

Maria Rosaria Carillo
Luca Pennacchio
(Eds)

Global Migration: Effects on Sending and Receiving Countries

CRISEI

ISBN: 978-88-98279-14-2

Contents

Introduction	V
Maria Rosaria Carillo and Luca Pennacchio	
Diversity and integration of immigrants in Italy.....	1
Vincenzo Lombardo	
The Effect of the Tariff War between Italy and France on the Big Surge of Mass Migration... ..	38
Mara Rosaria Carillo, Viviana De Falco and Luca Pennacchio	
The Long-Lasting Effects of Italian Mass Migration on Foreign Direct Investment Inflows	80
Mara Rosaria Carillo, Viviana De Falco and Luca Pennacchio	
Depopulation of Rural and Inner Areas in Southern Italy	130
Dina Moawad	
Family Reunification, Integration and Well-being Outcomes of Immigrants.....	139
Maria Rosaria Carillo and Tiziana Venittelli	

Contributors

Maria Rosaria Carillo, University of Naples Parthenope

Viviana De Falco, University of Naples Parthenope

Dina Moawad, University of Gastronomic Sciences of Pollenzo

Vincenzo Lombardo, University of Naples Parthenope

Luca Pennacchio, University of Naples Parthenope

Tiziana Venittelli, University of Naples Parthenope

Introduction

Maria Rosaria Carillo and Luca Pennacchio

This volume brings together contributions by CRISEI scholars and colleagues whose research addresses migration from complementary economic, social, historical, and territorial perspectives. Migration is examined both as a contemporary process that shapes receiving societies and as a historical phenomenon with long-lasting effects on places of origin. The chapters focus on three closely connected themes: the integration and well-being of immigrants in destination countries; demographic decline and migration's role in the development of rural and inner areas; and the causes and long-term consequences of Italian mass migration. Taken together, they show that migration outcomes depend not only on individual characteristics and economic incentives, but also on local social environments, public policies, institutions, and the networks that connect origin and destination areas over time.

The contribution by Vincenzo Lombardo examines how the ethnic composition of Italian municipalities affects immigrants' cultural integration. The analysis combines survey data for more than 12,000 immigrants living in 233 municipalities with official information on local population composition. The use of Italian at home is the main integration indicator, supplemented by measures of attachment to Italy, satisfaction with living in the country, and attitudes towards intermarriage. To address immigrants' non-random location choices, the paper uses a two-stage least squares strategy with shift-share instruments. The results indicate that greater ethnic diversity promotes the adoption of the host-country language and identity, whereas a higher concentration of an immigrant's own ethnic group reduces assimilation. The evidence points to inter-group contact as the main mechanism: diversity encourages mixed friendship networks and reduces reliance on segregated ethnic spaces, while own-group concentration strengthens social and residential isolation. The findings therefore suggest that integration policies should pay close attention to local conditions that facilitate cross-group contact and limit the formation of ethnic enclaves.

Maria Rosaria Carillo and Tiziana Venittelli consider another central dimension of integration by studying the relationship between family reunification, subjective well-being, and immigrants' attachment to the host country. Using microdata from the 2011 ISTAT Survey on the Social Integration of Immigrants in Italy, they first compare married non-EU immigrants who live with their spouses in Italy with those whose spouses remain abroad. The authors find that cohabitation with a spouse is associated with significantly higher life satisfaction, even after accounting for a broad set of demographic, socioeconomic, migration, and residential characteristics. They then examine

immigrants who experienced reunification before and after the introduction of the Bossi-Fini Law in 2002. Reunification after the reform is associated with longer periods of spousal separation, lower life satisfaction, and a lower probability of feeling at home in Italy. The results suggest that the benefits of family unity may be weakened by lengthy and burdensome procedures. The chapter consequently highlights how the design and implementation of reunification policies can affect both immigrants' well-being and their integration into the receiving society.

Dina Moawad addresses the territorial dimension of migration, examining depopulation in rural and inner areas of Southern Italy and discussing policy responses that can support their demographic and economic development. The policy brief draws on 678 questionnaires collected under the “COVID-19 and Migration: A Local Development Perspective for Inner Areas” project, with particular attention to the municipality of Lioni in the Alta Irpinia area. The evidence shows that the pandemic temporarily encouraged migrants to return to their places of origin as remote work and distance learning reduced the need to remain in large urban centers. Most returnees subsequently migrated again when in-person activities resumed or when suitable local employment opportunities remained unavailable. At the same time, many respondents reported a willingness to settle permanently in their home area if adequate conditions for remote work or study were provided. On this basis, the chapter advances two main policy proposals: using the settlement of international migrants and refugees to alleviate labor shortages and population decline and investing in broadband access and public coworking spaces to retain or attract remote workers and students. Migration policy and place-based development policy are thus presented as mutually reinforcing components of a strategy for inner areas.

The contribution by Maria Rosaria Carillo, Viviana De Falco, and Luca Pennacchio turns to the historical causes of Italian mass migration. It investigates whether the protectionist policies adopted in the late nineteenth century, particularly the tariff war between Italy and France from 1888 to 1898, contributed to the sharp rise in emigration to the New World. The authors construct an original regional dataset that combines migration flows and agricultural production between 1881 and 1913. Their pre-post and difference-in-differences estimates show that regions specialized in export-oriented crops most exposed to the loss of the French market, especially wine and olive oil, subsequently experienced substantially higher growth in emigration. The estimated increase attributable to specialization in the affected crops ranges from 3.5 to 4.8 percentage points, while migration growth in the regions suffering the largest export losses was approximately twice that recorded elsewhere. The effects emerged with a delay, consistent with the time required for agricultural producers and workers to adjust to declining foreign demand. The chapter interprets emigration as an adjustment mechanism in regions where industrialization could not absorb the labor

displaced from agriculture and shows how trade policy may also have reinforced the economic divide between Northern and Southern Italy.

A second chapter by Carillo, De Falco, and Pennacchio examines whether Italian mass migration also left a lasting imprint on contemporary international investment. The authors digitize historical records of emigration from 16 Italian regions to 26 destination countries between 1880 and 1925 and link them to the regional stock of foreign-participated firms in 2017. Within an extended gravity framework, historical emigration is positively and significantly associated with present-day inward foreign direct investment in each of the three periods considered, with the strongest relationship for 1901-1913, when Italian emigration reached its peak. The chapter also investigates a possible mechanism of persistence by considering Italian institutions established abroad before 1927. The interaction between historical migration and these institutions is positively associated with inward FDI, particularly for economic and commercial organizations. The findings are consistent with the view that diaspora communities and their institutions preserved information, trust, and cultural and economic connections between destination countries and Italian regions of origin.

The contributions collected in this volume span different periods, units of analysis, and methodological approaches, yet they converge on a common conclusion: migration is a relational and institutionally embedded process. In receiving countries, immigrants' integration and well-being are influenced by patterns of local interaction and by the rules governing family life. In areas facing population decline, migration can support demographic renewal only when employment opportunities, services, and infrastructure make settlement sustainable. In sending regions, trade shocks and uneven structural transformation can drive emigration, while the communities and institutions migrants create may shape economic relations more than a century later. By connecting these dimensions, the volume contributes to a fuller understanding of how migration affects both sending and receiving countries and which policies can influence those effects.

DIVERSITY AND INTEGRATION OF IMMIGRANTS IN ITALY

Vincenzo Lombardo*

University of Naples Parthenope & GLO

June 15, 2024

Abstract

This paper analyzes the effects of diversity of the municipalities and integration of the immigrants in Italy. Using a 2SLS identification strategy with shift-share instruments, I find that greater ethnic group diversity in a municipality increases the integration. I validate these results across a large set of robustness checks and specifications. These findings suggest that while diversity facilitates the adoption of a common identity across immigrants.

