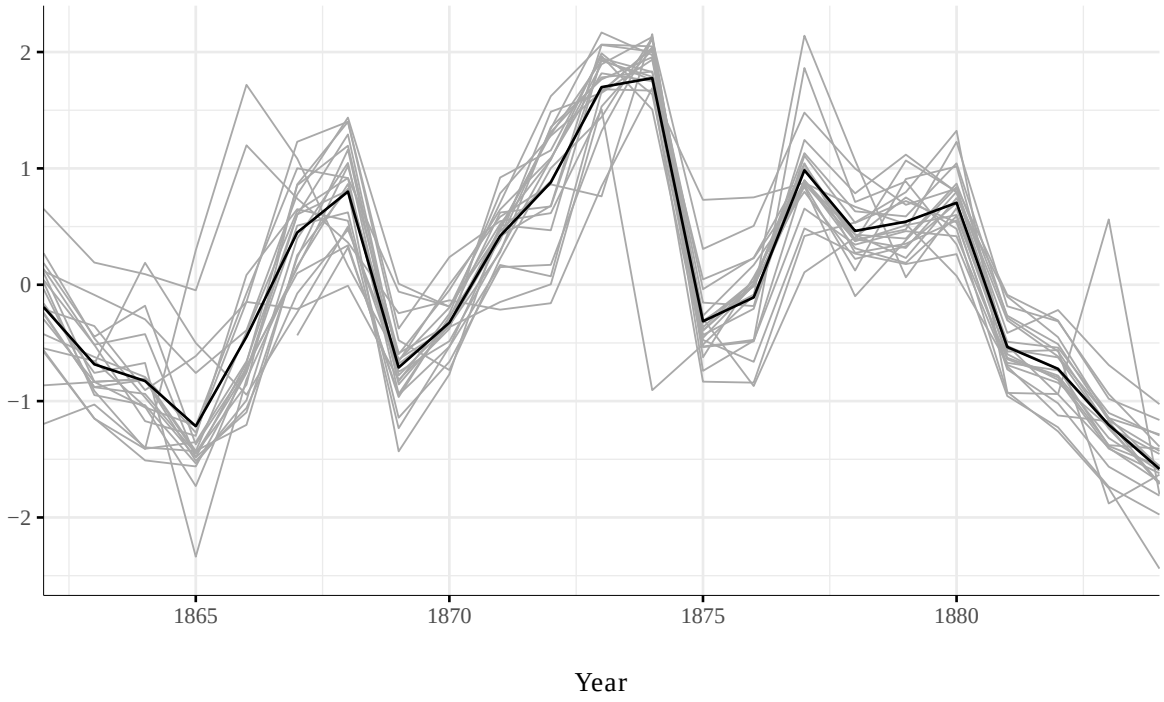


Figure A2.3: Co-movements of wheat prices in Italian major markets.
Source: *Movimento dei prezzi di alcuni generi alimentari dal 1862 al 1885* (Direzione Generale della Statistica, 1886).
Notes: This graph shows how wheat prices co-moved across the main Italian city markets. Prices have been standardized. The darker line represents the average.

Table A2.1: Full specification - Output

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE (output)	3.7540***	2.6574**	1.1362	5.1657***
	[1.399]	[1.094]	[1.071]	[1.770]
GPE (output), 5-year vol.	0.3873***	0.5484***	0.1742	-0.0639
	[0.144]	[0.155]	[0.112]	[0.162]
Provincial industrialization	0.1036	-0.3256	-1.2644***	0.2339
	[0.453]	[0.534]	[0.238]	[0.523]
GPE * Share of non-landowners	-0.0480***	-0.0379***	-0.0119	-0.0627***
	[0.016]	[0.013]	[0.013]	[0.021]
Global migration network	0.9227***			
	[0.056]			
Transoceanic migration network		0.8489***		
		[0.041]		
Ellis Island (US) migration network			0.2363***	
			[0.047]	
European migration network				0.9861***
				[0.058]
Observations	1,374	1,368	1,051	1,355
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.763	0.646	0.722	0.751

Notes: This table reports the estimates of equation (2.3) for all possible destinations, using output data rather than acreage shares as in the main analysis. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.2: Full specification - 1879-1883 acreage

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE (79-83)	3.2901**	1.7922	1.3214	5.6742***
	[1.499]	[1.310]	[1.042]	[1.713]
GPE (79-83), 5-year volatility	0.1693	0.0240	-0.1075	-0.0107
	[0.149]	[0.135]	[0.122]	[0.205]
Provincial industrialization	-0.0747	-0.5804	-1.2920***	0.2705
	[0.448]	[0.525]	[0.225]	[0.503]
GPE * Share of non-landowners	-0.0383**	-0.0176	-0.0123	-0.0667***
	[0.017]	[0.016]	[0.013]	[0.021]
Global migration network	0.9320***			
	[0.057]			
Transoceanic migration network		0.8668***		
		[0.043]		
Ellis Island (US) migration network			0.2405***	
			[0.047]	
European migration network				0.9879***
				[0.058]
Observations	1,374	1,368	1,051	1,355
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.758	0.636	0.721	0.751

Notes: This table reports the estimates of equation (2.3) for all possible destinations, using 1879-1883 acreage shares rather than 1876-1881 as in the main analysis. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.3: Full specification - National prices series

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE (national prices)	2.9941*	2.8866***	1.1423*	5.3157**
	[1.769]	[0.898]	[0.576]	[2.530]
GPE (national prices), 5-year volatility	2.9673***	2.5261**	-0.5326	-0.3921
	[1.076]	[1.230]	[1.137]	[1.445]
Provincial industrialization	-0.0123	-0.5424	-1.3054***	0.2208
	[0.463]	[0.535]	[0.228]	[0.507]
GPE * Share of non-landowners	-0.0371*	-0.0289**	-0.0088	-0.0647**
	[0.019]	[0.011]	[0.006]	[0.027]
Global migration network	0.9468***			
	[0.059]			
Transoceanic migration network		0.8647***		
		[0.046]		
Ellis Island (US) migration network			0.2374***	
			[0.048]	
European migration network				1.0038***
				[0.058]
Observations	1,374	1,368	1,051	1,355
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.761	0.640	0.721	0.751

Notes: This table reports the estimates of equation (2.3) for all possible destinations using national price series rather than international ones. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.4: Full specification - One-year lags

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	4.5683***	4.5357	2.3221	8.2302***
	[1.666]	[2.962]	[1.560]	[1.667]
GPE, 5-year volatility	0.4571**	0.6304***	-0.0182	0.0524
	[0.203]	[0.229]	[0.131]	[0.278]
Provincial industrialization	-0.0684	-0.7583	-1.2286***	0.1609
	[0.500]	[0.601]	[0.231]	[0.561]
GPE * Share of non-landowners	-0.0587***	-0.0515*	-0.0173	-0.0787***
	[0.017]	[0.027]	[0.019]	[0.017]
Global migration network	0.7620***			
	[0.075]			
Transoceanic migration network		0.6590***		
		[0.076]		
Ellis Island (US) migration network			0.2411***	
			[0.038]	
European migration network				0.8338***
				[0.067]
Observations	1,306	1,301	1,049	1,288
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.692	0.493	0.732	0.699

Notes: This table reports the estimates of equation (2.3) for all possible destinations, lagging all variables by one year. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.5: Full specification - Global migration only (3-to-7-year volatility windows)

VARIABLES	(1) Global emigration rate	(2) Global emigration rate	(3) Global emigration rate	(4) Global emigration rate	(5) Global emigration rate
GPE	3.8139**	3.7386**	3.6227**	2.9989**	2.9115**
	[1.547]	[1.464]	[1.409]	[1.405]	[1.377]
Provincial industrialization	-0.0815	-0.0645	-0.0207	0.2074	0.2082
	[0.453]	[0.454]	[0.456]	[0.446]	[0.445]
GPE * Share of non-landowners	-0.0437**	-0.0417**	-0.0404**	-0.0335**	-0.0325**
	[0.017]	[0.016]	[0.016]	[0.016]	[0.016]
Global migration network	0.9377***	0.9369***	0.9342***	0.9016***	0.9002***
	[0.057]	[0.056]	[0.057]	[0.062]	[0.062]
GPE, 3-year volatility	0.1905				
	[0.130]				
GPE, 4-year volatility		0.3037*			
		[0.156]			
GPE, 5-year volatility			0.4641***		
			[0.171]		
GPE, 6-year volatility				0.4567**	
				[0.208]	
GPE, 7-year volatility					0.4441**
					[0.215]
Observations	1,374	1,374	1,374	1,306	1,306
Time-frame	1881-1912	1881-1912	1881-1912	1881-1912	1881-1912
No. of Provinces	69	69	69	69	69
Adjusted R-squared	0.759	0.760	0.761	0.735	0.735

Notes: This table reports the estimates of equation (2.3) for global migration. It includes both our GPE index and GPE volatility measures over – based on equations (2.1) and (2.2) respectively – computed over 3-to-7-year windows and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.6: Full specification - Transoceanic migration only (3-to-7-year volatility windows)

VARIABLES	(1) Transoceanic emigration rate	(2) Transoceanic emigration rate	(3) Transoceanic emigration rate	(4) Transoceanic emigration rate	(5) Transoceanic emigration rate
GPE	2.4306*	2.3798*	2.2618*	1.3490	1.2063
	[1.321]	[1.277]	[1.222]	[1.435]	[1.432]
Provincial industrialization	-0.5821	-0.5675	-0.5222	-0.2352	-0.2294
	[0.524]	[0.525]	[0.531]	[0.536]	[0.536]
GPE * Share of non-landowners	-0.0267*	-0.0253*	-0.0237	-0.0127	-0.0110
	[0.015]	[0.015]	[0.015]	[0.017]	[0.017]
Transoceanic migration network	0.8670***	0.8672***	0.8653***	0.8653***	0.8644***
	[0.044]	[0.044]	[0.044]	[0.049]	[0.049]
GPE, 3-year volatility	0.0932				
	[0.145]				
GPE, 4-year volatility		0.1935			
		[0.175]			
GPE, 5-year volatility			0.4055**		
			[0.183]		
GPE, 6-year volatility				0.5249**	
				[0.202]	
GPE, 7-year volatility					0.5477**
					[0.218]
Observations	1,368	1,368	1,368	1,302	1,302
Time-frame	1881-1912	1881-1912	1881-1912	1881-1912	1881-1912
No. of Provinces	69	69	69	69	69
Adjusted R-squared	0.637	0.637	0.639	0.574	0.574

Notes: This table reports the estimates of equation (2.3) for transoceanic migration. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – computed over 3-to-7-year windows and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.7: Full specification - US-bound migration only (3-to-7-year volatility windows)

VARIABLES	(1) US emigration rate	(2) US emigration rate	(3) US emigration rate	(4) US emigration rate	(5) US emigration rate
GPE	1.1084	1.1612	1.1432	1.2002	1.1952
	[1.076]	[1.049]	[1.048]	[1.044]	[1.049]
Provincial industrialization	-1.2858***	-1.2911***	-1.2911***	-1.2876***	-1.2907***
	[0.226]	[0.227]	[0.228]	[0.227]	[0.228]
GPE * Share of non-landowners	-0.0080	-0.0089	-0.0085	-0.0091	-0.0091
	[0.013]	[0.012]	[0.012]	[0.012]	[0.012]
Ellis Island (US) migration network	0.2373***	0.2376***	0.2377***	0.2372***	0.2377***
	[0.048]	[0.048]	[0.048]	[0.048]	[0.048]
GPE, 3-year volatility	-0.0873				
	[0.105]				
GPE, 4-year volatility		-0.0908			
		[0.143]			
GPE, 5-year volatility			-0.0706		
			[0.165]		
GPE, 6-year volatility				-0.1078	
				[0.164]	
GPE, 7-year volatility					-0.0802
					[0.165]
Observations	1,051	1,051	1,051	1,051	1,051
Time-frame	1892-1912	1892-1912	1892-1912	1892-1912	1892-1912
No. of Provinces	69	69	69	69	69
Adjusted R-squared	0.721	0.721	0.721	0.721	0.721

Notes: *Notes*: This table reports the estimates of equation (2.3) for US-bound migration using Ellis Island data. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – computed over 3-to-7-year windows and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.8: Full specification - European migration only (3-to-7-year volatility windows)

VARIABLES	(1) European emigration rate	(2) European emigration rate	(3) European emigration rate	(4) European emigration rate	(5) European emigration rate
GPE	5.9067***	5.8890***	5.8793***	5.5021***	5.5722***
	[1.737]	[1.684]	[1.667]	[1.481]	[1.503]
Provincial industrialization	0.2363	0.2556	0.2606	0.3423	0.3322
	[0.495]	[0.500]	[0.505]	[0.526]	[0.525]
GPE * Share of non-landowners	-0.0671***	-0.0668***	-0.0665***	-0.0622***	-0.0631***
	[0.020]	[0.019]	[0.019]	[0.017]	[0.017]
European migration network	0.9867***	0.9892***	0.9898***	0.9614***	0.9606***
	[0.057]	[0.057]	[0.057]	[0.062]	[0.063]
GPE, 3-year volatility	-0.1381				
	[0.184]				
GPE, 4-year volatility		-0.0419			
		[0.237]			
GPE, 5-year volatility			-0.0129		
			[0.245]		
GPE, 6-year volatility				0.0321	
				[0.324]	
GPE, 7-year volatility					-0.0390
					[0.317]
Observations	1,355	1,355	1,355	1,288	1,288
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912	1881-1912
No. of Provinces	69	69	69	69	69
Adjusted R-squared	0.752	0.751	0.751	0.736	0.736

Notes: This table reports the estimates of equation (2.3) for European migration. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – computed over 3-to-7-year windows and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.9: Full specification - Pre-1894

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) European emigration rate
GPE	5.9119**	1.6035	6.7087**
	[2.749]	[2.608]	[3.217]
GPE, 5-year volatility	0.6549**	0.0647	0.1072
	[0.309]	[0.404]	[0.314]
GPE * Share of non-landowners	-0.0522*	-0.0351	-0.0713**
	[0.028]	[0.028]	[0.033]
Global migration network	0.7303***		
	[0.090]		
Transoceanic migration network		0.6412***	
		[0.092]	
European migration network			0.8726***
			[0.256]
Observations	339	333	330
Time-frame	1881-1894	1881-1894	1881-1894
No. of Provinces	69	69	69
Adjusted R-squared	0.282	0.501	0.101

Notes: This table reports the estimates of equation (2.3) for all possible destinations, using pre-1894 data. We excluded US-bound emigration as this series starts in 1892. We include both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces). Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.10: Full specification - Post-1894

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	3.1409*	4.1831**	0.9666	3.3458*
	[1.698]	[1.656]	[0.811]	[1.891]
GPE, 5-year volatility	0.2636	-0.1079	0.0226	0.3639
	[0.231]	[0.236]	[0.161]	[0.399]
Provincial industrialization	0.3225	0.1067	-1.0945***	0.3954
	[0.447]	[0.567]	[0.204]	[0.575]
GPE * Share of non-landowners	-0.0408**	-0.0500**	-0.0101	-0.0422*
	[0.020]	[0.020]	[0.010]	[0.022]
Global migration network	0.5772***			
	[0.073]			
Transoceanic migration network		0.6013***		
		[0.098]		
Ellis Island (US) migration network			0.2713***	
			[0.044]	
European migration network				0.6757***
				[0.083]
Observations	1,035	1,035	1,001	1,025
Time-frame	1894-1912	1894-1912	1894-1912	1894-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.607	0.327	0.709	0.641

Notes: This table reports the estimates of equation (2.3) for all possible destinations, using post-1894 data. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.11: Full specification - Welfare ratios from 1905 onward

