

LIKE AN INK BLOT ON PAPER:
TESTING THE DIFFUSION HYPOTHESIS OF MASS MIGRATION, ITALY 1876-1920

Yannay Spitzer
Ariell Zimran

Working Paper 30847
<http://www.nber.org/papers/w30847>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2023, Revised March 2024

For helpful comments, we thank Ran Abramitzky, Brian A'Hearn, Marcella Alsan, Michael Amior, Brian Beach, Levi Boxell, Thomas Chaney, Carlo Ciccarelli, William Collins, Giuseppe De Arcangelis, Giovanni Facchini, Joseph Ferrie, Michela Giorcelli, Walker Hanlon, Timothy Hatton, Mushfiq Mobarak, Joel Mokyr, Jose Moya, Kaivan Munshi, Federico Nastasi, Cormac Ó Gráda, John Parman, Hannah Postel, Sarah Quincy, Hillel Rapoport, Marco Tabellini, Gaspare Tortorici; seminar participants at Harvard University, the Hebrew University of Jerusalem (the Department of Economics, the Department of Environmental Economics and Management, and the Migration Workshop), the London School of Economics, Northwestern University, Oxford University, Queen's University Belfast, Stanford University, Tel Aviv University, the University of California Davis, the University of Michigan, the University of Mississippi, the University of Nottingham, the University of St. Andrews, the University of Warwick, Vanderbilt University (the Department of Economics and the Robert Penn Warren Institute), Yale University, and the Economics of Migration Senior Migration Seminar; and conference participants at the 2019 Galatina Summer Meetings, the 2020 Allied Social Sciences Association Meetings, the 2020 H2D2 Research Day at the University of Michigan, the 2020 Economic History Association of Israel Meeting, the 2021 Economic History Society Conference, the 2021 Migration and Development Conference, the 2022 Economic History Association Conference, and the 2022 Columbian Shock Conference at Yale University. Christine Chang, Jenny Liu, Jared Katz, Paul Mitalipov, Elizabeth Nelson, Ilan Pargamin, Sarah Robinson, Stephania Stavropoulos, Claire Whittaker, and Danielle Williamson provided excellent research assistance. This project was supported by a Vanderbilt University Research Scholar Grant. Work on this paper was completed while Ariell Zimran was a W. Glenn Campbell and Rita Ricardo-Campbell National Fellow and the William C. Bark National Fellow at the Hoover Institution, Stanford University and while he was a Visiting Scholar at the Opportunity and Inclusive Growth Institute at the Federal Reserve

Bank of Minneapolis; funding from both institutions is gratefully acknowledged. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or of the Federal Reserve System. This material is based upon work supported by the National Science Foundation under Grant No. SES-1425598. This research was supported by the Israel Science Foundation (grant No. 3121/21). Additional financial support was provided by the Northwestern University Center for Economic History, The Falk Institute for Economic Research, the Hebrew University of Jerusalem, the German-Israeli Foundation, and Vanderbilt University. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Yannay Spitzer and Ariell Zimran. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Like an Ink Blot on Paper: Testing the Diffusion Hypothesis of Mass Migration, Italy 1876-1920
Yannay Spitzer and Ariell Zimran
NBER Working Paper No. 30847
January 2023, Revised March 2024
JEL No. F22,J61,N33,N34

ABSTRACT

To explain the evolution of mass migration from Italy (1876--1920), we introduce the diffusion hypothesis, which augments traditional models of migration by allowing the spatial diffusion of migration over social networks in the country of origin. We test and validate the predictions of this theory using a new and comprehensive municipality- and district-level panel of emigration data over four decades. Italian migration began in a few separate epicenters from which it expanded gradually over time in an orderly spatial movement that dominated local push factors. While the main patterns that we document are parsimoniously explained by the diffusion hypothesis, they are difficult to rationalize by alternative theories. Confirming the diffusion hypothesis in the Italian case suggests that networks may be far more important than is commonly understood in shaping other movements of mass migration and may shape them in ways that have not previously been understood.

Yannay Spitzer
Department of Economics
The Hebrew University of Jerusalem
Mt. Scopus
Jerusalem 9190501
Israel
and CEPR
yannay.spitzer@mail.huji.ac.il

Ariell Zimran
Department of Economics
Vanderbilt University
2301 Vanderbilt Place
Nashville, TN 37235
and NBER
ariell.zimran@vanderbilt.edu

A data appendix is available at <http://www.nber.org/data-appendix/w30847>

1 Introduction

International migration largely takes place within social networks (Massey et al. 1993), in which friends and relatives reduce the costs and enhance the payoffs of migration by providing information, inspiration, and material assistance. Augmenting simple push-pull migration models with networks can therefore rationalize features of migration flows that are otherwise difficult to explain (Munshi 2020), including their magnitude, composition, destination choices, and migrants’ labor market outcomes in the destination (Beaman 2012; Becker et al. 2023; Buggle et al. 2023; Chiquiar and Hanson 2005; Fernández-Huertas Moraga 2013; McKenzie and Rapoport 2010; Orrenius and Zavodny 2005; Stuart and Taylor 2021).

In this paper, we introduce the *diffusion hypothesis* to explain the broad patterns of one of the largest, most important, and best-documented movements of voluntary mass migration in history—the mass emigration of Italians between 1876 and 1920. This theory augments basic models of migration with social networks by allowing migration to diffuse over space through personal contacts between localities in the country of origin—that is, the process by which emigrants within the network sequentially support their contacts’ emigration can span local borders.¹ We define, test, and verify this theory with novel and detailed emigration data and show that the new feature of spatial diffusion is indispensable for understanding the Italian migration. Unlike traditional push-pull frameworks that account for local networks but do not permit spatial diffusion, we show that the diffusion hypothesis can easily and parsimoniously explain the most important macro-patterns of the Italian migration—how emigration grew in numbers at the national level, how regions switched from having no migration to mass migration within short periods of time, why this switch occurred at different times in different places with large swaths of the country waiting for decades before suddenly entering mass migration, and how destination choices varied across regions and over time.

As we operationalize the diffusion hypothesis, emigration shares the epidemiological properties of an infectious disease.² Just as individuals do not contract one unless exposed to someone else who has been infected, individuals generally do not emigrate unless one of their close contacts has already emigrated, regardless of economic rewards and other push factors. Scaled up to the locality level, this implies that, even in places with high emigration potential,³ emigration is limited in the absence of connections to networks of prior migrants. The key idea is that, in most cases, emigration from a locality begins through personal con-

¹We think of a social network as a network of personal connections between individuals, such as close relatives, friends, neighbors, and colleagues, that are sufficiently strong to enable one individual’s migration to increase the future migration probabilities of his contacts.

²To be clear, we do not intend the use of an epidemiological model or analogy to imply any normative judgements regarding migration. It is instead a useful analogy to describe the spread of a phenomenon within a population.

³By high emigration potential, we mean strong incentives for emigration, exclusive of network effects.

tacts with recent migrants in neighboring places rather than spontaneously through local pioneer migration. Further scaled up to the national level, emigration initially emanates from only a few epicenters, which may or may not be selected at random, and as individuals from these places emigrate, their contacts in nearby places become connected to friends and relatives abroad. In this way, neighboring places infect one another, leading “the contagion of emigration [to] spread over the map much like an ink blot on paper” (Moya 1998, p. 113).⁴ The result of this process in Italy was that regions closer to the epicenters produced mass emigration early on, whereas regions farther away, many of which were characterized by similar push factors and thus had similar potential for emigration, had to wait, sometimes for several decades, before their potential was abruptly unleashed when they were reached by the expansion of migration over the national network.

We draw motivation for the diffusion hypothesis from preliminary evidence provided by both local histories and by broad patterns in the data. First, the diffusion hypothesis is historically plausible: we survey the literature on the sociology of the Italian migration and find numerous indications that it depended on local networks that often extended into neighboring localities. Furthermore, maps that show the evolution of the geography of Italian emigration to North America also provide *prima facie* supporting evidence.⁵ The maps in Figure 1 report the district-level rates of North America-bound emigration by half decade from the late 1870s.⁶ In Figure 2, municipality-level maps do the same at a finer administrative level, but with data starting in 1884.⁷ Migration rates were initially high in only a small number of epicenters. Over time, migration spread in an orderly and spatially consecutive manner to the rest of the country. Many regions that had at first produced no migration at all turned out to be enthusiastic participants once this movement reached them; conversely, by the end of the Age of Mass Migration, the initial epicenters no longer stood out in terms of their emigration rates, inconsistent with their leadership being the product of a particular proclivity for emigration.

Our analysis builds on this preliminary evidence in three steps. First, we formalize the diffusion hypothesis by proposing a model that combines a push-pull framework of migration with an underlying process of diffusion over a spatial network. Whereas research on migration networks tends to assume that they are purely local, the key new element in this model is that it allows networks to span nearby localities within the sending country. The micro-foundations of the model treat migration as a technology that becomes available to individuals once a person to whom they are linked has migrated. Having gained the option of

⁴Gould (1980, p. 283) used a similar analogy: “One might describe this process as one of ‘diffusion,’ at least in the mechanical sense in which a drop of ink on a small piece of blotting paper gradually ‘diffuses’ over the whole area.”

⁵The data underlying these maps are discussed in section 4.

⁶By district, we refer to the Italian *circondario* or *distretto*.

⁷By municipality, we refer to the Italian *comune*.

emigration, individuals can then choose whether to move to the same destination as their contact based on typical push and pull factors, such as income differences or demographic pressures. If they migrate, their connections subsequently gain access to the technology of migration to the same destination, and so on. The model permits both *intra*-place diffusion, as in the traditional conceptualization of networks, and *inter*-place diffusion—the new feature by which migration spreads to nearby places. Pioneers (individuals who migrate without being linked to a prior migrant) and long-distance contacts are allowed; spatial diffusion will dominate the evolution of emigration if they are sufficiently rare, in which case migration will be characterized by the following patterns.⁸

1. *Convergence*. The variation in emigration rates across places diminishes over time (σ -convergence) as laggards catch up with leaders (β -convergence).
2. *S-shaped local time trends*. The migration trend of the typical place is characterized at first by little or no migration, followed by a sharp take-off and ultimately a plateau.
3. *Correlated destinations*. The similarity between the destination choices of migrants from any two places is decreasing in the distance between them and increasing over time.
4. *Frontier effect and spatial expansion*. Defining the frontier of mass emigration as the contour of places that had experienced mass migration by the previous period, the probability that a place outside the frontier will enter mass emigration is negatively related to its distance from the frontier; places farther from the epicenters of emigration enter mass emigration later.

Second, we test these predictions using official Italian emigration statistics, which we digitized for this project, with a particular focus on migration to North America.⁹ Augmented by a broad array of high-resolution official statistics and geographic data from post-unification Italy, the emigration data provide fine detail at the district and municipality level, covering more than 280 districts and 8,000 municipalities over four decades. We discover a series of stylized facts corresponding to the patterns predicted by our model. Importantly, these stylized facts are more than simply predictions that we verify—they dominate the data and define how the Italian emigration evolved. Among them are the following results.

⁸On the other hand, if pioneers are not rare or long-distance contacts are common, the model collapses, at the limit, to a standard push-pull one in which spatial relations to past migration play no direct role.

⁹We focus on migration to North America because it is the only flow that we can observe from its early stages to its peak. The other two main streams of emigration were to other countries in Europe and to South America. We find and report similar patterns in emigration to all destinations. But emigration to Europe was often characterized by extremely high rates of repeat migration that disproportionately inflate the emigration counts, potentially more so in places closer to the border, which might artificially create a spatial trend. Emigration to South America started earlier, and so we can only observe the later stages of its development.

1. The coefficient of variation in municipality-level emigration rates to North America fell to less than a third of its initial value—evidence of σ -convergence. This pattern was driven by laggards catching up with leaders in a pattern of β -convergence—municipalities in the bottom quartile of pre-1900 emigration rates to North America experienced a 200-fold greater increase in average annual emigration rates after 1900 as opposed to only a 4-fold increase among municipalities in the top quartile.
2. The average municipality, regardless of when it first entered mass migration, experienced an S-shaped time trend, increasing in its migration rate from essentially zero to about 25 per thousand over a period of 20 years, after which its emigration rate plateaued.
3. A one-standard deviation increase in distance between two provinces was associated with an increase in the dissimilarity index of their destinations of just under half of a standard deviation, and this dissimilarity declined by about one standard deviation, from 1876–1880 to 1911–1914.
4. An increase in distance from the previous half-decade’s frontier of mass migration from 0 to 150 km (about one standard deviation) was associated with an 80-percent decline in the average emigration rate from just under 5 per thousand to about 1 per thousand. This frontier effect clearly translated into a pattern of spatial expansion, as a one-standard deviation increase in distance from epicenters was associated with a decline in the hazard of entering mass migration of about 40 percent.

In the third step of our analysis, we confirm *spatial contagion*—emigration from one place causing emigration from neighboring places—as the most likely mechanism underlying these factors. Identifying this causal effect is an empirical challenge. It could be that neighboring places shared unobserved characteristics that affected the timing or magnitude of mass migration, or that a time-trend in such characteristics was correlated among neighbors. It could also be that there was a process of spatial diffusion but that the diffusing characteristic was not personal links to previous migrants but rather some other migration-inducing characteristic, such as local policies (Andrews and Seguin 2015), industrialization (Franck and Galor 2022), or demographic trends (in the spirit of Spolaore and Wacziarg 2009). To identify the causal effect of spatial contagion, we estimate a spatial lag model of migration using a novel instrumental variables strategy that leverages differences across similar places in their neighboring populations’ proximity to emigration epicenters, which, in turn, imply differential likelihoods of entering early into a phase of large-scale emigration.¹⁰ This source of variation is exogenous under the assumption that the orientation of the spatial distribution of neighboring populations (i.e., whether neighboring places are on average closer or farther from the source)

¹⁰Although our approach is novel, it is based on standard approaches to the estimation of spatial lag models, and our results are strengthened by the use of more typical methods for estimating such models.

is independent of a place’s own characteristics. Our results yield an estimated elasticity of own emigration with respect to the emigration of nearby municipalities of greater than 0.50, thus confirming the existence of the spatial contagion mechanism.

Could these stylized facts of the Italian emigration be explained by an alternative theory? The diffusion hypothesis stands in contrast to a class of *internalist* explanations, in which factors internal to a location alone determine the timing and extent of migration,¹¹ without spatial contagion effects. This class of explanations includes standard models of migration with local networks, though these cannot explain the spatial patterns that we document. It also encompasses models that attribute migration patterns primarily to local economic and demographic fundamentals, such as the dominant paradigm in the literature on the Age of Mass Migration—the *modernization hypothesis*. Traced back at least as far as Thistlethwaite (1960 [1991]) and canonized by Hatton and Williamson (1998), this hypothesis argues that mass migration was triggered by aspects of economic modernization, such as the release of poverty traps, the loosening of connections to the land, and the rise of demographic pressures. In this view, networks mattered, but they were purely local and were one factor among many. The timing of mass migration did not depend on proximity to already infected places but on the timing of modernization.¹²

To evaluate the relative validity of the diffusion hypothesis and internalist explanations based on local fundamentals, we estimate hazard models for the timing of a district’s entry into mass emigration as a function of its observable economic characteristics and its exposure to spatially diffusing migration networks measured by distance to epicenters. Unsurprisingly, some internal characteristics, including indicators of modernization, were correlated with the timing of emigration. But we show that their roles were small relative to the consistent first-order role of diffusion. Moreover, the directions of the correlations of some modernization indicators run contrary to the modernization hypothesis. Furthermore, we do not find that internal factors can explain strong convergence of emigration rates over time. We also find that diffusion processes were specific to the destination, documenting separate evolutions of the migrations to North America, South America, and Europe. While this is not strictly inconsistent with internalist explanations, a complex version of such explanations is required to rationalize it, yet it derives directly from a simple version of the diffusion hypothesis. Internalist explanations in general, and the modernization hypothesis in particular, thus fail to offer a plausible, simple, and complete explanation for the stylized facts that we document.

¹¹Internal conditions include both factors incentivizing emigration and local factors that may have been responsible for temporarily restraining emigration from places where the incentives for migration were strong. We think of these all collectively as push factors.

¹²Returning to our epidemiological analogy above, internalist explanations view emigration as having the epidemiology of cancer—its prevalence is determined solely by whether the conditions for the disease are right, and spatial correlation might arise from spatial correlation in these conditions.

The accumulated evidence leads us to view the diffusion hypothesis as the best available explanation for the set of stylized facts of the Italian emigration that we document. Although some of them could plausibly be generated by alternative explanations, the diffusion hypothesis is the most parsimonious theory that can explain them all.

To be clear, the diffusion hypothesis does leave certain important facts unexplained. Most importantly, we do not explain why certain places became original epicenters or why certain people became pioneers. Instead, we take it as given that some individuals and locations had taken initial leadership and that their selection may not have been random. Nevertheless, we note that the epicenter of transalpine migration was the northern border, and most epicenters of transatlantic migration were near major ports. This suggests that places that had pre-existing cross-border connections—perhaps through historical trade patterns—were more likely to pioneer the mass migration movement. However, since the number of epicenters is small, we cannot test this conjecture statistically, and regardless of the actual mechanism, the validity of the diffusion hypothesis depends on the nature of the expansion of migration from the initial epicenters, not on these epicenters being selected at random.

This paper contributes to the literature on the economics of migration (Bansak, Simpson, and Zavodny 2020; Borjas 2014), and specifically to the literature on migration networks (Beaman 2012; Becker et al. 2023; Buggle et al. 2023; Chay and Munshi 2015; Chiquiar and Hanson 2005; Dai et al. 2020; Fernández-Huertas Moraga 2013; McKenzie and Rapoport 2010; Munshi 2003, 2011, 2020; Munshi and Rosenzweig 2016; Orrenius and Zavodny 2005; Stuart and Taylor 2021). It shows that spatial diffusion of migration in the origin may, in some cases, be a powerful mechanism that dominates the evolution of mass migration. This is a departure from the standard approach to modeling migration networks, in which networks are generally thought of as applying to specific and fixed communities in the origin (Munshi 2020). In the language of our model, the standard approach incorporates only intra-place diffusion, whereas we demonstrate the theoretical and empirical importance of inter-place diffusion to understanding the dynamic geography of emigration. To be clear, we do not argue that spatial diffusion is a common feature of all migration movements. The goal of this paper is to offer a proof of existence—by showing that spatial diffusion was an important force in one major migration movement, and that modifying a model of migration to allow for such diffusion is necessary to parsimoniously explain all of the features of this movement, we show that spatial diffusion is a potentially important force that must be considered when studying other cases of mass migration.

Our results also provide empirical support for predictions regarding the role of migrant networks that may also emerge from models that view networks as important but do not include a role for spatial diffusion. While

it is broadly accepted that networks support and contribute to migration (e.g., Beaman 2012; Munshi 2003, 2011), the magnitude of networks’ role remains the subject of debate (Munshi 2020). We show that their role in the Italian case was so important that their presence alone could switch a region from little or no migration to mass migration within a short period of time, independent of any internal structural changes.¹³ Conversely, emigration was dormant in some regions of Italy for several decades only because existing networks were too geographically remote. Such a fundamental role for networks as drivers of migration, over and above economic fundamentals, has been documented in the case of occupational mobility (Kerr and Mandorff 2023; Munshi 2011, 2020),¹⁴ but to our knowledge we are the first to document it systematically in the case of international migration. This implies that the absence of mass migration need not indicate that the circumstances in the places of origin are not conducive for it, and that a potential for a rapid switch to mass migration may exist.

As with other studies of the Age of Mass Migration (Abramitzky and Boustan 2017), the near total absence of restrictive migration policy enhances these contributions. Whereas in other contexts restrictive migration policy may artificially amplify the effects of networks,¹⁵ the open borders policy during this period obviates such concerns.

This paper also contributes to understanding two of the most important questions in the economic history of the Age of Mass Migration. First, what were the key determinants of the underlying demand for emigration? We show that as the Italian emigration reached a state of near saturation, in the sense that all places had already been exposed to migration, the map of Italian emigration took on a new shape, resembling that of a topographic map of the country. Based on the diffusion hypothesis, we interpret this map to be the best reflection available of the spatial variation in push factors, no longer obscured by geographic variation in exposure to migration. If true, this hitherto unnoticed regularity suggests that something associated with the highland economy was a primary factor causing emigration. We leave exploration of the specific factor to future research, but note that recognizing the role of diffusion is crucial in identifying this regularity.

Second, we contribute to better understanding the *delayed migration puzzle*—why was mass migration from Italy and the rest of southern and eastern Europe so slow to develop before suddenly surging around the

¹³That is, it is well understood that liquidity constraints pose a significant impediment to migration from developing economies, and often the conclusion is that extreme poverty must be alleviated before mass migration is generated (Burghard, Chaney, and Hassan 2019; Gray, Narciso, and Tortorici 2019; McKenzie and Rapoport 2007). The lesson that we draw from the Italian migration is that social liquidity constraints trump financial ones, in the sense that the former are the real bottleneck, and that they can solve the latter.

¹⁴“...in general, a fortuitous confluence of circumstances is required to jump-start the network. Thus, different origin communities will start sending migrants at different points in time even when the underlying fundamentals that determine the payoffs at the origin and the destination are the same” (Munshi 2020, pp. 507–508).