Keywords: Diversity, Integration, Migration

JEL: F22, J15, O15, R23, Z10, Z13

*Department of Business and Economics, University of Naples Parthenope. E-mail: vincenzo.lombardo@uniparthenope.it. The research is funded by the Ministry of University and Research of Italian Government under the PRIN project “The Role of Skill-Biased Technological Change in Immiserizing Growth, Spatial Polarization, and Product Quality Upgrading” (Prin-2022-Carillo, ID project 2022ZKA4PZ, CUP I53D23002850006). The views and opinions expressed are solely those of the author and do not necessarily reflect those of the Ministry of University and Research, and the Italian Government cannot be held responsible for them.”

1 Introduction

Political and social debates in Western societies are increasingly centered on the integration of first and successive generations of immigrants. These concerns are often fueled by protests or riots in segregated areas across Europe and the United States, which many associate with rapid transformations in the ethnic composition of local populations. This paper empirically investigates the relation between ethnic diversity influences the integration of immigrants at the municipality level in Italy.

Italy provides a unique setting for this study because, while it has a relatively short history as a host country, it underwent a massive shift in migration inflows starting around 2004, with the foreign-born population growing from 1.33 million in 2001 to 3 million by 2008. This period was characterized by an unprecedented heterogeneity in growth across different ethnic groups and geographic areas, radically altering the composition of the immigrant population and the whole local social fabric. Certain groups, particularly from Eastern Europe, experienced a growth rate exceeding 193%, while other regions like East and South Asia grew by over 70%. This national influx translated into localized changes at the municipality level; provinces such as Trento, Campobasso, and Chieti saw their municipalities' average diversity levels and ethnic shares multiply significantly. This sharp and varied increase in both the number of co-ethnics (shares) and the variety of groups (diversity) provides a powerful source of variation for identifying how rapid environmental changes shape the cultural identity and assimilation of newcomers.

The central inquiry of this study focuses on ethnic group diversity, measured by a fractionalization index. I explore whether interactions among diverse groups reduce prejudice and stimulate the adoption of a common identity. The primary data source for this study is a unique survey on the integration and economic conditions of immigrants (ISMU), which includes interviews with over 12,049 migrants across 233 Italian municipalities in 2008 and 2009. This dataset is particularly valuable as it includes both regular and irregular immigrants and provides detailed information on language use, social networks, and frequented locations. I supplement this with official ISTAT register and census data to construct municipality-level measures of ethnic fractionalization and population density. To proxy for integration, I use the Italian language use at home, a choice that signals deep beliefs and preferences toward a specific identity rather than a mere desire to improve fluency for economic reasons. My empirical analysis reveals that greater ethnic group diversity in a municipality leads to a significantly higher assimilation rate of immigrants.

To address the endogeneity of immigrant settlement patterns, I employ a 2SLS identification strategy utilizing a set of shift-share instruments. Immigrants are not randomly assigned across municipalities and often self-select into locations based on existing social networks. I

exploit the tendency of new immigrants to settle in places where their co-ethnics previously resided by using 2003 official register data, the first year available, to predict 2008 ethnic shares and diversity. This strategy instruments the current ethnic composition with predicted counterparts, ensuring that the results are not driven by unobserved factors simultaneously affecting immigration and assimilation. The 2SLS estimates remain robust across various specifications, including the inclusion of detailed individual controls, municipality-level characteristics, and various fixed effects.

This paper contributes to several strands of the literature. First, I contribute to the study of identity formation and cultural assimilation ([Abramitzky et al., 2020](#); [Akerlof and Kranton, 2000](#); [Battu and Zenou, 2010](#); [Bisin et al., 2016](#); [Carillo et al., 2023](#); [Constant et al., 2024](#); [Fouka, 2020](#); [Fouka et al., 2022](#)). Existing studies often focus on the economic incentives for learning a host country’s language [Bleakley and Chin \(2010\)](#) or on national-level trends. My contribution lies in providing fine-grained empirical evidence at the municipality level in a “new” host country like Italy. By using the Italian language use at home as a proxy for identity rather than just economic proficiency, I provide a deeper look at how local ethnic environments shape deep-seated cultural choices and the preservation of ethnic identity.

Finally, I provide new insights into how the nature of diversity matters. Much of this literature focuses on the differentiated impact of diversity on economic growth, public goods, and conflict ([Alesina et al., 2003](#); [Alesina and La Ferrara, 2005](#); [Ashraf and Galor, 2013a,b](#); [Esteban et al., 2012](#); [Fearon, 2003](#); [Montalvo and Reynal-Querol, 2005](#)). I shift the focus toward the social and cultural outcomes of the newcomers and demonstrate that ethnic fractionalization acts as a catalyst for integration.

The remainder of the paper is organized as follows. Section 2 describes the data, the measures of assimilation, and local ethnic composition. Section 3 presents the empirical approach, the 2SLS identification strategy, and the construction of the shift-share instruments. Section 4 presents the main results, the analysis of their robustness, and the tests on the validity of the identification strategy. Section 5 concludes.

2 Data and Descriptive Statistics

Survey data and assimilation measures To measure (social) identities and assimilation rate of the immigrants in Italy, the empirical analysis uses survey data collected by the ISMU Foundation by interviewing 12,049 immigrants between 2008 and 2009. The survey provides specific questions on immigrants’ identity and self-reported feelings of belonging to the host country, and information on social, cultural, political and economic conditions of the respondents.¹

¹The data employed in this paper with the full Italian coverage are quite unique and used also by [Carillo et al. \(2023\)](#). The only other source of data for analysis on identity and assimilation in Italy is the survey developed by the National Institute of Statistics (ISTAT) and used by [Bisin and Tura \(2025\)](#).

Interviewed immigrants come from 125 different countries of origin, defined from the declared country of citizenship, mainly poorer than Italy, with most of them from Eastern Europe, Northwest Africa and Asia (Fig. A.1 in the Appendix). They are located in 233 Italian municipalities distributed in both the North and South of Italy (Fig. A.2).² The high correlation in the municipality-ethnic group shares of immigrants between the survey data (ISMU) and official registry (ISTAT) ensures that the survey data are highly representative of the actual distribution of the ethnic groups across the Italian municipalities (Tab. A.1 and Fig. A.3).

The main outcome for immigrants' assimilation is an indicator capturing how frequently immigrants speak Italian at home. Specifically, I exploit the survey question "On a daily basis, how much do you use the Italian language at home/in the family?" to build the dummy *Italian language at home* equal to one if immigrants answer 3 or more on a scale from 1 to 5, where 1 corresponds to "Never" and 5 to "Always". The extent to which immigrants use the host-country language at home, when allegedly they are not required to do so, is widely viewed as a reliable indicator of their assimilation and integration (Abramitzky et al., 2020; Bazzi et al., 2019; Bleakley and Chin, 2010; Bisin and Tura, 2025; Clots-Figueras and Masella, 2013; Fouka, 2020).³ This is consistent with a tradition in the literature that links language use to social and ethnic identity because language defines membership to specific groups and determines the exclusion of nonmembers (Alesina et al., 2003; Fearon, 2003; Lazear, 1999; Desmet et al., 2017).⁴ In the migration context, speaking the host-country language at home or within the family is therefore interpreted as a signal of identification with the host society, rather than merely the result of a desire to improve fluency or of particular household structures.