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	2.5285***	1.7503	0.5750	2.2565**
	[0.841]	[1.357]	[0.586]	[1.063]
GPE, 5-year volatility	-0.1694	0.1582	-0.0607	-0.5517*
	[0.264]	[0.422]	[0.293]	[0.328]
Provincial welfare ratio	-0.0435	-0.0654	0.0981	-0.2256
	[0.088]	[0.148]	[0.082]	[0.152]
Provincial industrialization	0.5882**	0.7514*	0.2437	0.2803
	[0.223]	[0.383]	[0.155]	[0.380]
GPE * Share of non-landowners	-0.0298***	-0.0183	-0.0075	-0.0251*
	[0.010]	[0.017]	[0.007]	[0.013]
Global migration network	0.2466**			
	[0.122]			
	[0.167]	[0.288]	[0.205]	[0.245]
Transoceanic migration network		0.3277		
		[0.200]		
Ellis Island (US) migration network			0.2931***	
			[0.063]	
European migration network				0.0211
				[0.121]
Observations	552	552	552	552
Time-frame	1905-1912	1905-1912	1905-1912	1905-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.291	0.290	0.620	0.140

Notes: This table reports the estimates of equation (2.3) for all possible destinations, accounting for provincial welfare ratios generously provided by Federico et al (2018). Unfortunately this series is only available from 1905 onward. This specification includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.12: Full specification - Migration rates computed using 1881 population figures

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	3.4416***	2.5396*	0.1508	5.6878***
	[1.255]	[1.349]	[1.178]	[1.391]
GPE, 5-year volatility	0.5251***	0.5214**	-0.0543	0.0898
	[0.178]	[0.211]	[0.192]	[0.249]
Provincial industrialization	0.2419	-0.3803	-1.0576***	0.4822
	[0.447]	[0.494]	[0.224]	[0.506]
GPE * Share of non-landowners	-0.0386***	-0.0258	0.0047	-0.0630***
	[0.014]	[0.016]	[0.014]	[0.016]
Global migration network	0.9566***			
	[0.052]			
Transoceanic migration network		0.8591***		
		[0.052]		
Ellis Island (US) migration network			0.2441***	
			[0.043]	
European migration network				1.0239***
				[0.054]
Observations	1,984	1,973	1,282	1,949
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.750	0.633	0.742	0.745

Notes: This table reports the estimates of equation (2.3) for all possible destinations. Out-migration rates are computed as the ratio between absolute migration figures and population in 1881 rather than using data yearly data from the *Annuario Statistico Italiano*. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.13: Internal Migration

VARIABLES	(1) Internal emigration rate
GPE	1.7880
	[1.128]
GPE, 5-year volatility	-0.5426
	[0.356]
Provincial industrialization	0.4944
	[0.349]
Internal migration network	-0.5014
	[0.391]
GPE * Share of non-landowners	-0.0221
	[0.014]
Observations	345
Time-frame	1906-1912
No. of Provinces	69
Adjusted R-squared	0.089

Notes: This table reports the estimates of equation (2.3) for total internal migration. It includes both our GPE index and GPE volatility measure – based on equations (2.1) and (2.2) respectively – and an additional set of controls. This series starts in 1906. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces). Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.14: Full specification - Bridge regressions

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(4) European emigration rate
GPE	4.1170**	4.2050**	4.3812**
	[1.565]	[1.613]	[1.730]
GPE, 5-year volatility	0.0374	-0.2169	-0.0546
	[0.210]	[0.192]	[0.371]
Provincial industrialization	0.2874	-0.0651	0.3641
	[0.449]	[0.526]	[0.548]
GPE * Share of non-landowners	-0.0457**	-0.0445**	-0.0498**
	[0.018]	[0.019]	[0.020]
Global migration network	0.8576***		
	[0.084]		
Transoceanic migration network		0.8364***	
		[0.075]	
European migration network			0.8311***
			[0.069]
Observations	1,171	1,170	1,155
Time-frame	1892-1912	1892-1912	1892-1912
No. of Provinces	69	69	69
Adjusted R-squared	0.722	0.496	0.697

Notes: This table reports the estimates of equation (2.3) for global, transoceanic and European migration, using the time frame of the US series, 1892-1912. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All regressions include province and year fixed effects. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.15: Full specification - Provinces producing less wheat/corn

VARIABLES	(1) Global emigration rate	(2) Transoceanic emigration rate	(3) Ellis Island (US) emigration rate	(4) European emigration rate
GPE	4.3891**	2.8506*	0.1854	4.7653**
	[1.975]	[1.672]	[1.083]	[2.254]
GPE, 5-year volatility	0.3131	0.0773	-0.2165	0.0547
	[0.222]	[0.212]	[0.217]	[0.315]
Provincial industrialization	-0.7372	-1.3121**	-1.3236***	0.2032
	[0.653]	[0.513]	[0.302]	[0.971]
GPE * Share of non-landowners	-0.0506**	-0.0349*	0.0019	-0.0508*
	[0.022]	[0.020]	[0.013]	[0.027]
Global migration network	0.9078***			
	[0.082]			
Transoceanic migration network		0.8090***		
		[0.060]		
Ellis Island (US) migration network			0.2376***	
			[0.061]	
European migration network				0.9874***
				[0.080]
Observations	714	708	551	705
Time-frame	1881-1912	1881-1912	1892-1912	1881-1912
No. of Provinces	69	69	69	69
Adjusted R-squared	0.780	0.714	0.759	0.727

Notes: This table reports the estimates of equation (2.3) for all possible destinations, using restricting the focus to provinces whose acreage for wheat or corn was below 20 per cent. It includes both our GPE index and GPE volatility measures – based on equations (2.1) and (2.2) respectively – and an additional set of controls. Migrant networks do not account for return/temporary migration. Non-landowners include sharecroppers, day laborers, salaried laborers as well as renters. All variables are logged. Standard errors are clustered at the provincial level (69 provinces).
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Chapter 3

Post-Disaster Migration

Evidence from the 1908 Messina and Reggio Calabria Earthquake

Yannay Spitzer,[*] *Gaspare Tortorici,*[‡] *Ariell Zimran*[¶]

March 2019

Abstract

This paper revolves around a major historical disaster that took place in southern Italy during the age of mass migration: the earthquake that shook Messina and Reggio Calabria on December 28th 1908. Having claimed as many as 100 thousand lives and destroyed entire cities, this event is often regarded as the most devastating natural disaster in recent European history. We evaluate the municipality-level effect of the earthquake on total and US-bound out-migration, based on the complete Ellis Island administrative records and official Italian statistics. We find that the earthquake had a differential effect in Calabria and Sicily, only determining a transitory, statistically significant, decline of out-migration rates in the former.

Keywords: *Age of mass migration; earthquake-induced migration; post-disaster population dynamics; determinants of international migration.*

JEL Codes: F22, J61, N33, 015

[*]Hebrew University of Jerusalem,[‡] Trinity College Dublin,[¶] Vanderbilt University and NBER.

Gaspare Tortorici acknowledges that this research project has been funded by the Irish Research Council. Funding was also provided by the National Science Foundation (Grand number: SES-1425598), the Exploring Travel and Data Award from the Economic History Association, the Northwestern University Center for Economic History, the Eisner Fund and the Balzan Foundation.

Outstanding research assistance was provided by Elizabeth Nelsen, Sarah Robinson, Danielle Williamson, and Claire Wittacker.

We are greatful to seminar participants at Vanderbilt University and Trinity College Dublin for helpful comments.

3.1 Introduction

Natural disasters have rife consequences on human lives and the environment. As urban areas expand - in terms of geographical scope and population - and climate changes, the exposure to natural hazards is predicted to rise in both developed and developing countries (United Nations, 2015). Modified environmental and economic conditions might induce migration - which could therefore be thought as an adjusting mechanism.

Natural disasters can be interpreted as exogenous negative shocks that directly or indirectly affect households' income and wealth. The short-run migration response to catastrophic events is therefore tied to income elasticity. The consequences of income shocks on aggregate migration rates are *a priori* ambiguous for countervailing economic forces operate in opposite directions: on one hand, negative income shocks might decrease the opportunity cost of migrating; on the other, a sharp income decline could tighten liquidity constraints, preventing would-be migrants from relocating elsewhere. For instance, during the Great Irish Famine of the mid-1840s and early 1850s, out-migration to death ratios were lower in poorer counties (O Gràda and O'Rourke, 1997). In other words, the most vulnerable share of the population could not finance transoceanic crossings and remained caught in a poverty trap despite the relatively affordable fixed up-front costs of the voyage, open border policies and the support of networks.[1] On the other hand, income shocks entangle heterogeneous individual responses that echo in aggregate figures. Similarly, the radius of migration flows - internal or international - and their temporal horizon - permanent or transitory - could be determined by both household-specific choices and aggregate factors such as legal restrictions (Stark and Bloom, 1985).

At its core, this paper evaluates the local short-run migration response to one of the most devastating natural disasters in recent European history: the earthquake that leveled Messina,

[1] More recent studies based on theoretical extensions of the standard Roy-Borjas model point toward a non-linear, inverted-U-shaped, function linking aggregate income (or wealth) and migration rates. For example, Bazzi (2017) includes wealth heterogeneity into a standard migration choice model and tests how a positive income shock influences international migration from Indonesia. See also Cattaneo and Peri (2016).

Reggio Calabria and their surroundings in 1908 - MRC earthquake henceforth for simplicity. Despite having claimed the lives of more than 100,000 individuals (Dickie, 2014), it is an overshadowed historical event occurred within a unique historical and institutional framework. Indeed it took place during the midst of the age of mass migration, while out-migration from both Calabria and Sicily morphed into a semi-universal phenomenon involving virtually every municipality. Moreover, pre-WWI international migration remained largely unfettered, with major transoceanic and European destinations implementing next-to-open border policies.[2] This characterizing element might help informing current debates as international migration is currently limited by much more binding legal restrictions.

Recent studies have highlighted how Italian international migration displayed a surprising degree of variability: while sharing some common traits, provincial and municipal patterns exhibit diverse trajectories in terms of magnitude (see Chapter 1), self-selection (Spitzer and Zimran, 2018), destination choices (Fauri, 2015), and timing (Gray, Narciso and Tortorici, 2018). Understanding the roots of this heterogeneity is a challenge that might require exploring the role of non-economic forces like natural disasters (Abramitzky and Boustan, 2017).

Our paper is the first attempt to reconstruct the dynamics of municipality-level migration flows following a severe natural disaster through an econometric approach, within the context of the age of mass migration. We implement a DD analysis and complement it with event studies; in both cases, the geo-temporal unit of reference is the Italian municipality-year.

Our study relies on a distinctive combination of sources. Total and US-bound municipal out-migration figures are obtained from geo-located Ellis Island administrative records and newly transcribed Italian official statistics. Local earthquake intensities are based on the systematic recording of casualties, injured, damaged buildings and infrastructure in the immediate aftermath of the quake.[3] The analysis - consistent across alternative measures

[2]For instance, the United States only introduced a series of progressively more restrictive Immigration Acts in 1917, 1921 and 1924.

[3]This document later flowed into a more recent publication had originally been compiled by the most noted Italian geologists of the time, Prof Mario Baratta and Giuseppe Mercalli, in collaboration with local

of earthquake-related damage - shows that the earthquake determined a temporary decline of out-migration from Calabria while evidencing no statistically significant, robust, effect in Sicily. We examine two possible mechanisms to explain the differential effect of the earthquake on out-migration rates in the two regions. First, we test whether network effects interacted with earthquake intensity, to examine the hypothesis that larger, tighter or more vintage networks might foster migration in the aftermath of a natural disaster - for instance, by releasing liquidity constraints (see, among other studies, McKenzie and Rapoport 2007). Second, we explore the damage accumulation hypothesis exploiting a historical accident: Calabria was previously affected by two violent earthquakes in 1905 and 1907. Given that severe damage was local, mainly clustered around the epicenters in the center and south of the region, these quakes caused widespread destruction in several Calabrian municipalities that would later be affected by the MRC disastrous quake. Neither quakes determined substantial economic losses across the strait, in Sicily.

We find no evidence that larger networks were associated with a higher propensity to migrate from severely affected municipalities. The lack of such an effect is somehow surprising in light of recent contributions: for instance, according to Mahajan and Yang (2017), the deterioration of local amenities following hurricanes spurs increased migration toward the United States. Based on restricted US census data on migrants' origin countries, the paper argues that larger migrant networks in the US tend to amplify migration responses, facilitating legal immigration. Nevertheless, the role of migration networks might evolve over time and be non-linear in size meaning that after a certain threshold it only marginally influences future would-be migrants' propensity to relocate because most affected individuals are already linked to migration networks. Indeed, when the earthquake hit the area, emigration had already become a consolidated phenomenon in every municipality: on average, the communal pre-1904 emigration to 1901 population ratio (per cent) was about 2.6, with a minimum of 1.7.

authorities.

There is no robust statistical evidence of out-migration rates declining or increasing as a consequence of neither the 1905 nor the 1907 earthquakes. However, when interacting these quakes with the 1908 MRC disaster (separately or jointly), we find that out-migration rates decreased comparatively more in localities that were affected by both the MRC and a previous quake. In other words, while a violent event might not immediately influence out-migration, the resulting accumulation of damage to physical capital over time might translate into an actual restraining force once a greater catastrophe strikes. Taken together, our findings suggest that catastrophic events might produce heterogeneous effects within the boundaries of a relatively confined area. On the other hand, they do not point toward any obvious spikes in international out-migration following a severe natural disaster.

The contribution of this paper is three-fold and relates to different strands of literature. First, it evaluates the short-term migration response to sudden adverse economic shocks in the absence of legal barriers: it therefore contributes to the flourishing literature on how post-disaster adjustments operate through population dynamics.[4] There are several events that might fall under the umbrella term of natural disaster: floods, droughts, earthquakes, eruptions, tornadoes, famines etc. Each of these phenomena might affect income levels and employment opportunities differently, thus influencing migration responses heterogeneously. For instance, Tse (2012) documents that earthquakes, volcanic eruptions and floods reduce migration rates through diverse channels. In particular: volcanic eruptions increase land productivity and its value, incentivizing would-be migrants to stay; on the other hand, earthquakes reduce earnings and entail a considerable physical asset destruction that prevent would-be migrants from relocating.

In a study centered on the Dust Bowl, Hornbeck (2012) argues that more eroded counties experienced a large population decline as a result of both higher out-migration and lower in-migration. Such population movements served as a major economic adjustment mechanism aimed at re-equilibrating local labor markets throughout the 1930s and 1950s. Indeed,

[4]See Millock (2015) or Gröschl and Steinwachs (2016) for comprehensive overviews.

severely affected areas were associated with lower labor-capital ratios in agriculture, lower unemployment and higher (proxied) wages - relative to low-erosion counties. Drawing from the same natural disaster, Long and Siu (2016) reconstruct post-Dust-Bowl migration patterns and highlight the importance of population inflows in shaping its demographic effects. While the series of storms increased out-migration rates from affected areas relative to anywhere else in the US during the 1930s, the relative depopulation of these regions can only be explained through a significant decline of migrant inflows.[5]

Second, it deepens the blooming debate revolving around the driving forces of mass migration from Europe at the turn of the 20$^{\text{th}}$ century.[6] Early contributions (Gould, 1979, 1980) pointed out that separately identifying the effects of permanent income differentials and short-run income fluctuations is a key issue. Indeed, while out-migration from Europe was very volatile and followed US business cycles, migrants did not always hail from lower-income countries - especially between the 1870s and the 1890s. Spitzer (2015, 2016) dynamically analyzes pre-WWI Jewish emigration from the Pale of Settlement and shows how it underwent a process of spatial diffusion; its pro-cyclicality with respect to US business cycles reflected aggregate timing choices rather then long-run income differentials. Karadja and Prawitz (forth.) show how the initial wave of international out-migration from Sweden to the United States during the age of mass migration was sparked by a series of severe local frosts, interacted with proximity to the nearest port - a proxy for transportation costs. Such events prompted a process of path-dependency that self-sustained itself through networks. Abramitzky, Boustan and Eriksson (2012) suggest that wealth discouraged Norwegian out-migration toward the United States before WWI.