¹⁵For instance, the role of networks may be amplified in cases where migration policy prioritizes family reunification.

turn of the century?¹⁶ Gould (1980) originally proposed a diffusion-based explanation, according to which Italy, as well as other peripheral countries, was well suited for mass emigration already in the 1870s, but this potential took decades to materialize into mass migration because of the protracted process of spatial diffusion of migration networks that determined when this potential would be unleashed. The gradual rise of Italian migration, documented in Figure 3, reflected sequential local surges of mass emigration, which had spread and reached a sufficiently large part of Italy around the turn of the century to culminate in a wave of mass emigration to the United States. Although this explanation has been at least partly accepted by social historians (Baines 1995; Lowell 1987; Moya 1998), it has remained a minority view in the economic history literature, which has instead favored the modernization hypothesis (Hatton and Williamson 1998). Moreover, it has never been rigorously tested, in part because the data requirements are difficult to meet and in part because a complete and testable theoretical framework to guide analysis was lacking. In this paper, we meet these challenges and are able, for the first time, to test and confirm the diffusion hypothesis with rich geographically disaggregated data and within a clear theoretical framework.¹⁷

2 Background

2.1 Networks in the Economics of Migration

It is a fundamental tenet of the economics of migration that migration generally occurs within social networks. In nearly every econometric study of migration, the past stock or flow of migrants from the same place is a strong predictor of current migration (Beine, Bertoli, and Fernández-Huertas Moraga 2016; Beine, Docquier, and Özden 2011; Bertoli and Fernández-Huertas Moraga 2015; Docquier, Peri, and Ruysen 2014). The role of networks has been substantiated through direct quasi-experimental methods (e.g., Beaman 2012; McKenzie and Rapoport 2007, 2010; Munshi 2003; Woodruff and Zeteno 2007) and by showing that a number of facts, such as the limited explanatory power of wage gaps as a predictor of migration flows and perplexing selection patterns (e.g., Chiquiar and Hanson 2005; Fernández-Huertas Moraga 2011, 2013; Ibarrran and Lubotsky 2007; McKenzie and Rapoport 2007; Munshi 2020; Orrenius and Zavodny 2005), which standard migration models struggle to rationalize, can be parsimoniously explained by augmenting these models with social networks.

¹⁶See Online Appendix Figure B.1.

¹⁷Our results complement recent evidence on Jewish migration from the Pale of Settlement, which is strongly suggestive of a diffusion process being the primary determinant of its macro trends (Spitzer 2021). We do not suggest that there was a single continent-wide diffusion process; instead, by validating this hypothesis in Italy, we show that, in explaining the delayed migration puzzle, economic historians need not look for a change in the economic or demographic circumstances that were specific to the time emigration surged in order to explain it.

In a recent review of the subject, however, Munshi (2020) argues that, although the causal role of networks in affecting migration choices is well established, the literature has yet to show that “the magnitude of their role is substantial” (Munshi 2020, p. 518).¹⁸ Moreover, and particularly important for our case, all models and empirical inquiries implicitly or explicitly assume that networks are static and local (e.g., are measured at the level of a municipality or a country)—that is, that networks apply solely to a prespecified social group. Such models cannot capture spatial diffusion. A single exception to this rule is Barthel and Neumayer (2015), who show that a network of asylum seekers from one country increases the probability that persons from a neighboring country also seek asylum at the same destination country, possibly due to sharing a network at the destination based on ethnic and linguistic similarity.

Our testing and confirmation of the diffusion hypothesis thus leads to three main insights for the economics of immigration. First, it demonstrates the potential magnitude of the effect of networks, assigning a far more consequential role to them than has been previously documented empirically. Rather than simply one factor that facilitates migration (i.e., a “cost shifter” or “payoff enhancer”), networks take center stage. They are the first-order determinant of the geographic origins of migration and of their evolution of time. Their absence can inhibit mass emigration not only for years but even several decades, as conditions that are conducive to migration may not be enough to generate migration spontaneously. And when this condition reverses and mass migration begins, this need not be the result of any transformative economic or demographic event. This lesson may stand even in very different contexts without spatial diffusion. Such a role for networks—one in which they are so important that they may cause groups with similar fundamentals to enter into a movement at very different times—has been suggested theoretically (Munshi 2020) but this paper is the first to demonstrate it empirically.¹⁹

Second, our confirmation of the diffusion hypothesis in the Italian case shows that understanding both the temporal and the spatial evolution of networks is crucial in predicting the future course of mass migration movements—which places are likely to experience migration surges in the near future, and how destination choices are likely to change. The importance of both the standard local network and the new feature of networks reaching over space will depend on the context. In the case of the Italian migration, including such a feature in the model will prove crucial in parsimoniously explaining fundamental patterns of this movement in a way that no other model can. But in other cases the spatial aspects may prove less important. Regardless, such a feature of networks is one that must be considered as a force potentially shaping the evolution of

¹⁸Dai et al. (2020) and Munshi and Rosenzweig (2016) provide such quantifications. See also Munshi (2014). Munshi (2020) also proposes a general procedure to identify the effects of migrant networks, which our analysis broadly follows.

¹⁹Kerr and Mandorff (2023) and Munshi (2011) provide empirical support for such a phenomenon for occupational mobility, but it has not been demonstrated in the case of spatial mobility.

migration.

Finally, a proper recognition of the spatio-temporal evolution of networks can focus the attention of scholars away from largely random patterns towards those that provide useful information, thus revealing the underlying incentives for migration. That is, while dominating the evolution of migration, networks may mask underlying push factors until the country approaches saturation.

2.2 Italian Migration in the Age of Mass Migration

The mass emigration of Italians during the Age of Mass Migration is one of the largest voluntary migration flows in history. In total, from 1877–1920, over 14 million Italians emigrated relative to a 1901 Italian population of about 32.5 million. This flow was divided into three major destinations—nearly 5 million traveled to North America, while South America attracted nearly 3 million and over 6 million traveled to other European destinations. Figure 3 describes these three movements over time, and also divides them by their geographic origins within Italy. The migration to North America, which is our primary focus, began at low levels in the late 1870s,²⁰ grew modestly until about 1900, and then surged, reaching its peak in the 1910s before being cut off by World War I; at its peak, the annual flow from Italy to North America regularly exceeded 200 thousand people, or about 6 per thousand of the Italian population, and was primarily composed of individuals from southern Italy. Migration to South America also began at low levels (though higher than those for North America), but achieved its peak much earlier. It also exhibited a change in its geographic origins over time from north to south. Finally, European migration, primarily a northern Italian phenomenon (given that region’s proximity to these destinations) was already substantial when the official statistics became available in the late 1870s. A surge in this movement is evident around 1900, though in this case there is serious concern that the surge may be the spurious product of changes in data collection and that much of the measure is comprised of repeated seasonal migrants, as we discuss in more detail below.

Major questions remain open regarding the Italian migration. There is a scholarly consensus that it was primarily driven by large and persistent gaps in standards of living between Italy and the destination countries (e.g., Hatton and Williamson 1998).²¹ But no widely accepted theory exists that is capable of explaining either the geographic variation in emigration within Italy or its geographically staggered rollout.

The first challenge is that economic development was not consistently correlated with emigration rates. As we show in Figure 4, the cross-province correlation between 1876 income (specifically, real wages as

²⁰The United States was at first so unimportant as a migration destination that Italian emigration statistics did not report it separately from Canada until 1880.

²¹The report of the Dillingham Commission pointed out that “practically all emigration from Italy is primarily due to purely economic causes” (US Congress 1911a, p. 153).

measured by Federico, Nuvolari, and Vasta 2019) and emigration to all destinations was positive in the first stages of the Italian migration, only turning to the expected negative sign in the 1890s. Moreover, Italian regions that were seemingly comparable in terms of conditions conducive to mass emigration had widely different timings of its onset,²² a fact that did not go unnoticed by contemporary observers (US Congress 1911a, p. 164).²³ For example, as can be seen in Figures 1 and 2, mass emigration gradually spread south through the western *Mezzogiorno*, from one neighboring province to another: Salerno in Campania in the late 1870s, Cosenza in northern Calabria in the 1880s, Catanzaro and Reggio Calabria in southern Calabria during the 1890s, and finally Messina, across the strait, around the turn of the century.²⁴ Real wages were not far apart in these provinces, and at the very least, their ranking was orthogonal to the order in which they entered mass emigration. Our general point, applied to this case, is that this 25-year trickle south of emigration across this rather equally poverty-stricken region cannot be explained by variation in economic conditions. This suggests that the *underlying causes* of emigration and the *trigger* that caused the potential for emigration to actually materialize in any given area were separate factors.²⁵

Consistent with the modernization hypothesis, Foerster (1919) suggested that emigration from poorer places began later as a result of more binding liquidity constraints,²⁶ and Faini and Venturini (1994) have found statistical evidence supporting this view. Hatton and Williamson (1998) rejected a role for liquidity constraints and found instead that the determinants of migration, both across European countries and across Italian provinces, were real wage gaps relative to destinations, demographic pressures, and the level of employment in agriculture (a negative measure of industrialization). Following Thistlethwaite (1960 [1991]), they concluded that “mass emigration in Europe had to await the forces of industrialization at home and a glut in the mobile age cohort driven by a demographic transition that industrialization produced” (Hatton and Williamson 1998, p. 46).²⁷

However, a number of important patterns of European migration challenge such an explanation for the causes of Italy’s geographically varied emigration experience. In general, studies show neither positive nor

²²See Online Appendix Figure B.3 for an example.

²³Foerster (1919, p. 104) pointed out that “It is significant that emigration should not have originated where misery was greatest.”

²⁴This geographic progression was described in detail by Foerster (1919, pp. 102–104). Online Appendix Figure B.4 provides maps of Italian regions and provinces for reference.

²⁵In the words of Foerster (1919, p. 48), “The fact that emigration from Campania was abundant before it became so in Calabria, and that it only as much as ten or fifteen years later assumed large proportions in Sicily, need signify merely that the *occasion which turned a passive into an active cause* arose earlier in one compartment than in another” (emphasis added).

²⁶“[Emigration] began where there was the chance of saving enough money for passenger fares and has best maintained itself where wages were at a minimum level” (Foerster 1919, p. 104).

²⁷According to MacDonald (1963) and MacDonald and MacDonald (1964), the propensity to emigrate was a result of different constellations of agricultural organization and communal relations. This view, somewhat similar to Hirschman’s (1970) *Exit, Voice, and Loyalty*, has gained traction in the socio-historical literature (e.g., Baily 1999; Barton 1975; Silverman 1968; Sturino 1990; Yans-McLaughlin 1977), but came under criticism by Gabaccia (1984a,b, 1988).

negative systematic correlations between economic conditions and emigration in the Age of Mass Migration (Baines 1995). More specifically, evidence that mass emigration and economic or demographic modernization emerged in the same places is inconsistent and contested at best. For instance, Jewish emigration from the Pale of Settlement in the Russian Empire began in a few impoverished provinces in the northwest, only later spreading to nearby centers of Polish industrialization, and much later to the relatively well-off communities of central and eastern Ukraine (Spitzer 2021). Similarly, Ireland’s early leadership in migration predated its industrialization, and there is no evidence of rising demographic pressure there when its emigration first began to surge before the Great Famine (Cohn 2009; Mokyr 1983; Mokyr and Ó Gráda 1982).

Modernization indicators have also been shown to perform poorly in explaining Italian emigration, both in time-series analysis (Ardeni and Gentili 2014) and when accounting for multiple destinations (Moretti 1999). As noted by the Dillingham Commission,²⁸ the geographic distribution of early Italian emigration bears little resemblance to that of early Italian industrialization, which occurred in concentrated geographic pockets, primarily in the northwest (Ciccarelli and Fenoaltea 2013; Federico, Nuvolari, and Vasta 2019; Iuzzolino, Pellegrini, and Viesti 2013). Some of these industrialization hotspots, such as in Liguria and in the Alpine slopes, were indeed emigration leaders, but the nearby Po Valley would never develop mass emigration despite its leading role in the industrialization movement. Moreover, other emigration epicenters were in the northeast, in the center, and in the south. Some had existing traditional and extractive industries, yet they were generally not a part of the modern industrialization movement. Neither were broad demographic trends associated with the evolution of emigration. The north led Italy’s Demographic Transition (Del Panta 1997, p. 10; Vecchi 2011, Table S6), yet the early sources of mass emigration were scattered throughout the country. Recent attempts to assess the relationship between demographic pressures and emigration have been inconclusive, lacking statistical power or credible identification (Ardeni and Gentili 2014; Faini and Venturini 1994; Gomellini and Ó Gráda 2013; Hatton and Williamson 1998).

Evidence on the diffusion hypothesis is even scarcer. Gould (1980, Figure 1) highlighted σ -convergence patterns across provinces within regions in Italy, as well as in Hungary and Portugal. He also informally argued that there existed β -convergence and S-shaped sub-national time series in Italy (pp. 282–288). Spitzer (2021) found both σ - and β -convergence in the case of Jewish migration from the Pale of Settlement. Qualitative evidence similarly indicated diffusion in Scandinavian (Lowell 1987) and Spanish (Moya 1998) emigration. In the first econometric test of the diffusion hypothesis, Hatton and Williamson (1998) found persistence

²⁸“It will be seen that as a rule the heaviest emigration originated in the compartimenti where the proportion of industrial workers was the smallest . . . and it is well known that comparatively little Italian emigration originates in the large cities” (US Congress 1911a, p. 175).

in the emigration rates of Italian provinces over time (see also Gomellini and Ó Gráda 2013)—evidence that networks were important in determining the size of migratory flows—but failed to find a relationship between literacy—which they viewed as a factor that could have facilitated the spread of information—and emigration rates. They concluded that diffusion “offers few empirical predictions and says nothing about why emigration rates eventually declined” (p. 15) and that “while such forces [as diffusion] mattered, there is little evidence that persistence or literacy dominated [Italian] provincial emigration rates with anything like the force often assigned to them in the qualitative literature” (p. 121).

Prior attempts to test the diffusion hypothesis have been limited by two missing factors—a complete theoretical framework from which to make testable predictions, and a sufficiently long, rich, and geographically disaggregated panel dataset with which to identify the *inter*-place transmission of emigration. On these grounds, we view the diffusion hypothesis as one that is plausible and capable of explaining a fundamental puzzle of the economics of the Age of Mass Migration, but which has yet to be rigorously tested.

2.3 The Role of Networks in Italian Emigration

Is the diffusion hypothesis plausible within the social context of post-unification Italy? Is the available historical evidence consistent with it? Although the notion that emigration had epidemic-like features was widely recognized by contemporaries,²⁹ the full implication of the hypothesis—that diffusion was the *primary* determinant of the timing of the onset of mass migration—is absent in any contemporary account, including Foerster (1919) and the Dillingham Commission Report (US Congress 1911a), arguably the two most comprehensive contemporary inquiries into the causes of Italian emigration.

For such an explanation to be plausible, the social structures that supported emigration must have had certain non-trivial characteristics. They had to be sufficiently strong to support chain migration. They had to be local, yet occasionally crossing community boundaries. And when they did cross community boundaries, they had to reach primarily over short distances, only rarely spanning longer distances. Furthermore, alternative mechanisms that supported migration, such as direct recruitment by foreign governments and businesses or poaching by shipping agents, had to be either negligible or themselves operating within migration networks. Finally, pioneers had to be rare. In what follows, we survey the relevant evidence from the historical literature to evaluate the plausibility of these conditions in Italy during the Age of Mass Migration.

To what extent did Italians engage in chain migration? Much of the debate concerning the sociology of Italian emigration evolved as a reaction to Banfield (1958) and Handlin (1951), who cast doubt on the viability

²⁹The usage of metaphors such as “migration fever” prevailed in virtually every sending country (Moya 1998, pp. 95–96).

of strong personal and communal relations among south Italian immigrant peasants in the United States, and by implication also on the prospects of strong migrant networks. Subsequent literature has modified this view of weak social links, showing that both kin- and municipality-based ties played an important and constructive role during and after migration (Bell 1979; Briggs 1978; Gabaccia 1984b; Nelli 1967; Vecoli 1964; Yans-McLaughlin 1977). It portrayed the Italian migration to the United States as being dominated by characteristics of chain migration: early migrants provided funding, information, accommodations, assistance in the labor market, and close examples of successful migration to their friends and kin, who in turn would do the same for theirs (Baily 1999; Cinel 1982; MacDonald and MacDonald 1964; Sturino 1990). As one immigrant put it, “Immigrants almost always came to join others who had preceded them—a husband, or a father, or an uncle, or a friend” (quoted in Yans-McLaughlin 1977, p. 59). This assertion is supported by recent evidence from the Ellis Island arrival records (Spitzer and Zimran 2018).³⁰ Further support for the importance of chain migration comes from the ubiquity of town-to-town migration—the specialization of specific towns or small regions in Italy in migration to specific towns in the United States.³¹ This pattern was also noticed by contemporary observers, such as the Dillingham Commission, which particularly emphasized the role that letters and the ubiquity of return migrants played in enabling migration.³²

While the importance of social networks in the Italian migration is documented beyond doubt in the historical literature, the diffusion hypothesis crucially depends on one particular feature of these networks—that they spread gradually across municipalities. For this, there had to exist some (though not necessarily many) short-distance contacts across municipalities, while long distance contacts had to be scarcer or weaker. What historical evidence exists supporting the existence of such contacts? Small-region networks were documented in detail among immigrants in Cleveland (Barton 1975) and among immigrants in Chicago from the Calabrian Rende region (Sturino 1990). Similarly, studies of many smaller US cities found small-region clusters of Italian settlement,³³ as well as a tendency of Italian organizations to divide along provincial lines.³⁴ Weaker evidence to the same effect is the tendency of Italian American communities within the

³⁰In a sample of 31,476 adult Italian passengers arriving at Ellis Island between 1907 and 1925, 33 percent of all males and 72 percent of all females reported joining an immediate family member already present in the United States. Almost all of the rest named other relatives and friends, such that the share of passengers not reporting any contact in the United States was only 5 percent (Spitzer and Zimran 2018, Table A.1). In fiscal years 1908–1910, only 5.9 percent of North Italian and 1.1 percent of South Italian immigrants to the United States did not report joining either a friend or a relative (US Congress 1911b, p. 363, Table 40).

³¹For examples see cases listed by MacDonald and MacDonald (1964, Appendix II) and Cinel (1982, p. 28).

³²All of these features of chain migration are clearly illustrated in the case study of Antonio Squadrito (Online Appendix C), an early migrant from the Sicilian town of Gualtieri-Sicamino. Within less than a decade, a small number of early migrants were followed by “more than one tenth of the population” (Brandenburg 1904, p. 109).

³³For example, in Buffalo (Yans-McLaughlin 1975, pp. 25–26), St. Louis (Mormino 1986 [2002]), Tampa and Ybor City (Pizzo 1981, pp. 128–130), and Pittsburgh (Bodnar, Simon, and Weber 1982, p. 47).

³⁴For example, in San Francisco (Cinel 1982) and Buffalo (Yans-McLaughlin 1975, p. 125).

great metropolitan centers, such as New York, Chicago, and Toronto, to cluster by small areas of origin, thus forming “many Little Italies” (Baily 1999; Nelli 1967; Park and Miller 1921; Sturino 1990; Vecoli 1983; Zucchi 1985). In the case study of Antonio Squadrito from the small Sicilian town of Gualtieri-Sicamino (Online Appendix C), his followers included residents from four or five different neighboring localities. Outside of the literature on emigration, Lecce, Ogliari, and Orlando (2022) show that social contacts across nearby Italian towns existed in the context of trade, marriage, and linguistic ties.

Were there viable alternatives to chain migration that were independent of geographic proximity to previous migrants? Recruitment by labor agents was another method on which some Italians relied in their migration to the United States, in particular under the *padrone* system (Iorizzo 1966; Koren 1897; Nelli 1964; Peck 2000). However, it was not altogether disconnected from social networks; instead, it depended on them.³⁵ Even as some agents recruited workers from across Italy, “the emigrant relied on his townspeople to get in touch with the network of agencies and sub-agencies which eventually would lead to a job and cash” (Zucchi 1985, p. 121).³⁶ Some governments, such as Argentina and Brazil, and later Australia, New Zealand, and certain Canadian provinces, had policies of assistance and subsidies for immigrants (Baines 1995; Kelley and Trebilcock 1998). But subsidized emigration was ultimately banned in Italy by the 1902 Prinetti Decree (Baily 1999; Foerster 1919; Gould 1980), and even before that assisted migration was a rarity, particularly among US-bound immigrants. When assisted migration did exist, it was rarely independent of social networks (US Congress 1911a, pp. 61–64). In brief, insofar as overseas emigration was facilitated by such alternatives to chain migration networks, there is little evidence that they were capable of inducing the migration of Italians who were not yet part of these networks. The alternatives were not substitutes but complements to social networks.

3 Theoretical Model

Our model builds on the traditional push-pull framework (Sjaastad 1962; Todaro 1969), in which individuals’ incentives for migration are determined by push and pull factors, such as real wage gaps between the origin

³⁵The recruiting *padrone* had sub-agents who would travel back to “collect a work force in their home town in Italy” (MacDonald and MacDonald 1964, p. 86), and he “kept his paesani [fellow townsmen] together” (MacDonald and MacDonald 1964, p. 86). The *padrone* banker was “generally a paesano” (Foerster 1919, p. 391), and the US-based labor boss was “an extension of the informal networks” (Baily 1999, p. 98). The Dillingham Commission agreed with this assessment: “actual and direct contract-labor agreements cannot be considered as the direct or immediate cause of any considerable portion of the European emigration . . . immigrants, or at least newly arrived immigrants, are substantially the agencies which keep the American labor market supplied with unskilled laborers from Europe. . . as a rule, each immigrant simply informs his nearest friends that employment can be had and advises them to come. It is these personal appeals which, more than all other agencies, promote and regulate the tide of European emigration to America” (US Congress 1911a, p. 61).