To validate this variable and its interpretation, I also use, as a robustness check, other more subjective measures of assimilation. The first is a self-identification proxy of the individuals' attachment to the host country, derived by the survey question "How much do you feel you belong to Italy?". I build the dummy *Host identity* equal to one if the interviewee responds "Very Much" or "Enough" and zero if the answer is "Far Too Little" or "Little" (Battu and Zenou, 2010; Bisin et al., 2016; Carillo et al., 2023; Constant and Zimmermann, 2008). Another subjective measure of assimilation is the dummy *Glad to live in Italy* equal to one if the respondent answers "Good" or "Very Good" to the survey question "Overall, how are you in Italy?" and zero if the answer is "Neither good nor bad", "Bad", or "Very bad" (Piracha et al., 2023). I also construct the dummy *Daughter Intermarriage* equal to one if the respondent answers "Very

²Municipalities (Comuni) correspond to LAU level 2 (formerly NUTS level 5) in the Eurostat definition. In the sample, they are distributed across 13 of the 20 Italian regions: Piedmont, Lombardy, Trentino-South Tyrol, Veneto, Emilia-Romagna, Tuscany, Marche, Abruzzo, Lazio, Campania, Molise, Apulia and Sicily.

³Bisin and Tura (2025) have recently recovered a variable for the use of Italian at home from another Italian survey, the Condition and Social Integration of Foreign Nationals Survey (2011-2012), and interpreted it "as a proxy for (negative) socialization rates to the culture of origin, that is, an indication that parents are not exercising much effort in transmitting their culture to their children".

⁴For instance, Alesina et al. (2003), Fearon (2003), and Desmet et al. (2017) use the language spoken at home to code the ethnic identity of the individuals who do not explicitly declare it.

much” or “Enough” to the survey question “If your daughter were to marry an Italian citizen, how much would you approve of such a union?” and zero if the answer is “Far too little”, or “Little”. This measure captures parental attitudes toward intermarriage, often used as an indicator of social assimilation (Battu and Zenou, 2010; Bazzi et al., 2019; Bisin et al., 2016; Fouka et al., 2022; Meng and Gregory, 2005).⁵ Finally, I construct an index by extracting the first factor from a principal component analysis of all the measures, and then define the *Identity index* as a dummy equal to one for observations with values above the median. One advantage of the use of Italian at home as a measure of social assimilation is that, despite being self-reported like the other measures, it captures a relatively more objective attitude and manifestation of the immigrants toward the host country.

Depending on the measure used, between 50 and 70% of the sample can be classified as assimilated into Italian culture, with nearly 50% of the immigrants declaring that they speak Italian language at home. Reassuringly, a preliminary analysis shows that the use of Italian at home is associated with the attachment and sense of belonging to Italy. Moreover, around 91% of immigrants retains a strong ethnic identity, declaring to feel to belong to their country of origin “very much” or “enough” (*Home identity*).

In addition, the survey collects extensive information on individual characteristics. The full sample includes both regular (about 90%) and irregular (10%) immigrants, aged 18 or older at the time of interview and with an average age at arrival in Italy of about 28 years old (Tab. A.2). Respondents, mostly males (52%) and married with children (57%), spend many years in Italy (on average 7.8 years) and have quite a high level of education; 62% of immigrants report at least a high school degree, with 19% of them stating they have a bachelor’s degree or higher level of education. These characteristics also explain their mastery of Italian language; the average score of *Proficiency*, measured as the self-reported ability in speaking and reading, is 3.5 on a scale from 1 to 5. Finally, 27% of the sample are Catholic, 21% Orthodox and 40% Muslim; the religious minorities are represented by Coptic, Evangelical, Buddhist, Hindus, Sikh and those professing other religions, while about 7% state no religious affiliation. About 82% of respondents are employed, 70% of them have a regular and about 10% an irregular job. Finally, immigrants’ labor income is also reported as an eight-class measure. Among those with a positive income, the majority (67%) state a monthly net income from labor of between 600 and 1200 Euros, while 13% and 20% of the sample report an income lower than 600 and higher than 1200 Euros, respectively.⁶

⁵I cannot use actual intermarriage as one of the outcomes because it would imply to lose many observations due to data limitations. I, then, follow Battu and Zenou (2010) and Bisin et al. (2016) in using the attitude toward intermarriage as a proxy for identity and social assimilation. I exploit the actual intermarriage rate in the robustness analysis for the small subsample in which this information is available.

⁶The income classes with the corresponding share of immigrants are: No Income (23.02%), < 600 (10.11%), 600-799 (16.43%), 800-999 (18.59%), 1000-1199 (16.20%), 1200-1499 (9.97%), 1500-2000 (3.93%), > 2000 (1.74%).

Municipality data and diversity measures I combine the survey data with municipality level data obtained from different official sources. The first source is the official register (ISTAT) from which I recover, for each municipality and country of origin, the population and shares of immigrants by municipality-country of origin. I use these data to build the two main municipality-level variables that capture the degree of diversity and the concentration of each ethnic group in each municipality. Assigning natives to group $o = 1$, I compute the base diversity measure as the fractionalization index:

$$Diversity_m = \sum_{o=1}^O Share_{om}(1 - Share_{om}) \quad (1)$$

where $Share_{om} = N_{om}/P_m$ indicates the concentration of each ethnic group in each municipality. It is calculated as the population share of immigrants from country of origin o in municipality m , given by the total number of members of ethnic group o , N_{om} , divided by the overall municipality population, $P_m = \sum_o N_{om}$. I compute these measures from data of the population register at the beginning of the 2008, some months before the submission of the survey employed to measure the assimilation rate of immigrants. Figure A.4 in Appendix illustrates the distribution of the two measures across the municipalities.

I retrieve additional information on several characteristics of the municipalities from the 2001 official census, and on financial data (i.e., per-capita income) from the Ministry of Interior. In particular, from the 2001 official census, I recover data on the education level using the share of high educated individuals, the unemployment rate, the number of per-capita non-profit organizations to account for the social capital of the municipalities (Guiso et al., 2008; Nannicini et al., 2013), the economic and industrial structure exploiting information on the share of workers in agriculture, industry, service sector, commerce, and in skill intensive jobs. Finally, I also control for some geographical features, among which the size of the municipality in terms of area (per km^2), its altitude, and coastal location. Table A.3 in the Appendix summarizes the municipality variables used in the empirical analysis.

3 Empirical strategy

3.1 Baseline specification

The main specification used to investigate the relationship between immigrants' assimilation and ethnic composition of the municipalities is

$$y_{iom(p)} = \beta_0 + \beta_1 Diversity_m + \mathbf{X}'_{iom} \delta + \mathbf{\Gamma}'_m \gamma + \alpha_o + \alpha_p + \lambda_{ay} + \mu_{dplace} + \varepsilon_{iom} \quad (2)$$

where the main outcome of interest y_{iom} is the dummy *Italian language at home* of immigrant i from country of origin o in the municipality m (in province p). The coefficients of interest, β_1

and β_2 , capture the effects of the two dimensions of local ethnic composition. In particular, β_1 measures the effect of overall ethnic diversity as defined by the fractionalization index in (1), that measures the probability that two randomly selected individuals in the municipality belong to different ethnic groups, and thus captures the overall ethnic heterogeneity of the municipality's population and how evenly groups are distributed.