Third, it draws attention on a neglected natural disaster with extensive demographic and

[5]Similarly, the Great Mississippi Flood of 1927 intensified black out-migration from the Delta, an area traditionally characterized by segmented labor markets and racial institutions; following the decline in the supply of black labor, planters invested in capital-intense technologies that promoted the relative agricultural development of flooded areas - despite a certain landowners' reticence to part from the *status quo* (Hornbeck and Naidu, 2014).

[6]See Hatton and Williamson (1998) for a seminal contribution.

economic consequences.[7]

The remaining of the paper is organised as follows. Section 3.2 is concerned with the historical background; section 3.3 describes our main data sources; section 3.4 traces the econometric framework; section 3.5 describes results and threats to validity; section 3.6 concludes.

3.2 Historical Background

3.2.1 The earthquake and its aftermath

December 28th 1908, 5.21 a.m.: a violent earthquake shook the strait of Messina and its surroundings (figure 3.1, at page 107) for more than half a minute, with seismographs reporting a 7.1 magnitude on the Richter scale. A severe tsunami[8] hit the coastline shortly thereafter with waves as high as 40 ft while fires rapidly spread across the area. The damage was unprecedented: about 100,000 casualties,[9] countless injured and chaotic streams of refugees, with survivors being searched among the rubble for more that 20 days (De Santo, Bisaccia and De Santo, 2017).

Messina, Reggio Calabria, and a number of neighboring towns were almost entirely annihilated by what is regarded as the most destructive natural disaster having affected Europe in the past few centuries - certainly Italy after its unification in 1861 (Dickie, 2014).

In comparison, the death toll caused by the 1906 San Francisco earthquake is estimated at around 3,000 individuals - despite widespread destruction and economic losses. Figure 3.2 - at page 108 - depicts the original isoseismal map drawn by *Giuseppe Mercalli*, the renowned seismologist-volcanologist who created the widely used scale that measures the intensity of material damage wrought by an earthquake. After having visited these regions, he included an extra intensity degree, XI, "*catastrofe*".

[7] Based on a comprehensive panel of all natural disasters that took place between 1920 and 2010 in the US, Boustan et al. (2017) show that the migration response to natural disasters is evident only in the most severe cases.

[8] See Risk Management Solutions (2008) for a detailed description.

[9] Estimates vary from ca. 90,000 to 120,000 victims. See Bertolaso et al. (2008) and Caminiti (2009) for more details.

Figure 3.1: Earthquake epicentre (Latitude: 38.15; Longitude: 15.683) and 150-Km-radius-circle. *Source*: Guidoboni et al (2007), *Istituto Nazionale di Geofisica e Vulcanologia*.

Given that telegraph and railway lines were disrupted, the Italian government only had fragmentary news about the earthquake in its immediate aftermath. Official authorities did not realize the magnitude of what had happened in the area until the evening of December 29th: indeed, rescue and relief operations were initiated by British and Russian soldiers, whose ships docked in Messina before the Italian Navy.[10]

[10]Note that the timing of the disaster did not facilitate a rapid intervention because it took place shortly after Christmas. Russian ships were temporary moored in the port of Augusta - in southern Sicily - while part of the British fleet in the Mediterranean see was stationed in Malta. See, among others, Sciarrone (2013) for a reconstruction of the events.

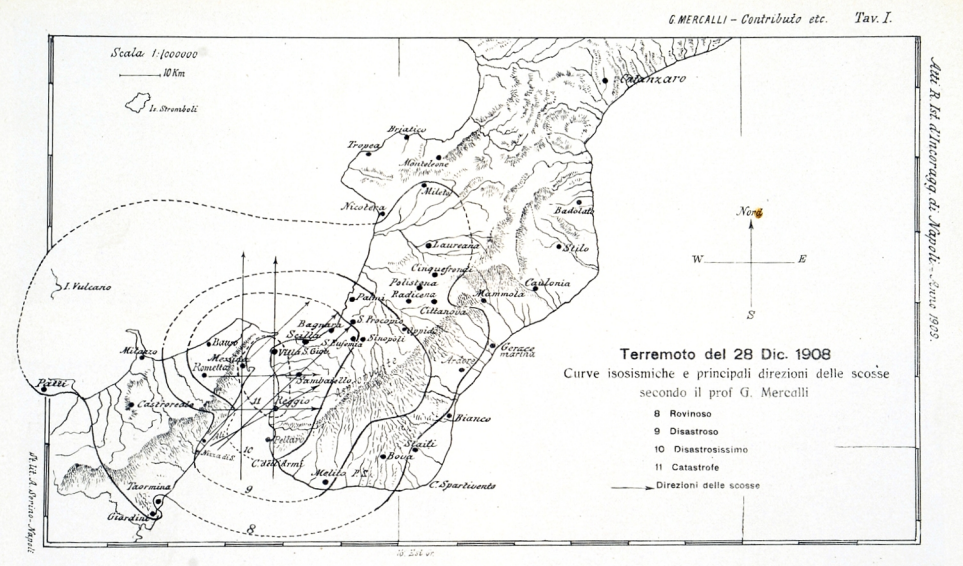


Figure 3.2: Original Mercalli drawing of isoseismal lines.
Source: *Atti del R. Istituto d'Incoraggiamento di Napoli (1909).*

Destruction was widespread enough to trigger debates on whether Messina should be rebuilt on the same site or simply leveled by the Italian military, with survivors being relocated elsewhere. On January 12th 1901, the Italian Parliament passed a law that approved the reconstruction of the city on the same site and allocated 30 million *lire* to the provision of barracks (Vianello, 2014). In order to restore basic municipal services as soon as possible, a barrack camp was built - and named "*Michelopoli*", after the Italian Member of Parliament *Giuseppe Micheli*. The Italian government - headed by *Giovanni Giolitti* - declared the State of siege, militarized the area and forcibly evacuated it to facilitate relief efforts and rescue operations. General *Francesco Mazza*, the special commissioner of the operation, established the curfew and ordered to summarily apply the death penalty to contrast and prevent looting. Aside from public order considerations, the measure had been justified by a widespread fear of epidemics and the necessity to protect bank vaults. When the state of emergency ended - on February 14th - there were more that 11,000 soldiers in Messina.

Because of its intensity and reach, this earthquake represents a watershed in terms of scientific research, urban planning and anti-seismic regulations.[11] According to Dickie (2014),

[11] Taylor (2016) argues that the the earthquake is a turning point in reportage photography too because

about 90 per cent of all building in Messina did not withstand the quake and 30,000 to 60,000 thousand individuals died, out of a population of 147,000 in 1901.[12] Reggio Calabria, the other commercial hub of the area, shared a similar fate.

Early observers (Omori, 1909; Hobbs, 1909) highlighted that the catastrophic tragedy resulted from an unfortunate combination of natural risk factors and human malpractice. The site exhibits congenital issues as topography itself made the area very vulnerable. While lying on a strategic sickle-shaped plain on the Sicilian side of the strait, Messina is located at the convergence between the African and Eurasian tectonic plates; the alluvial soil composition of the area allows seismic waves to diffuse through it without virtually any resistance, making it prone to landslides (Baratta, 1910).

Despite anti-seismic regulations had already been introduced after the disastrous quake of 1783, construction techniques were poor and security measures largely ignored.[13] The new architectural standards were based on the idea of *casa baraccata*:[14] in order to respond as an interconnected unit to quakes, each construction should have been supported by multiple wooden frames embedded in the brickwork; it should have had a maximum of 3 floors as well as a symmetrical blueprint and no vaults (Tobriner, 1983). The photographs taken in the aftermath of the quake clearly show how buildings were entirely stripped of external walls: the vast majority of them were too tall and had been erected with small stones held together by low-quality mortar.

No fewer than 40,000 refugees left the area in the immediate aftermath, spontaneously or forced by the government. These streams were mainly directed toward other large Sicilian cities and Naples - possibly because they were railway or port hubs - while smaller groups were welcomed by the provincial Rescue Committees that mushroomed across the peninsula

photographs acquired both a forensic and an immediate emotional value. These photographs often flowed into magazines, periodicals and journals like *Nature* and *The National Geographic*. Interestingly, the same pictures circulated across different media.

[12]The 1908 MRC earthquake is often referred to as, simplistically, the Messina earthquake as the city bore the highest cost in terms of absolute number of deaths.

[13]Moreover, note that after the 1894 earthquake, a violent event that hit the same area, the Italian government did not secure damaged buildings.

[14]In the spirit of the Portuguese *gaiola pombalina* introduced after the 1755 Lisbon earthquake.

(Parrinello, 2015). According to the official refugee census conducted in 1909, there were more than 20,000 refugees in Catania; about 11,000 refugees were recorded in Palermo, 2,600 in Syracuse and 8,000 in Naples. Official statistics suggests that this movement was, at least in part, short-lived, as the majority of refugees returned to the affected areas after a few weeks. Indeed, the population of Messina bounced back to 127,000 individuals in 1911, a figure that implies a large in-migration of newcomers. Such counter-movement was probably strengthened by temporary internal migration: anecdotal evidence suggests that individuals from neighboring towns might have temporarily migrated to Messina - and Reggio Calabria - in expectations of better employment opportunities.

The enormous amount of rubble produced by the quake had to be removed before the reconstruction could start. Given that Messina was the main citrus market in the Mediterranean Sea[15] and its port was the fourth largest in Italy, any permanent damage to this infrastructure would arguably be particularly detrimental for the economy of the area. There is anecdotal evidence suggesting that the main docks were not significantly damaged neither by the earthquake nor the *tsunami* and repaired quickly (Royal Society for the Encouragement of Arts and Commerce, 1910). Indeed, the port resumed its activities soon after the disaster - with the first citrus load departing during the second week of January. Bosworth (1981) and Mortara (1913) - in an article eloquently titled "*The economic revival of Messina*" - argue that trade and banking indices - e.g. cargo tonnage and value as well as commercial discounts and advances - were already back to pre-1908 levels in 1912. According to the author, Messina was bound to be re-built because of it strategic position on the strait.[16]

It is interesting to notice that Italian authorities stopped soup kitchens from distributing food in May 1909 because it was argued that those who wished to support themselves and their families through an employment had already found one in the citrus, port or construction

[15]Intensive commercial ties were interwoven with all major European countries. Interestingly, in the 1904-5, Russia was major citrus importer and trade routes connected Messina to Odessa.

[16]It is interesting to note that there were both an English and an American consulate in Messina. The British interest in re-building the city was probably connected to the sizable economic presence of British and British-Italian firms operating in the area since decades (Di Paola and Savasta, 2005).

industries. According to the Royal Society for the Encouragement of Arts and Commerce (1910), "*during the early part of 1909 relief stores, timber, and materials for the erection of temporary buildings were poured into the district, creating such a demand for manual labor that wages rose considerably*".[17]

International aid was swift. According to the Bank of Italy, private donations and international aid combined amounted to about 21 million *lire* by the end of February 1909 - roughly equivalent to 82 million current euro. International aid from the US translated into both funds and material support - e.g. wood and prefabricated barracks to be transported across the Atlantic Ocean. Provincial Rescue Committee also provided food and cloths while the Red Cross remained in the area for several months. However, the reconstruction process proved to be neither rapid nor sufficiently funded, with the anti-seismic urban regulation plan - *Piano Luigi Borzì* in Messina; *Piano Pietro De Nova* in Reggio Calabria - only being approved in 1911.[18] After the end of WWI these regions were still gigantic construction sites.[19] While huts were supposed to be an emergency measure, by 1922 there were more than 70,000 people living in temporary lodgings (Parrinello, 2015).

The Italian government provided assistance to refugees and established a centralized subsidy system to rescue commercial, artisanal and industrial activities.[20] In particular, the Bank of Italy - together with the Central Rescue Committee - appointed inspectors that would audit whether or not businesses were eligible for financial aid. However, Di Paola and Savasta (2005) argue that the complex application process delayed any rapid flows of

[17] Kirchberger (2017) documents the resilience of Indonesian labor markets after the 2006 Yogyakarta earthquake. The event determined a positive shock in the construction sector, raising demand and its marginal product. This prompted a downward shift of the labor supply in agriculture, thus increasing the marginal product of labor in this sector - as well as stimulating wage growth. See also, among others, Yang (2008).

[18] See Passalacqua (2007) for an interesting introduction to the architectural rationale of the reconstruction.

[19] According to US Censuses, in 1900 San Francisco's population was about 342,000; the figure increased to ca. 417,000 in 1910 despite the 225,000 refugees that were initially displaced by the event. There is evidence that a fifth of evacuees never returned, suggesting that internal migration from elsewhere played a key role. According to US Census Bureau figures, New Orleans resettlement after Katrina seems to be quite slow in comparison with its population today being still lower than in 2005, when the hurricane hit.

[20] Public investment - in the form of subsidies to local firms and reconstruction funds - might generate an additional labor demand and reduce would-be migrants incentives to relocate. See Boustan, Kahn and Rhode (2012) and Barone and Mocetti (2014).

resources to the area. According to the Bank of Italy, only 866 businesses were deemed eligible by April 1909 - over about 30,000 applications received by both the Central Rescue Committee and the Prefecture of Messina. A prerequisite to access funding was that the applicant returned in the quake-affected area and resided there. However, given that most of Messina was covered in rubble, those who lost their house and physical capital could not re-build their activities *in situ* and had to apply for the concession of state-owned land first.

On the other hand, subsidies were not supposed to be grants but, rather, loans granted at particularly favorable conditions. While widows, students and handicapped individuals were also granted subsidies, the official governmental guidelines explicitly stated a preference for male *capifamiglia*.

The allocation of monetary funds to business activities rather than individuals was highly debated but eventually motivated on the premise that, as the Duke of Aosta - the King's cousin and head of a relief committee - put it, *"it is immoral to look after a wave of vagabonds and to make men lazy"* (Bosworth, 1981). This idea was also supported by the Rescue Committee and agreed upon already on January 11th, during a general meeting.[21]

There is anecdotal evidence that the spatial and sectoral allocation of subsidies was influenced by governmental decisions. For example, Di Paola and Savasta (2005) document how silk production - a strategic industry - was prioritized over hospitality. Despite silk production was widespread in both the province of Messina and Reggio Calabria, the Central Rescue Committee opted to channel more resources to the latter because producers were already grouped in a consortium and several municipalities were reliant on the activity.

Restifo (1995) and Parrinello (2012) offer a descriptive overview of the population movements in the aftermath of the MRC quake, highlighting the desirability of further, disaggregate and evidence-based, research. However, little-to-no-attention has been paid to international migration. This natural disaster took place in the midst of a process of mass migration, when

[21]It is interesting to notice that many commentators refereed to southern Italians as indolent - even in the immediate aftermath of the quake. In contrast to what happened in San Francisco, the phoenix-like re-birth narrative didn't apply to neither Messina nor Reggio Calabria despite both cities promptly recover.

most of the affected individuals were likely to have overseas options: given that migration networks were already well-established and legal restrictions blunt, international migration was likely to represent a straightforward coping strategy.

3.2.2 Mass migration from Sicily and Calabria

Migration abroad from Calabria and Sicily bore resemblance to the *mezzogiorno* as a whole. Emigration from both regions remained relatively contained until the turn of the 20th century; it spiraled and morphed into a mass phenomenon throughout the 1900s. Unlike north Italian migration, the United States became a primary destination, absorbing the majority of migrants. The port of New York was by far the major gateway. The qualitative literature on internal migration across Italian provinces suggests that, while present, such flows were dwarfed by international migration, especially after 1900 (Gallo, 2012). Arguably, returns-to-migration and employment opportunities were greater in overseas destinations and it is likely that reaching these countries involved an easier and quicker voyage.

Figure 3.3 - at page 115 - displays US-bound (panel *a*) and total (panel *b*) provincial cumulative (1901-1914) out-migration rates while figure 3.4 - at page 116 - disaggregates stocks into their municipal dimension. Taken together, these maps highlight both a positive correlation between total and US-bound rates and a notable degree of variability across provinces and municipalities.

While providing a key role in relief operations through financial aid, barracks provision and reconstruction, the United States pointed out that assisted individuals could not relocate on American soil, meaning that neither provincial Rescue Committees nor the government could fund the voyage of would-be migrants.