³⁶Such was the case of four boys whose departure was assisted by Antonio Squadrito (Online Appendix C).

and the destination. We capture the role of networks and of spatial diffusion by nesting the decision of whether or not to migrate within a Susceptible-Infectious-Recovered (SIR)-like epidemiological model, with an underlying geographic network structure.³⁷ This model is capable of describing a world in which the role of migrant networks can take one of three broad forms. At one extreme, networks may be unimportant, with migration determined entirely by push and pull factors. Alternatively networks may be important but purely local, as in the standard network framework. In the third, spatial diffusion exists and is the primary determinant of the timing of mass migration. We then determine the model’s testable predictions under a parameterization that produces the third form. These predictions form the basis of our empirical analysis.

We do not expect evidence supporting any single prediction to individually validate the diffusion hypothesis, as each of them could potentially be rationalized by some combination of alternative explanations (some with more difficulty than others) as we discuss in more detail in section 5.5. Our goal is to document a number of new stylized and striking facts about the Italian emigration, all of which can be parsimoniously explained by the diffusion hypothesis alone.

3.1 Basic Setup

Individuals may be in one of three states. They begin as *unlinked*. These individuals are not able to migrate regardless of the incentive to do so. *Linked* individuals have access to the migration technology—a necessary condition for migration. Unlinked individuals may switch to being *linked* in one of two ways—when one of their contacts migrates,³⁸ or spontaneously. Every period, linked individuals make a choice of whether or not to migrate based on push and pull factors. If they migrate, they become a *migrated* individual and their unlinked contacts become linked. If the migrated individual had become linked spontaneously (rather than through the emigration of one of his contacts) and none of his contacts migrated before he did, then he is a *pioneer*. Individuals have both *intra-place connections* to other individuals in their same place and *inter-place connections* to individuals in other places. For simplicity, the following discussion will focus on the case of a single destination. When there are multiple destinations, the progress of individuals from unlinked to linked to migrated is separate for each destination and individuals linked to multiple destinations decide whether to migrate to one of them or to remain in the origin.

The diagram in Figure 5 illustrates the model and its main concepts using a hypothetical chain of events.

³⁷The SIR model is originally due to Bernoulli (1776) and Kermack and McKendrick (1927), and has been applied in economics by Burnside, Eichenbaum, and Rebelo (2016) and Eichenbaum, Rebelo, and Trabandt (2021), among others.

³⁸This switch can capture prior migrants providing material support to potential migrants and the provision of information by prior migrants to potential migrants, among others. All are consistent with the spirit of our model, which requires only that an individual’s contacts’ migration somehow enable his own. These mechanisms are indistinguishable in our data and we are agnostic as to which one of them carried more weight.

There are three municipalities, A , B , and C . The first individual to migrate was a_1 from municipality A . He was a pioneer, in the sense that he migrated after switching spontaneously from susceptible to linked without contact with a prior migrant. He was connected to other residents of municipality A , a_2 and a_3 , and his migration converted them from unlinked to linked. Eventually, a_2 and a_3 also decided to migrate, converting four more unlinked individuals in municipality A to being linked. This is a case of *intra-place diffusion* of the migration technology. As the process proceeds, municipality A is likely to quickly become *saturated*, in the sense that all individuals would either become linked or will have already migrated, leaving no more unlinked individuals. At this point, municipality A 's migration rate is determined solely by push and pull factors and not by the rate at which the migration technology diffuses.

The migration of individual a_3 also linked b_1 , an out-of-town contact in neighboring municipality B . This is a case of *inter-place diffusion*, which caused *spatial contagion* of migration from municipality A to municipality B . If individual b_1 were eventually to migrate, municipality B would likely advance towards saturation with some time lag relative to municipality A , transmit the migration technology to other neighboring municipalities, and so on. Municipality A is an *epicenter*, since migration was already common there before arriving in its neighboring municipalities. Municipality C , on the other hand, is further from A , and without receiving the migration technology through an inter-place linkage, one of its residents, individual c_1 , spontaneously gained the option to migrate. If he migrates, he becomes a pioneer and is likely to start a new chain of migration spreading from municipality C .

The main state variables of the model for place i in period t are $\mathfrak{S}_{it} = \{U_{it}, L_{it}, M_{it}, N_{it}\}$, where U_{it} , L_{it} , and M_{it} denote the share of individuals within the municipality who belong to each of the three states (unlinked, linked, and migrated) and N_{it} is a measure of the exposure of municipality i to emigrants in all other municipalities. The latter can be thought of as the probability that any out-of-municipality contact of an individual in municipality i is a migrated person. It takes the form

$$N_{it} = \frac{\sum_{j \neq i} M_{jt} P_j d_{ij}^\pi}{\sum_{j \neq i} P_j d_{ij}^\pi}, \quad (1)$$

where P_j is the population of municipality j , d_{ij} is the distance between municipalities i and j , and $\pi < 0$ is the rate at which the likelihood that an individual in municipality i has a contact in municipality j decays with distance.³⁹ By definition, $U_{it}, L_{it}, M_{it}, N_{it} \in [0, 1]$ and $U_{it} + L_{it} + M_{it} = 1$.

The main parameters of the model are the set $\Theta = \{\lambda, \delta, \alpha, \pi\}$, where $\lambda > 0$ is the number of individuals

³⁹ Greater interactions with individuals in closer spatial proximity are consistent with Coleman (1988) and Jackson, Rodriguez-Barraquer, and Tan (2012).

in the same municipality to which each individual is connected, which determines the rate of intra-place diffusion; $\delta > 0$ is the number of individuals in other municipalities to which each individual is linked, which determines the rate of inter-place diffusion; and $\alpha > 0$ is the rate at which individuals spontaneously gain the option to emigrate, which governs the prevalence of potential pioneers.

Let m_{it} denote the probability that a linked individual from municipality i chooses to migrate in period t . The variation across municipalities in the probability m_{it} reflects the underlying variation in push factors in migration internal to municipality i .⁴⁰

In each period, the timeline is as follows. First, individuals who were linked in the previous period decide whether or not to migrate; then, new links are created, caused by individuals who emigrated in the first part of the period, or by spontaneous generation. The implied laws of motion for the state variables are then

$$\Delta M_{it} = m_{it} L_{it},$$

which is the change in the fraction of the population that has already migrated. The fraction of individuals exiting the susceptible state is the rate of those who become linked,

$$\Delta U_{it} = -[1 - (1 - \alpha)(1 - \lambda \Delta M_{it})(1 - \delta \Delta N_{it})] U_{it};$$

that is, new linked individuals are created either spontaneously, from linkages to newly migrating individuals in the municipality, or from linkages to newly migrating individuals in other municipalities. These individuals are added to the fraction linked, for which the law of motion is

$$\Delta L_{it} = -m_{it} L_{it-1} + [1 - (1 - \alpha)(1 - \lambda \Delta M_{it})(1 - \delta \Delta N_{it})] U_{it};$$

that is, those who become linked are added and those who migrate are lost.

3.2 Discussion

Under different parameterizations of Θ , our model can be calibrated so as to reflect the three different cases of the role of networks. To the extent that pioneers are rare (α is small), out-of-municipality contacts are frequent (δ is large), and very remote out-of-municipality contacts are rare (π is large in absolute value), migration will be dominated by a process of spatial diffusion. As a rule, mass emigration will not take off

⁴⁰These push factors are distinct from linkage status, and capture such features as the income level or the degree of economic modernization.

before the migration technology arrives through the diffusion of short-distance inter-place contacts. Even regions in which the incentives for emigration are high and no internal characteristic hinders migration may be prevented from producing mass emigration for a long period of time. Moreover, no change in local push factors is necessary for mass emigration to suddenly be ignited.

Eliminating out-of-municipality contacts entirely ($\delta = 0$) produces a standard model with networks but without spatial diffusion. If the arrival rate of pioneers (α) is sufficiently small, an absence of networks can still prevent migration for long periods from places with strong migration incentives, and the rare arrival of pioneers may drive a place into mass migration without a local change in push factors.⁴¹

Finally, our model can be reduced to a simple push-pull model without a role for networks if the number of contacts, both within and outside a place, approaches 0 ($\lambda = \delta = 0$) or when pioneers are very common (α large).⁴² Under such a parameterization, if a place does not produce mass emigration, it must be because the internal characteristics of that place are not conducive to emigration (Hatton and Williamson 1998, p. 39). Networks supporting migration will spontaneously be generated wherever local factors are conducive to emigration. Thus, conditional on local characteristics, the timing of the onset of mass emigration in a place is independent of whether or not migration was already present among its neighbors.

3.3 Predictions

When the parameters of the model are such that spatial diffusion is dominant, the model makes several predictions that can be evaluated in the data, as demonstrated in the simulations of Online Appendix D.

Prediction 1 (Convergence). The overall cross-place variation in the rates of emigration caused by underlying variation in push factors is initially augmented by the variation in access to the emigration technology. As a growing number of places are infected and approach saturation, the latter source of variation is gradually

⁴¹This framework shares common ground with the standard analysis of migration, but with one notable difference. In particular, our conceptualization of the friends and relatives effect at the micro level is different in that it views a network connection to be necessary for migration rather than as simply a continuous cost shifter or payoff enhancer. But at the aggregate level this difference is largely immaterial, or at most a matter of a different arbitrary choice of functional form: in both models, current aggregate migration from a municipality is some function of past migration. The substantive difference is that our model (under the first parameterization) allows the inter-place diffusion of migrant networks—the feature that enables the formalization of spatial diffusion. Spatial relationships in migration might still be observed due to spatial correlation in push factors.

⁴²When the bulk of the country has achieved saturation, the differences between the parameterizations with and without rare pioneers are largely eliminated. As a result, our model can also capture a phenomenon in which improvements in standards of living eventually reduce migration—when everyone is linked, the migration decision is based solely on push and pull factors, and smaller wage differences reduce the incentive to emigration. Hatton and Williamson (1998) argue that the typical curve of migration has an inverse-U shape, as emigration eventually eliminates real wage gaps, thus reducing the incentive to emigrate. They reject the diffusion hypothesis for failing to predict the downward-sloping side of the curve (p. 15), which is seen in the cases of German and Scandinavian emigration. But this insight regarding the drivers of emigration at saturation shows that, if indeed real wage gaps are eliminated, then a decline following saturation is perfectly consistent with the diffusion hypothesis. The fact of the matter is that very little reduction of real wage gaps occurred between Italy and the United States before World War I.

eliminated, such that the overall variation levels off around a lower rate, reflecting only variations in push factors. This leveling is manifested in a pattern of σ -convergence—cross-sectional measures of dispersion of migration rates will decline steadily, until they stabilize when the entire country is saturated.⁴³ Second, this process is generated by β -convergence. Places that are latecomers to migration due to an initial absence of linkage to prior migrants experience rapidly rising migration rates shortly after linkage, whereas places that are already saturated have higher rates but little or no growth. With convergence in migration rates coming from laggards catching up, the β -convergence prediction implies a strong negative relationship between past migration rates and future growth in migration. To be clear, the diffusion hypothesis does not predict that all places will converge to a uniform rate of emigration. Due to variation in local push factors, a significant amount of variation may remain even when the entire country is saturated.⁴⁴

Prediction 2 (S-Shaped Local Trends). Before any individual in a place is linked, the place’s emigration rate will be zero. Once the first individuals in a place are exposed, intra-place diffusion will generate a rapid increase in the emigration rate as individuals become linked, emigrate, and link their connections. Eventually, the place will reach saturation when nearly everyone is linked, and the rate of emigration will stabilize around a level determined by push factors. When combined, these three phases will create an S-shaped local time series of emigration rates. A steadily and gradually rising trend in national emigration rates will be a result of the accumulation of many successive sharply rising local S-curves will also generate the convergence in Prediction 1. This, too, was an observation linked by Gould (1980) to diffusion, and is also a common prediction of the technology-adoption literature (e.g., Bass 1969; Jovanovic and Lach 1989).⁴⁵

Prediction 3 (Correlated Destinations). Two neighboring places will typically share the same destination-specific networks due to their proximity. Therefore, they should have a similar menu of migration options and a similar distribution of destination choices. On the other hand, two distant places are more likely to be part of different networks, potentially leading to different destinations and thus generating different distributions of destination choices. Moreover, as the network of migration to each destination spreads across the country, the set of potential destinations of any two places will likely become increasingly similar.

⁴³This prediction was first suggested and assessed by Gould (1980), who measured cross-regional Gini coefficients in Italy, Portugal, and Hungary.

⁴⁴Gould (1980, p. 314) points out that “The process of diffusion . . . did not guarantee that pioneer migration would be followed by a mass movement increasing in some predetermined mathematical progression. If the conditions were not propitious: if the income gain was insufficiently large, for example, or the conditions of the migrant community unacceptable in some other way, the pioneer movement would prove still-born.” Such was the case, for instance, in the migration from the Spanish province of Málaga to Argentina, which was initiated by migration subsidies, but faded after these were removed (Sánchez Alonso 2015).

⁴⁵The standard SIR model follows the S with a declining portion of the curve coming from immunity due to prior exposure. In this case, we do not predict such a decline for two reasons. First, the rates of emigration, even where they were the highest, were never high enough to completely deplete the population. Second, the continual entry of individuals into the age cohorts associated with emigration would keep the pool of potential emigrants well stocked. Indeed, despite 14 million departures from Italy over our study period, the population grew from nearly 27 million in 1871 to nearly 38 million in 1921.

Therefore, the similarity in the distribution of migration destinations of two places should increase with the proximity between them and increase over time.

Prediction 4 (Frontier Effect and Spatial Expansion). The mechanism that undergirds the diffusion process is *spatial contagion*—the transmission of the migration technology from places already engaged in emigration to neighboring places. The immediate prediction that follows from it is the *frontier effect*: defining the *frontier* in any given period to be the boundary of an area that has already crossed a certain threshold level of emigration, the probability that a place enters mass emigration in the current period is positively related to its proximity to the frontier in the previous period. This ultimately results in *spatial expansion* of mass emigration, such that, starting from the early sources of mass emigration (the epicenters), successive places will enter mass emigration in spatial order from near to far. Thus, proximity to epicenters is expected to be an important determinant of migration rates early in the migration process and to become less important as the diffusion process plays out.

To be clear, patterns satisfying this prediction can be the product of other mechanisms, and section 6 is dedicated to showing that in our case it was most likely the product of spatial contagion. However, beyond being consistent with the basic mechanism, the frontier effect in itself is a crucial component of the diffusion hypothesis. Places not already experiencing mass migration must only rarely begin to do so unless they are close enough to places where the migration technology has already arrived. If spatial contagion occurs but it is not strong enough to dominate the evolution of emigration and produce the frontier effect, then emigration does not spread primarily by spatial diffusion, and the diffusion hypothesis fails. The same is true in that the frontier effect must generate spatial expansion from epicenters. Therefore, the frontier effect and spatial expansion are not only predictions but also necessary conditions that must be satisfied.

Predictions 1 and 2 do not require spatial diffusion to occur, and indeed will also arise in the case of purely local networks and rare pioneers. Alone, they are indicators of the importance of networks but not necessarily of spatial diffusion. But spatial diffusion is the simplest and most plausible case that generates all four predictions together.

4 Data

4.1 Sources and Construction

Our main data source is the *Statistica della Emigrazione Italiana per l'Estero*.⁴⁶ This series of volumes was published approximately every two years from the 1870s to the 1920s by the Italian *Direzione Generale della Statistica*. We digitized three data panels from this source. The first is a panel of annual emigration counts spanning the period 1884–1920 at the level of the municipality (*comune*), of which there were more than 8,300 in Italy.⁴⁷ The second is a series of annual emigration counts at the district level (*circondario* or *distretto*), of which there were 284, which enables us to extend our temporal coverage to begin 8 years earlier in 1876. The last is an annual panel (1877–1920) of emigration counts for 28 consistently defined destinations (usually countries) at the level of the province (*provincia*), of which there were 69 in Italy.⁴⁸ We focus in most of our analysis on three aggregated main destinations—North America, South America, and Europe, which together comprised 96.8 percent of all Italian emigration during the period 1877–1920.⁴⁹

The *Statistica della Emigrazione per l'Estero* has previously been used by a number of studies of Italian emigration (e.g., Ardeni and Gentili 2014; Faini and Venturini 1994; Gould 1980; Hatton and Williamson 1998; Moretti 1999), but none has used data at a level finer than the province-decade. The high resolution of the municipality-level Italian emigration data that we collected yields perhaps the most detailed data in terms of geographic disaggregation and temporal coverage available on a migration flow as large, as geographically varied in origin and destination, and as historically important as that from Italy during the Age of Mass Migration.⁵⁰ These features are essential to our study of the spatio-temporal expansion of migration and thus to our evaluation of the diffusion hypothesis—the spread of emigration over space simply cannot be observed at a sufficiently fine level with data at the level of the province or higher, and smaller countries with available locality-level data, such as Sweden or Denmark (Boberg-Fazlić, Lampe, and Sharpe 2021; Karadja and Prawitz 2019; Lowell 1987) are characterized by distances too short to observe a lengthy diffusion process.

⁴⁶Detailed citations for this and the other historical statistical publications that we use are given in Online Appendix E.

⁴⁷We have data for 8,317 municipalities, though a lack of population counts in some cases limits our sample of municipalities with known emigration rates to 8,029.

⁴⁸The publications omit tables for 1879 for the district and province-by-destination data and for 1888 for the province-by-destination data. In 1916 and 1917, there was virtually no transatlantic migration because of World War I, and consequently there were no volumes published for these years.

⁴⁹These data are based on contemporary jurisdictional boundaries, which experienced some changes during our study period, as well as in the century since. Online Appendix F describes how we harmonized these data to fit consistently defined geographic units.

⁵⁰Karadja and Prawitz (2019) and Lowell (1987) use highly detailed data on emigration from Sweden—a country with less than one-sixth of Italy's population. Boberg-Fazlić, Lampe, and Sharpe (2021) use individual and parish level data on Danish emigrants. Fernández-Sánchez (2021) uses detailed data from a single region of Spain. Work in progress by Fontana et al. (2021) also uses the Italian emigration data.

The emigration counts are based on passports issued to would-be emigrants by local administrative officials. Although it provides the most comprehensive data available on Italian migration, there are some known issues with this source, such as inaccurate reporting by the mayors (*sindaci*) of the Italian municipalities. The most concerning issue is a 1901 change in Italian law.⁵¹ Prior to this law, passports were helpful, but costly and not strictly required; after 1901, they became free and compulsory when departing from Italy for trans-Atlantic destinations.⁵² It is therefore potentially concerning that the figures document a surge in emigration, in particular to the United States, between 1900 and 1901.⁵³ But US arrival data (Barde, Carter, and Sutch 2006) show growth in Italian arrivals from 1900–1901 that closely matches the increase in our data, reducing the concern that this surge was spurious or that many emigrants avoided taking out passports prior to the policy change.⁵⁴ Another issue is that the municipality-level data for 1884–1903 aggregate some municipalities with low but non-zero emigration rates into a single figure for each district. Municipalities included in this aggregation will appear to have an emigration rate of zero in these years. We address this concern in Online Appendix G, where we repeat our main municipality-level analyses assigning the aggregate emigration to unlisted municipalities, with similar results. This concern does not affect the district-level data.

Another concern raised by Foerster (1919, ch. 2) and Hatton and Williamson (1998, ch. 6) is that the distinction in the emigration data between temporary and permanent immigration, when it is made, is unreliable. We agree, but we do not view this as a deficiency affecting our analysis. Return and repeat migration were frequent (Bandiera, Rasul, and Viarengo 2013), but the intended duration of migration upon departure was subject to unpredictable changes (Ward 2017). Our goal is to explain the total flow of labor, permanent or temporary, and therefore we ignore this distinction and count both cases equally. However, the issue of return and repeat migration becomes acute in the northern border regions, where seasonal migration across the border was so frequent that in several municipalities the total number of leavers throughout the period far exceeded the total population. While this is an encouraging indication that even easy overland exits were documented in the data, it leads us to treat border-region emigration counts in particular, and, more generally, emigration counts to Europe as a whole, with caution.

Our benchmark specifications use emigration rates based on 1901 population as the denominator, and

⁵¹See Foerster (1919, pp. 11, 21) and Hatton and Williamson (1998, p. 98).

⁵²See the discussion of the accuracy of the Italian emigration data by Foerster (1919, pp. 10–22).

⁵³According to Foerster (1919, p. 21), the Italian official statistics were less precise than the American immigration data, and that around 1901 there was a change from under- to over-enumeration of Italian emigrants. (Hatton and Williamson 1998, ch. 6) describe this surge as spurious, but this is largely due to their distinction between temporary and permanent migration, which we address below.

⁵⁴This is shown in Online Appendix Figure B.5. Although there is little difference around 1901, larger differences emerge later in the study period.

we verify robustness to using 1881 population instead (Online Appendix H).⁵⁵ Since the smallest geographic unit for which destination data are available is the province, we impute destination-specific emigration rates for each municipality and district based on the province-year-specific weights of destinations.⁵⁶ As described above, our main focus is on migration to North America (the United States and Canada).⁵⁷

In addition to the emigration data, we draw, from a wide range of sources, a battery of municipality- and district-level post-unification characteristics that are potentially relevant for determining emigration rates. Their purpose is two-fold. First, they serve as control variables in the various statistical tests for the predictions of the diffusion hypothesis. Second, some of them proxy for local features that, according to internalist explanations, might have determined the timing of mass emigration, and we use them in order to assess their validity relative to that of the diffusion hypothesis. Municipality-level data include geographic characteristics (elevation, distance to the coast, and distance to land borders), distance to the nearest railway line in 1881 (Ciccarelli and Groote 2017), birth and death rates in 1881, per capita membership in mutual aid societies in 1878, and per capita deposits in postal savings banks in 1886.⁵⁸ District-level data are digitized from the 1881 census—the earliest with the data that we require—and include demographic and labor-force composition—the fraction of the male labor force employed in agriculture or industry (including traditional cottage industries), the fraction of the population younger than 15 years old, and the fraction of males aged 15 or older who were literate. Altogether, these variables account for market access, industrialization, demographic pressures, social capital, and financial development; their summary statistics are presented in Online Appendix Tables B.1 and B.2.