All specifications always include country of origin α_o and province α_p fixed effects to account for (omitted) time-invariant characteristics of each ethnic group and province. I also control for individual and municipality characteristics included in the vectors $\mathbf{X}_{iom}$ and $\mathbf{\Gamma}_m$, respectively. The individual level covariates account for the overall proficiency in Italian language, gender, the presence of children, and include dummies for marital status, education, income, and religion. The vector of municipality controls, $\mathbf{\Gamma}_m$, features total population, overall foreign density, per-capita income, and the share of high educated as a measure of local human capital. In the baseline specifications, it also includes geographic controls such as average elevation, slope, and coastal location, and the per-capita number of no-profit associations to control for a supply side measure of the local acceptance and integration of immigrants as associated to different levels of municipality social capital. In addition, I add to the right-hand side a full set of age at arrival $\times$ years in Italy fixed effects (λ_{ay}) to account for the possibility that the time spent in the place of destination has a differentiated effect on the identity formation and assimilation process depending on the age at arrival of the immigrants.⁷ Finally, in the most preferred specification, I also condition on day of week $\times$ place of interview fixed effects (μ_{dplace}) to wash out part of the random measurement errors induced by the use of self-reported measures of identity as well as to account for potential selection of migrants in particular places of interview on specific days of the week (e.g., immigrants with strong ethnic identity interviewed particularly in religious or ethnic places in specific days).⁸

Identification To address the endogeneity of the fractionalization index, I follow the standard approach in the literature by implementing a 2SLS estimator and constructing a set of shift-share instruments to predict both the index of diversity and the share of immigrants (Card, 2001). This approach, largely employed in the immigration literature, exploits the

⁷For instance, younger children learn languages more easily than older individuals, and this has been shown to be correlated also with the identity formation (Bleakley and Chin, 2010; Clots-Figueras and Masella, 2013). Likewise, the years spent in the host country are a powerful predictor of the assimilation rate of the immigrants (Abramitzky et al., 2020). Figure A.5 in Appendix shows the not monotonic effects of the interplay of two variables and how the assimilation rates of the immigrants changes according to variation in the two measures.

⁸The survey indicates the following possible twelve places of interview: centers providing services and assistance (reception, work, health, counseling service, refectory, public offices), training centers (Italian courses, professional training courses, schools, universities), worship (churches, mosques, temples), ethnic shops (kebabs, Islamic butchers, take-aways, food products), entertainment (cinema, discos, sports facilities, bars, restaurants), shopping centers, meeting places (stations, squares, parks, lakes), markets (municipal markets, flower market, fruit and vegetable), workplaces or workforce recruitment (construction sites, textile laboratories, restaurants and hotels, gatehouses, agricultural fields and farms), associations and cultural centers, service centers (phone centers, money transfer agencies), private residences.

propensity of immigrants to settle in areas where previous immigrants of the same origin have chosen to migrate (Adao et al., 2019; Borusyak et al., 2022, 2025; Bove et al., 2023; Campo et al., 2022; Card, 2001; Fouka et al., 2022; Goldsmith-Pinkham et al., 2020; Jaeger et al., 2018; Tabellini, 2020). Based on this logic, I construct the instrument of $Diversity_m$ starting by predicting the number of members of the different ethnic groups o located in each municipality m , $\hat{N}_{om}$. I take the 2003 as the reference (i.e., initial) year because this is the first year for which official register data on the number of immigrants by municipality-country of origin are available. Then, I employ a leave-out formulation to construct the predicted population for each municipality and country of origin at the beginning of the 2008 as it follows:

$$\hat{N}_{om} = N_{om}^{2003} + \left(\frac{N_{om}^{2003}}{N_{o(IT)}^{2003}} \right) \times \Delta N_{o(IT-R(m))} \quad (3)$$

where $\Delta N_{o(IT-R(m))} = N_{o(IT-R(m))}^{2008} - N_{o(IT-R(m))}^{2003}$ is the leave-out version of the aggregate ‘shift’ component given by the change in the stock of immigrants from country of origin o in the whole Italy excluding the immigrants in the region R of the municipality m . This approach ensures that region-specific shocks that may influence the ethnic composition of the municipalities within a region are not included and accounted for by the instrument (Adao et al., 2019; Campo et al., 2022; Fouka et al., 2022; Tabellini, 2020). Then, the predicted number of immigrants $\hat{N}_{om}$ at the beginning of the 2008 is obtained by apportioning this aggregate shift across the municipalities according to the initial distribution of the immigrants’ ethnic groups as reflected in the initial shares ($N_{om}^{2003}/N_{o(IT)}^{2003}$) of each country of origin group o in each municipality m with respect to the total number of immigrants of each ethnic group who were already in Italy in 2003 ($N_{o(IT)}^{2003}$). Adding this component to the initial level of the population of each group o in each municipality m , N_{om}^{2003} , I obtain the predicted municipality-country of origin population at the beginning of the 2008. With this predicted variable in hand, I construct the predicted shares of immigrants from country o in municipality m , and then the predicted diversity given by the fractionalization index applied to the predicted shares as:

$$\widehat{Diversity}_m = \sum_o \widehat{Share}_{om} (1 - \widehat{Share}_{om}) \quad (4)$$

I employ the measure in equation (4) as instrument for the endogenous variables of interest, $Diversity_m$. Section 4.2 and the Appendix provide detailed discussion of instrument validity and identification strategy. In addition to the evidence supporting the shift-share instruments employed in the analysis, I report robustness checks using also alternative instruments that exploit other sources of quasi-random variation in the allocation of migrants from different countries of origin across municipalities.

Table 1 presents the first stage results for both the endogenous variables across different

TABLE 1: ETHNIC COMPOSITION, IDENTITY, AND ASSIMILATION: FIRST STAGE

	Dependent variable							
	Diversity				Own ethnic group share			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted diversity	0.975*** (0.037)	0.985*** (0.038)	0.993*** (0.037)	0.928*** (0.032)	0.218 (0.545)	0.407 (0.509)	0.277 (0.391)	-0.182 (0.397)
Predicted own ethnic share	-0.001 (0.000)	-0.001 (0.001)	-0.001* (0.000)	-0.001* (0.000)	0.865*** (0.059)	0.856*** (0.058)	0.863*** (0.043)	0.862*** (0.043)
Country of origin FE	yes	yes	yes	yes	yes	yes	yes	yes
Province FE	yes	yes	yes	yes	yes	yes	yes	yes
Individual controls	no	yes	yes	yes	no	yes	yes	yes
DoW $\times$ Place FE	no	no	yes	yes	no	no	yes	yes
Age at arrival $\times$ Years in Italy FE	no	no	yes	yes	no	no	yes	yes
Municipality controls	no	no	no	yes	no	no	no	yes
Observations	11534	9419	9289	9288	11534	9419	9289	9288
Countries of origin (#)	125	118	118	118	125	118	118	118
Municipalities (#)	231	225	225	224	231	225	225	224
R^2	0.965	0.964	0.972	0.975	0.946	0.946	0.962	0.963
F-test of excluded instruments	398.60	393.70	411.17	524.00	301.76	294.13	435.76	281.37
S & W weak identification test	760.70	746.18	895.07	891.33	224.25	235.87	439.30	430.48

Notes. All regressions include country of origin and province fixed effects. Individual controls are: *proficiency in Italian language, gender (male dummy), having children, marital status, compulsory school, high school, BA degree +, religion and income dummies*. Municipality controls are: *population (log), foreign density (log), per-capita income (log), per-capita number of no-profit associations (x1000), mean elevation, slope, coastal location*. Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

specifications. Columns (1)-(4) report the estimates of the first stage when the endogenous variable is $Diversity_m$, while columns (5)-(8) those when the variable is the own ethnic group share, $Share_{om}$. The instruments display very high predictive power and the first-stage relationships remain robust to the inclusion of additional controls. The baseline specifications, including only country of origin and province fixed effects, are firstly presented in columns (1) and (5). Notably, across all estimations, $Diversity_m$ is predicted exclusively and strongly by its instrument, *Predicted diversity*, while the own ethnic group $Share_{om}$ is similarly predicted only by its corresponding instrument, *Predicted own ethnic share*. The high correlation between the instruments and endogenous variables remain remarkably stable in magnitude and statistical significance as the specifications become sequentially more saturated, adding to the right-hand side the set of individual controls in columns (2) and (6), the day of week by place of interview and age at arrival by years in Italy fixed effects in (3) and (7), and finally the municipality controls in (4) and (8). Throughout the specifications, the estimated coefficients barely change in size, indicating the strength and consistency of the instruments. The power of the instruments is further highlighted by the large values of the F-test on the excluded instruments and by the [Sanderson and Windmeijer \(2016\)](#) weak identification test, all of which well above the conventional threshold of 10 ([Stock and Yogo, 2005](#)).