The 1907 Immigration Act had already explicitly defined as *assisted* any person "*whose ticket or passage is paid for with the money of another, or who is assisted by others to come, unless it is affirmatively and satisfactorily shown that such person does not belong to one of the foregoing excluded classes, and that said ticket or passage was not paid for by*

any corporation, association, society, municipality, or foreign government, either directly or indirectly".[22] Given that refugees benefited from subsidies and relief funds, they fit this definition. However, as mentioned in the previous section, the procedure to access subsidies was very inefficient and only a minor percentage of applications had been processed in the year following the earthquake. The approach adopted by US authorities was neither new nor a consequence of the earthquake. Indeed, it is in line with the general principle that guided of pre-WWI immigration policies: the United States did not attempt to numerically restrict flows but, rather, focused on admitting passengers who could support themselves, i.e. *who were* - simply - *not likely to become public charge*.

[22]It also prevented unaccompanied children under 16 years of age from being admitted to the US, raised the head-tax to be paid by steamship companies on each passenger from 2 to 4 dollars and excluded imbeciles, feeble-minded persons, epileptics, insane persons, and professional beggars.

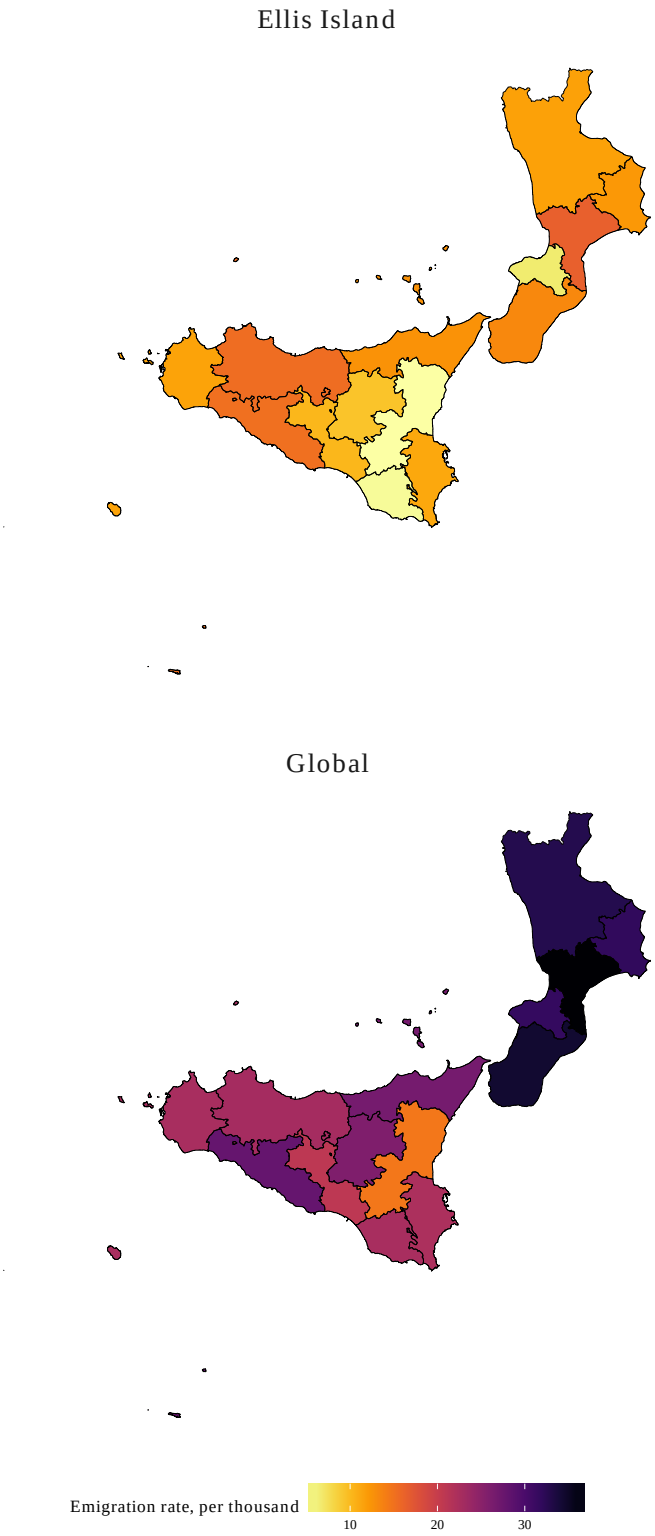


Figure 3.3: Average provincial emigration from Calabria and Sicily, 1901-1914.
Source: Ellis Island administrative records and official Italian statistics published by the *Direzione Generale della Statistica* (*ad annum*).
Notes: Panel *a* at the top of the page; panel *b* at the bottom of the page.

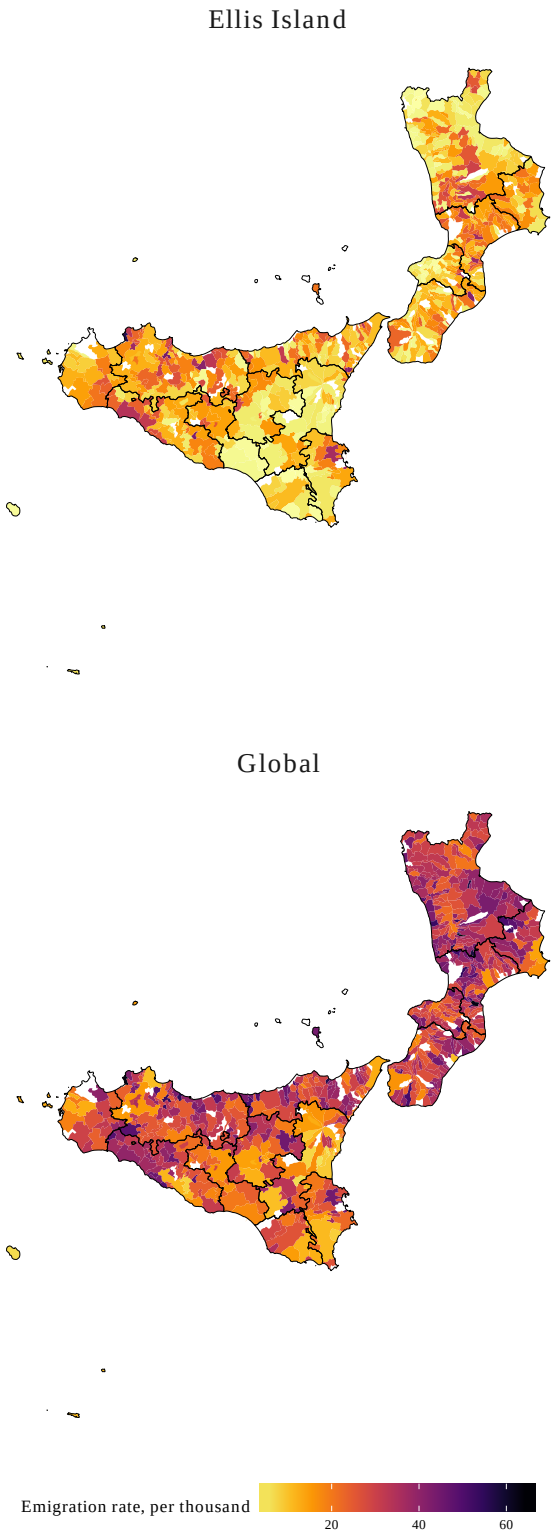


Figure 3.4: Average municipal emigration from Calabria and Sicily, 1901-1914.
Source: Ellis Island administrative records and official Italian statistics published by the *Direzione Generale della Statistica* (*ad annum*).
Notes: Panel *a* at the top of the page; panel *b* at the bottom of the page.

3.3 Data

3.3.1 Ellis Island Administrative Records

US-bound emigration data trace Ellis Island administrative records, a recently digitized collection of passenger manifests encompassing the vast majority of individuals travelling through the Port of New York between 1892 and 1924.[23] As far as Italian migration is concerned, the archive contains information on roughly 4.8 million passengers identified by name, marital status, age, date of arrival, ethnicity, nationality, and last place of residence. We utilize the latter piece of information to compute aggregate emigration flows from Italian municipalities. While places of birth might be more appropriate in other contexts, knowing passengers' last residences helps us reconstructing international population movements in the aftermath of the quake.

It should be noted that this data set is slightly different from the one detailed in Chapter 1. However, while resulting from partially different matching processes independently developed by the authors,[24] the two data sets exhibit very nuanced differences - especially throughout the 1900s, when migration from Italy peaked and the bulk of observations is clustered. Indeed, there is an almost perfect spatio-temporal correlation at the province-year level in terms of absolute migration. The key advantage of using this version of the data set is the availability of migration rates rather than absolute migration figures at the municipal level.

3.3.2 Overall migration rates

We supplement the data from the Ellis Island arrival records with data from the *Statistica della Emigrazione Italiana per l'Estero*, which was published biannually by the Italian *Direzione Generale della Statistica*. From this source, we digitized annual emigration data (which are not disaggregated by destination) for every Italian municipality for the period 1886-1920. We also associated this data with the geographic locations of modern municipalities. For the

[23] After the introduction of the Steerage Act (1819), shipping companies were legally bound to produce a list of passengers on board - and held responsible for any aliens entering the country illegally.

[24] See Spitzer and Zimran (2018) and section 1.2 in Chapter 1 for more details.

purposes of this paper, we restrict attention to the data for Calabria and Sicily for the years 1905-1912.

These official emigration statistics were compiled from the passports that Italians were required to obtain before departing from an Italian port. While overland emigration would be expected to generate an under-enumeration of emigrants from areas in northern Italy, where such emigration was common, the fact that most emigration in the south was to the Americas implies that it was likely mostly captured by this mechanism.

3.3.3 Earthquake intensity

Our data on earthquake intensities[25] are taken from the *Catalogue of Strong Earthquakes in Italy (461 B.C. - 1997) and the Mediterranean Area (760 B.C. - 1500)* (Guidoboni et al., 2007). Sponsored by the Italian Institute of Geophysics and Volcanology (INGV), the collection contains detailed geo-coded information on 1739 earthquakes based on original historical sources.[26]

As far as the MRC earthquake is concerned, aside from reporting the Richter magnitude and exact date/location of the epicenter, the archive provides us with both qualitative and quantitative accounts of damages, deaths, and injured individuals at the municipal level. Figures are based on a contemporary state-of-the-art study by Baratta (1910), an eminent Italian seismologist of the time. Intensity is measured using the Mercalli scale[27] and varies from a minimum of II (hardly felt) to a maximum of XI (catastrophic), including halves.

[25] According to the USGS, intensity is defined as "the effect of an earthquake on the Earth's surface".

[26] Note that, while insightful, cross-country analyses might entail an attenuation bias for they do not capture the circumscribed nature of disasters; on the other hand, data might not be entirely comparable across borders.

[27] Kirchberger (2017) suggests that reported figures might, in general, be biased. In order to address the issue, she suggests utilizing an exogenous measure of ground motions that captures the severity and distribution of shaking. The United States Geological Survey (USGS) does supply ShakeMaps on several events but, unfortunately, the MRC earthquake has not been covered. However, given the inherent quality of our main source, we believe that our data should not suffer from this issue: first, the compilation had been done by a single, highly competent, individual; second, he based his judgement on both official data gathered from local authorities and his own, direct, experience.

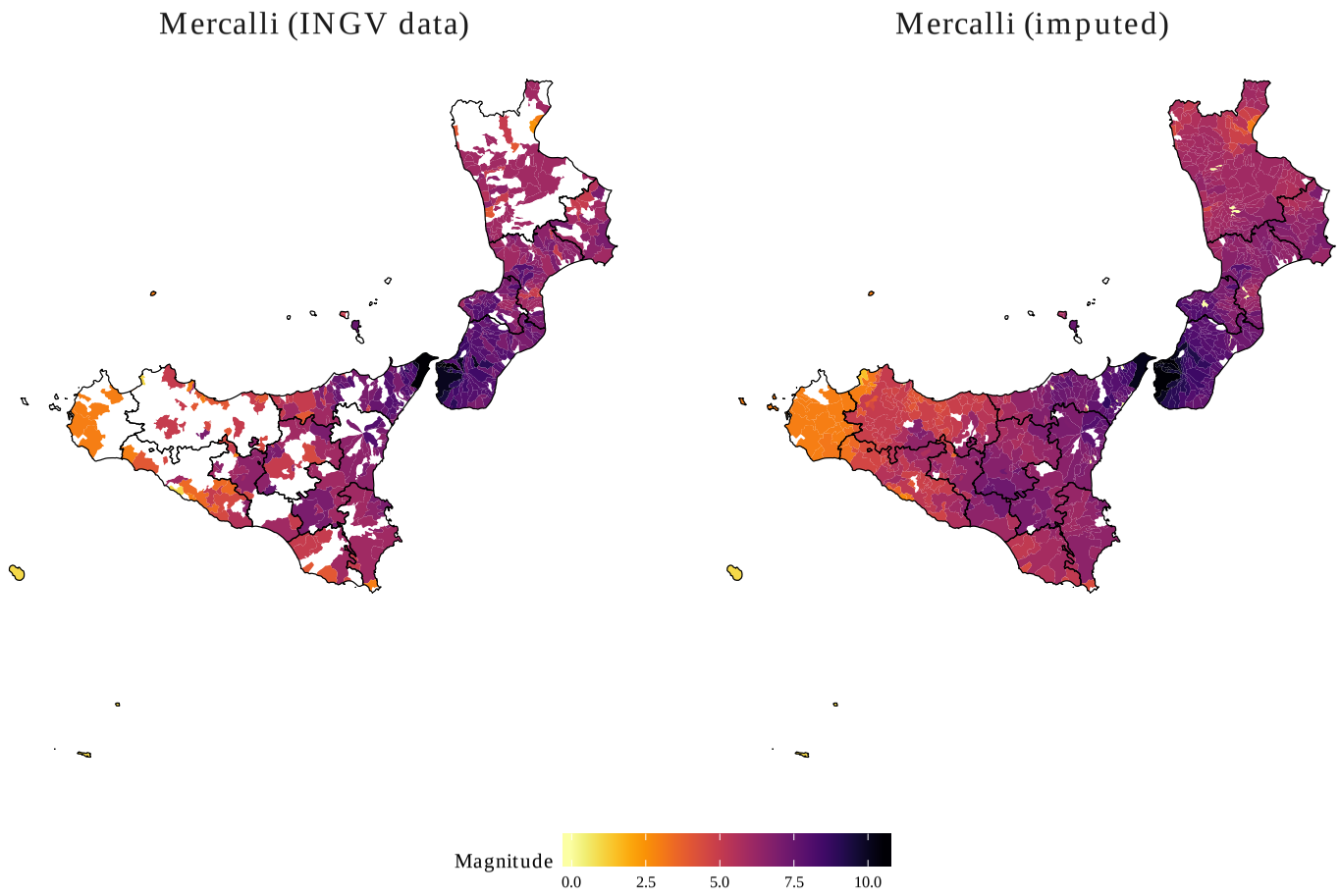


Figure 3.5: Mercalli intensities as reported in Guidoboni et al (2007).
Source: Guidoboni et al (2007), *Istituto Nazionale di Geofisica e Vulcanologia.*
Notes: We impute missing data making use of available observed Mercalli itensities and an inverse distance weighting technique based on their latitude and longitude. Panel *a* on the left; panel *b* on the right.

Figure 3.5 (panel *a*) - at page 119 - maps all the available raw information on earthquake intensities, with darker shaded areas being more affected.[28] Despite the comprehensive coverage of the earthquake area, we impute missing values using the following methodology.

Because there were some municipalities where no intensity measures were recorded, we made an imputation. In particular, we used the data on severity by latitude and longitude to impute the severity at unobserved latitude and longitude using an inverse distance weighting technique. Then, for municipalities where no severity measures were observed, we used the maximum of the imputation as the measure of severity. For municipalities where severity measures were observed, we retain the observed data rather than using the imputation. In practice, the impact of this approach on our preferred measure - the binary indicator of experiencing at least a severity of 8 on the Mercalli scale - is to set any municipality with unobserved data to have not experienced the earthquake severely. This is consistent with the fact that missing data are largely for municipalities further from the epicenter.