⁵⁵We use 1901 population because it is the population reported in the 1904–1905 volume of the emigration statistics, which is the first volume in which the emigration of all municipalities is reported regardless of magnitude. It thus enables us to secure population measures that are most comparable to those of our emigration data. A related issue to the distinction between the legal and actual population is internal migration. While mass internal migration is largely a phenomenon of the Fascist and post-World War II period (A’Hearn and Venables 2013, p. 625), there is some evidence of significant population movements in other settings, such as along the former borders after unification (A’Hearn and Rueda 2022). But it is clear that this was overall a minor phenomenon (and a poorly documented one) relative to international migration. We are grateful to Brian A’Hearn for a helpful discussion on this topic. See also Spitzer, Tortorici, and Zimran (2022) for a discussion of the relative magnitudes of internal and international migration.

⁵⁶A potential consequence of this is that there may be artificial within-province-period correlation in municipalities’ emigration to a particular destination that does not extend past provincial borders. We address this issue by repeating our main results using data for all destinations in Online Appendix I. Our results are qualitatively unaffected in general.

⁵⁷We add Canada because the volumes for early years do not distinguish between the two countries in the provincial counts. Canada comprised 2.4 percent of all migrants to North America in an average year in which counts for Canada and the United States were separately reported, and never more than 7.5 percent.

⁵⁸Sources for these data are described in Online Appendix E. There are a number of potential measures of the presence of credit, savings, or liquidity, as described in more detail by Spitzer, Tortorici, and Zimran (2022). Using the postal savings banks data from 1886 combines a variety of desirable features: the banks are sufficiently small so as to potentially have been relevant to individuals and the data are sufficiently close to 1881, which is the year in which we observe most other local characteristics; thus, the postal savings data enable us to capture the presence of local savings and liquidity without introducing too many distinct variables.

4.2 Summary Statistics

Figures 1 and 2 present maps of emigration rates for municipalities and districts by half decades.⁵⁹ In Figure 1, these are at the district level and begin in 1876. In Figure 2 they are at the municipality level and begin in 1884.⁶⁰ As discussed above, they provide visual evidence of the existence of spatial expansion, as well as for the approach to saturation in the latter periods. Interestingly, the map of emigration to all destinations in the final period bears a remarkable resemblance to the topographical map of the country (compare Online Appendix Figure B.11, panel f and Online Appendix Figure B.12): the coefficient of correlation between emigration and elevation is 0.37 in the 1911–1914 half decade at the municipality level. Though investigating its causes is beyond the scope of this paper,⁶¹ this strong correlation is clear indication that elevation was, in a sense, a first-order determinant for a place’s potential for emigration, yet its importance became visible only after the diffusion process was nearly complete. This attests to the power of the absence of networks in constraining migration from otherwise emigration-prone places.

We define an epicenter of emigration to North America to be a district that had an average annual emigration rate of at least 1 per thousand to North America during the period 1876–1883, and which did not have a neighboring district with a greater annual average emigration rate to North America in this period. This criterion defines 6 epicenter districts, marked in Figure 6.⁶² We define the distance from the epicenter as the distance of each municipality from the nearest capital municipality of an epicenter district. For South America, we devise a similar definition but with a higher threshold of 5 per thousand because South America-bound emigration started before the late 1870s and we first observe it when it was already well developed.⁶³ Because the Europe-bound emigration was clearly greatest in the districts sharing a land border with neighboring European countries, we use the distance to this border as the measure of distance to the European epicenter. Finally, we define the *frontier of mass migration* to a destination to be the contour of districts that had ever achieved an emigration rate of at least 5 per thousand by a given half decade.

As shown in Figure 6, the evolution of the frontiers diverged meaningfully across the three major destination groups. While there is some similarity between the geographic origins of the three flows—for instance, emigration rates from Tuscany and Latium were low regardless of the destination, and both the North and

⁵⁹This length of period smooths out short-term fluctuations in emigration without obscuring longer-term trends.

⁶⁰Online Appendix Figures B.6–B.11 present analogous figures for migration to South America, to Europe, and to any destination.

⁶¹The relationship between elevation and emigration has previously been remarked upon by Gould (1980, pp. 290–291), Cinel (1982, p. 31), and Sturino (1990, p. 14). But our understanding is that this is the first time that this phenomenon has been quantified and shown to be so striking and consistent throughout the country.

⁶²These districts and the provinces in which they lie are Sala Consilina (Salerno), Isernia (Campobasso), Corleone (Palermo), Chiavari and Albenga (Genova), and Pozzuoli (Napoli).

⁶³The epicenters of South American emigration that we identify are Lagonegro (Potenza), Chiavari (Genoa), Asiago (Vicenza), and Gemona (Udine), and are also marked in Figure 6.

South American flows had epicenters in Campania, Basilicata, and Calabria—there are also major differences. The contrasting north-south divides in the European and North American flows is clear, but may have been the product of lower migration costs to Europe in the north and the greater prevalence of repeat migration where distances to the border were shorter. Other differences, however, are harder to explain by some destination-specific regional advantages. For example, Veneto in the northeast had extremely high rates of migration to South America, but low rates to North America, and to some extent the same was true in the northwest. Similarly, Sicilian migration to South America came primarily from the southern half of the island, whereas migration to North America spread out from its northwest.

5 Patterns

In this section, we check whether each of the four predictions laid out in section 3 are substantiated in the data, after which we test the relative efficacy of diffusion and internalist explanations in determining the timing of the onset of mass migration. Throughout this section, the geographic units of analysis are the district and the municipality, and, except where otherwise indicated, the temporal unit of analysis is the half decade. The main destination of interest is North America. We present analogous results for migration to all destinations in Online Appendix I

5.1 Convergence

Figure 7 examines the trend in the annual cross-sectional dispersion of emigration at the district and the municipality levels, for emigration to North America. Our preferred measure is the coefficient of variation because it is both normalized by scale and, unlike the standard deviation of log-emigration, can account for cases of zero migration. There is clear evidence of steady, almost perfectly monotone, σ -convergence, with the coefficient of variation decreasing from around 4 in 1880 to just over 1 in 1910.⁶⁴

Figure 8 shows that this σ -convergence was not the product of a tendency of all places to regress towards the mean, but of β -convergence—new areas entering mass emigration, and areas that had already experienced migration achieving saturation and stabilizing emigration rates. We compute the average annual emigration rate for each place for the periods before and after 1900 and plot the relative change between the two

⁶⁴Due to the tendency to not report the specific migration counts of municipalities producing little migration in the early years, one might expect to see an exaggeration of the downward trend in cross-municipality variation. However, the district totals were almost always fully reported, and the trends of the decline in variance based on either municipality or district data is very similar. This suggests, first, that the problem of omission of low migration counts is not detrimental; and second, that the decline in variance occurred across larger units rather than within them. This is consistent with the idea that the arriving tide of migration lifted all boats in the same area.

periods against the rate in the first period. As the β -convergence prediction implies, there is a strong, negative, and nearly monotonic relationship between these variables. Places with the highest emigration rates in the first period experienced almost no growth. The average municipality in the top quartile of pre-1900 migration experienced approximately a quadrupling of emigration rates, while emigration from bottom-quartile municipalities grew more than 200-fold. Importantly, not a single district and only very few municipalities reduced their emigration rates at all. This is clear evidence that the β -convergence is not a simple case of mean-reversion or of churning of leading and lagging places due to random shocks. Instead, the rate of migration in the early period was the effective lower bound for the rates in the later period. Considering that real wage gaps relative to destination countries were roughly stable, this is strongly consistent with the notion of saturation—emigration rates plateaued around their full potential when the diffusion of migration was completed.⁶⁵

A natural concern is that the β -convergence patterns are spuriously generated by measurement errors or by idiosyncratic random shocks, as the pre-1900 emigration rates enters positively into the right-hand side and negatively into the left-hand side. In Online Appendix J, we present a variety of exercises showing that this correlation may have been partly augmented by such bias, but that it is nevertheless real and strong.

5.2 S-Shaped Local Trends

Prediction 2 is that the typical course of the evolution of emigration at the local level followed an S-shaped curve. Figure 9 plots a smoothed time series of migration rates for district (panels a and c) or municipalities (panel b and d) with the time measure normalized so that year zero is the first year in which the area reached an emigration rate of at least 5 per thousand. Panels (a) and (b) present both the average migration rate for each year-since-mass-migration bin alongside the median and quartiles of the distribution, smoothed over time. A clear S-shape is evident for both the municipality and district levels for all quartiles and for the mean, the latter with tight confidence intervals.⁶⁶ The average municipality took about 25 years to make the transition from little migration to saturation, and once a place contracted migration, the surge was rapid and largely irreversible.

Panels (c) and (d) of Figure 9 divide the municipalities and districts by the period in which they first reached the mass emigration threshold. Regardless of when the migration surge in a place began, it followed roughly the same path, except that in the late 1890s the surge appears to have been somewhat faster. This

⁶⁵This also means that no place had an inverse-U-shaped emigration curve, regardless of the length of its emigration experience.

⁶⁶That is, the confidence intervals are such that only an S-shaped curve for the average municipality can be drawn within them.

is inconsistent with the notion that the leading places in the migration movement had characteristics that made them particularly well suited to migration: if the latecomers were places with weaker push factors, we should expect their emigration path to be lower and to reach a lower peak. Summing up the evidence on the convergence and the S-shaped time series, the continuous national surge in emigration during the period from unification to World War I was not the product of a rising tide that lifted all boats. Rather, places were sequentially lifted from no migration to their mass migration potential at different times but along the same path.

5.3 Correlated Destinations

The data on emigration by destination at the province level enable us to test Prediction 3 regarding correlated destinations. We compute a dissimilarity index V_{ijt} between the destination-country distributions of the emigration flows of every province pair ij in every half decade t (Duncan and Duncan 1955). The dissimilarity index has the convenient feature of being interpretable as the fraction of the emigration flow from province i that would have to have been rerouted to match the destination distribution of province j (or vice versa). Panel (a) of Figure 10 presents a nonparametric regression for each half decade of the dissimilarity indices of each province pair against the distance between the provinces. The relationship is clearly positive, as expected: pairs farther from one another had more dissimilar destination choices. Moreover, conditional on distance, dissimilarity indeed seems to diminish from period to period. Dissimilarity that increases with distance could also be the result of greater dissimilarity between the characteristics of provinces, which might affect destination choices, rather than between their migration networks. It is therefore important to control for within-pair differences in characteristics in a formal test of this pattern. The prediction of diminishing dissimilarity over time, however, does not have such a straightforward alternative explanation.

We estimate an equation of the form

$$V_{ijt} = \alpha_t + \beta \log(d_{ij}) + \mathbf{x}_{ij}'\gamma + \varepsilon_{ijt},$$

where α_t are half-decade fixed effects and d_{ij} is the distance between the capitals of provinces i and j and the controls \mathbf{x}_{ij} are absolute differences between provinces in their individual-level 1881 agricultural employment share, industrial employment share, literacy rate, and fraction under age 15.⁶⁷ Panel (b) of Figure 10 plots

⁶⁷We present the estimates of β in Online Appendix Table B.3, using all destinations or different combinations of destinations, with and without control variables. The upward slope of the dissimilarity-distance relationship is robust to the inclusion of controls and to a focus on different time periods or sets of destinations, casting doubt on the notion that increasing differences in local characteristics over distance can explain our results.

the half-decade fixed effects from this regression with the 1876–1880 half decade as the excluded category, as well as the fixed effects from a similar regression limiting attention to province pairs less than 300 kilometers apart. There is a clear decline in the fixed effects over time, and the decline is particularly monotone for province pairs within the 300-kilometer range.⁶⁸

5.4 The Frontier Effect and Spatial Expansion

Figure 11 shows the main evidence regarding the frontier effect. In panel (a), the rates of emigration at the district level are plotted in a non-parametric regression against the distance from the frontier a half-decade earlier, including in the sample in each period only places that had not yet produced mass migration to North America.⁶⁹ When pooling all periods together, the expected trend is sharp and clear (notice that the scale of the vertical axis is logarithmic). At a distance of 25 kilometers, the rate of emigration is on average 3.2 per thousand. It then decays rapidly to 2.3 at 50 kilometers, 1.0 at 100 kilometers, and 0.3 at 150 kilometers; beyond that, the effect of the frontier weakens, as should be expected at distances that are unlikely to allow for personal contacts. The same pattern is apparent, albeit with some volatility, in each half-decade separately as well. Distance from the frontier is more than a characteristic that correlates with emigration rates; the general absence of high emigration rates beyond 100 kilometers from the expanding frontier of mass migration suggests that proximity to this frontier of mass migration was, in most cases, *necessary* for the onset of mass emigration.⁷⁰

A formal test for the frontier effect is presented in panel (b) of Figure 11. We regress emigration rates on period-specific functions of distance from the previous period’s frontier. To account for places with zero recorded emigration, we use the binomial maximum likelihood estimation described in Appendix A. The estimates are indeed universally negative, and are statistically significant until the half decade beginning 1915. Moreover, adding controls for local characteristics does not reduce, and sometimes increases, the estimates. This suggests that the frontier effect is not driven by spatial trends in observed characteristics.⁷¹

The diffusion hypothesis also requires the frontier effect to have led to another observable pattern—that

⁶⁸In more intuitive terms, the baseline specification suggests that moving from the 25th to the 75th percentile of distance between provinces entails a 0.17-standard deviation greater dissimilarity. Over the period between 1876–1880 and 1911–1914, the dissimilarity of any pair of provinces diminished by 0.20 standard deviations of the 1876–1880 distribution.

⁶⁹Online Appendix Figure B.13 presents analogous results with the municipality as the unit of analysis.

⁷⁰Pooled over all periods, the share of districts entering North American mass migration for the first time was 36.4 percent when the previous half-decade’s frontier was less than 50 kilometers away, as opposed to 2.2 percent when it was over 100 kilometers away.

⁷¹Online Appendix Figure B.15 shows that the frontier effect was destination specific, formalizing the visual evidence in Figure 6. It repeats the analysis of panel (b) of Figure 11, but in addition to the relationship between emigration to North America and distance to the North American frontier, it also shows the relationship between migration to European and the South American frontiers. Out of the three main streams, only migration to North America is systematically negatively correlated with distance to the North American frontier.

places situated farther from the initial epicenters contracted emigration later than nearby places. To validate this pattern of spatial expansion from epicenters, panel (a) of Figure 12 plots non-parametric regressions of the logarithm of half-decade average annual North American emigration rates against distance from the nearest epicenter, which is fixed over time (as opposed to distance from the frontier in Figure 11, which changes each period).⁷² Districts closer to the epicenters were indeed emigration leaders throughout the study period, as evidenced by the negative slopes of the curves. There is also a gradual leveling of the curve as more distant areas entered into emigration, to the point that it is nearly flat by the last half decade. Emigration rates from districts under 50 kilometers from the epicenter were 5.8 times greater than from districts 50–100 kilometers away, and 20.8 times greater than those in the range 100–200 kilometers in the period 1876–1880. By 1911–1914 these ratios had shrunk to 0.9 and 2.3, respectively.

We test this pattern formally in panel (b) of Figure 12 by regressing emigration rates on half-decade-specific functions of distance from the nearest epicenter, using the binomial maximum likelihood regression. At both the district and the municipality level, the coefficients on distance from epicenter are initially negative and monotonically decline in magnitude over time. As in the frontier regressions, this pattern is robust to controls, and it is notable that after adding them the coefficients of the last periods for the municipality-level data become indistinguishable from zero. In other words, distance from the epicenters was highly predictive of emigration rates at the beginning of the Italian migration, and ceased to be so by the time the country reached saturation.

Finally, to formalize the notion that distance from the epicenters determined the timing of entry into mass emigration, Table 1 presents the results of semiparametric Cox proportional hazard models (Zeng, Mao, and Lin 2016) for the timing of entry into the frontier of mass emigration to North America, focusing on the district as the unit of observation (since that is the unit of analysis that we consider when determining whether a place has entered the frontier of mass migration). Coefficients below 1 indicate that an increase in the variable in question is associated with a lower hazard, and thus with a later onset of mass emigration. Column (1) shows that a one-standard deviation increase in distance from the epicenter was associated with a 42 percent lower hazard of achieving mass migration at any time. Columns (2) and (3) show that this pattern is robust to controlling for broad regions (i.e., north, south, and center) and to controlling for the various district-level characteristics that we observe. Thus, distance from epicenters was a determinant of the timing of the onset of mass emigration as well as of its rate.

In sum, our findings complement the β -convergence prediction in showing that the laggards who caught

⁷²Online Appendix Figure B.14 presents analogous results with the municipality as the unit of analysis.

up with the leaders were really the more distant places narrowing the gap relative to those closer to the epicenters. In order to enter mass emigration, places had to be situated close to the recent frontier of mass emigration. This, in turn, generated a pattern of spatial expansion from the epicenters outwards, whereby farther places experienced a later onset of mass migration. Moreover, the robustness of these findings to local controls suggests the spatial trends in emigration were not likely a result of systematic spatial trends in underlying characteristics. As a rule, emigration from distant places was delayed, sometimes by decades, for no apparent reason other than their location relative to the epicenters.

5.5 Alternative Theories

The key advantage of the diffusion hypothesis, and among the strongest evidence of its validity, is that it can parsimoniously explain all of the patterns that we have documented above. Nevertheless, some of these patterns “might be explained in other ways” than by diffusion (Hatton and Williamson 1998, p. 99). In this section we assess the degree to which internalist explanations that view local economic and demographic factors as the key determinants of the timing of mass migration can explain the stylized facts that we find.⁷³ We pay particular attention to the modernization hypothesis, which is the incumbent explanation in the literature on the Age of Mass Migration for migration. Before beginning this analysis, we emphasize that, in order to rationalize the spatial patterns documented above, internalist explanations require strongly spatial patterns in the drivers of migration to be present. We find such patterns both unlikely and inconsistent with the available historical evidence. Nonetheless, this section and section 6 are devoted to providing more concrete evidence on the matter.

The key feature of any plausible internalist explanation is that factors internal to a location were the primary determinants of the timing of a place’s onset of mass migration. In the case of the modernization hypothesis, these factors must be specifically modernization indicators. To test for such patterns, we repeat the hazard regressions on a set of district-level characteristics around 1881. We present the results in Table 2, where all the explanatory variables are normalized for comparability and the rates of emigration are for all destinations.⁷⁴ Column (1) repeats the basic regression from Table 1, but using distance from any emigration epicenter. Columns (2), (3), and (4) report the hazard ratios associated with a one-standard deviation increase in each of several proxies of modernization—the share of workers in agriculture (a negative proxy for economic development), the birth rate or the fraction of the population under age 15 (proxies for demographic

⁷³Because they cannot explain the spatial patterns that we document, we do not consider in this section more standard models that view networks as the fundamental drivers of the timing of mass migration without spatial diffusion.

⁷⁴We do not separate emigration by destinations because this separation is not implied by the modernization hypothesis. Destination-specific regressions yield the same qualitative outcomes.

pressures), literacy, and urbanization measured as the share of the population living in municipalities with over ten thousand residents—as well as measures of mortality, social capital, and financial development. Columns (5) and (6) then reintroduce distance to epicenters, allowing diffusion to “compete” with these observable characteristics.

These results are not consistent with local characteristics being the main determinants of the timing of the onset of mass migration. Whether compared across specifications (column 1 as compared to columns 2, 3, and 4) or within specifications (columns 5 and 6), the roles of the local characteristics are in all cases second-order when compared to the consistently strong role of diffusion, with the mass emigration hazard more than doubling with each standard deviation reduction in distance in columns (5) and (6). Notably, two of the modernization proxies seem to act in a direction opposite to that predicted by the modernization hypothesis: districts that were more agricultural and less urbanized developed mass emigration earlier. Literacy is positively associated with earlier mass emigration in columns (2) and (3), as expected, although more weakly than the distance to the epicenter, and the direction of its impact changes in columns (5) and (6). The results for our proxies for demographic pressure are mixed: only the fraction of population under age 15 is positively associated with earlier mass emigration, as expected, though again more weakly than distance to the epicenter.⁷⁵ The evidence on the importance of modernization factors in determining the timing of the onset of mass migration is at best mixed and weak.

Internalist explanations also struggle to explain convergence in migration rates. Such explanations would rationalize convergence as a result of a process in which an impediment to migration—such as a liquidity constraint or lack of modernization—is sequentially removed, allowing places to experience surges in emigration.⁷⁶ If that were the case, then the internal factors should predict higher emigration early on, but at a diminishing rate as time goes by, similar to the evidence on the diminishing importance of the distance to the epicenter (Figure 12). In Figure 13 we examine whether a dynamic in which local characteristics diminish over time as predictors of emigration rates is borne out in the data. In each period, we regress district-level emigration to any destination on distance to epicenters and the local characteristics used in Table 2 (all normalized, for comparability). In the early periods, the coefficients are qualitatively the same as in the hazard regressions and the distance to the epicenter is by far the strongest predictor. Over time, all coefficients indeed converge towards zero, but those on the local economic characteristics do so from a

⁷⁵For this reason, we use the fraction under age 15 as our preferred proxy for demographic pressure in this table.