4 Results

4.1 Main results

Table 2 reports the estimates of Equation (2) when the indicator of immigrants' assimilation, the dummy *Italian language at home*, is regressed on the main variables of interests, $Diversity_m$, and the immigrants' own ethnic group share, $Share_{om}$. Panel A presents the OLS estimates, while Panel B the 2SLS results. In columns (1), I begin by introducing the two measures in the most parsimonious specification accounting only for country of origin and province fixed effects, to highlight the importance of estimating the conditional effects of diversity and one's own ethnic group share. As emphasized previously, fractionalization and immigrant's own ethnic group share capture fundamentally different aspects of the composition of the local populations. The former quantifies the overall ethnic heterogeneity of the municipality's population and how evenly groups are distributed, while the latter gauges the effects of the local concentration of the immigrant's co-ethnic community on the assimilation of the individuals. In line with the theoretical predictions, the coefficient on diversity is positive, while that on the immigrants' own ethnic group share is negative. These estimates indicate that, conditional on the relevance of one's own ethnic group, greater diversity across ethnic groups encourages the assimilation of the immigrants, as reflected in the increase of the use at home of the language of the country of destination. Conversely, greater concentration of immigrants' own ethnic group dampens the propensity to assimilate to the host country's common language.

Following the structure of the first-stage results (Tab. 1), in columns (2) to (4), I sequentially introduce a more demanding set of controls, and the coefficients remain remarkably stable in both magnitude and statistical significance. In column (2), I add the baseline set of individual characteristics, including a measure of overall proficiency in Italian language employed to disentangle the identity component from the learning incentive of the language use at home, along with a set of dummies for gender, having children, marital status (*single*, *married*, *widowed*, *divorced*), education (*no education*, *compulsory*, *high school*, *BA degree+*), religion, and income classes. Next, column (3) strengthens the identification by exploiting only very small variation across individuals who arrived at the same age and spent the same number of years in Italy, and were interviewed in the same type of location on the same weekday. Results remain highly stable and, if anything, the precision of the estimates improves. Finally, columns (4) adds the baseline set of municipality-level controls, none of which turns out to be statistically significant (see Tab. A.6 in Appendix). In the robustness analysis, I provide further evidence that the results do not depend on omitted features of the municipality, employing an even larger set of municipality-level variables.

Across all the specifications, the 2SLS estimates in Panel B are only slightly larger than the corresponding OLS coefficients, suggesting that the instruments correct some possible down-

TABLE 2: ETHNIC COMPOSITION, IDENTITY AND ASSIMILATION: MAIN RESULTS

	Dependent variable: Italian language at home			
	(1)	(2)	(3)	(4)
Panel A: OLS estimates				
Diversity	0.504** (0.252)	0.522** (0.247)	0.492*** (0.175)	0.504** (0.211)
Own ethnic group share	-0.050*** (0.016)	-0.044*** (0.013)	-0.036*** (0.012)	-0.035*** (0.012)
R^2	0.133	0.268	0.472	0.474
Panel B: 2SLS estimates				
Diversity	0.529* (0.274)	0.552** (0.269)	0.527*** (0.193)	0.577** (0.246)
Own ethnic group share	-0.045*** (0.017)	-0.042*** (0.014)	-0.030** (0.014)	-0.029** (0.014)
Kleibergen-Paap F-statistic	225.13	213.45	448.90	446.57
Country of origin FE	yes	yes	yes	yes
Province FE	yes	yes	yes	yes
Individual controls	no	yes	yes	yes
DoW $\times$ Place FE	no	no	yes	yes
Age at arrival $\times$ Years in Italy FE	no	no	yes	yes
Municipality controls	no	no	no	yes
Observations	11534	9419	9289	9288
Countries of origin (#)	125	118	118	118
Municipalities (#)	231	225	225	224
Mean dependent	0.451	0.445	0.445	0.445

Notes. The dependent variables is the dummy equal to one for the use of Italian language at home. All regressions include country of origin and province fixed effects. Individual controls are: *proficiency in Italian language*, *gender (male dummy)*, *having children*, *marital status*, *compulsory school*, *high school*, *BA degree* +, *religion* and *income dummies*. Municipality controls are: *population (log)*, *foreign density (log)*, *per-capita income (log)*, *per-capita number of no-profit associations (x1000)*, *mean elevation*, *slope*, *coastal location*. Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

ward bias, while being also very powerful in accounting for the sources of this bias, as reflected in the very large Kleibergen-Paap F-statistics. Focusing on the most preferred 2SLS specification in column (6), the point estimates imply that one standard deviation increase in *Diversity* increases the use of Italian language at home by 4.3 percentage points corresponding to an increase of 9.6% of the average use of Italian language at home across the immigrants. Likewise, a one standard deviation increase in the *Own ethnic group share* in the municipality decreases

the use of Italian language at home by 2.7 percentage points corresponding to a decrease of about 6% of the average use of Italian language at home across the immigrants.

Overall, these findings suggest that a larger local variety of ethnic communities strengthens immigrants' incentives to converge toward a shared identity, while a stronger concentration of one's own group instead tends to hinder the adoption of the host country's common language.

4.2 Testing the identification strategy and instrument validity

In this section, I discuss the conditions for the validity of the identification strategy and the main evidence in support of the shift-share instruments employed in the analysis; the details are in Appendix A and B. The identification strategy and shift-share instruments adopted in the paper rely on the exogeneity of the share component, the initial (i.e., 2003) local settlements of the different groups of immigrants by their country of origin, and the exogeneity of the shift factor, the nation-wide changes in the stock of immigrants from the country of origin o between the years 2003 and 2008 (Adao et al., 2019; Borusyak et al., 2022; Goldsmith-Pinkham et al., 2020; Jaeger et al., 2018).

On the first side, the key identification assumption is that conditional on the province and country of origin fixed effects, and municipality controls, no other unobserved factors simultaneously affect the assimilation rate, the initial distribution of the immigrants from different origin countries, and the ethnic diversity of the municipalities in Italy. This assumption implies that there should not be other municipalities' characteristics that are correlated with the 2003 ethnic shares, and may have affected the subsequent movements, distribution, diversity, and the assimilation rate of the immigrants across the municipalities. I deal with these concerns in several ways. I begin by examining how initial ethnic shares, ethnic population level, and diversity, all measured in 2003, together with predicted diversity correlate with a broad set of predetermined municipal characteristics measured in 2001, the first census year before the survey for which municipality data are available.⁹ These characteristics capture a wide range of demographic factors (such as population and education), economic conditions (including income and unemployment), industrial structure (the distribution of employment across sectors), cultural features (the number of non-profit organizations), and geographic attributes. The results reported in Table A.4 in the Appendix support the hypothesis that violations of the exclusion restriction are unlikely. None of the predetermined municipal characteristics measured in 2001, except for municipal population, which is always included in the baseline set of controls, is statistically significant or correlated with either the initial ethnic shares or the (log) size of the municipality-country of origin population. Moreover, the predetermined local factors are also uncorrelated with both the initial diversity in 2003, and the predicted diversity.