Figure 3.5 (panel *b*) depicts the result of the imputation. After having assigned a Mercalli value to each municipality, we construct a dummy variable taking value 1 whenever the intensity of the quake is greater than or equals 8. The rationale behind this approach draws on the USGS Modified Mercalli Intensity Scale classification[29] as there might be a non-linearity in the migration response if the earthquake exceeds a certain severity threshold. Figure 3.6 - at page 121 - is based on the above-mentioned dummy and confirms historical accounts of the earthquake having been particularly destructive in the surroundings of Messina and Reggio Calabria, both in terms of human losses and physical capital.

[28] Given that maps retrace current administrative borders, white municipal polygons could arise from different causes: first, there might be a mismatch between current and past geographical borders; second, information is missing because the quake was not felt - which might be possible in northern Calabria and Western Sicily - or information was not recorded at all.

[29] Available here: https://geomaps.wr.usgs.gov/parks/deform/gmercalli.html

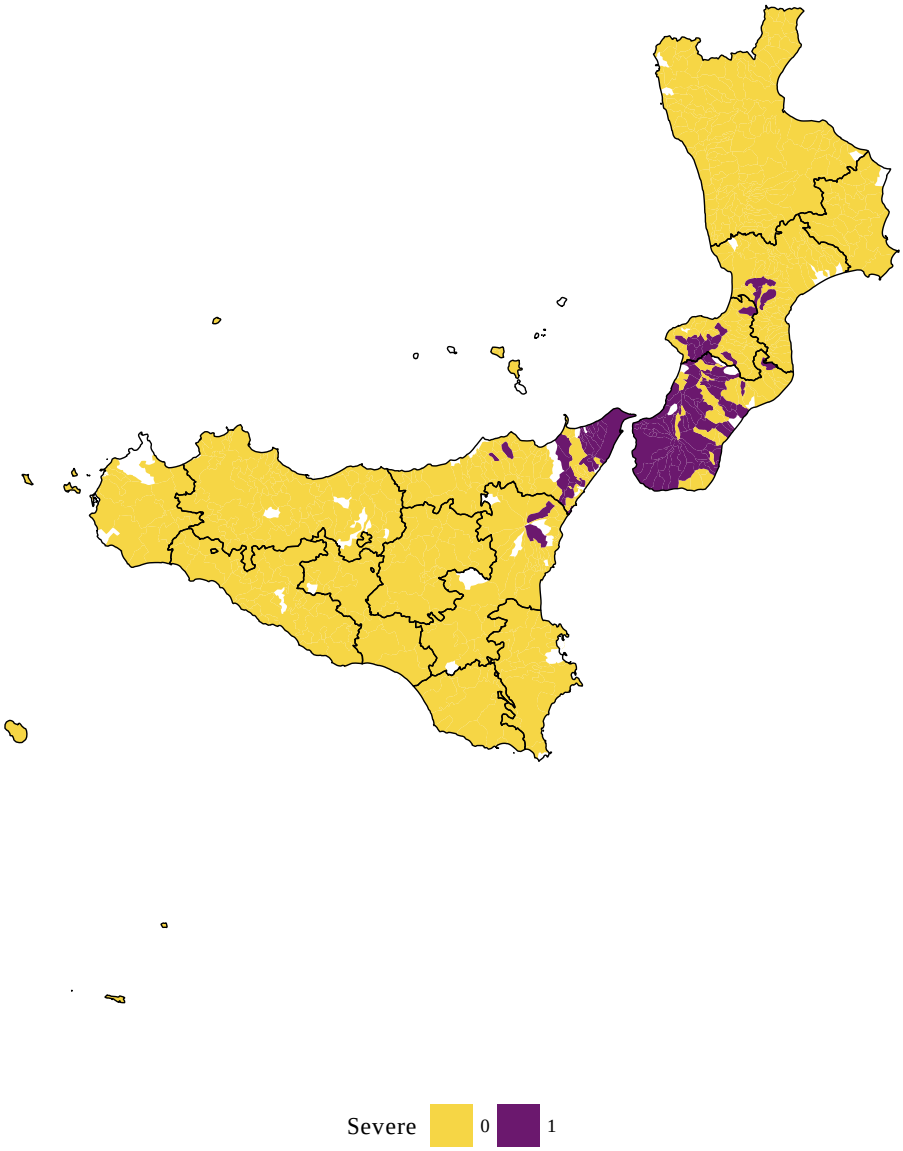


Figure 3.6: Severely affected municipalities (Mercalli intensity > 8).
Source: Guidoboni et al (2007), *Istituto Nazionale di Geofisica e Vulcanologia.*
Notes: We impute missing data making use of available observed itensities and an inverse distance weighting technique based on their latitude and longitude.

3.3.4 Summary Statistics

Table 3.1 - at page 124 - presents the main summary statistics (1905-1912), splitting the sample into Calabria and Sicily and differentiating between severely and non-severely affected municipalities. Figure 3.7 - at page 123 - shows the temporal evolution of migration flows from Calabria and Sicily, evidencing how migration pre-trends in both severely and non-severely affected municipalities largely coincided - at least on aggregate.

Note that there is an evident, short-lived, decline of out-migration rates in 1908. Despite out-migration from Italy plunged following the *Panic of 1907*, it had already resumed by mid-1909, arguably because would-be migrants simply postponed their voyage. This is especially true if transoceanic migration is regarded as a temporary investment.[30] While the Panic took place in October, the 1907 Immigration Act had already been passed in February - alongside the gentleman's agreement with Japan. This piece of legislation is unlikely to have caused the 1908 blip in migration rates from Italy *per se*. Although the Act re-asserted and broadened the definition of unwelcome individuals, aggregate Italian arrivals peaked in 1907, with November being the busiest month on record between 1900 and 1914. The plunge of migration rates in 1908 is therefore likely to mirror the fact that US-bound out-migration reacted pro-cyclically with respect to its business cycle - meaning that the labor supply was indeed quite elastic and information flowed well across the Atlantic. Indeed, although the consequences of the 1907-8 crisis spilled over sectors generally reliant on cheap immigrant labor such as construction and mining (Escamilla-Guerrero and Lopez-Alonso, 2018), aggregate migration flows recovered to pre-1907 levels by the late 1909 - alongside with the US economy as a whole (Noyes, 1909).[31]

[30] Note that Keeling (2012) depicts 1908 as a year of losses for transoceanic shipping companies as East-bound journeys did not compensate the drop in West-bound ones. According to Storhaug (2002), in 1908, returns (167,000) outnumbered arrivals (135,000).

[31] As detailed in the following section, we include province-by-year fixed effects in our specifications in order to capture year-to-year swings in international labor demand.

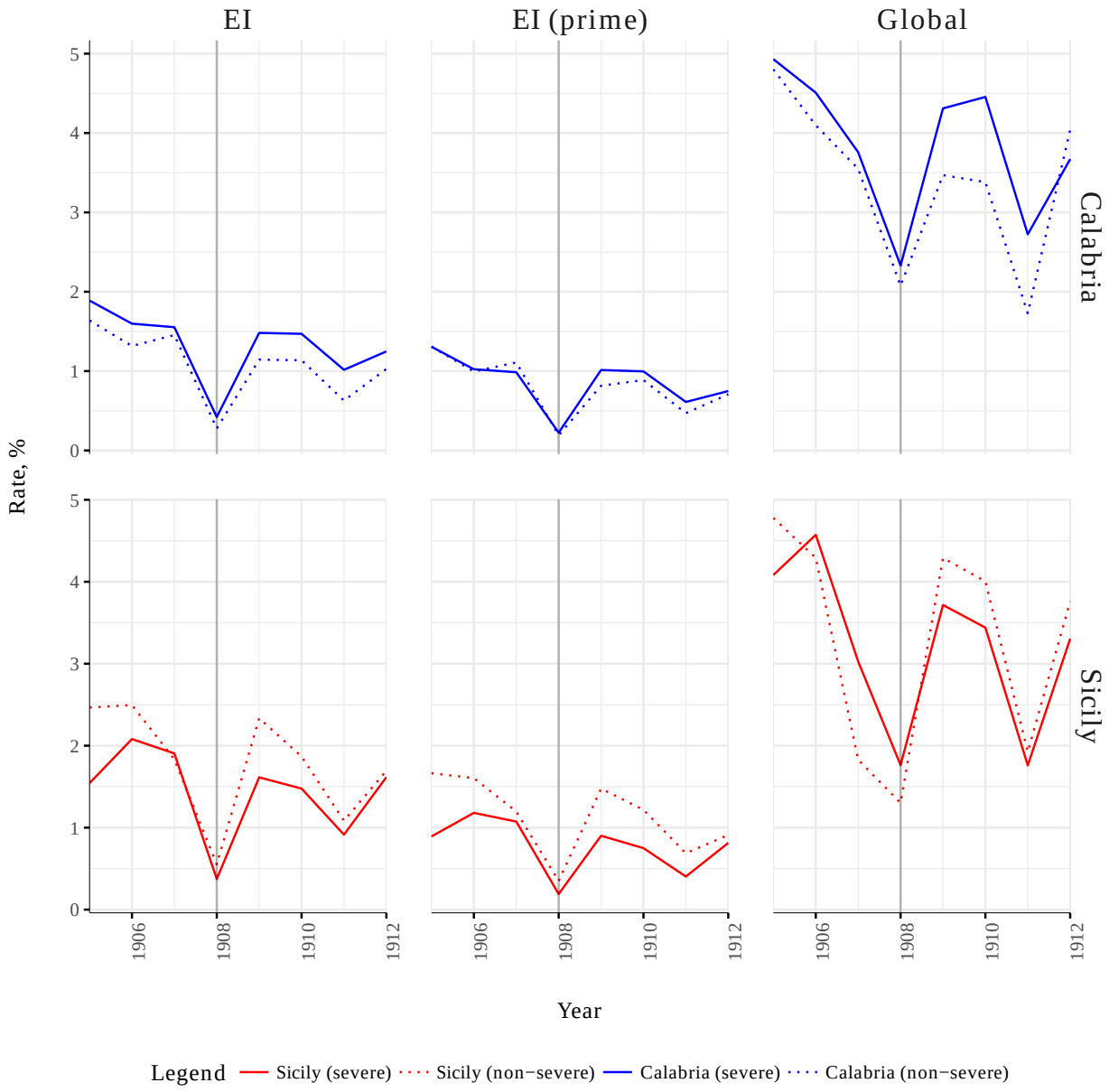


Figure 3.7: Yearly out-migration trends before and after the 1908 earthquake from severely and non-severely affected municipalities.
Source: Ellis Island administrative records and official Italian statistics published by the *Direzione Generale della Statistica* (*ad annum*).
Notes: Severe indicates Mercalli intensities greater than or equal to 8.

Table 3.1: Summary Statistics

	Calabria (Non-Severe)	Calabria (Severe)	Sicily (Non-Severe)	Sicily (Severe)
Emigration rate (Total)	3.72	3.39	3.13	3.27
	(1.82)	(1.9)	(2.11)	(2.03)
	[0; 16.39]	[0; 16.66]	[0; 18.4]	[0; 12.49]
Emigration rate (EI)	1.37	1.08	1.4	1.79
	(1.22)	(1.4)	(1.43)	(1.46)
	[0; 9.04]	[0; 12.5]	[0; 11.85]	[0; 8.61]
Emigration rate (EI, prime)	0.92	0.81	0.76	1.14
	(0.86)	(1.01)	(0.78)	(0.93)
	[0; 7.21]	[0; 8.37]	[0; 4.98]	[0; 6.19]
Distance from the epicenter	-4.65	-3.26	-4.76	-3.49
	(0.43)	(0.85)	(0.53)	(0.42)
	[-5.43; -3.09]	[-4.62; -1.66]	[-5.97; -3.2]	[-4.34; -2.74]
Mercalli intensity (INGV)	6.33	8.58	5.65	8.38
	(0.82)	(0.85)	(1.49)	(0.75)
	[2.5; 7.5]	[8; 10.5]	[1; 7.5]	[8; 10.5]
Mercalli intensity (Imputed)	6.41	8.68	5.79	7.98
	(1.16)	(1.35)	(1.67)	(1.53)
	[0; 8.87]	[0; 10.56]	[0; 7.93]	[0; 10.07]
Population	3476.75	4065.97	13320.22	9903.41
	(3400.23)	(5306.28)	(27511.28)	(27875.11)
	[621; 31824]	[546; 44415]	[462; 309694]	[833; 149778]
N. Obs.	1728	752	1368	264

Notes: Means are displayed without brackets; SD in round brackets; *Minima* and *maxima* in squared brackets. Severe indicates Mercalli intensities greater than or equal to 8.

3.4 Econometric Framework

3.4.1 DD

The core of our paper relies on evaluating how out-migration flows from Italian municipalities varied in response to the MRC earthquake intensity. We utilize years as preferred temporal units of analysis and estimate the following equation in a Difference in Difference framework:

$$Y_{m,t} = \beta(Treat_m \times Post_t) + \delta Post_{p,t} + \gamma_m + \varepsilon_{m,t} \qquad (3.1)$$

$Y_{m,t}$ represents the municipal migration rate from municipality m at time t. Migration rates are defined as the ratio migrants and population in 1901.[32] $Treat_m$ corresponds to different raw and imputed measures of local earthquake intensity, defined as in the previous section. Standard errors are clustered at the municipality level.[33] Earthquake-related casualties are accounted for as follows:

$$\widetilde{Y_{m,t}} = \beta(Treat_m \times Post_t) + \delta Post_{p,t} + \gamma_m + \varepsilon_{m,t} \qquad (3.2)$$

where:

$$\widetilde{Y_{m,t}} = \begin{cases} log(\frac{migrants}{populationin1901}) & t \leq 1908 \\ log(\frac{migrants}{populationin1901 - deaths}) & t \geq 1909 \end{cases}$$

We include both municipal (γ_m) and province-year ($Post_{p,t}$) fixed effects - which are therefore embodied in a provincial trend. We do not use weights and report population-weighted estimates in the appendix as robustness checks (see tables A3.1, A3.2 and A3.3). The rationale behind using weights is that municipalities contribute equally to the estimation: each of them is thought as an independent entity that might be affected by and react to a given shock idiosyncratically. Finally, we run various sets of regressions on the entire sample

[32] While having yearly population data would be desirable, we rely on figures taken from the last available census; inter-census reconstructions are not available.

[33] Note that there are cumulatively 778 municipalities in the two regions.

before splitting it between Calabria and Sicily in order to explore whether the earthquake had differential effects in the two regions.[34]

3.4.2 Event Studies

In order to test for pre-trends, we complement our DD analysis with event studies. The approach is based on the following equation and includes data from 1905 to 1912:[35]

$$Y_{m,t} = \alpha_c + \beta_t \, Treat_m + \delta_{p,t} + \varepsilon_{m,t} \qquad (3.3)$$

where α_m and $\delta_{p,t}$ are respectively municipality and province-year fixed effects. Standard errors are clustered at the municipality level.

3.5 Results

Note that the disposition of columns is consistent across the paper: we present death-adjusted estimates in even-numbered columns and non-death-adjusted estimates in odd-numbered columns; in the interest of space, we only report estimates based on the severe/non-severe dummy, constructed as explained above.

Tables 3.2 to 3.4 - at pages 127 and 128 - report unweighted[36] DD estimates ($Treat_c \times Post_t$ only) from equations (3.1) and (3.2), focusing on, respectively, both regions, Calabria, and Sicily. The earthquake did not produce a statistically significant shift in total out-migration rates from the affected area as a whole (table 3.2, Column 1) - even accounting for casualties (table 3.2, Column 2). Splitting the sample into Calabria and Sicily reveals that severely affected areas reacted differently to the quake: while total out-migration rates declined in Calabria (table 3.3, Columns 1 and 2), the response is opposite in Sicily (table 3.4, Columns

[34] Note that we also estimate equations (3.1) and (3.2) making use of an alternate definition of $Treat_m$: $d_c = -\,log(distance\ from\ the\ epicentre)$ where the *distance from the epicenter* is limited to a radius of 150 Km. Results are presented in tables A3.7, A3.8 and A3.9 in the appendix. We also use actual Mercalli intensities for each municipality and presents results in tables A3.4, A3.5 and A3.6.