⁷⁶A naïve version of the modernization hypothesis, such as is implied by Hatton and Williamson’s (1998) analysis, would argue that modernization factors simply augmented other push factors in increasing the demand for migration. This view, however, fails to offer a simple explanation for the patterns of convergence in emigration rates, since the factors triggering early emigration would also be associated with higher rates overall. Leaders would simply be situated on a higher emigration path. As shown above, it is hard to make the case that, as a rule, early adopters had the highest emigration potential.

relatively low starting point, and in some cases the convergence is barely noticeable.⁷⁷ We conclude from this that spatial diffusion, captured by distance from epicenters, is the most likely source for the extraordinarily strong convergence in emigration rates that we document.

Another way in which the diffusion hypothesis outperforms internalist explanations is by providing a straightforward explanation for why migration streams to different destinations were at least partly independent of each other (Figure 6), and why similarity in destinations increased with proximity and over time (section 5.3). These patterns are not strictly inconsistent with internalist explanations—it could be argued that the overall level of emigration was not affected by a diffusion process, but that the destination was chosen based on the networks that were available in the vicinity. Such an explanation does, however, still emphasize the importance of inter-place networks and raises questions as to how, given their dominance, emigration can start in their absence. That is, the most generous interpretation is that internalist explanations require a more complex or an ad hoc explanation to account for these patterns.

To be clear, we do not argue that internal factors, including different aspects of modernization, did not affect emigration. What we do learn from this analysis is that they do not appear to offer a plausible, simple, and complete explanation for the set of stylized facts of the Italian emigration. To the extent that internal factors did play a role in determining the rates of emigration, their impact was small relative to the consistent first-order role of the diffusion process and they perform poorly in explaining the convergence in emigration rates. Moreover, evidence supporting the modernization hypothesis is mixed and partly contradictory.

6 The Spatial Contagion Mechanism

The fundamental building block of the diffusion hypothesis is the spatial contagion mechanism. In this section we set out to establish that spatial contagion was indeed the mechanism that caused the spatial diffusion of Italian emigration. Verifying the causal power and the economic significance of spatial contagion is important because it is a key differentiating feature that is not inherent to any internalist explanation. The challenge is to show that the correlation between a place’s rate of emigration and its neighboring places’ lagged rates of emigration reflects a causal relationship rather than simply a correlation driven by other spatially correlated and potentially unobserved local factors that cause emigration.

⁷⁷Results are similar for the variables excluded from this figure for clarity—1881 birth rate, mutual aid members per capita, and postal savings credit per capita.

6.1 The Rationale of the Instrumental Variables Approach

The baseline estimation equation is a spatial lag model of the form

$$\log(e_{it}) = \alpha_t + \beta \log(e_{-it-1}) + \mathbf{x}_i' \delta_t + \varepsilon_{it}, \quad (2)$$

where e_{it} is the rate of emigration from municipality i in half decade t , α_t are period (half-decade) fixed effects, \mathbf{x}_i is a vector of controls, and δ_t is a period-specific vector of coefficients. We specify equation (2) in logarithmic form rather than in levels in order to make the effects proportional to the level of migration.⁷⁸ We cluster standard errors at the district level, which permits correlation between any municipality-half decade observation within the same district, either over space, over time, or both.⁷⁹

The regressor of interest in equation (2) is e_{-it-1} , which we refer to as *lagged emigration exposure*. This is a lag of a weighted average of emigration rates of all other municipalities, with greater weight exerted by nearer and more populous municipalities. In a manner analogous to equation (1), we define this object as

$$e_{-it} = \frac{\sum_{j \neq i} e_{jt} N_j d_{ij}^\theta}{\sum_{j \neq i} N_j d_{ij}^\theta}, \quad (3)$$

where N_j is the population of municipality j , d_{ij} is the distance between municipalities i and j , and θ is the rate at which the influence of other municipalities' emigration rates on that of municipality i decays over distance. As in equation (1), the numerator is a measure of effective proximity to emigration, whereas the denominator is a measure of effective proximity to population. The measure e_{-it} is thus a population- and distance-weighted average emigration rate in the neighborhood of municipality i in period t . Like N_{it} in our model, it can be thought of as the emigration rate among an individual's out-of-municipality contacts in the previous period.⁸⁰ A value of $\theta < -2$ is consistent with an individual's frequency of contacts diminishing with distance. For computational tractability we separate the estimation of θ from that of other parameters.

⁷⁸This is consistent, for example, with Mahajan and Yang (2020), who find in linear regressions that the effect of hurricanes on migration increases in the size of the network, which represents the base level of migration. This assumption appears to fit the data better, at least by the criterion of generating seemingly normal distributions. Spitzer (2021) and Spitzer, Tortorici, and Zimran (2022) also focus on the logarithm of the emigration rate instead of its level.

⁷⁹Such spatial lag models have recently been used to study diffusion in other settings (e.g., Aidt and Leon-Ablan 2022; Aidt, Leon-Ablan, and Satchell 2022). Ours, however, is the first to pair this model with an instrumental variables strategy based on the epidemiological intuition of diffusion.

⁸⁰The measure e_{-it} satisfies two desirable conditions. The first is that it is robust to splitting municipalities. The second is that it is robust to uniform changes in population density. If the measure of exposure were a function of the number of emigrants (rather than the local *rate* of emigration), then doubling the population everywhere near a municipality (with an accompanying doubling of the number of emigrants) would double an individual's emigration exposure. To reflect the limited number of connections that a person can have, our measure is robust to population density, which would have no impact on the proximity-weighted emigration rate. The implied assumption is that the number of links that any individual has outside of his municipality is fixed and independent of the population density in the neighborhood.

We estimate equation (2) by NLS and arrive at a value of $\theta = -2.83$, which we later use throughout our main analysis.⁸¹ The coefficient of interest in equation (2) is β , which can be interpreted as an elasticity of emigration with respect to lagged emigration exposure.

The obvious challenge in estimating equation (2) by OLS is that local determinants of emigration, including unobserved ones, are likely to be spatially correlated, biasing upwards the estimate of β . Identification of the effect of emigration exposure on emigration therefore requires that we find a source of variation in neighbors' emigration that is independent of a place's own internal demand for emigration.

The diagram in Figure 14 illustrates the intuition behind our instrumental variables approach. Consider two municipalities, A and B , that are identical in all of their internal characteristics. In particular, they are equidistant from a source of emigration (an epicenter or a frontier) at a distance d_2 . The only difference between the two municipalities is that the neighboring population of municipality A is distributed such that on average it is closer to the source than is municipality A , whereas the neighboring population of municipality B is on average farther from the source than municipality B itself. The average neighbor of municipality A is therefore likely to encounter the spreading wave of emigration earlier than is the average neighbor of municipality B , not because of any feature that is correlated with their internal characteristics, but simply due to the different spatial orientation of their neighbors with respect to the source. Therefore, we can construct an instrumental variable for emigration exposure based on the weighted distance of a municipality's neighbors to the source.

6.2 Implementation of the Instrumental Variables Approach

In practice, the instrumental variable for emigration exposure \tilde{e}_{-it} is constructed in a manner analogous to the actual emigration exposure measure e_{-it} . It is defined as

$$\tilde{e}_{-it} = \frac{\sum_{j \neq i} \tilde{e}_{jt} N_j d_{ij}^\theta}{\sum_{j \neq i} N_j d_{ij}^\theta}, \quad (4)$$

but instead of being a weighted average of actual emigration, e_{jt} , it is a weighted average of predicted emigration, \tilde{e}_{jt} , which is only based on municipality j 's distance to nearest emigration source. In particular, to construct \tilde{e}_{jt} , we estimate a non-parametric regression of the form

$$\log(e_{jt} + \varepsilon) = f_t(z_j) + u_{jt}; \quad (5)$$

⁸¹As we show in Online Appendix L, our findings are robust to using alternate values of θ selected by estimating equation (2) with different sets of controls or fixed effects.

that is, a period-specific non-parametric regression of log emigration on distance from the nearest emigration source.⁸² We then set $\tilde{e}_{jt} = \exp[\hat{f}_t(z_j)]$ and estimate equation (2) using $\log(\tilde{e}_{-it})$ as an instrument for $\log(e_{-it})$.⁸³ The key identifying assumption underlying this strategy is that, conditional on a municipality's distance from the nearest emigration source, the spatial orientation of its neighbors (i.e., whether on average they are closer or farther from the source) is random. Importantly, the distance to the emigration source need not be exogenous; the strategy merely uses the observed fact that a municipality's distance to the emigration source is correlated with its emigration rates. Returning to the illustration in Figure 14, since the neighbors of municipality A are on average closer to the emigration source, we expect that the actual emigration exposure of municipality A , which is a weighted average of the neighbors' actual emigration, will be greater than that of municipality B , whose neighbors are on average farther from the emigration source. For the same reason, the predicted emigration of the neighbors of municipality A is on average greater than is that of municipality B 's neighbors, and as a result the weighted average of these predicted measures will also be greater.⁸⁴ The latter is then used as an instrument for the former. Notice that this source of variation, stemming from random differences in the geographic orientation of the neighboring population, which are likely to be small, is at risk of suffering from low statistical power; as we discuss below, this poses difficulties in some of our specifications.

In all of our specifications, we control for a municipality-period's own value of predicted emigration \tilde{e}_{it} , computed as in equation (5), allowing the coefficient to vary by period. This is the most straightforward way of controlling for the municipality's own location with respect to the source, as our strategy requires; the identifying variation is in predicted emigration exposure conditional on own predicted emigration, and using \tilde{e}_{it} ensures that both expected values are calculated in the same way. Our baseline specifications restrict the sample to municipalities situated between 50 and 250 kilometers from the sources of emigration.⁸⁵

⁸²The addition of $\varepsilon = 0.0001$ on the left-hand side of equation (5) is needed to ensure that municipality-half decades with no emigration are included in the construction of emigration exposure.

⁸³It is not necessary to adjust standard errors for the use of this generated instrument (Wooldridge 2002, pp. 116–117).

⁸⁴Put differently, municipalities A_H and B_L are predicted to have greater emigration than municipalities A_L and B_H because of their distances to the epicenter. Because A_H and B_H constitute a greater share of neighbors of municipalities A and B , respectively, than do municipalities A_L and B_L , municipality A 's estimated emigration exposure is greater than municipality B 's.

⁸⁵When using the epicenter-based instrument, removing the range restriction strengthens the results substantially. For the frontier-based instrument, the restriction is necessary in order to avoid an undue influence of municipalities actually within the frontier of mass migration. Since an estimate of the migration rates of all municipalities is necessary in constructing the instrument, it is necessary to take a stance on the appropriate distance from the frontier of municipalities already within it, and we set that distance to zero. However, this creates a mass point in the distribution, in which all municipalities have the same estimated migration, limiting the available variation. The range restriction assures that we are looking sufficiently far away so that these municipalities do not exert an excessive influence.

6.3 Hidden Threats to the Validity of the Instrument

The baseline specifications also include controls for the geometry of national borders and for population density. Each of these are crucial in addressing potential risks to the validity of the instrument which are not plainly visible. First, consider a municipality that is located on the contours of Italy—either on the coast or on the land border. For geometric reasons (and because we do not observe municipalities directly on the other side of the land border), any source of emigration within the country will tend to be closer to the municipality’s average neighbors than to the municipality itself.⁸⁶ This regularity will cause a positive correlation between predicted emigration exposure and position along the coast or the land border. If such locations have systematically different emigration potentials, this would amount to an endogeneity problem. We address this threat by controlling for half decade-specific functions of distance to the coast and to the land border.

The second hidden threat comes from variation in population density around a municipality. Consider a case in which the predicted emigration of a municipality is a downward sloping convex function of distance to the source (which in our case, it is). Then a reduction of the population density around the municipality would be associated with a greater average predicted emigration of its neighbor, while keeping own predicted emigration unchanged.⁸⁷ If population density is correlated with unobserved determinants of emigration, this would create endogeneity in the instrument. A straightforward way to address this issue is to control for a measure of population density, for which we use the denominator of the right-hand side of equation (3).

Another threat comes from the fact that the predicted values of municipality i ’s neighbors \tilde{e}_{jt} are in small part based on municipality i ’s own realized rate of emigration, which creates a small possibility of endogeneity in predicted emigration exposure.⁸⁸ This is accounted for by controlling for the municipality’s own lagged predicted emigration \tilde{e}_{it-1} . An alternative way to remove this source of endogeneity altogether is to estimate equation (5) while excluding the source catchment into which a municipality falls. For example, the emigration exposure for municipalities whose nearest emigration epicenter is the district of Corleone is constructed by estimating equation (5) for all municipalities except those for which the nearest epicenter is Corleone and then applying the prediction to all municipalities in the Corleone catchment. For distance

⁸⁶For a simple example, consider a municipality located at the corner of a grid. When the source of emigration is located strictly inside the grid, there exists a positive radius around the corner municipality within which every other municipality is closer to the source than itself. Exceptions to this rule are conceivable in real-life data, but are unlikely to occur systematically.

⁸⁷It is easy to see this in the simple example in which the population is positioned along a line; a proportional expansion of the neighboring population away from the municipality would leave the weight of each municipality unchanged, but due to Jensen’s inequality, the increase in predicted emigration of the municipalities closer to the source will be greater than the reduction in predicted emigration of municipalities farther from the source.

⁸⁸Consider municipality i at a distance d from the source. The predicted rate of emigration for municipalities $j \neq i$ that are at a distance of just above or below d is a weighted average of actual emigration rates that includes that of municipality i .

from the frontier, a municipality’s estimated emigration is constructed by estimating equation (5) excluding a municipality’s own province.⁸⁹

6.4 Results

Table 3 presents the results of our estimation.⁹⁰ Panel A uses an instrument based on distance from emigration epicenters. Column (1) includes no controls beyond those described above, and only period fixed effects. The first-stage F -statistic clearly passes the Staiger and Stock (1997) threshold. The estimated elasticity of a municipality’s own emigration with respect to the portion of lagged neighbors’ emigration driven by variation in their own distance from epicenters is positive (approximately 1) and statistically significant, as is to be expected in the presence of a spatial contagion mechanism. Column (2) adds controls for period-specific functions of latitude, longitude, elevation, exposure to cities,⁹¹ and the local characteristics used in section 5.5. Doing this only slightly attenuates the coefficient on lagged emigration exposure. Columns (3)–(8) add increasingly fine fixed effects at the region, region-period, province, province-period, district, and district-period, respectively. The magnitude of the coefficient is further attenuated, and the precision of the estimate and the strength of the instrument (as measured by the first-stage F -statistic) diminish, particularly when controlling for period-specific regional fixed effects (column 4). As mentioned above, the identifying variation is likely very small, and thus as the control becomes very tight statistical power is eroded. Nevertheless, except for the specifications that control for district fixed effects, in which statistical power is all but lost, the estimated elasticities are in the range 0.5–1 and are at least marginally statistically significant.

⁸⁹The solution is different because the frontier is not a small number of places, each with a well-defined catchment.

⁹⁰Analogous OLS results are in Online Appendix Table B.4. Online Appendix Table B.5 presents results of using a more standard approach of estimating a spatial lag model of the form

$$\log(e_{it}) = \alpha_t + \beta \log(e_{-it-1}) + \gamma \log(e_{it-1}) + \mathbf{x}'_i \delta_t + \varepsilon_{it},$$

instrumenting for lagged emigration exposure using half decade-specific functions of neighbors’ characteristics (e.g., Aidt and Leon-Ablan 2022). This is operationalized by using the binomial maximum likelihood estimator of Appendix A to estimate neighbors’ emigration as a function of all observables (including distance from epicenter) and then constructing an instrument as in equation (4). The results are qualitatively similar to the main results, albeit with far *greater* statistical power. However, although this type of instrument can capture diffusion in the sense that it can document a causal effect of neighbors’ emigration on own emigration, the use of all covariates in constructing the instrument implies that it does not necessarily capture the diffusion of migration coming from the gradual spread of networks over the country. Our preferred approach, on the other hand, by basing identification solely on the orientation of neighboring population relative to the oncoming tide of diffusion, ensures that we capture only diffusion of that type, which is what we are primarily interested in. Put differently, the local average treatment effect arising from the more conventional approach is less helpful in making our argument than is the local average treatment effect coming from our preferred approach.

⁹¹That is, we construct a measure analogous to equation (3), but use city population instead of number of emigrants. The cities in question are Bologna, Catania, Firenze, Genova, Messina, Milano, Napoli, Palermo, Roma, Torino, and Venezia. We include this measure because urban areas are likely to be differentially responsive when they are reached by the expanding tide of emigration. Returning to Figure 14, if municipality A is neighbored by urban areas and municipality B is not, we expect differences in the degree to which they are exposed to emigration by their neighbors.

Panel B of Table 3 repeats the same estimation, but uses the distance to the dynamic frontier rather than the static epicenter to construct the instrument. Since, as shown above, the predictive power of the distance to the epicenter diminishes over time whereas that of the distance to the frontier does not, this is likely to increase the strength of the instrument without violating its exogeneity. As in the regressions that estimate the frontier effect (section 5.4), the sample changes for each period to include only municipalities that had not yet reached the frontier. The results are qualitatively the same as in panel A, but with a meaningful improvement in the strength of the instrument and in statistical power.

In Online Appendix Table B.6, we repeat the analyses of Table 3, but, as described above, we construct the instrument based on estimating equation (5) while excluding a municipality’s own catchment area. Both the estimates and the statistical power are somewhat sensitive, but the qualitative results remain unchanged: estimates that have sufficient statistical power point at an elasticity in the range of 0.5–1.

To be sure, the results of our instrumental variables analysis are not perfectly satisfying in terms of precision or stability, likely due to the limited variation in the instrument, leaving little statistical power under a large enough set of fixed effects. Nevertheless, despite the demanding constraints, the outcomes of this exercise broadly point at an economically significant contagion effect, most likely at an elasticity of 0.5 or more. While we cannot pinpoint the precise reasons behind this effect, it appears that it had to do with physical proximity between populations. Based on our reading of the sociology of the Italian migration (section 2.3) and of other similar movements, we argue that the most plausible explanation is that there existed inter-place diffusion of migration networks.⁹² The evidence on the spatial contagion mechanism is not meant to stand alone as a proof of the diffusion hypothesis. Rather, we view it as a part of a wider body of evidence, together with the results of section 5.

7 Summary of Robustness Checks

The Online Appendix presents a variety of robustness checks for the main results. Online Appendix G repeats the main results incorporating municipalities that are not listed in the earlier emigration statistics volumes, but which may have been included in the data for “Other Municipalities” provided for each district, by allocating this extra emigration equally to all of these unlisted municipalities. Online Appendix H uses 1881

⁹²An alternative explanation could be that emigration in one municipality caused some change—such as in the economy or the culture of the place—and that this change spilled over to neighboring municipalities and caused emigration there. The local effects of emigration in that period is indeed a subject that lacks quantitative evidence. While we suspect that such effects did exist, we find it hard to believe that their spillovers to neighboring communities were sufficiently strong to be a major cause for emigration. Yet even if they were, such spillover effects would have been part of a slightly different version of the diffusion hypothesis, which would have kept all of the important implication of our preferred version.

population as the basis for computing emigration rates. Online Appendix I repeats the results using data on emigration to all destinations instead of only to North America. Among other issues, this addresses the concern that the correlation of emigration from municipalities in the same province may have been inflated by the fact that the emigration-by-destination data (which are not used in this case) are available only at the province level. Online Appendix K repeats the results including municipalities with no emigration in a particular half decade, which are otherwise excluded due to the use of the logarithm of emigration as the main outcome in many analyses. Online Appendix L repeats the results of section 6 using different values of θ to compute lagged emigration exposure and the instrument. Finally, Online Appendix M repeats the estimates of section 6 using increasingly fine geographic fixed effects as in Barsbai et al. (2017). As a rule, the results are not qualitatively affected.

8 Conclusion

In this paper, we introduce and validate the diffusion hypothesis in the context of Italy during the Age of Mass Migration. This hypothesis views the presence or absence of migrant networks as the chief determinant of the timing of the onset of mass migration, but it departs from standard models of migration with social networks by permitting migration to diffuse over space in the country of origin. We formalize this hypothesis by developing an epidemiological model of migration within a spatial network and deriving its testable predictions, and validate it using Italian emigration statistics. The predictions of our model are all borne out in these data and we use an instrumental variables strategy to show that the most likely mechanism to have generated these stylized facts was the key building block of the diffusion hypothesis—spatial contagion. Moreover, our findings strongly suggest that diffusion was not merely one factor among many that affected migration. Instead, diffusion can parsimoniously explain all of the major stylized facts of this migration, which are difficult to rationalize through models that view networks as crucial but do not allow them to span local boundaries.

This is the first paper to include spatial networks within a theoretical model of migration and the first to document in detail patterns of spatial diffusion of migration in the origin and to show that they can be substantial in magnitude. Notably, we do not argue that spatial diffusion is an important feature of all episodes of mass migration. But from the confirmation that it dominated one of the most important episodes of migration in human history, we draw the conclusion that spatial diffusion should be part of the toolbox of migration scholars when considering other cases of mass migration. It may help to explain or

predict them—how they grow, how their geographic origins and destination choices evolve, and how they reach maturity.