⁹For clarity, I refer to variables measured in 2001 as "predetermined", and to those computed in 2003 as "initial".

A related concern is that the power of the instruments and the main results may be driven by a specific subset of immigrants, according to their country of origin, whose location choices and assimilation rates are influenced by omitted, predetermined municipal characteristics. This does not appear to be the case. Predetermined municipal characteristics are uncorrelated with the initial municipal shares of the ten most representative immigrant groups nationwide (Table A.5). Moreover, the results are robust to sequentially dropping each of the 118 countries of origin, as the 2SLS coefficients on *Diversity* and *Own ethnic group share* remain virtually unchanged relative to the baseline estimates (Figure A.6). The findings are also robust to the inclusion of predetermined municipal controls measured in 2001, with point estimates only marginally different from the baseline (column 3 in Table A.6). Taken together, these results suggest that omitted local characteristics are unlikely to have jointly affected immigrant movements, municipal composition, and assimilation outcomes.

The second set of threats to the identification comes from potential violations of the exogeneity of the nationwide aggregate shocks, ‘the shift’ used to apportion the initial shares across the municipalities. The leave-out formulation plausibly guarantees that the shift component should be independent of unobserved local factors. Notwithstanding, one concern might still be that the change in the number of immigrants from each country of origin has been affected by specific shocks in each municipality. To deal with this unlikely but potential threat, I perform a further series of robustness checks to ensure that the size of the ethnic groups is not systematically related to cross-municipalities pull factors. In particular, I show that the main results remain robust, and the coefficients slightly change in magnitude when using different versions of the instruments (Table A.7). First, as a sanity check, I use a modified version of the instruments where the aggregate shift component does not consider the leave-out formulation but is computed by the change in the stock of immigrants from the country of origin o throughout Italy. Next, I compute the shares of immigrants, diversity measures, and their corresponding instruments by aggregating countries of origin into the 22 world geographical sub-regions defined by the United Nations, plus the native category.¹⁰ Results are remarkably consistent with the main specifications, suggesting also that potential spatial correlation across countries of origin of the immigrants does not influence the identification strategy and estimates (Adao et al., 2019; Borusyak et al., 2022).

As an additional test, I exploit the EU successive enlargements quasi-experiment in several ways (Adda et al., 2025; Bove et al., 2023). Within the time window of the analysis, the Czech Republic, Cyprus, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, the Slovak Republic,

¹⁰The UN classifies the countries according to the following sub-regions: Australia and New Zealand; Caribbean; Central, Northern, and South America; Central, Eastern, South-Eastern, Southern, and Western Asia; Eastern, Middle, Northern, Southern, and Western Africa; Eastern, Northern, Southern, and Western Europe; Melanesia; Micronesia; Polynesia. The ISMU sample of immigrants covers all but the following six sub-regions: Australia and New Zealand, Northern America, Western Europe, Melanesia, Micronesia, Polynesia.

and Slovenia joined the EU in 2004, and Romania and Bulgaria became EU members in 2007. The access to the EU gave EU ‘new’ citizens the right to reside and work in any other EU country, stimulated large migration flows, particularly from ‘new’ to ‘old’ Member States, and then exogenously modified also in Italy the size and composition of the population of immigrants both at the national and local level. Thus, I modify the diversity and predicted diversity measures by sequentially excluding the countries that joined the EU in 2004, 2007, and the whole period 2004-2007. I further drop these countries from the sample. In all cases, results remain robust and consistent with the main findings.

Lastly, to ensure that the number of immigrants from each country of origin is not systematically related to cross-municipalities pull factors, I build another set of instruments based only on push factors in the country of origin of the immigrants (Alesina et al., 2016; Bove et al., 2023; Fouka et al., 2022).¹¹ Using a gravity model of migration, I first predict the number of immigrants entering Italy from each country of origin o in the year 2008. I then apportion these nationally predicted inflows across the municipalities using the initial local settlements. The resulting shares and diversity measures are then different from those used in the baseline analysis because they are derived only from push factors that predicted the movements of the immigrants out of their country of origin towards Italy. Despite these differences, the 2SLS coefficients on *Diversity* and *Own ethnic group share* obtained when substituting the baseline instruments with their push versions are statistically significant and remain remarkably close in magnitude to the main findings (column 9 in Table A.7). Hence, these results reassure that even potential biases generated by the correlation between the national aggregated shocks and local pull factors are unlikely.

These robustness checks support the exogeneity of the shift-share instruments and the validity of the identification strategy. As a final omnibus test, I also perform a falsification exercise based on placebo first-stage regressions. I randomly reassign the two instruments, *Predicted diversity* and *Predicted own ethnic share*, across the municipalities and the municipality-country of origin groups, respectively. I replicate this reshuffling ten thousand times, and in each iteration, I re-estimate the preferred specification of the first-stage regressions (i.e., cols. 5 and 10 in Table 1). Figure A.7 in Appendix plots the resulting placebo coefficients and shows that the simulated point-estimates are normally distributed around zero and lie very far from the baseline first-stage estimates. These results reinforce the credibility of the identification strategy by confirming that the observed first-stage relationships are not driven by mechanical correlations or spurious patterns because the strong first-stage coefficients in the data do not emerge when the instruments are randomly reassigned.

¹¹All the details of the exact procedure are in Appendix B.

TABLE 3: OTHER OUTCOMES

	Dependent variable				
	Host identity	Glad to live in Italy	Intermarriage daughter	Host identity Index	Home Identity
	(1)	(2)	(3)	(4)	(5)
Diversity	0.534* (0.290)	0.813*** (0.269)	0.614** (0.257)	0.819*** (0.283)	0.045 (0.147)
Own ethnic group share	-0.012 (0.011)	-0.022 (0.014)	-0.031*** (0.011)	-0.043*** (0.013)	0.013** (0.006)
Observations	9494	9636	8764	8284	9603
Kleibergen-Paap F-statistic	446.18	446.33	500.19	505.91	451.16
Mean dependent	0.570	0.694	0.717	0.499	0.909

Notes. Two-stage least square (2SLS) regressions. The dependent variables are dummies for the host identity (1), how glad to live in Italy (2), daughter intermarriage approval (3), the host identity index (4), and the home identity (5). The specifications of all regressions are based on column 6 of Table 2 and include country of origin, province, day of week $\times$ place of interview and age at arrival $\times$ years in Italy fixed effects, as well as individual and municipality controls. Individual controls are: *proficiency in Italian language*, *gender (male dummy)*, *having children*, *marital status*, *compulsory school*, *high school*, *BA degree +*, *religion* and *income dummies*. Municipality controls are: *population (log)*, *foreign density (log)*, *per-capita income (log)*, *per-capita number of no-profit associations (x1000)*, *mean elevation*, *slope*, *coastal location*. Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.3 Other outcomes

The main results presented thus far reveal that the local ethnic composition of the municipalities influences immigrants' assimilation, as proxied by their use of Italian at home. Evidence in Table 3 further corroborates the interpretation that these patterns reflect the role of diversity and ethnic concentration in shaping immigrants' identity. At this end, I examine a set of additional measures that more directly, although more subjectively, capture the immigrants' broader sense of attachment and belonging to the host country or their country of origin.