[35] Note that changing the time frame of the analysis yields similar results.

[36] As already hinted at, we also provide population-adjusted estimates: tables A3.1, A3.2 and A3.3 are the weighted equivalents of tables 3.2 to 3.4.

Table 3.2: Calabria and Sicily - Unweighted estimates

VARIABLES	(1) Total	(2) Total (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	-0.0710	-0.0552	-0.1908	-0.1763	-0.2385*	-0.2237*
	[0.088]	[0.087]	[0.123]	[0.122]	[0.126]	[0.126]
Observations	2,862	2,862	2,731	2,731	2,693	2,693
Adjusted R-squared	0.371	0.371	0.476	0.476	0.440	0.441

Notes: This table reports the estimates of equations (3.1) and (3.2) for Calabria and Sicily, jointly, between 1907 and 1910. The key explanatory variable is defined as in section 3.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

Table 3.3: Calabria - Unweighted estimates

VARIABLES	(1) Total	(2) Total (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	-0.2798**	-0.2614**	-0.3474**	-0.3302**	-0.3966**	-0.3788**
	[0.113]	[0.113]	[0.165]	[0.165]	[0.182]	[0.182]
Observations	1,525	1,525	1,404	1,404	1,381	1,381
Adjusted R-squared	0.232	0.230	0.478	0.478	0.452	0.453

Notes: This table reports the estimates of equations (3.1) and (3.2) for Calabria only, between 1907 and 1910. The key explanatory variable is defined as in section 3.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

1 and 2).[37] As far as out-migration rates toward the United States are concerned, there is a statistically significant reduction in Calabria only (table 3.3, Columns 3 and 4). Restricting the analysis to men aged 18-55 yields consistently negative estimates for Calabria (table 3.3, Columns 5 and 6).

Figure 3.8 to 3.10 - at pages 129 to 131 - show estimated changes in migration rates for severely affected municipalities relative to changes for non-severely affected ones, accounting/not for casualties and setting 1907 as a reference year[38] - equation (3.3). Event studies

[37] Note that the absolute value of the coefficients is about the same but signs are opposite.
[38] We opted for 1907 because out-migration in 1908 was unusually low.

Table 3.4: Sicily - Unweighted estimates

VARIABLES	(1) Total	(2) Total (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	0.2706**	0.2824***	0.0156	0.0264	-0.0331	-0.0221
	[0.110]	[0.105]	[0.178]	[0.176]	[0.162]	[0.161]
Observations	1,337	1,337	1,327	1,327	1,312	1,312
Adjusted R-squared	0.433	0.433	0.473	0.473	0.427	0.427

Notes: This table reports the estimates of equations (3.1) and (3.2) for Sicily only, between 1907 and 1910. The key explanatory variable is defined as in section 3.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

convincingly support our DD analysis for Calabria. Indeed, consistent with the identification assumption, out-migration rates changed similarly before 1907 in severely and non-severely affected areas. Severely affected areas in Calabria experienced a statistically significant, short-lived, decline of both total and US-bound out-migration rates between 1908 and 1909. After 1909, there is evidence of out-migration rates recovering - despite yearly coefficients are rarely significant.

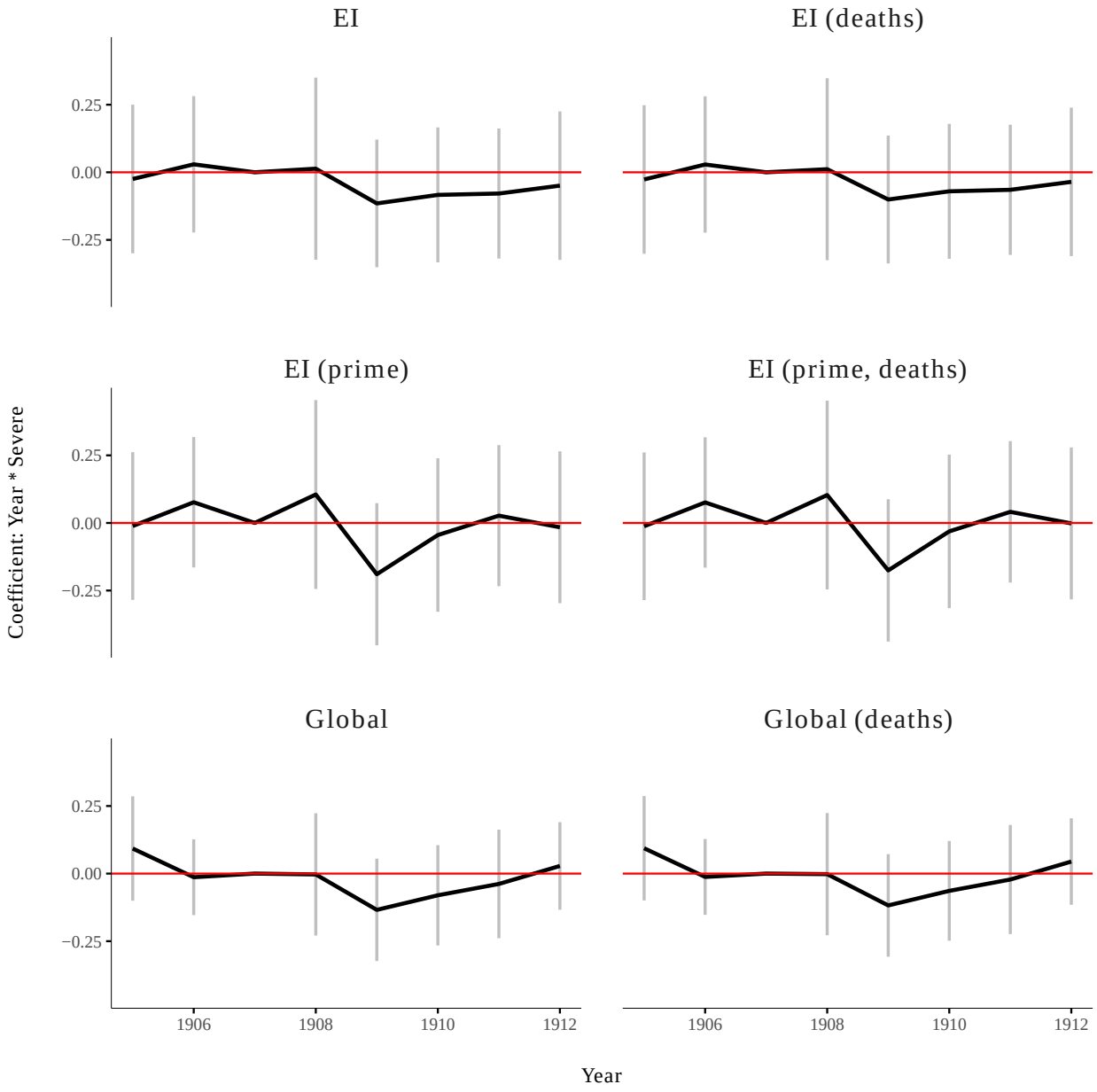


Figure 3.8: Event studies; Calabria and Sicily. These graphs report estimated differences in out-migration rates between severely (Mercalli intensity > 8) and non-severely affected municipalities based on equation (3.3). Segments indicate 95 per cent confidence intervals based on standard errors clustered at the municipality level.

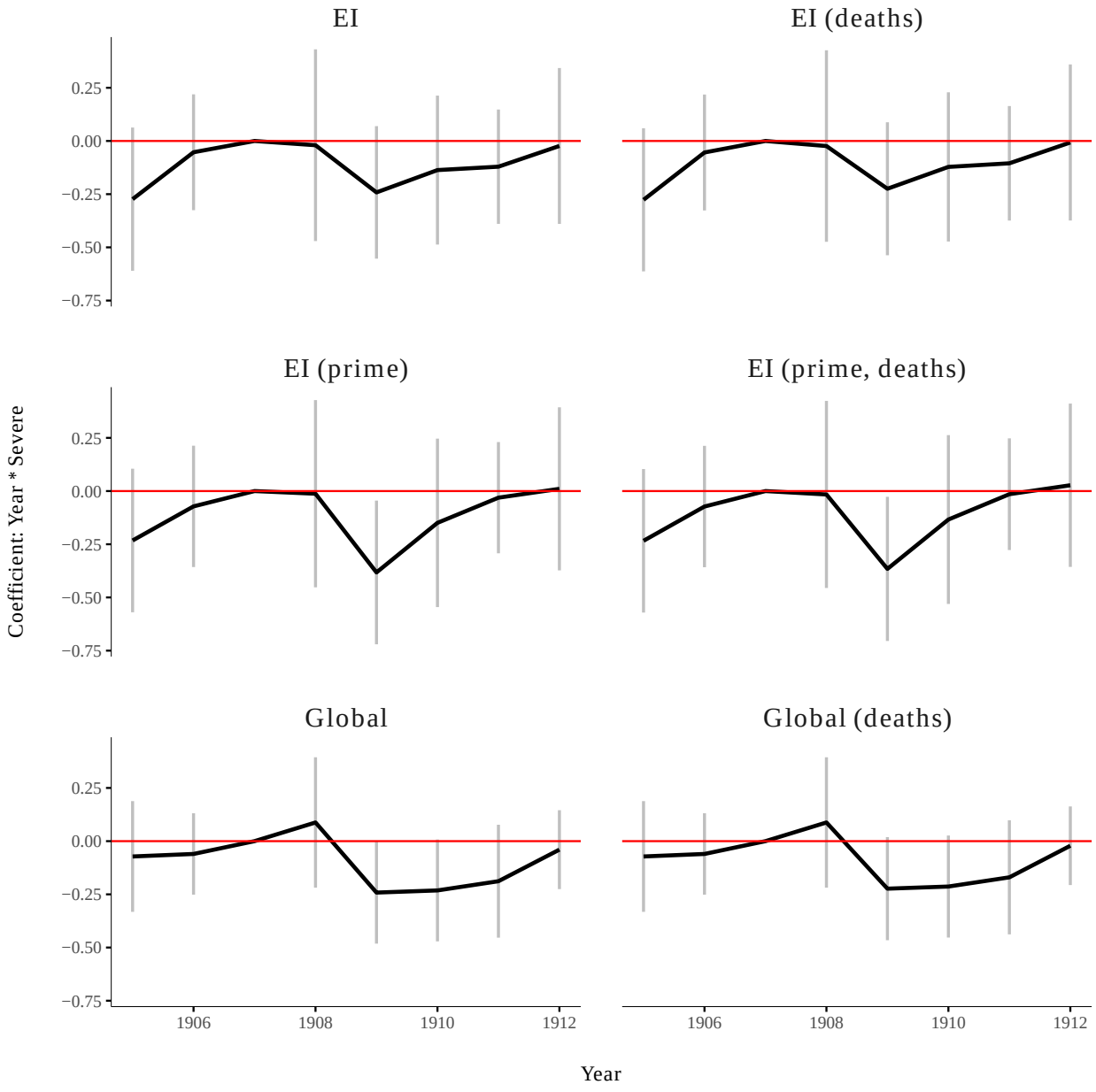


Figure 3.9: Event studies; Calabria only. These graphs report estimated differences in out-migration rates between severely (Mercalli intensity > 8) and non-severely affected municipalities based on equation (3.3). Segments indicate 95 per cent confidence intervals based on standard errors clustered at the municipality level.

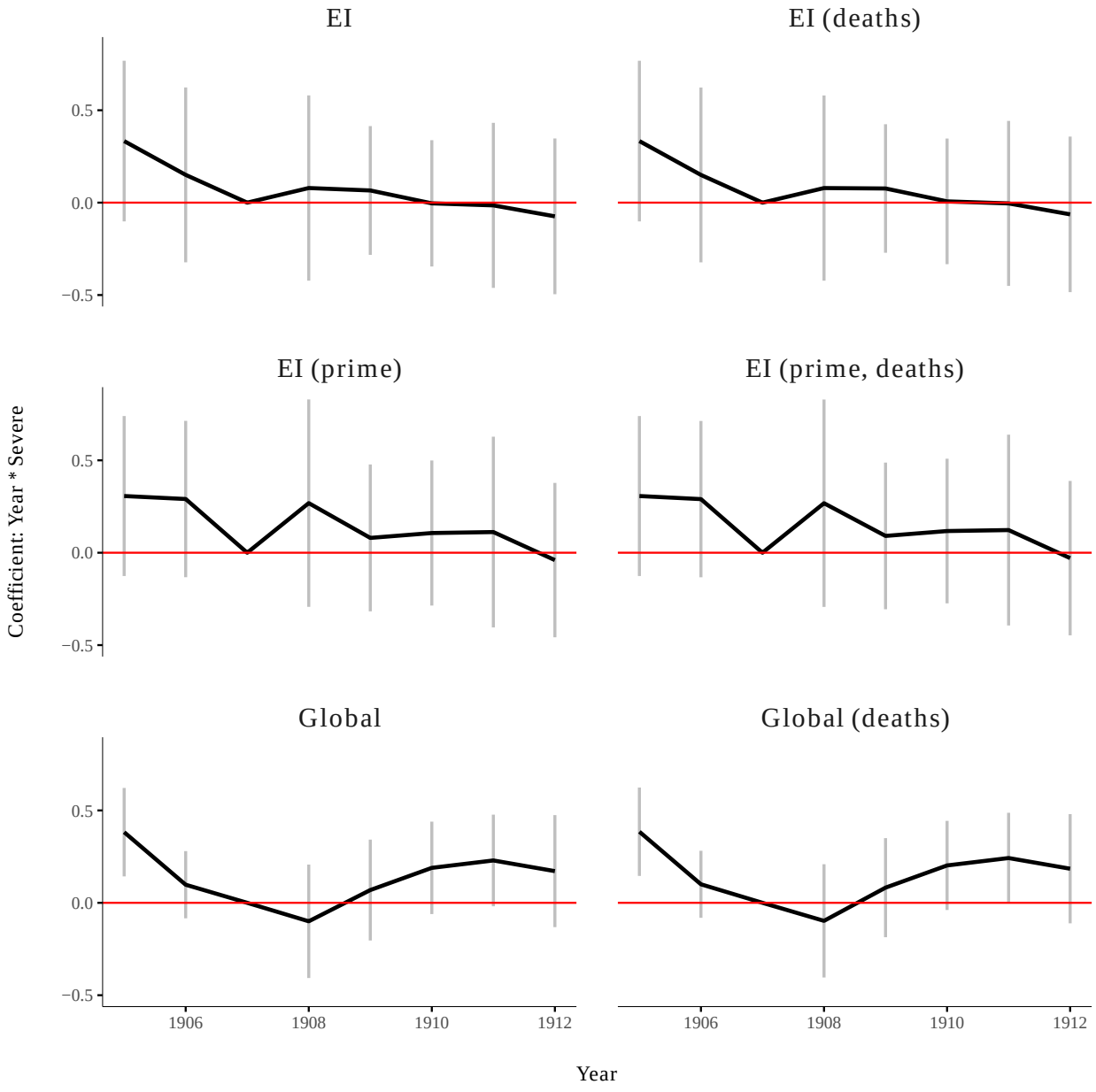


Figure 3.10: Event studies; Sicily only. These graphs report estimated differences in out-migration rates between severely (Mercalli intensity > 8) and non-severely affected municipalities based on equation (3.3). Segments indicate 95 per cent confidence intervals based on standard errors clustered at the municipality level.

3.5.1 Mechanisms

3.5.1.1 Pre-1904 migration

A differential response of out-migration rates to the MRC earthquake might be related to the vintage, dimension and cohesiveness of migration networks. Intuitively, networks might provide would-be migrants with financial support, logistic information and practical help once the decision to migrate has been taken. Given that US policies did not allow assisted migrants to relocate on American soil, the role of networks might have been crucial because only non-subsidized, economically self-sufficient, passengers could enter the country.