Our findings are also important for cases in which migration takes place within networks without exhibiting significant patterns of spatial diffusion. In particular, we show that the effect of migrant networks may be so important that they surpass economic fundamentals in determining the timing of migration. Such a strong role for networks has been suggested in theory, but has not previously been documented empirically. This may have interesting implications for modern migration flows. In particular, it may mean that developing countries that currently produce little migration are not abstaining from it due to economic underdevelopment or liquidity traps, but due to the lack of links to previous migrants. Once those links are formed, migration might suddenly be unleashed. Importantly, unlike in standard push-pull models of migration, such quick surges may occur even without discernible local economic, demographic, or political triggers. It predicts that what may appear to be a continuously rising trend in migration from a source country as a whole may in fact be a series of small regions taking off and quickly reaching a steady high rate of migration (S-shaped local migration trends). Thus, a steady rising trend of migration from a given country may stabilize when the country reaches saturation (as we believe happened in Italy during the first decade of the twentieth century), again without any economic trigger for a change in the trend.

References

- A'Hearn, Brian and Valeria Rueda (2022). "Internal Borders and Population Geography in the Unification of Italy." *Journal of Economic History* Forthcoming.
- A'Hearn, Brian and Anthony J. Venables (2013). "Regional Disparities: Internal Geography and External Trade." In *The Oxford Handbook of the Italian Economy Since Unification*. Gianni Toniolo (ed.). New York: Oxford University Press. Chap. 21, pp. 599–630.
- Abramitzky, Ran and Leah Platt Boustan (2017). "Immigration in American Economic History." *Journal of Economic Literature* 55:4, pp. 1311–1345.
- Aidt, Toke and Gabriel Leon-Ablan (2022). "The Interaction of Structural Factors and Diffusion in Social Unrest: Evidence from the Swing Riots." *British Journal of Political Science* 52, pp. 869–885.
- Aidt, Toke, Gabriel Leon-Ablan, and Max Satchell (2022). "The Social Dynamics of Collective Action: Evidence from the Diffusion of the Swing Riots, 1830–1831." *Journal of Politics* 84:1, pp. 209–225.
- Andrews, Kenneth T. and Charles Seguin (2015). "Group Threat and Policy Change: The Spatial Dynamics of Prohibition Politics, 1890–1919." *American Journal of Sociology* 121:2, pp. 475–510.
- 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, pp. 452–471.
- Baily, Samuel L. (1999). *Immigrants in the Lands of Promise: Italians in Buenos Aires and New York City, 1870–1914*. Cornell University Press.
- Baines, Dudley (1995). *Emigration from Europe 1815–1930*. Cambridge: Cambridge University Press.
- 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, pp. 23–47.
- Banfield, Edward C. (1958). *The Moral Basis of a Backward Society*. Free Press.
- Bansak, Cynthia, Nicole Simpson, and Madeline Zavodny (2020). *The Economics of Immigration*. 2nd ed. New York: Routledge.
- Barde, Robert, Susan B. Carter, and Richard Sutch (2006). "Table Ad106–120: Immigrants, by country of last residence—Europe, 1820–1997." In *Historical Statistics of the United States*. Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright (ed.). Cambridge: Cambridge University Press, pp. 1.560–1.563.
- Barsbai, Toman, Hillel Rapoport, Andreas Steinmayr, and Christoph Trebesch (2017). "The Effect of Labor Migration on the Diffusion of Democracy: Evidence from a Former Soviet Republic." *American Economic Journal: Applied Economics* 9:3, pp. 36–69.
- Barthel, Fabian and Eric Neumayer (2015). "Spatial Dependence in Asylum Migration." *Journal of Ethnic and Migration Studies* 41:7, pp. 1131–1151.
- Barton, Josef J. (1975). *Peasants and Strangers: Italians, Rumanians, and Slovaks in an American City, 1890–1950*. Cambridge: Harvard University Press.
- Bass, Frank M. (1969). "A New Product Growth for Model Consumer Durables." *Management Science* 15:5, pp. 215–227.
- Beaman, Lori (2012). "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the US." *Review of Economic Studies* 79, pp. 128–161.
- Becker, Sascha O., Volker Lindenthal, Sharun Mukand, and Fabian Waldinger (2023). "Persecution and Escape: Professional Networks and High-Skilled Emigration from Nazi Germany." *American Economic Journal: Applied Economics* Forthcoming.
- Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga (2016). "A Practitioner's Guide to Gravity Models of International Migration." *World Economy* 39:4, pp. 496–512.
- Beine, Michel, Frédéric Docquier, and Çağlar Özden (2011). "Diasporas." *Journal of Development Economics* 95, pp. 30–41.
- Bell, Rudolph M. (1979). *Fate and Honor, Family and Village: Demographic and Cultural Change in Rural Italy since 1800*. Chicago: University of Chicago Press.
- Bernoulli, Daniel (1776). "Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir." In *Mémoires de mathématique et de physique*. Academie Royale des Sciences.
- Bertoli, Simone and Jesús Fernández-Huertas Moraga (2015). "The Size of the Cliff at the Border." *Regional Science and Urban Economics* 51, pp. 1–6.
- Boberg-Fazlić, Nina, Markus Lampe, and Paul Sharpe (2021). "The Sleeping Giant Who Left for America: The Determinants and Impact of Danish Emigration During the Age of Mass Migration." EHES Working Paper 213.

- Bodnar, John, Roger Simon, and Michael P. Weber (1982). *Lives of Their Own: Blacks, Italians, and Poles in Pittsburgh, 1900–1960*. Urbana and Chicago: University of Illinois Press.
- Borjas, George J. (2014). *Immigration Economics*. Cambridge: Harvard University Press.
- Brandenburg, Broughton (1904). *Imported Americans: The Story of the Experiences of a Disguised American and His Wife Studying the Immigration Question*. New York: Frederick A. Stokes Company.
- Briggs, John Walker (1978). *An Italian Passage: Immigrants to Three American Cities, 1890–1930*. New Haven: Yale University Press.
- Buggle, Johannes, Thierry Mayer, Seyhun Orcan Sakalli, and Mathias Thoenig (2023). “The Refugee’s Dilemma: Evidence from Jewish Migration out of Nazi Germany.” *Quarterly Journal of Economics* 138:2, pp. 1273–1345.
- Burchardi, Konrad B., Thomas Chaney, and Tarek A. Hassan (2019). “Migrants, Ancestors, and Foreign Investments.” *Review of Economic Studies* 86:4, pp. 1448–1486.
- Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo (2016). “Understanding Booms and Busts in Housing Markets.” *Journal of Political Economy* 124:4, pp. 1088–1147.
- Chay, Kenneth and Kaivan Munshi (2015). “Black Networks After Emancipation: Evidence from Reconstruction and the Great Migration.” Mimeo., Yale University.
- Chiquiar, Daniel and Gordon H. Hanson (2005). “International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States.” *Journal of Political Economy* 113:2, pp. 239–281.
- Ciccarelli, Carlo and Stefano Fenoaltea (2013). “Through the Magnifying Glass: Provincial Aspects of Industrial Growth in Post-Unification Italy.” *Economic History Review* 66:1, pp. 57–85.
- Ciccarelli, Carlo and Peter Groote (2017). “Railway Endowment in Italy’s Provinces, 1839–1913.” *Rivista di Storia Economica* 33, pp. 45–88.
- Cinèl, Dino (1982). *From Italy to San Francisco: The Immigrant Experience*. Stanford: Stanford University Press.
- Cohn, Raymond L. (2009). *Mass Migration Under Sail: European Immigration to the Antebellum United States*. New York: Cambridge University Press.
- Coleman, James S. (1988). “Social Capital in the Creation of Human Capital.” *American Journal of Sociology* 94, S95–S120.
- Dai, Ruochen, Dilip Mookherjee, Kaivan Munshi, and Xiaobo Zhang (2020). “The Community Origins of Private Enterprise in China.” Mimeo., Yale University.
- Del Panta, Lorenzo (1997). “Infant and Child Mortality in Italy, Eighteenth to Twentieth Century: Long-term Trends and Territorial Differences.” In *Infant and Child Mortality in the Past*. Alain Bideau, Bertrand Desjardins, and Héctor Pérez Brignoli (ed.). New York: Oxford University Press. Chap. 1, pp. 7–21.
- Docquier, Frédéric, Giovanni Peri, and Ilse Ruyssen (2014). “The Cross-Country Determinants of Potential and Actual Migration.” *International Migration Review* 48:1, S37–S99.
- Duncan, Otis Dudley and Beverly Duncan (1955). “A Methodological Analysis of Segregation Indexes.” *American Sociological Review* 20:2, pp. 210–217.
- Eichenbaum, Martin, Sergio Rebelo, and Mathias Trabandt (2021). “The Macroeconomics of Epidemics.” *Review of Financial Studies* 34:11, pp. 5149–5187.
- Faini, Riccardo and Alessandra Venturini (1994). “Italian Emigration in the Pre-War Period.” In *Migration and the International Labor Market 1850–1939*. Timothy J. Hatton and Jeffrey G. Williamson (ed.). London: Routledge, pp. 72–90.
- Federico, Giovanni, Alessandro Nuvolari, and Michelangelo Vasta (2019). “The Origins of the Italian Regional Divide: Evidence from Real Wages, 1861–1913.” *Journal of Economic History* 79:1, pp. 63–98.
- Fernández-Huertas Moraga, Jesús (2011). “New Evidence on Emigrant Selection.” *Review of Economics and Statistics* 93:1, pp. 72–96.
- (2013). “Understanding Different Migrant Selection Patterns in Rural and Urban Mexico.” *Journal of Development Economics* 103, pp. 182–201.
- Fernández-Sánchez, Martín (2021). “Mass Emigration and Human Capital over a Century: Evidence from the Galician Diaspora.” Mimeo., LISER.
- Foerster, Robert F. (1919). *The Italian Emigration of Our Times*. New York: Russell & Russell.
- Fontana, Nicola, Marco Manacorda, Gianluca Russo, and Marco Tabellini (2021). “Emigration and Long-Run Economic Development: The Effects of the Italian Mass Migration.” Mimeo., HBS.
- Franck, Raphaël and Oded Galor (2022). “Technology-Skill Complementarity in Early Phases of Industrialization.” *Economic Journal* 132:642, pp. 618–643.
- Gabaccia, Donna (1984a). *From Sicily to Elizabeth Street: Housing and Social Change among Italian Immigrants, 1880–1930*. SUNY Press.

- Gabaccia, Donna (1984b). "Kinship, Culture, and Migration: A Sicilian Example." *Journal of American Ethnic History* 3:2, pp. 39–53.
- (1988). "The Transplanted: Women and Family in Immigrant America." *Social Science History* 12:3, pp. 243–253.
- Gomellini, Matteo and Cormac Ó Gráda (2013). "Migrations." In *The Oxford Handbook of the Italian Economy Since Unification*. Gianni Toniolo (ed.). New York: Oxford University Press. Chap. 10, pp. 271–302.
- Gould, John D. (1980). "European Inter-Continental Emigration: The Role of 'Diffusion' and 'Feedback'." *Journal of European Economic History* 9:2, pp. 267–315.
- Gray, Rowena, Gaia Narciso, and Gaspare Tortorici (2019). "Globalization, Agricultural Markets and Mass Migration: Italy, 1881–1912." *Explorations in Economic History* 74, 101276.
- Handlin, Oscar (1951). *The Uprooted: The Epic Story of the Great Migrations that Made the American People*. Boston: Little Brown.
- Hatton, Timothy J. and Jeffrey G. Williamson (1998). *The Age of Mass Migration: Causes and Economic Impact*. New York: Oxford University Press.
- Hirschman, Albert O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations and States*. Cambridge: Harvard University Press.
- Ibarraran, Pablo and Darren Lubotsky (2007). "Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census." In *Mexican Immigration to the United States*. George J. Borjas (ed.). University of Chicago Press. Chap. 5, pp. 159–192.
- Iorizzo, Luciano John (1966). *Italian Immigration and the Impact of the Padrone System*. Syracuse: Syracuse University Press.
- Iuzzolino, Giovanni, Guido Pellegrini, and Gianfranco Viesti (2013). "Regional Convergence." In *The Oxford Handbook of the Italian Economy Since Unification*. Gianni Toniolo (ed.). New York: Oxford University Press. Chap. 20, pp. 571–598.
- Jackson, Matthew O., Tomas Rodriguez-Barraquer, and Xu Tan (2012). "Social Capital and Social Quilts: Network Patterns of Favor Exchange." *American Economic Review* 102:5, pp. 1857–1897.
- Jovanovic, Boyan and Saul Lach (1989). "Entry, Exit, and Diffusion with Learning by Doing." *American Economic Review* 79:4, pp. 690–699.
- Karadja, Mounir and Erik Prawitz (2019). "Exit, Voice and Political Change: Evidence from Swedish Mass Migration to the United States." *Journal of Political Economy* 127:4, pp. 1864–1925.
- Kelley, Ninette and Michael Trebilcock (1998). *The Making of the Mosaic: A History of Canadian Immigration Policy*. Toronto: University of Toronto Press.
- Kermack, William O. and Anderson G. McKendrick (1927). "A Contribution to the Mathematical Theory of Epidemics." *Proceedings of the Royal Society of London, Series A, Containing Papers of a Mathematical and Physical Character* 115:772, pp. 700–721.
- Kerr, William R. and Martin Mandorff (2023). "Social Networks, Ethnicity, and Entrepreneurship." *Journal of Human Resources* 58:1, pp. 183–220.
- Koren, John (1897). *The Padrone System and Padrone Banks*. Washington: US Government Printing Office.
- Lecce, Giampaolo, Laura Ogliari, and Tommaso Orlando (2022). "Resistance to Institutions and Cultural Distance: Brigandage in Post-Unification Italy." *Journal of Economic Growth* 27, pp. 453–483.
- Lowell, Briant L. (1987). *Scandinavian Exodus: Demography and Social Development of 19th Century Rural Communities*. Boulder: Westview.
- MacDonald, John S. (1963). "Agricultural Organization, Migration and Labour Militancy in Rural Italy." *Economic History Review* 16:1, pp. 61–75.
- MacDonald, John S. and Leatrice MacDonald (1964). "Institutional Economics and Rural Development: Two Italian Types." *Human Organization* 23:2, pp. 113–118.
- Mahajan, Parag and Dean Yang (2020). "Taken By Storm: Hurricanes, Migrant Networks, and US Immigration." *American Economic Journal: Applied Economics* 12:2, pp. 250–277.
- Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J. Edward Taylor (1993). "Theories of International Migration: A Review and Appraisal." *Population and Development Review* 19:3, pp. 431–466.
- 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, pp. 1–24.
- (2010). "Self-Selection Patterns in Mexico-US Migration: The Role of Migration Networks." *Review of Economics and Statistics* 92:4, pp. 811–821.

- Mokyr, Joel (1983). *Why Ireland Starved: An Analytical and Quantitative Study of Irish Poverty, 1800–1851*. Boston: George Allen and Unwin.
- Mokyr, Joel and Cormac Ó Gráda (1982). “Emigration and Poverty in Prefamine Ireland.” *Explorations in Economic History* 19, pp. 360–384.
- Moretti, Enrico (1999). “Social Networks and Migrations: Italy 1876–1913.” *International Migration Review* 33:3, pp. 640–657.
- Mormino, Gray Ross (1986 [2002]). *Immigrants on the Hill: Italian-Americans in St. Louis, 1882–1982*. Columbia: University of Missouri Press.
- Moya, Jose C. (1998). *Cousins and Strangers: Spanish Immigrants in Buenos Aires, 1850–1930*. Berkeley: University of California Press.
- Munshi, Kaivan (2003). “Networks in the Modern Economy: Mexican Migrants in the US Labor Market.” *Quarterly Journal of Economics* 118:2, pp. 549–599.
- (2011). “Strength in Numbers: Networks as a Solution to Occupational Traps.” *Review of Economic Studies* 78, pp. 1069–1101.
- (2014). “Community Networks and the Process of Development.” *Journal of Economic Perspectives* 28:4, pp. 49–76.
- (2020). “Social Networks and Migration.” *Annual Review of Economics* 12, pp. 503–524.
- Munshi, Kaivan and Mark Rosenzweig (2016). “Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap.” *American Economic Review* 106:1, pp. 46–98.
- Nelli, Humbert S. (1964). “The Italian Padrone System in the United States.” *Labor History* 5:2, pp. 153–167.
- (1967). “Italians in Urban America: A Study in Ethnic Adjustment.” *International Migration Digest* 1:3, pp. 38–55.
- Orrenius, Pia M. and Madeline Zavodny (2005). “Self-Selection Among Undocumented Immigrants from Mexico.” *Journal of Development Economics* 78:1, pp. 215–240.
- Park, Robert E. and Herbert A. Miller (1921). *Old World Traits Transplanted*. New York: Harper & Brothers.
- Peck, Gunther (2000). *Reinventing Free Labor: Padrones and Immigrant Workers in the North American West, 1880–1930*. Cambridge: Cambridge University Press.
- Pizzo, Anthony P. (1981). “The Italian Heritage in Tampa.” In *Little Italies in North America*. Robert F. Harney and Vincenza Scarpaci (ed.). Multicultural History Society of Ontario.
- Sánchez Alonso, Blanca (2015). “Reflexiones sobre la emigración andaluza a América en los siglos XIX y XX.” In *Andalucía, España, Las Indias Pasión por la Historia: Homenaje a Antonie Miguel Bernal*. C. Martínez Shaw, P. Tedde de Lorca, and S. Tinoco (ed.). Madrid: Marcial Pons, pp. 205–224.
- Silverman, Sydel F. (1968). “Agricultural Organization, Social Structure, and Values in Italy: Amoral Familism Reconsidered.” *American Anthropologist* 70:1, pp. 1–20.
- Sjaastad, Larry A. (1962). “The Costs and Returns of Human Migration.” *Journal of Political Economy* 70:5, pp. 80–93.
- Spitzer, Yannay (2021). “Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914.” Mimeo., Hebrew University of Jerusalem.
- Spitzer, Yannay, Gaspare Tortorici, and Ariell Zimran (2022). “International Migration Responses Modern Europe’s Most Destructive Earthquake: Messina and Reggio Calabria, 1908.” NBER Working Paper 27506.
- 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, pp. 226–247.
- Spolaore, Enrico and Romain Wacziarg (2009). “The Diffusion of Development.” *Quarterly Journal of Economics* 124:2, pp. 469–529.
- Staiger, Douglas and James H. Stock (1997). “Instrumental Variables Regression With Weak Instruments.” *Econometrica* 65:3, pp. 557–586.
- Stuart, Brian A. and Evan J. Taylor (2021). “Migration Networks and Location Decisions: Evidence from US Mass Migration.” *American Economic Journal: Applied Economics* 13:3, pp. 134–175.
- Sturino, Franc (1990). *Forging the Chain: A Case Study of Italian Migration to North America, 1880–1930*. Toronto: Multicultural History Society of Ontario.
- Thistlethwaite, Frank (1960 [1991]). “Migration from Europe Overseas in the Nineteenth and Twentieth Centuries.” In *A Century of European Migrations, 1830–1930*. R. J. Vecoli and S. M. Sinke (ed.). Urbana: University of Illinois Press. Chap. 1, pp. 17–57.
- Todaro, Michael P. (1969). “A Model of Labor Migration and Urban Unemployment in Less Developed Countries.” *American Economic Review* 59:1, pp. 138–148.

- US Congress (1911a). *Reports of the Immigration Commission: Emigration Conditions in Europe*. Vol. 4. Washington, DC, 61st Congress, 3rd Session, Senate Document No. 748: Government Printing Office.
- (1911b). *Reports of the Immigration Commission: Statistical Review of Immigration 1820–1910—Distribution of Immigrants 1850–1900*. Vol. 3. Washington, DC, 61st Congress, 3rd Session, Senate Document No. 756: Government Printing Office.
- Vecchi, Giovanni (2011). *In ricchezza e in povertà. Il benessere degli italiani dall'Unità ad oggi*. Bologna: Il Mulino.
- Vecoli, Rudolph J. (1964). “Contadini in Chicago: A Critique of the Uprooted.” *Journal of American History* 51:3, pp. 404–417.
- (1983). “The Formation of Chicago’s ‘Little Italies’.” *Journal of American Ethnic History* 2:2, pp. 5–20.
- Ward, Zachary (2017). “Birds of Passage: Return Migration, Self-Selection and Immigration Quotas.” *Explorations in Economic History* 64, pp. 37–52.
- Woodruff, Christopher and Rene Zeteno (2007). “Migration Networks and Microenterprises in Mexico.” *Journal of Development Economics* 82:2, pp. 509–528.
- Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.
- Yans-McLaughlin, Virginia (1975). *Italian Women and Work: Experience and Perception*. Center for the Historical Study of Societies, State University of New York at Binghamton.
- (1977). *Family and Community: Italian Immigrants in Buffalo, 1880–1930*. Urbana and Chicago: University of Illinois Press.
- Zeng, Donglin, Lu Mao, and D. Y. Lin (2016). “Maximum Likelihood Estimation for Semiparametric Transformation Models with Interval-Censored Data.” *Biometrika* 103:2, pp. 253–271.
- Zucchi, John E. (1985). “Italian Hometown Settlements and the Development of an Italian Community in Toronto, 1875–1935.” In *Gathering Place: Peoples and Neighbourhoods of Toronto, 1834–1945*. Robert F. Harney (ed.). Toronto: Multicultural History Society of Ontario.