In particular, column (1) shows that greater diversity implies a stronger sense of belonging to the host country. Conversely, a higher concentration of one's own ethnic group is associated to a weaker sense of belonging, although the coefficient is imprecisely estimated and not statistically significant, likely reflecting the more subjective, and therefore noisier, nature of the *Host identity* indicator. Column (2) confirms these findings when inferring how glad immigrants report feeling to live in Italy, with the estimated coefficient for the co-ethnic share becoming more precise and not far from conventional statistical significance (p -value=0.130). The third column draws on the idea that parental approval of a daughter's heterogamous marriage with a native may signal an identity choice of the parents toward the host country (Bisin et al., 2016), consistent with the broader literature that interprets actual intermarriage as a revealed

expression of identity (Bazzi et al., 2019; Fouka et al., 2022; Meng and Gregory, 2005). The coefficients, very precisely estimated and statistically significant, indicate that in more diverse municipalities parents are more prone to approve their daughters' intermarriages with native-born individuals, whereas the opposite holds in communities characterized by greater ethnic concentration.

These findings anticipate and support the evidence presented in the next section highlighting the key role of intergroup contacts in explaining the effects of local ethnic composition on identity and assimilation. In particular, in line with the contact hypothesis theory (Allport et al., 1954; Lowe, 2024), the greater parental inclination toward daughter's intermarriage in more diverse communities may reflect the reduction in prejudice and stimulation of intergroup cooperation induced by the interpersonal contacts with diverse groups in more diverse communities.

Finally, the estimates remain robust and support an identity-based interpretation when the outcome is the host *Identity index*, constructed by combining all measures through principal factor analysis. Greater local diversity stimulates immigrants' broader attachment to the host country, whereas stronger ethnic concentration dampens the propensity to assimilate.

A complementary dimension of immigrants' social identification and assimilation is the degree to which they preserve their ethnic identity. If local ethnic composition shapes immigrants' assimilation through its influence on identity formation, one would likewise expect immigrants to retain a stronger attachment to their country of origin in communities with a larger presence of co-ethnics. The last column of Table 3 explores this possibility and shows that a higher concentration of one's own ethnic group strengthens immigrants' sense of belonging to their origin country, while greater diversity has no effect.

4.4 Further robustness checks

I additionally provide a wide range of robustness checks of the main results. Table A.8 in the Appendix begins by examining the role of individuals' characteristics. All specifications include a rich set of individual-level covariates, among which overall proficiency in the Italian language. On the one hand, controlling for Italian proficiency ensures that the use of Italian at home is not driven by instrumental motives, such as the need or desire to learn and practice the language to improve labor market outcomes, or by particular family structures, but rather reflects an identity-related choice. On the other hand, Italian proficiency may constitute a bad control, as it could itself be a proxy of the immigrants' assimilation (e.g., Battu and Zenou, 2010; Bisin et al., 2016; Bleakley and Chin, 2010; Clots-Figueras and Masella, 2013; Danzer et al., 2022); conditioning on it may therefore partially absorb the effects of local diversity and ethnic concentration. Reassuringly, excluding this variable leads to only a marginal increase in the

estimated coefficients, supporting the interpretation that immigrants' use of Italian at home is primarily driven by identity-related considerations rather than by instrumental motives aimed at improving proficiency in the host country's language. Likewise for the legal status of the immigrants, as the main results are robust to the exclusion of the roughly 10% of irregular immigrants. Finally, the main findings remain virtually unchanged when the small number of second-generations immigrants is dropped from the sample.

The second set of robustness checks supports the identification strategy, which exploits very fine-grained variation across individuals. Table A.9 ensures that the main findings are not driven by cells with very few observations. The results remain robust when excluding countries, municipalities, and age at arrival by years in Italy cells with fewer than ten observations, or when dropping provinces with fewer than three municipalities. In line with these findings, Table A.10 shows that the main results are not sensitive to alternative administrative aggregations. As the analysis sequentially moves from the smallest to the largest administrative units and the diversity and the own ethnic group shares are correspondingly aggregated, the estimated standardized effects remain very close to the baseline results. This further reassures that the main findings are not driven by spatial correlation across neighboring municipalities.

Finally, I do not detect any notable patterns of heterogeneity or non-linearities in the data.

5 Concluding Remarks

This paper has empirically investigated how the ethnic composition of local populations shapes the identity and cultural assimilation of immigrants at the municipality level in Italy. By exploiting the rapid transformation of Italy's migration landscape between 2003 and 2008 and employing a 2SLS identification strategy based on shift-share instruments, I have identified a striking divergence between the effects of ethnic diversity and own-group concentration. My results demonstrate that while greater ethnic fractionalization acts as a catalyst for the adoption of the host country's language at home, a higher share of an immigrant's own ethnic group significantly curbs this assimilation process.

The primary mechanism driving these results is the role of inter-group contact. In more diverse municipalities, immigrants are less likely to frequent isolated "ethnic hubs", such as ethnic shops, service centers, and assistance centers, and are more likely to build mixed friendship networks that include both foreign and Italian individuals. Conversely, high own-ethnic concentration fosters "bonding" social capital, leading to social circles and habits that remain segregated within co-ethnic boundaries. This social isolation is further reflected in the high-resolution residential data; using an isolation index calculated at the census block level, I find that municipalities with greater spatial segregation exhibit lower rates of cultural integration. Diversity facilitates assimilation precisely by breaking down these ethnic enclaves and encour-

aging interactions with the broader society.

These findings are robust to a wide array of alternative explanations. I find no evidence that these patterns are driven by economic incentives, as diversity is not correlated with employment probability or specific sectoral outcomes. Similarly, the results are not explained by the marriage market or vertical transmission within families; the effect is strongest among couples of the same nationality, suggesting that the local environment shapes the identitarian choices of even the most culturally homogeneous households. However, the nature of diversity matters: while linguistic, religious, and genetic differences mimic the positive effects of fractionalization, extreme cultural distance in terms of fundamental values and beliefs can inhibit the interactions necessary for identity formation.

The broader implications of this study are particularly relevant for modern Western societies facing increasing concerns over integration and social cohesion. The results suggest that nation-building and integration are not necessarily hindered by diversity itself; rather, they are challenged by the spatial and social isolation that arises from high ethnic concentration and segregation. Public policies aimed at promoting diverse residential environments and facilitating inter-group contact may therefore be more effective in fostering a shared host identity than policies focused solely on economic incentives or language training in isolation. By providing evidence from a recent migration context, this paper sheds light on the complex interplay between local demographics and the cultural trajectories of newcomers in multi-ethnic societies.