Table 3.5 - at page 133 - reports estimates based on quations (3.2) and (3.3), including cumulative municipal migration before 1904 as an additional control variable. The table is split into 3 sections relative to, respectively, both regions, Calabria and Sicily only. However, none of the estimates are robustly statistically significant, suggesting that the differential effect of the MRC earthquake in Calabria and Sicily might be due to other factors. When framed within the contexts of the age of mass migration and the recent economic literature on networks (McKenzie and Rapoport, 2007; Mahajan and Yang, 2017), these results are somehow surprising. However, our study seem to point out that the effect of networks could be non-linear in size, in the sense that networks affect migration decisions marginally when most of would-be migrants have some connection at destination: after the 1908 MRC earthquake hit the area, emigration from these regions had already morphed into a mass phenomenon. Indeed, on average, the municipal pre-1904 emigration to 1901 population ratio (per cent) was about 2.6, with a minimum of 1.7

Table 3.5: Pre-1904 migration stock

VARIABLES	(1) Total	(2) Total (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	0.3476	0.0696	-0.0202	-0.0926	0.2480	0.1733
	[0.443]	[0.476]	[0.286]	[0.279]	[0.272]	[0.268]
Severe * Post * Pre-1904 emigration (Total)	-0.1773	-0.0544				
	[0.184]	[0.198]				
Severe * Post * Pre-1904 emigration (EI)			-0.0879	-0.0416	-0.2611**	-0.2138*
			[0.127]	[0.122]	[0.121]	[0.118]
Observations	2,862	2,862	2,659	2,659	2,622	2,622
Adjusted R-squared	0.377	0.376	0.466	0.467	0.429	0.430
Severe * Post	0.1245	-0.1875	-0.3691	-0.4347	0.0675	-0.0007
	[0.486]	[0.504]	[0.371]	[0.365]	[0.386]	[0.384]
Severe * Post * Pre-1904 emigration (Total)	-0.1644	-0.0308				
	[0.187]	[0.194]				
Severe * Post * Pre-1904 emigration (EI)			0.0192	0.0642	-0.2288	-0.1829
Observations	1,525	1,525	1,381	1,381	1,358	1,358
Adjusted R-squared	0.233	0.230	0.466	0.468	0.445	0.446
Severe * Post	0.4129	0.1694	0.8890	0.7944	0.7651	0.6706
	[0.739]	[0.592]	[0.633]	[0.619]	[0.563]	[0.548]
Severe * Post * Pre-1904 emigration (Total)	-0.0660	0.0472				
	[0.331]	[0.262]				
Severe * Post * Pre-1904 emigration (EI)			-0.4634*	-0.4085	-0.4452*	-0.3903
Observations	1,337	1,337	1,278	1,278	1,264	1,264
Adjusted R-squared	0.443	0.443	0.463	0.463	0.411	0.411

Notes: This table reports the estimates of equations (3.1) and (3.2) for Calabria and Sicily jointly (upper section), Calabria only (mid section) and Sicily only (lower section), between 1907 and 1910, including an interaction term that accounts for pre-1904 cumulative migration. The key explanatory variable is defined as in section 3.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

3.5.1.2 Multiple-earthquake hypothesis

A continuous series of earthquakes have plagued these areas across centuries, exposing both regions to highly seismic hazards. Although the last similarly devastating earthquake dated back to 1783, geographically localized quakes had been frequent. In particular, central and northern Calabria were shaken by severe earthquakes in, respectively, 1905 and 1907. Damage accumulation and the progressive weakening of buildings, coupled with a regulatory vacuum, contributed to magnify the consequences of the 1908 MRC earthquake, possibly reverberating on out-migration rates.

Given that both the 1905 and 1907 earthquakes only affected Calabria, we test whether a multiple-earthquake hypothesis can explain why we only find a statistically significant reduction of out-migration rates from severely affected municipalities on the continental side of the strait. Figure 3.11 - at page 135 - displays the distribution of municipalities that were affected by the 1905 earthquake. A rapid comparison with figure 3.6 - at page 121 - shows that the area severely damaged - Mercalli intensity greater than or equal to 8 - by the 1905 earthquake is more confined as its epicenter lies in correspondence of *Capo Vaticano*, close to the city of *Vibo Valentia*, slightly to the north.

First, we check whether the 1905 earthquake impacted on out-migration rates from Calabria using equations (3.1) and (3.2) - i.e. we run the same analysis as before using 1905 data. We report estimates in table 3.6 - at page 136. Although we find no evidence of a statistically significant effect, we include an interaction term between the MRC and the 1905 earthquake. Table 3.7 - at page 136 - shows that there are no statistically significant effects on total out-migration rates. However, when considering US-bound out-migration rates, we find that the interaction between the two earthquakes is consistently negatively signed - with the 1908 coefficients still not being significant. This result suggests that the MRC earthquake affected migration rates through the accumulation of previous damage.

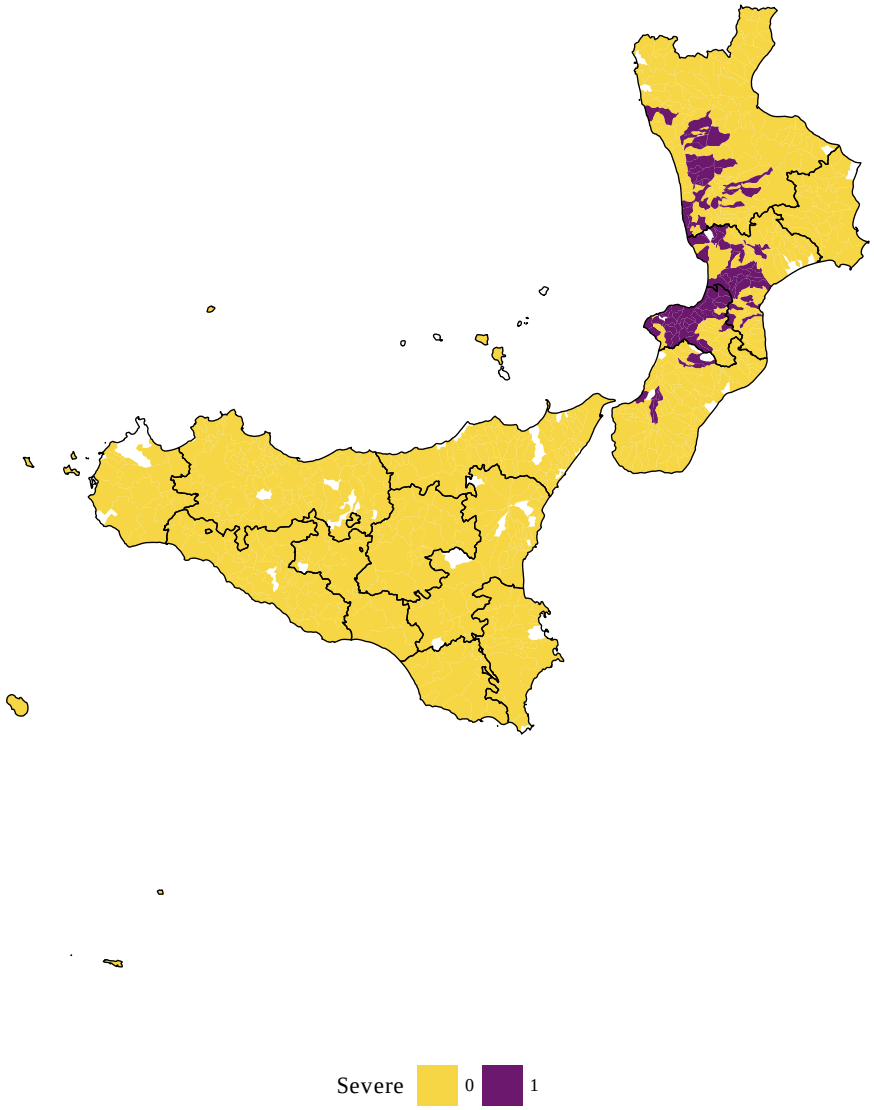


Figure 3.11: Severely affected municipalities in 1905 (Mercalli intensity > 8).
Source: Guidoboni et al (2007), *Istituto Nazionale di Geofisica e Vulcanologia*.
Notes: We impute missing data making use of available observed itensities and an inverse distance weighting technique based on their latitude and longitude.

Table 3.6: Calabria - 1905 earthquake

VARIABLES	(1) Total	(2) Total (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe 1905 * Post	0.0768	-0.0576	0.1383	0.1370	0.0950	0.0937
	[0.063]	[0.078]	[0.108]	[0.108]	[0.108]	[0.108]
Observations	1,521	1,525	1,404	1,404	1,381	1,381
Adjusted R-squared	0.198	0.226	0.476	0.477	0.450	0.451

Notes: This table reports the estimates of equations (3.1) and (3.2) for Calabria only, using 1905 earthquake data. The key explanatory variable is defined as in section 3.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

Table 3.7: Calabria - Interaction between the 1905 and 1908 earthquakes

VARIABLES	(1) Total	(2) Total (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	-0.1850	-0.1541	-0.1097	-0.0833	-0.1882	-0.1614
	[0.122]	[0.122]	[0.169]	[0.168]	[0.181]	[0.180]
Severe * Post * Severe 1905	- 0.2199	-0.2487	-0.7675***	-0.7915***	-0.6790**	-0.7032**
	[0.222]	[0.221]	[0.288]	[0.288]	[0.317]	[0.316]
Observations	1,525	1,525	1,404	1,404	1,381	1,381
Adjusted R-squared	0.232	0.231	0.481	0.482	0.454	0.455

Notes: This table reports the estimates of equations (3.1) and (3.2) for Calabria only, between 1907 and 1910, including an interaction term that accounts for the 1905 earthquake. The key explanatory variables are defined as in section 3.3, imputing missing values (they are a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

3.5.2 Threats to validity

When averaged within treatment and control groups, our key dependent variable - however defined - displays a large within-group variation over time. Stated differently, given that the relatively small number of treated municipalities (about 40), the idiosyncratic commune-year variation might be too large to be averaged out when comparing treatment and control groups. Moreover, affected municipalities are likely to be spatially correlated with each other.

An earthquake might indirectly influence local labor markets producing spillovers in non-severely affected municipalities, thus leading to interpreting our results as an under-estimation. Despite controlling for previous provincial out-migration trends, there might be concerns that non-severely affected municipalities are a weak control group for severely affected ones.

3.5.2.1 Falsification

We conduct a falsification exercise that consists in estimating a slightly modified version of equation (3.1), assuming that the earthquake affected southern Italy in 1906 rather than 1908. Given the way our key explanatory variable is constructed (i.e. it is a dummy variable), this is essentially the same as asking the data whether there are any significant differences between severely and non-severely affected municipalities beforehand: any significant coefficients would signal that emigration patterns from severely affected municipalities were following partially different patterns before 1908.

Coupled with event studies, table 3.8 - at page 138 - shows that these concerns are mild in the case of both regions analyzed jointly (first section) and Calabria (second section): indeed, the falsification test resulted in no statistically significant coefficients.[39] There are, however, some concerns about Sicily (third section): despite not being significant in the case of US-bound out-migration, coefficients are negative and statistically significant when considering total out-migration rates. On the other hand, especially if coupled with the

[39]For completeness, we run the same falsification exercise using the all the above-mentioned variants of our key explanatory variable and accounting for deaths, finding no other statistically significant coefficients.

Table 3.8: Falsification Exercise

VARIABLES	(1) Total	(2) EI	(3) EI (prime)
Severe * Post 1906	-0.0686	0.0907	0.0908
	[0.077]	[0.139]	[0.131]
Observations	2,860	2,715	2,683
Adjusted R-squared	0.449	0.488	0.451
Severe * Post 1906	-0.0095	0.1688	0.1623
	[0.063]	[0.130]	[0.125]
Observations	1,523	1,405	1,388
Adjusted R-squared	0.298	0.507	0.468
Severe * Post 1906	-0.3608***	-0.2227	-0.1805
	[0.118]	[0.221]	[0.196]
Observations	1,337	1,310	1,295
Adjusted R-squared	0.507	0.470	0.432

Notes: This table reports the estimates of equations (3.1) and (3.2) for Calabria and Sicily jointly (upper section), Calabria only (mid section) and Sicily only (lower section), assuming that the earthquake happened in 1906. The key explanatory variable is defined as in section 3.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

evidence already presented in the event studies using Ellis Island data (figure 3.10, page 131), it's hard to discern a pre-trend that would invalidate our estimates: the fact that such pre-trend is more visible only when using official statistics might be due to the lower quality of official statistics with respect to Ellis Island data. Moreover, the results of the falsification test can be partially explained by the fact that the southeast of Sicily - especially the province of Syracuse - which is also further away from the epicenter - was a late comer and migration rates from this area appear to have raised faster during the 1900s.

3.6 Conclusion

The attention on climate change and migration is likely to draw the limelight on any forced or deliberate population movements that originate from nature-induced crises. While the repercussions of a natural disaster might be somehow specific to its type, intensity and location, these events often interact with social and economic fabrics, activating a wide range of coping mechanisms that influence the migration response of affected areas: individuals might react ex-ante, migrating in anticipation of a natural disaster; on the other hand they could opt for an ex-post strategy, smoothing household or personal income losses through internal or international migration. Whether or not migration is internal or international, temporary or permanent, depends - among other contingent factors - on when the disaster hits and what immigration policies other countries adopt. For instance, the potential of migration networks to attract migrants is certainly different during the initial phases of migration then it is when paths are well established and migrants' communities are settled in the destination countries. Similarly, external factors such as immigration policies determine what options are actually available to would-be migrants and evacuees. Nowadays, it is often the case that international migration is hampered by restrictive legal barriers. Moreover, a great deal of importance is placed on whether or not the migration is driven by economic variables or relates to refugees. Understanding how migration responds to natural disasters at the local level is therefore of prime interest to both researchers and policy makers.

This paper is framed within the context of one of the most destructive natural disasters in recent European history, the earthquake that hit the areas surrounding the strait of Messina in 1908. This catastrophe happened in the midst of a process of mass migration toward the United States, where most of would-be migrants could arguably rely on a network because local communities had already mushroomed in the country. Moreover, there were virtually no legal restrictions to such flows: the United States - as well as other overseas destination countries - implemented a next-to-open border policy. Overall, this paper contributes to the understanding of post-disaster international migration dynamics within a unique legal

framework during the age of mass migration. In particular, our results do not show an obvious surge of migration toward the most likely international destination in the immediate aftermath of an extreme natural disaster.

References

Abramitzky, Ran and Leah Boustan. 2017. "Immigration in American economic history." *Journal of Economic Literature* 55(4):1311–45.

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2012. "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." *The American economic review* 102(5):1832–1856.

Baratta, Mario. 1910. *La catastrofe sismica Calabro messinese (28 dicembre 1908)*. Società geografica italiana.

Barone, Guglielmo and Sauro Mocetti. 2014. "Natural disasters, growth and institutions: a tale of two earthquakes." *Journal of Urban Economics* 84:52–66.

Bazzi, Samuel. 2017. "Wealth heterogeneity and the income elasticity of migration." *American Economic Journal: Applied Economics* 9(2):219–255.

Bertolaso, Guido, Enzo Boschi, Emanuela Guidoboni and Gianluca Valensise. 2008. *Il terremoto e il maremoto del 28 dicembre 1908: analisi sismologica, impatto, prospettive*. Istituto Nazionale di Geofisica e Vulcanologia.

Bosworth, Richard JB. 1981. "The Messina earthquake of 28 December 1908." *European Studies Review* 11(2):189–206.

Boustan, Leah Platt, Matthew E Kahn and Paul W Rhode. 2012. "Moving to higher ground: Migration response to natural disasters in the early twentieth century." *The American Economic Review* 102(3):238–244.