Tables

Table 1: Survival time regressions

<i>Variables</i>	(1)	(2)	(3)
Distance to North American epicenter	0.583 ^a (0.054)	0.655 ^a (0.052)	0.562 ^a (0.064)
Observations	284	284	284
Broad region FE	No	Yes	Yes
Controls	No	No	Yes

Significance levels: ^a $p < 0.01$, ^b $p < 0.05$, ^c $p < 0.1$

Notes: This table presents estimated hazard ratios for semiparametric Cox proportional hazard models for the timing of entry into the frontier of mass migration to North America. Hypothesis testing is relative to a null hypothesis of a hazard ratio of 1. All variables are standardized to have mean 0 and standard deviation 1. The unit of observation is a district. The data run from the 1876–1880 half decade to the 1916–1920 half decade. The date of entering the mass migration frontier is intervalled by half decade. Broad region fixed effects are for the center and south (with the north as the excluded category). Controls are a district's share of agricultural employment, share of industrial employment, literacy rate, fraction under age 15, birth rate, and death rate, all in 1881, as well as its mean elevation, mutual aid society members per capita in 1878, and postal savings deposits per capita in 1886.

Table 2: Survival time regressions

<i>Variables</i>	(1)	(2)	(3)	(4)	(5)	(6)
Distance to epicenter	0.524 ^a (0.039)				0.470 ^a (0.057)	0.467 ^a (0.059)
Ag. labor share		1.165 ^b (0.075)	1.076 (0.067)	1.064 (0.068)	1.089 (0.077)	1.089 (0.077)
Literacy rate		1.283 ^a (0.073)	1.285 ^a (0.073)	1.279 ^a (0.088)	0.774 ^a (0.074)	0.788 ^c (0.101)
Share Urban 1881		0.792 ^a (0.069)	0.812 ^b (0.077)	0.828 ^c (0.080)	0.795 ^b (0.074)	0.793 ^b (0.076)
Fraction under 15			1.295 ^a (0.076)	1.306 ^a (0.087)	1.200 ^a (0.080)	1.204 ^a (0.083)
Birth Rate (1881)		0.912 (0.071)				
Death Rate (1881)				0.963 (0.069)	0.860 ^c (0.067)	0.857 ^b (0.067)
Mutual Aid Members per capita (1878)				1.063 (0.094)	1.100 (0.094)	1.108 (0.099)
log(Postal Savings Deposits per capita) (1886)				0.927 (0.069)	0.943 (0.065)	0.932 (0.083)
Observations	284	284	284	284	284	284
Broad region FE	No	No	No	No	No	Yes

Significance levels: ^a p<0.01, ^b p<0.05, ^c p<0.1

Notes: This table presents estimated hazard ratios for semiparametric Cox proportional hazard models for the timing of entry into the frontier of mass migration to any destination. Hypothesis testing is relative to a null hypothesis of a hazard ratio of 1. All variables are standardized to have mean 0 and standard deviation 1. The unit of observation is a district. The data run from the 1876–1880 half decade to the 1916–1920 half decade. The date of entering the mass migration frontier is intervalled by half decade. Broad region fixed effects are for the center and south (with the north as the excluded category).

Table 3: Spatial contagion results

<i>Variables</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: Epicenter-based instrument</i>								
Lagged Emigration Exposure	0.956 ^a (0.104)	0.878 ^a (0.122)	0.657 ^a (0.177)	0.641 ^a (0.217)	0.602 ^a (0.151)	0.475 ^c (0.244)	0.977 ^b (0.415)	0.176 (0.448)
Observations	31,463	31,463	31,463	31,463	31,463	31,462	31,463	31,427
F-statistic	30.71	42.75	19.33	12.07	21.57	8.552	23.50	6.559
<i>Panel B: Frontier-based instrument</i>								
Lagged Emigration Exposure	0.942 ^a (0.136)	0.881 ^a (0.157)	0.634 ^a (0.160)	0.749 ^a (0.166)	0.666 ^a (0.176)	0.632 ^a (0.162)	0.238 (0.588)	0.418 ^c (0.232)
Observations	11,207	11,206	11,206	11,205	11,206	11,196	11,206	11,170
F-statistic	64.51	55.20	60.20	46.14	49.19	44.85	19.58	28
Additional FE	None	None	C	CT	P	PT	D	DT
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Significance levels: ^a p<0.01, ^b p<0.05, ^c p<0.1

Notes: Panel A uses instruments constructed on the basis of a municipality's distance to the nearest epicenter of mass migration. Panel B uses instruments constructed on the basis of distance to the frontier of mass migration. Sample limited to municipality-half decades between 50 and 250km of the migration source (i.e., the epicenter or frontier). Standard errors clustered at the district level. All specifications include at least half-decade fixed effects and control for half decade-specific functions of own predicted lagged emigration based on distance from the emigration source, local population, distance to coast, and distance to the European frontier. Dependent variable is the log of the emigration rate to North America. Unit of observation is a municipality-half decade. Controls include half decade-specific functions of latitude, longitude, elevation, agricultural employment share, industrial employment share, literacy rate, fraction under age 15, distance to railroad, birth rate, death rate, mutual aid society membership per capita, and log postal savings deposits per capita. C denotes region (compartimento)-level fixed effects. P denotes province-level fixed effects. D denotes district-level fixed effects. CT, PT, and DT denote region-time, province-time, and district-time fixed effects.

Figures

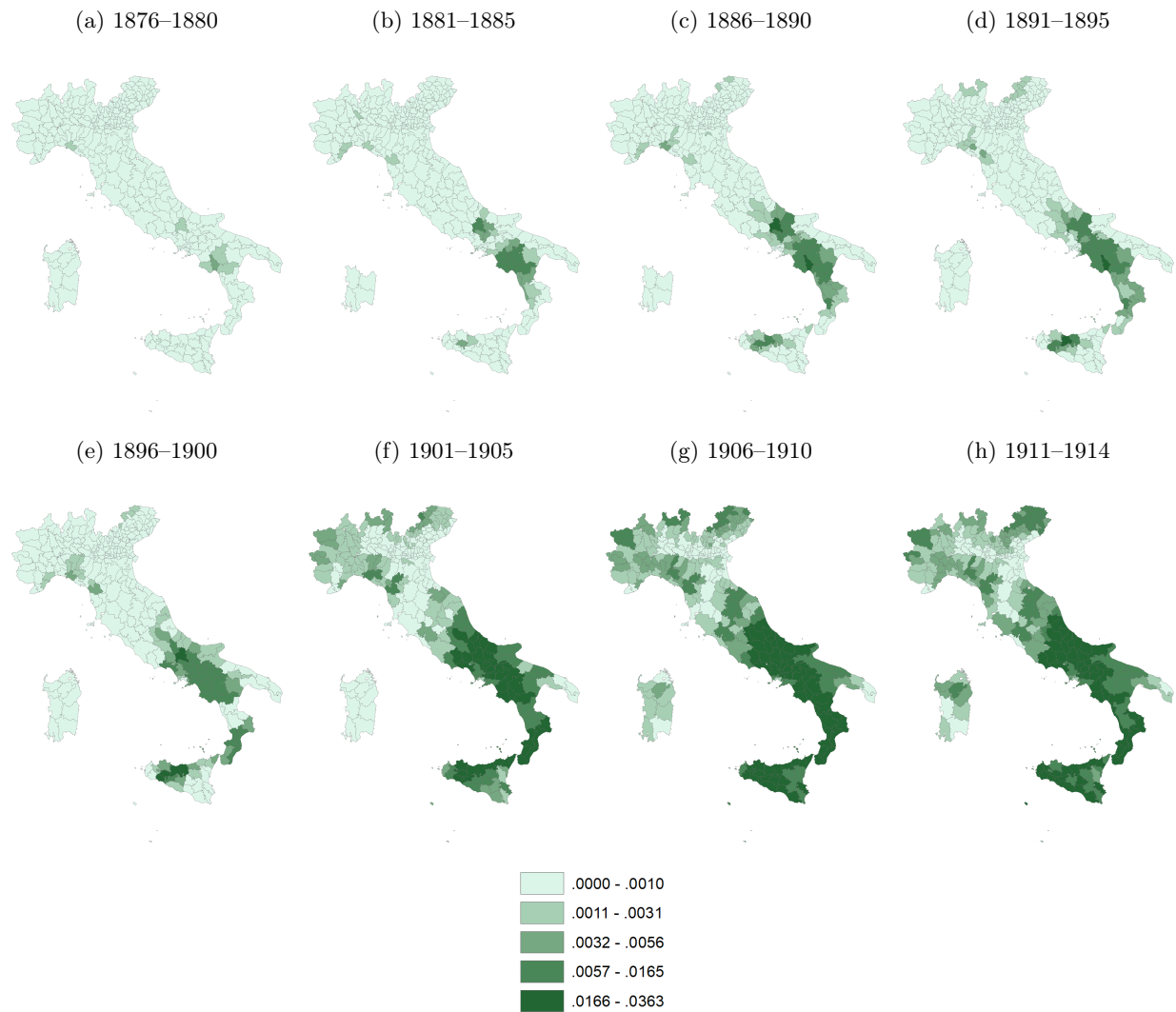


Figure 1: District-level emigration rates to North America

Note: Each panel presents a district's average annual emigration rate to North America in the period in question. Scale is based on quintiles of emigration rates in 1911-1914.

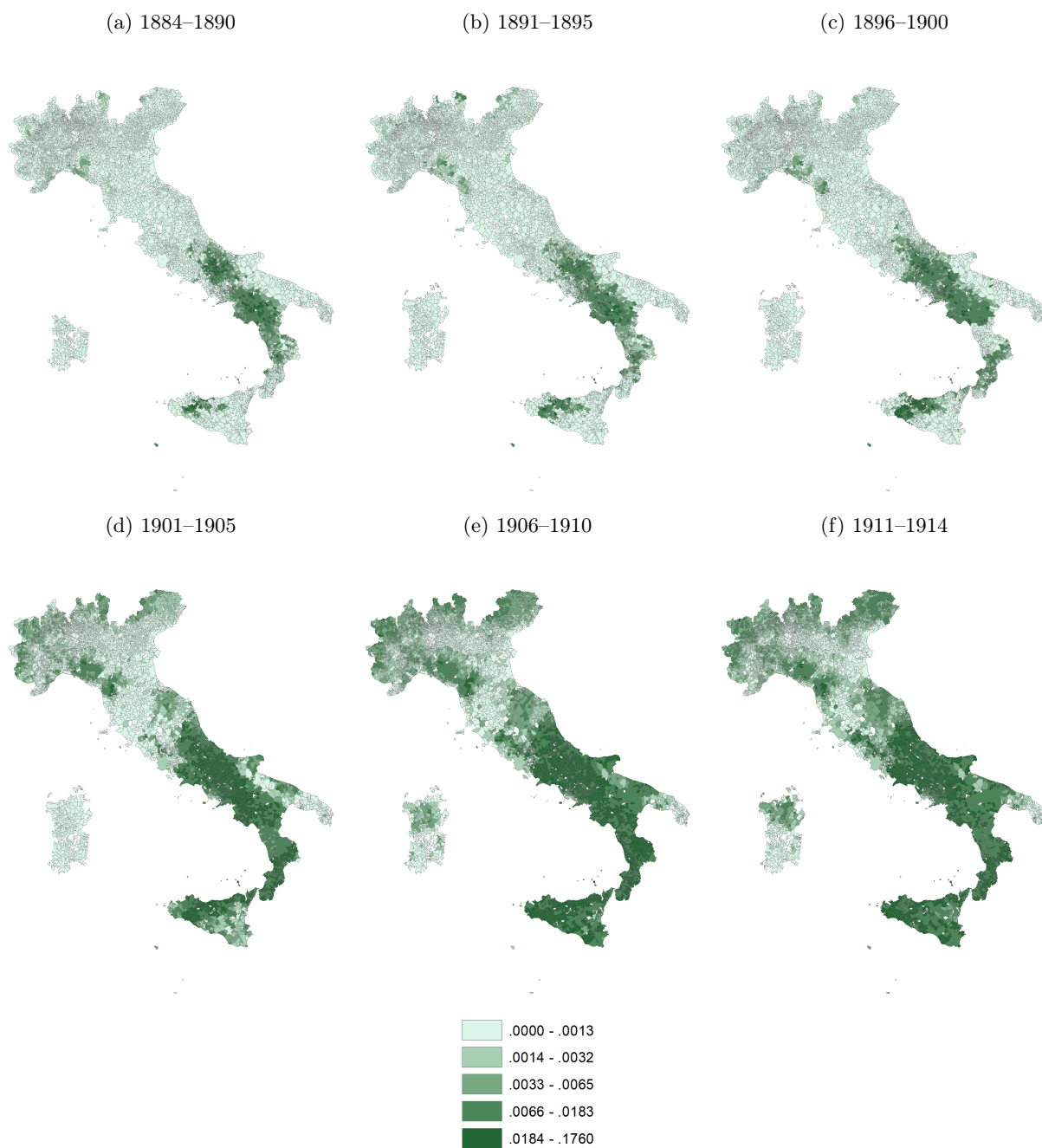


Figure 2: Municipality-level emigration rates to North America

Note: Each panel presents a municipality's average annual emigration rate to North America in the period in question. Scale is based on quintiles of emigration rates in 1911-1914.

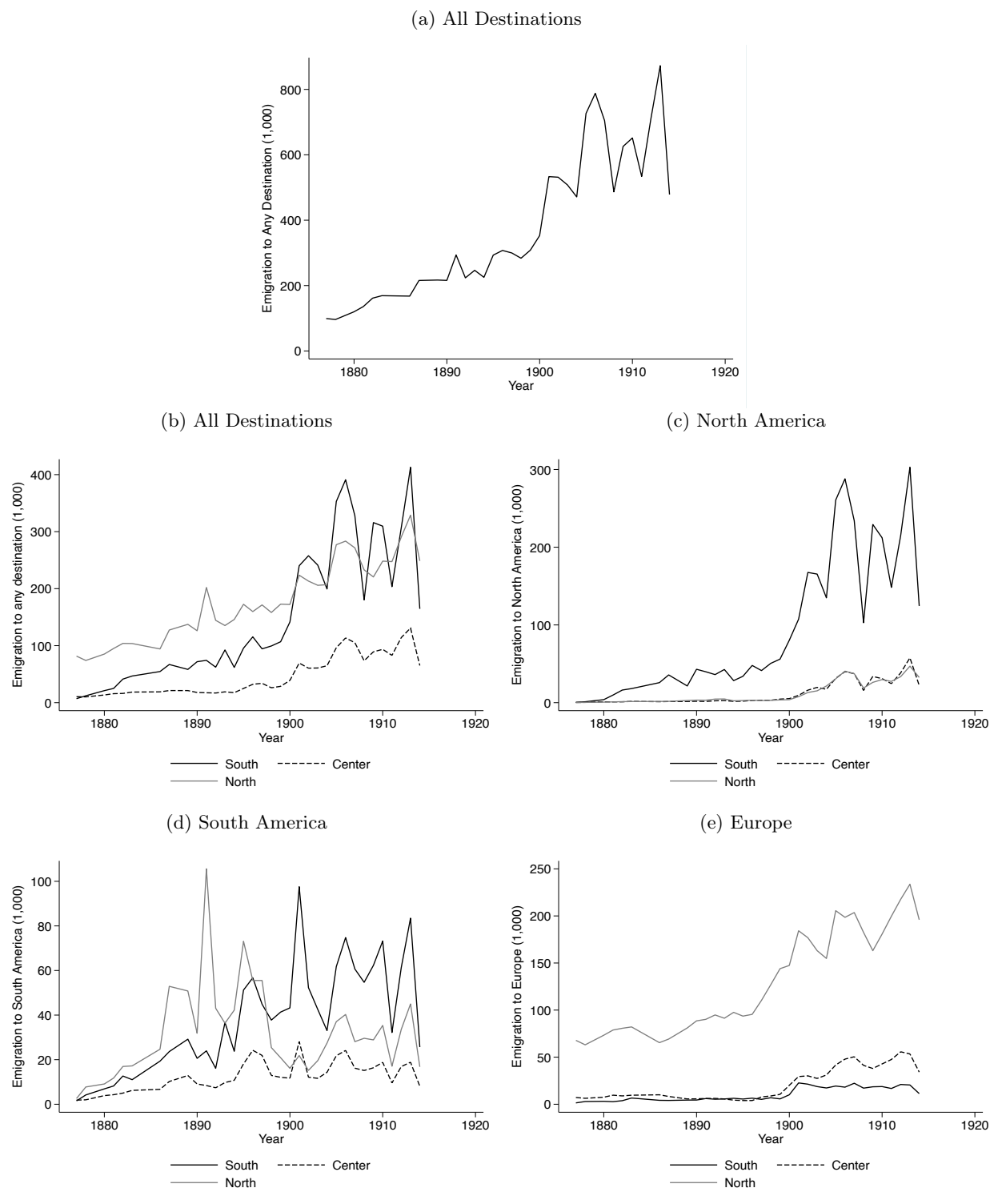


Figure 3: Emigration by origin and destination, 1876–1914

Note: These figures are based on our province-by-destination data. South includes the regions of Abruzzo, Campania, Puglia, Basilicata, Calabria, Sicilia, and Sardinia. Center includes the regions of Liguria, Toscana, Marche, Umbria, and Latium. North includes the regions of Piemonte, Lombardia, Veneto, and Emilia Romagna.

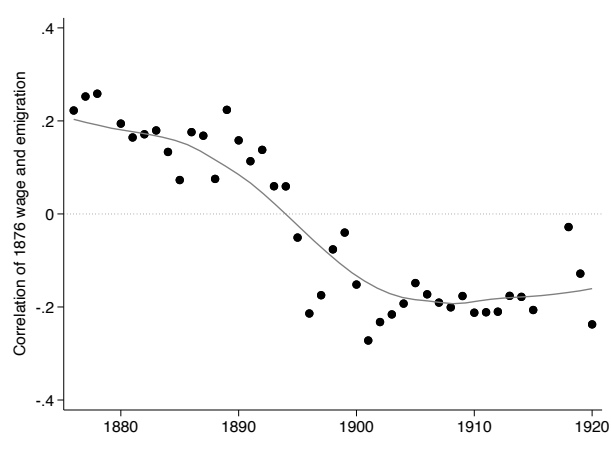


Figure 4: Correlation of province-level emigration and 1876 wages

Source: Real wage data are from Federico, Nuvolari, and Vasta (2019). The source of provincial emigration rates is described in text.

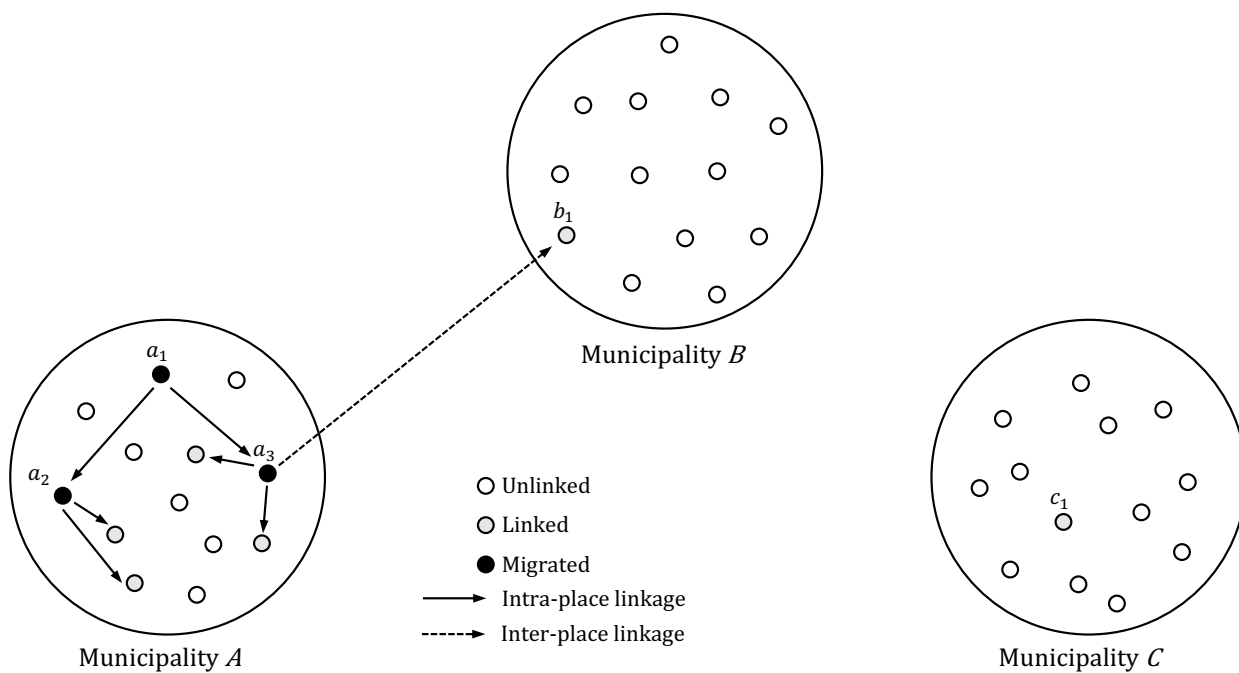


Figure 5: Intra- and inter-place transmission of the migration technology

Note: See explanation in section 3.2. Arrows indicate the direction of the diffusion of the migration “technology,” not the direction of linkage.