Boustan, Leah Platt, Matthew E Kahn, Paul W Rhode and Maria Lucia Yanguas. 2017. The Effect of Natural Disasters on Economic Activity in US Counties: A Century of Data. Technical Report No. 23410 NBER.

Caminiti, Luciana. 2009. *La grande diaspora: 28 dicembre 1908: la politica dei soccorsi tra carità e bilanci*. GBM.

Cattaneo, Cristina and Giovanni Peri. 2016. "The migration response to increasing temperatures." *Journal of Development Economics* 122:127–146.

De Santo, Natale Gaspare, Carmela Bisaccia and Luca Salvatore De Santo. 2017. "Maximum time-to-rescue after the 1908 Messina-Reggio Calabria Earthquake was 20 days: hints for disaster planning?" *Prehospital and disaster medicine* 32(3):249–252.

Di Paola, Maria Teresa and Sem Savasta. 2005. *Tra le righe. Il terremoto del 1908 attraverso i documenti inglesi e la presenza britannica nell'area dello Stretto*. Istituto di Studi Storici Gaetano Salvemini.

Dickie, John. 2014. *Una catastrofe patriottica. 1908: il terremoto di Messina*. Gius. Laterza & Figli, Roma-Bari.

Escamilla-Guerrero, D. and M. Lopez-Alonso. 2018. "Self-selection of Mexican migrants in the presence of random shocks: Evidence from the Panic of 1907." *Manuscript* .

Fauri, Francesca. 2015. *Storia economica delle migrazioni italiane*. Il mulino.

Gallo, Stefano. 2012. *Senza attraversare le frontiere: le migrazioni interne dall'unità a oggi*. Gius. Laterza & Figli Spa.

Gould, John D. 1979. "European inter-continental emigration 1815-1914: patterns and causes." *Journal of European Economic History* 8(3):593.

Gould, John D. 1980. "European International Emigration: The Role of Diffusion and Feedback." *Journal of European Economic History* 9(2):267.

Gray, Rowena, Gaia Narciso and Gaspare Tortorici. 2018. "Globalization, agricultural markets, and mass migration." Unpublished manuscript.

Gröschl, Jasmin and Thomas Steinwachs. 2016. "Do Natural Hazards Cause International Migration?" *CESifo Economic Studies* .

Guidoboni, E, G Ferrari, D Mariotti, A Comastri, G Tarabusi and Gianluca Valensise. 2007. "Catalogue of Strong Earthquakes in Italy (461 BC-1997) and Mediterranean Area (760 BC-1500)." *Annali di geofisica e vulcanologia* .

Hatton, Timothy J and Jeffrey G Williamson. 1998. *The age of mass migration: Causes and economic impact*. Oxford University Press on Demand.

Hobbs, William Herbert. 1909. "The Messina Earthquake." *Bulletin of the American Geographical Society* 41(7):409–422.

Hornbeck, Richard. 2012. "The enduring impact of the American Dust Bowl: Short-and long-run adjustments to environmental catastrophe." *The American Economic Review* 102(4):1477–1507.

Hornbeck, Richard and Suresh Naidu. 2014. "When the levee breaks: black migration and economic development in the American South." *The American Economic Review* 104(3):963–90.

Karadja, Mounir and Erik Prawitz. forth. "Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States." *Journal of Political Economy* .

Keeling, Drew. 2012. "Transatlantic shipping cartels and migration between Europe and America, 1880-1914." *Essays in Economic & Business History* 17.

Kirchberger, Martina. 2017. "Natural disasters and labor markets." *Journal of Development Economics* 125:40–58.

Long, Jason and Henry E Siu. 2016. Refugees from dust and shrinking land: Tracking the Dust Bowl migrants. Technical Report No. 22108 NBER.

Mahajan, Parag and Dean Yang. 2017. Taken by storm: Hurricanes, migrant networks, and US immigration. Technical Report No. 23756 NBER.

McKenzie, David and Hillel Rapoport. 2007. "Network effects and the dynamics of migration and inequality: theory and evidence from Mexico." *Journal of development Economics* 84(1):1–24.

Millock, Katrin. 2015. "Migration and environment." *Annual Review of Resource Economics* 7(1):35–60.

Mortara, Giorgio. 1913. "The Economic Revival of Messina." *The Economic Journal* 23(91):438–442.

Noyes, Alexander D. 1909. "A Year after the Panic of 1907." *The Quarterly Journal of Economics* 23(2):185–212.

O Gràda, Cormac and Kevin H O'Rourke. 1997. "Migration as disaster relief: Lessons from the Great Irish Famine." *European Review of Economic History* 1(1):3–25.

Omori, Fusakichi. 1909. "Preliminary report on the Messina-Reggio earthquake of Dec. 28, 1908." *Bulletin of the Imperial Earthquake Investigation Committee* 3(2).

Parrinello, Giacomo. 2012. "Post-disaster migrations and returns in Sicily: the 1908 Messina earthquake and the 1968 Belice Valley earthquake." *Global Environment* 5(9):26–49.

Parrinello, Giacomo. 2015. *Fault Lines: Earthquakes and Urbanism in Modern Italy*. Vol. 6 Berghahn Books.

Passalacqua, Federica. 2007. *Messina e Reggio prima e dopo il terremoto del 28 dicembre 1908*. GBM.

Restifo, Giuseppe. 1995. "Local administrative sources on population movements after the Messina earthquake of 1908." *Annals of Geophysics* 38(5-6).

Risk Management Solutions. 2008. The 1908 Messina Earthquake: 100-year retrospective. Technical report RMS.

Royal Society for the Encouragement of Arts, Manufactures and Commerce. 1910. "Messina since the Earthquake." *Journal of the Royal Society of Arts* 58(3015):919–920.

Sciarrone, Roberto. 2013. "Messina Earthquake, 1908." *Academic Journal of Interdisciplinary Studies* 2(2):317.

Spitzer, Yannay. 2015. "The Dynamics of Mass Migration: Estimating the Effect of Income Differences on Migration in a Dynamic Model of Discrete Choice with Diffusion." Unpublished manuscript, Brown University, Deptartment of Economics.

Spitzer, Yannay. 2016. "Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914." Unpublished manuscript, Norwestern University, Department of Economics.

Spitzer, Yannay and Ariell 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.

Stark, Oded and David E Bloom. 1985. "The new economics of labor migration." *The American Economic Review* 75(2):173–178.

Storhaug, Hans. 2002. Return migration: numbers, reasons and consequences. A European overview. In *Opening lecture to the annual AEMI meeting*.

Taylor, William M. 2016. "Assembling Ruin: Rubble Photography of the 1908 Messina Earthquake." *Transformations* .

Tobriner, Stephen. 1983. "La Casa Baraccata: earthquake-resistant construction in 18th-century Calabria." *Journal of the Society of Architectural Historians* 42(2):131–138.

Tse, C. 2012. "Do natural disasters lead to more migration? Evidence from Indonesia." *Social science research network publication. http://ssrn. com/abstract* 1906556.

United Nations. 2015. Global Report on Disaster Risk Reduction. Technical report United Nations Office for Disaster Risk Reduction.

Vianello, Giulia. 2014. Messina 1908: terremoto e ricostruzione. B.S. thesis Università Ca'Foscari Venezia.

Yang, Dean. 2008. "Risk, migration, and rural financial markets: Evidence from earthquakes in El Salvador." *Social Research: An International Quarterly* 75(3):955–992.

Appendix 3

Tables

Table A3.1: Calabria and Sicily - Weighted estimates

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	-0.2224	-0.0129	-0.2797**	-0.0676	-0.3093**	-0.0963
	[0.153]	[0.070]	[0.137]	[0.099]	[0.138]	[0.112]
Observations	2,862	2,862	2,731	2,731	2,693	2,693
Adjusted R-squared	0.540	0.538	0.510	0.515	0.448	0.454

Notes: This table reports the estimates of equations (1) and (2) for Calabria and Sicily, jointly, between 1907 and 1910 using population weights. The key explanatory variable is definied as in section 4.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3.2: Calabria - Weighted estimates

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	-0.2389**	-0.1802*	-0.2454*	-0.1836	-0.3325*	-0.2700
	[0.109]	[0.103]	[0.147]	[0.149]	[0.173]	[0.172]
Observations	1,525	1,525	1,404	1,404	1,381	1,381
Adjusted R-squared	0.354	0.337	0.528	0.533	0.482	0.487

Notes: This table reports the estimates of equations (1) and (2) for Calabria only, between 1907 and 1910 using poulation weights. The key explanatory variable is definied as in section 4.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3.3: Sicily - Weighted estimates

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	-0.2096	0.1157	-0.3035	0.0129	-0.2933	0.0237
	[0.2616]	[0.112]	[0.203]	[0.125]	[0.202]	[0.130]
Observations	1,337	1,337	1,327	1,327	1,312	1,312
Adjusted R-squared	0.538	0.538	0.504	0.511	0.432	0.438

Notes: This table reports the estimates of equations (1) and (2) for Sicily only, between 1907 and 1910 using population weights. The key explanatory variable is definied as in section 4.3, imputing missing values (it is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale). All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

Table A3.4: Calabria and Sicly - Mercalli intensities

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	-0.0765	-0.0562	-0.2112*	-0.1920	-0.2555**	-0.2358*
	[0.089]	[0.088]	[0.124]	[0.124]	[0.128]	[0.127]
Observations	1,858	1,858	1,783	1,783	1,762	1,762
Adjusted R-squared	0.381	0.381	0.480	0.481	0.449	0.450

Notes: This table reports the estimates of equations (1) and (2) for Calabria and Sicily, jointly, between 1907 and 1910. The key explanatory variable is based on the available Mercalli intensities reported in Guidoboni et al (2007). It is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale. All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3.5: Calabria - Mercalli intensities

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	-0.2632**	-0.2450**	-0.3504**	-0.3331**	-0.3975**	-0.3797**
	[0.114]	[0.114]	[0.165]	[0.165]	[0.181]	[0.182]
Observations	1,117	1,117	1,034	1,034	1,019	1,019
Adjusted R-squared	0.222	0.221	0.454	0.455	0.430	0.431

Notes: This table reports the estimates of equations (1) and (2) for Calabria only, between 1907 and 1910. The key explanatory variable is based on the available Mercalli intensities reported in Guidoboni et al (2007). It is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale. All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3.6: Sicily - Mercalli intensities

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
Severe * Post	0.2486**	0.2726**	-0.0124	0.0096	-0.0551	-0.0328
	[0.116]	[0.109]	[0.183]	[0.182]	[0.164]	[0.163]
Observations	741	741	749	749	743	743
Adjusted R-squared	0.490	0.491	0.514	0.515	0.470	0.470

Notes: This table reports the estimates of equations (1) and (2) for Sicily only, between 1907 and 1910. The key explanatory variable is based on the available Mercalli intensities reported in Guidoboni et al (2007). It is a dummy taking value 1 whenever the magnitude of the quake is equal or exceeds 8 on the Mercalli scale. All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

Table A3.7: Calabria and Sicily - Distance-based severity

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
- log(dist) * Post	-0.1356**	-0.0895	-0.2781***	-0.2359**	-0.3308***	-0.2884***
	[0.065]	[0.064]	[0.093]	[0.094]	[0.095]	[0.096]
Observations	1,816	1,816	1,718	1,718	1,696	1,696
Adjusted R-squared	0.384	0.383	0.474	0.474	0.456	0.456

Notes:
Notes: This table reports the estimates of equations (1) and (2) for Calabria and Sicily, jointly, between 1907 and 1910. The key explanatory variable is defined as $Treat_m$: $d_c = -\ log(distance\ from\ the\ epicentre)$ where the *distance from the epicenter* is limited to a radius of 150 Km. All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

Table A3.8: Calabria - Distance-based severity

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
- log(dist) * Post	-0.3563***	-0.3026***	-0.2546**	-0.2052*	-0.3677***	-0.3180**
	[0.091]	[0.095]	[0.114]	[0.115]	[0.125]	[0.127]
Observations	1,081	1,081	1,006	1,006	993	993
Adjusted R-squared	0.252	0.247	0.485	0.485	0.452	0.451

Notes: This table reports the estimates of equations (1) and (2) for Calabria only, between 1907 and 1910. TThe key explanatory variable is defined as $Treat_m$: $d_c = -\ log(distance\ from\ the\ epicentre)$ where the *distance from the epicenter* is limited to a radius of 150 Km. All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level. Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

Table A3.9: Sicily - Distance-based severity

VARIABLES	(1) Global	(2) Global (deaths)	(3) EI	(4) EI (deaths)	(5) EI (prime)	(6) EI (prime, deaths)
- log(dist) * Post	0.2652**	0.2977***	-0.3140**	-0.2830*	-0.2740*	-0.2428
	[0.117]	[0.101]	[0.157]	[0.159]	[0.149]	[0.151]
Observations	735	735	712	712	703	703
Adjusted R-squared	0.475	0.476	0.456	0.456	0.441	0.441

Notes: This table reports the estimates of equations (1) and (2) for Sicily only, between 1907 and 1910. TThe key explanatory variable is defined as $Treat_m$: $d_c = -\log(distance\ from\ the\ epicentre)$ where the *distance from the epicenter* is limited to a radius of 150 Km. All regressions include municipal and province*year FE. All variables are logged and SE are clustered at the municipal level.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.1

Conclusion

The current debate on migration is often filtered through aggregate perspectives. Where do migrants hail from? Would Africa or the Middle-East be a sufficient answer? Would Nigeria, or Afghanistan be a satisfactory one? In its essence, Chapter I puts forward a perspective that has already been suggested by many researchers: migration should be analyzed through the lenses of granular data because its dynamic components are likely to operate at the local level. Making use of a recently digitized demographic collection, I use Ellis Island administrative records to reconstruct municipality-level US-bound out-migration flows from Italy between 1892 and 1912. I provide descriptive evidence of how migration followed a process of path-dependency and progressively spread across municipalities, fostered by ex-ante networks. I also pinpoint the relevance of cyclical flows, an overlooked element in current analyses seeking to grasp the effect of immigration on economic variables in destination countries.

How do migration flows react to globalization - thought as the progressive integration of commodity markets across the globe? While this question is apt to be studied within the current economic landscape, analyzing the first instance of globalization could offer precious insights. Indeed, technological change in steamship coupled with liberal trade and immigration policies made the world a much more inter-connected environment from the mid-1840s. Joint with Prof Rowena Gray and Prof Gaia Narciso, Chapter II revolves around the idea that competition-induced agricultural commodity price shocks might affect the propensity to out-migrate internationally, based on local production structures. We carry out a province-level analysis and find that higher agricultural commodity prices tended to release liquidity constraints thus promoting out-migration. Similarly, higher price volatility incentivized out-migration as it de-stabilized income sources.

Natural disasters have overarching consequences and their relevance is predicted to rise in the future. However, there is no consensus on how they influence migration patterns from affected

areas. Joint with Prof Yannay Spitzer and Prof Ariell Zimran, Chapter III investigates the migration response to the 1908 Messina and Reggio Calabria earthquake, one of the deadliest and most disruptive natural disaster in recent European history. Our findings do not point toward any obvious increase in international out-migration rates. Rather, we find that the quake had heterogeneous effects and only produced a transitory decline of out-migration rates from affected municipalities in Calabria. Networks did not increase the propensity to out-migrate because the earthquake took place when US-bound migration had already become endemic - meaning that their effect could vary at different stages of the migration process. Moreover, we show how migration decreased more in municipality that had been affected by another earthquake in their recent past, thus testing a damage accumulation hypothesis.

Although pre-WWI Italian out-migration seems to be a particularly good example of multi-dimensional diversity, it was not the only one during the age of mass migration: anecdotal evidence suggests that micro-level dynamics widely varied within other European countries too. Further disaggregate research should voice such heterogeneity and shed light on it. Similarly, a comprehensive dynamic framework cannot abstract from considering the spatial and occupational distributions of migrants at arrival. Novel studies based on linking passengers lists to US Censuses should evaluate the evolution of networks in destination countries (from ex-ante to ex-post), assess their impact on location choices, individual economic outcomes and, eventually, the decision to stay, return, or go back and forth multiple times.