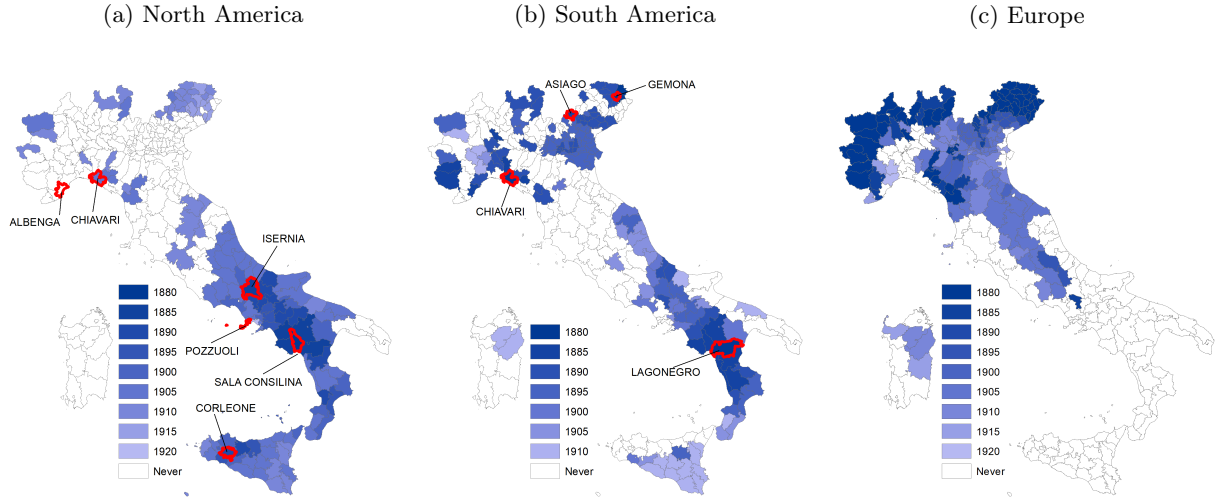


Figure 6: Epicenters and frontiers of mass migration by destination

Note: Districts are shaded according to the half decade in which they first achieved an average annual emigration rate to the listed destination of at least 5 per thousand. Darker districts entered the frontier earlier. In panels (a) and (b), the districts labeled and highlighted with a bold outline are epicenter districts. For migration to North America, the epicenters are Albenga and Chiavari in Liguria, Isernia in Abruzzi e Molise, Pozzuoli and Salerno in Campania, and Corleone in Sicily. For migration to South America, the epicenters are Albenga in Liguria, Asiago and Gemona in Veneto, and Lagonegro in Basilicata. We define the epicenter of migration to Europe to be the European land border.

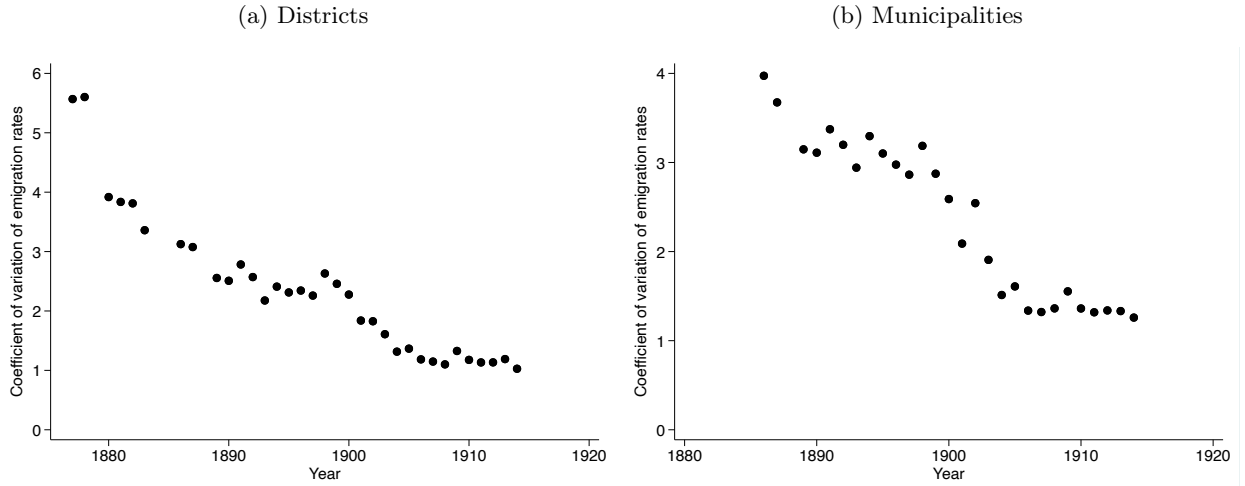


Figure 7: σ -convergence in emigration rates to North America

Note: Each point represents the coefficient of variation in emigration rates to North America in a particular year.

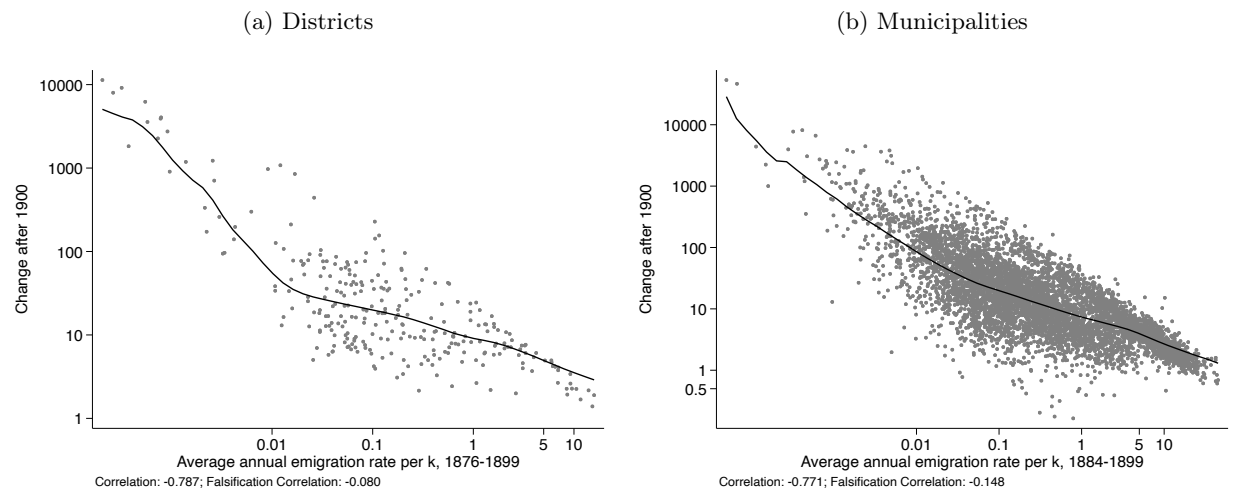


Figure 8: β -convergence in emigration rates to North America

Note: Each point represents a municipality or district. The x -axis is the average annual emigration rate for a district for 1876–1899 or a municipality for 1884–1899 on a log scale. The y -axis is the ratio of the average emigration rate before and after 1900, also on a log scale. The falsification correlation is the correlation of the change in emigration and emigration after 1900; that it is not positive indicates that the negative relationship shown in the graphs is unlikely to be spurious, as explained in Online Appendix J.

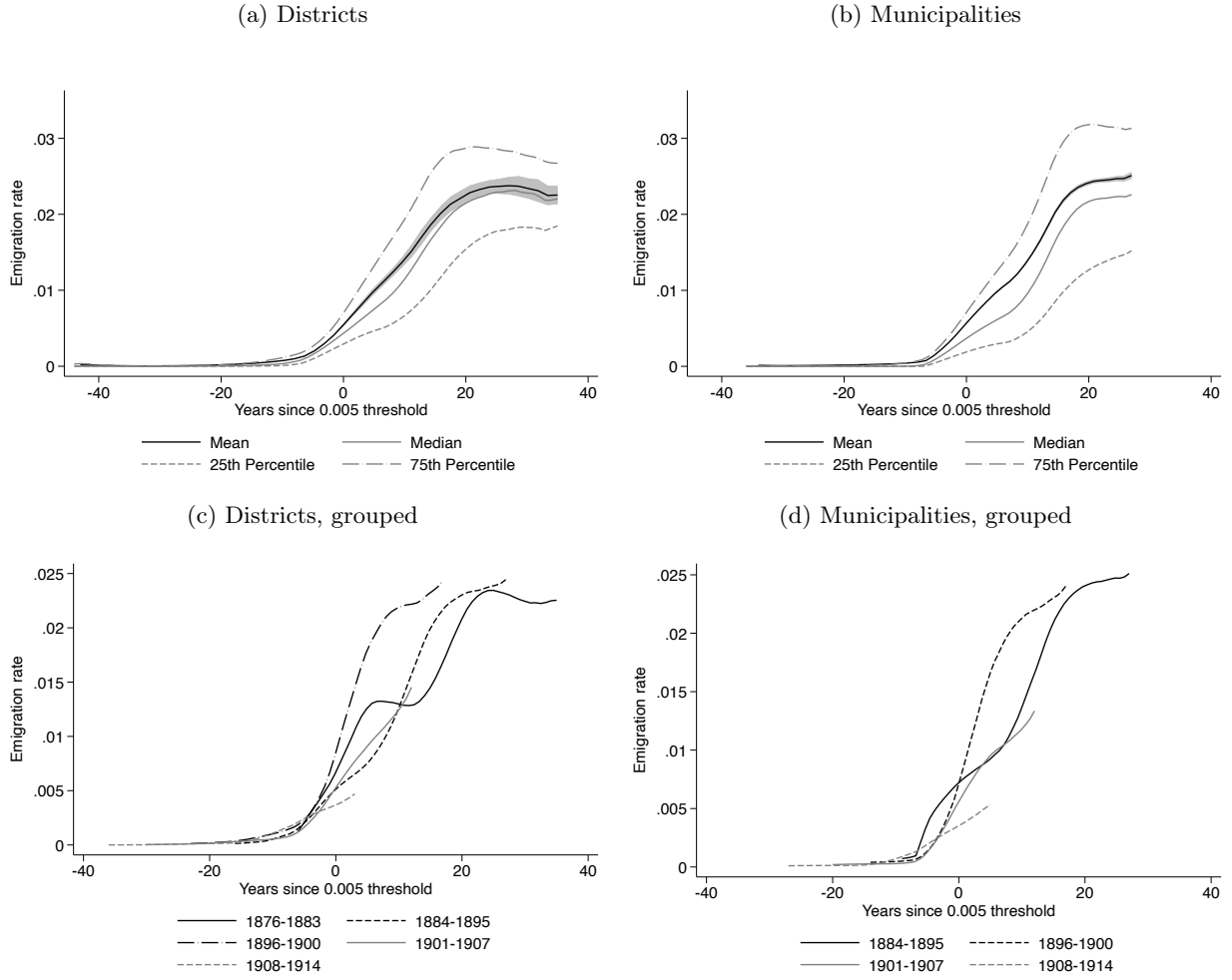


Figure 9: S-shaped time series of migration to North America

Note: Panels (a) and (b) plot a non-parametric regression (the mean), as well as quartiles of emigration rates to North America against time, normalized so that year 0 is the first year in which a place had an emigration rate of at least 5 per thousand. Shaded areas are 95-percent confidence intervals for the mean. Panels (c) and (d) are the same as (a) and (b) but divide areas according to the half decade in which they crossed the threshold.

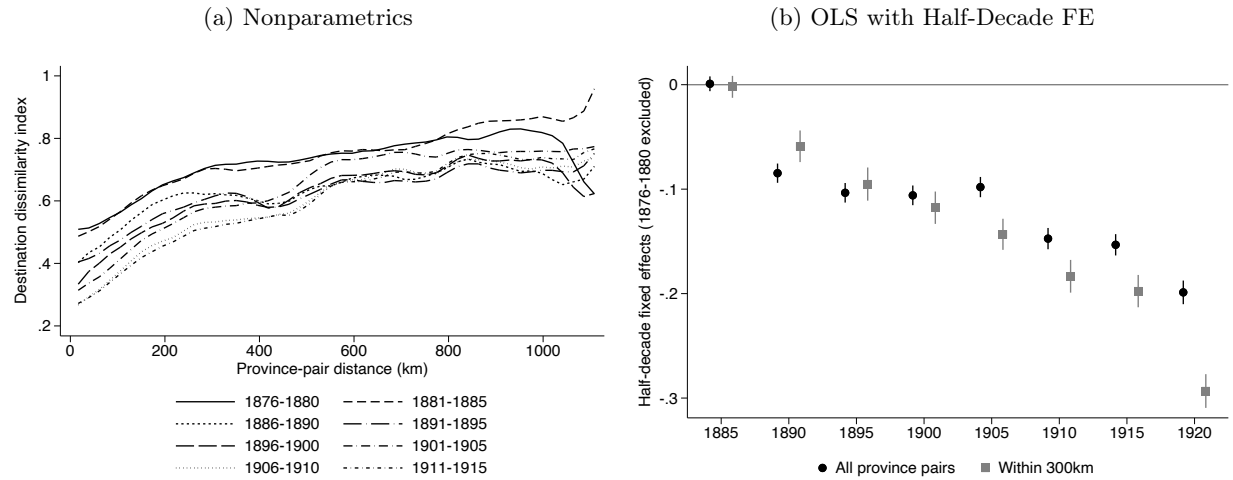


Figure 10: Destination dissimilarity by distance and half decade

Note: Panel (a) plots non-parametric regressions for each half decade of the dissimilarity index between two provinces' emigration and the distance between them. Panel (b) plots half-decade fixed effects from a regression of dissimilarity on province-pair distance and these fixed effects, excluding the 1876–1880 half decade.

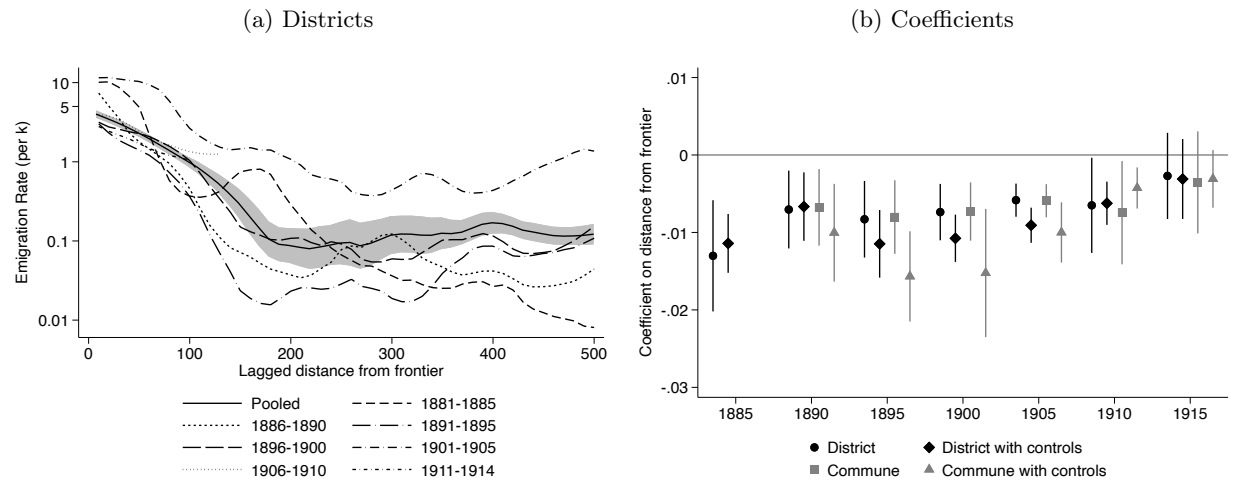


Figure 11: Emigration rates to North America by distance to the mass migration frontier (km)

Note: Panel (a) presents non-parametric regressions of the log of average annual migration rates for the whole sample and for each half decade on the distance from a district that had ever achieved an average annual migration rate of at least 5 per thousand by the previous half decade, limiting the sample to districts that had not yet achieved this threshold. Shaded areas are 95-percent confidence intervals. Panel (b) estimates a binomial maximum likelihood regression of emigration rates on half decade-specific functions of lagged distance from the frontier of mass migration to North America and plots the coefficients on lagged distance from the frontier. Panel (b) also includes regressions controlling for half decade-specific functions of various controls.

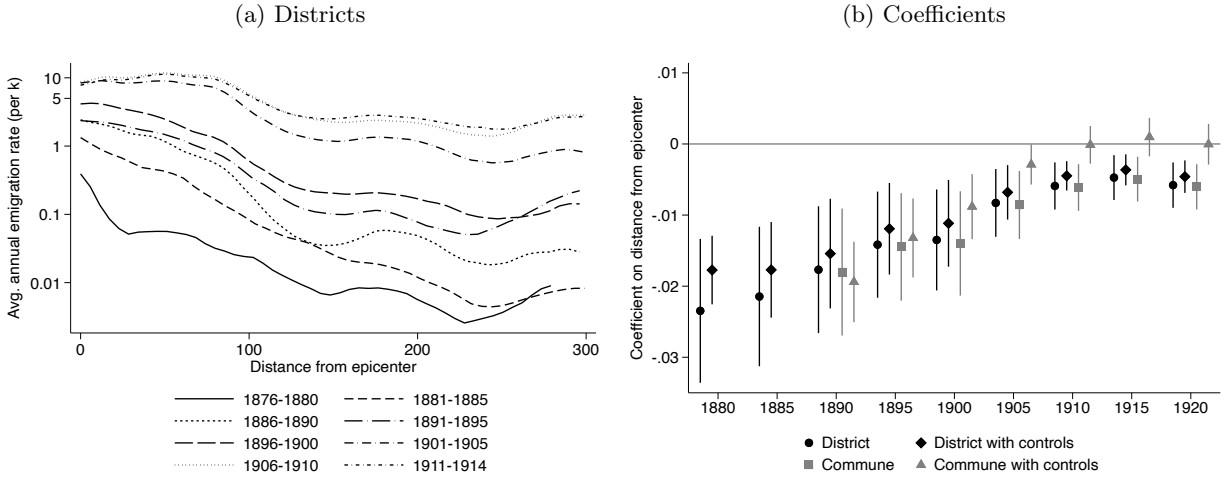


Figure 12: Emigration rates to North America by distance to epicenter (km)

Note: Panel (a) plots non-parametric regressions of the log of the average annual emigration rate for each half decade against distance to the nearest epicenter of emigration to North America. Panel (b) estimates a binomial maximum likelihood regression of emigration rates on half decade-specific functions of distance from the nearest epicenter of emigration to North America and plots the coefficients on distance from epicenter. Panel (b) also includes regressions controlling for half decade-specific functions of various controls.

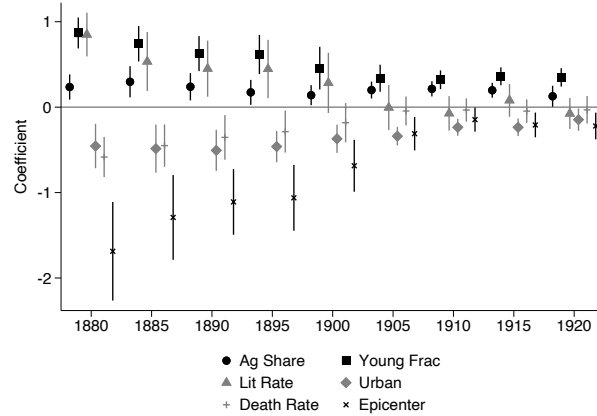


Figure 13: Relationship of migration to various local characteristics

Note: This figure presents the results of a regression of emigration to any destination on year-specific functions of various district characteristics. All explanatory variables are standardized to have mean zero and standard deviation one.

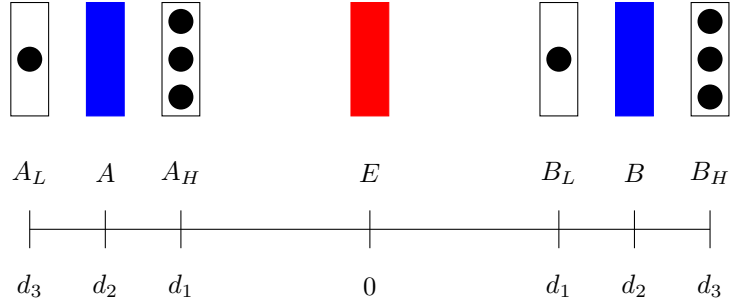


Figure 14: Illustration of the identification strategy

Note: Rectangles indicate municipalities and circles indicate population. The population of municipalities A and B and of the epicenter municipality E are unimportant to the example and are not specified. Municipalities A_H and B_H are more populous than municipalities A_L and B_L , respectively. The number line indicates each municipality's distance from the epicenter municipality E ; for instance, municipalities A and B are both at distance d_2 from the epicenter.

A Binomial Maximum Likelihood Regression

In Figures 11, 12, 13, and B.15, we estimate a regression in which municipality or district i 's emigration rate in period t , p_{it} , is expressed in the logit form

$$p_{it} = \frac{\exp(\nu_t + \eta_t z_i + \mathbf{x}_i' \gamma_t)}{1 + \exp(\nu_t + \eta_t z_i + \mathbf{x}_i' \gamma_t)},$$

where z_i is municipality i 's distance from the nearest epicenter of emigration to North America, ν_t is a period-specific intercept, η_t and γ_t are period-specific coefficients, and \mathbf{x}_i are controls. This method is intended to address observations of zero migration by treating these as cases in which all individuals have a strictly positive migration probability, p_{it} , but the realization of every resident of the municipality is to stay. After determining this logit migration demand, we use the binomial distribution to determine the probability that a given number of emigrants are observed from municipality i in time period t given p_{it} and the municipality's baseline population N_i , which enables us to estimate ν_t , η_t , and γ_t by maximum likelihood. The log-likelihood function after removing constants is

$$\mathfrak{L} = \sum_i \sum_t e_{it} N_i \log(p_{it}) + (1 - e_{it}) N_i \log(1 - p_{it}),$$

where e_{it} is the realized (i.e., observed) rate of emigration and the model is estimated by maximum likelihood. Notably, this likelihood is identical to the one that would result from estimating a standard logit model at the individual level for the entire population of Italy while retaining inference at the level of the municipality.