References

- Abramitzky, R., L. P. Boustan, and K. Eriksson (2020). Do Immigrants Assimilate More Slowly Today Than in the Past? *American Economic Review: Insights* 2(1), 125–141.
- Adao, R., M. Kolesár, and E. Morales (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics* 134(4), 1949–2010.
- Adda, J., P. Pinotti, and G. Tura (2025). There’s More to Marriage than Love: The Effect of Legal Status and Cultural Distance on Intermarriages and Separations. *Journal of Political Economy* 133(4), 1276–1333.
- Akerlof, G. A. and R. E. Kranton (2000). Economics and Identity. *Quarterly Journal of Economics* 115(3), 715–753.
- Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003). Fractionalization. *Journal of Economic Growth* 8(2), 155–94.
- Alesina, A., J. Harnoss, and H. Rapoport (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth* 21(2), 101–138.
- Alesina, A. and E. La Ferrara (2005). Ethnic Diversity and Economic Performance. *Journal of Economic Literature* XLIII, 762–800.
- Allport, G. W., K. Clark, and T. F. Pettigrew (1954). *The nature of prejudice*, Volume 2. Addison-wesley Reading, MA.
- Ashraf, Q. and O. Galor (2013a). Genetic diversity and the origins of cultural fragmentation. *American Economic Review* 103(3), 528–533.
- Ashraf, Q. and O. Galor (2013b). The “Out of Africa” Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *Americam Economic Review* 103(1), 1–46.
- Battu, H. and Y. Zenou (2010). Oppositional Identities and Employment for Ethnic Minorities: Evidence from England. *Economic Journal* 120, F52–F71.
- Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2019). Unity in diversity? How inter-group contact can foster nation building. *The American Economic Review* 109(11), 3978–4025.
- Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2016). Bend It Like Beckham: Ethnic Identity and Integration. *European Economic Review* 90, 146–164.
- Bisin, A. and G. Tura (2025). Marriage, Fertility, and Cultural Integration in Italy. RF Discussion Paper 73/25, Rockwool Foundation.
- Bleakley, H. and A. Chin (2010). Age at arrival, English proficiency, and social assimilation among US immigrants. *American Economic Journal: Applied Economics* 2(1), 165–92.
- Borusyak, K., P. Hull, and X. Jaravel (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies* 89(1), 181–213.
- Borusyak, K., P. Hull, and X. Jaravel (2025). A practical guide to shift-share instruments. *Journal of Economic Perspectives* 39(1), 181–204.

- Bove, V., L. Elia, and M. Ferraresi (2023). Immigration, fear of crime, and public spending on security. *The Journal of Law, Economics, and Organization* 39(1), 235–280.
- Campo, F., M. Mendola, A. Morrison, and G. Ottaviano (2022). Talents and Cultures: Immigrant Inventors and Ethnic Diversity in the Age of Mass Emigration. *Journal of the European Economic Association* 20(5), 1971–2012.
- Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics* 19(1), 22–64.
- Carillo, M. R., V. Lombardo, and T. Venittelli (2023). Social identity and labor market outcomes of immigrants. *Journal of population economics* 36(1), 69–113.
- Clots-Figueras, I. and P. Masella (2013). Education, Language and Identity. *The Economic Journal* 123(570), F332–F357.
- Constant, A. F., S. Schüller, and K. F. Zimmermann (2024). Ethnic spatial dispersion and immigrant identity. *Journal of Chinese Economic and Business Studies* 22(2), 205–230.
- Constant, A. F. and K. F. Zimmermann (2008). Measuring Ethnic Identity and its Impact on Economic Behaviour. *Journal of the European Economic Association* 6, 424–433.
- Danzer, A. M., C. Feuerbaum, M. Piopiunik, and L. Woessmann (2022). Growing up in ethnic enclaves: language proficiency and educational attainment of immigrant children. *Journal of Population Economics* 35(3), 1297–1344.
- Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2017). Culture, ethnicity, and diversity. *American Economic Review* 107(9), 2479–2513.
- Esteban, J., L. Mayoral, and D. Ray (2012). Ethnicity and conflict: An empirical study. *American Economic Review* 102(4), 1310–42.
- Fearon, J. D. (2003). Ethnic and Cultural Diversity by Country. *Journal of Economic Growth* 8(2), 195–222.
- Fouka, V. (2020). Backlash: the Unintended Effects of Language Prohibition in US Schools after World War I. *The Review of Economic Studies* 87(1), 204–239.
- Fouka, V., S. Mazumder, and M. Tabellini (2022). From Immigrants to Americans: Race and Assimilation during the Great Emigration. *The Review of Economic Studies* 89(2), 811–842.
- Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review* 8(110), 2586–2624.
- Guiso, L., P. Sapienza, and L. Zingales (2008). Social capital as good culture. *Journal of the European economic Association* 6(2-3), 295–320.
- Jaeger, D. A., J. Ruist, and J. Stuhler (2018). Shift-Share Instruments and The Impact of Immigration. NBER Working paper 24285, National Bureau of Economic Research.
- Lazear, E. P. (1999). Culture and Language. *Journal of Political Economy* 107(S6), S95–S126.
- Lowe, M. (2024). Has intergroup contact delivered? *Annual Review of Economics* 17(17), 321–344.

- Meng, X. and R. G. Gregory (2005). Intermarriage and the economic assimilation of immigrants. *Journal of Labor economics* 23(1), 135–174.
- Montalvo, J. G. and M. Reynal-Querol (2005). Ethnic polarization, potential conflict, and civil wars. *American economic review* 95(3), 796–816.
- Nannicini, T., A. Stella, G. Tabellini, and U. Troiano (2013). Social capital and political accountability. *American Economic Journal: Economic Policy* 5(2), 222–250.
- Piracha, M., M. Tani, Z. Cheng, and B. Z. Wang (2023). Social Assimilation and Immigrants' Labour Market Outcomes. *Journal of Population Economics* 36(1), 37–67.
- Sanderson, E. and F. Windmeijer (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of econometrics* 190(2), 212–221.
- Simons, G. F. and C. D. Fennig (2017). *Ethnologue: Languages of the World* (20th ed.). Dallas, Texas: SIL International.
- Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference in Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Chapter 5, pp. 80–108. Cambridge: Cambridge University Press.
- Tabellini, M. (2020). Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration. *The Review of Economic Studies* 87(1), 454–486.

APPENDIX

A Additional Figures and Tables

FIGURE A.1: MIGRANTS' SHARE BY COUNTRY OF ORIGIN

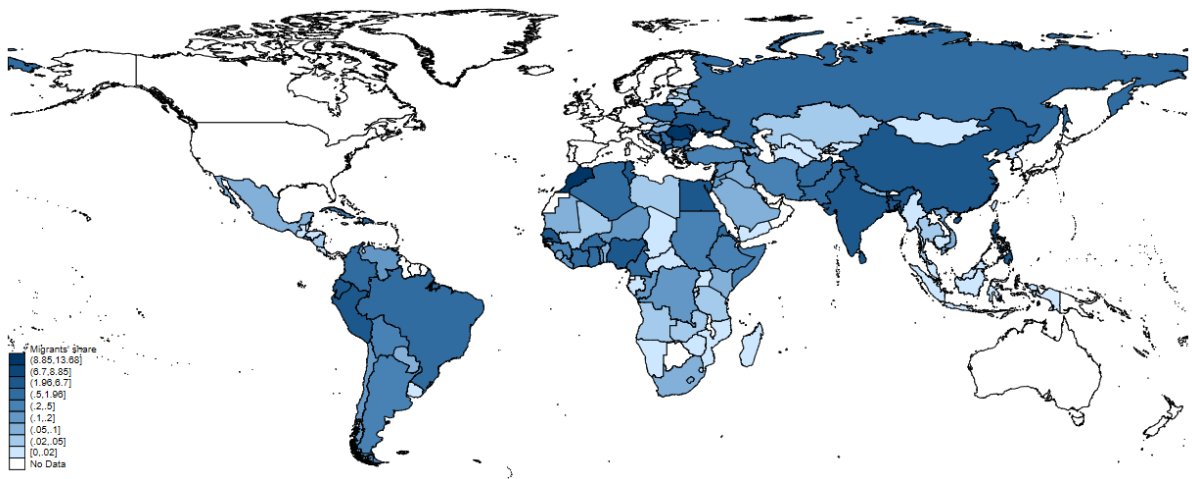


FIGURE A.2: MUNICIPALITY LOCATION AND MIGRANTS' SHARE

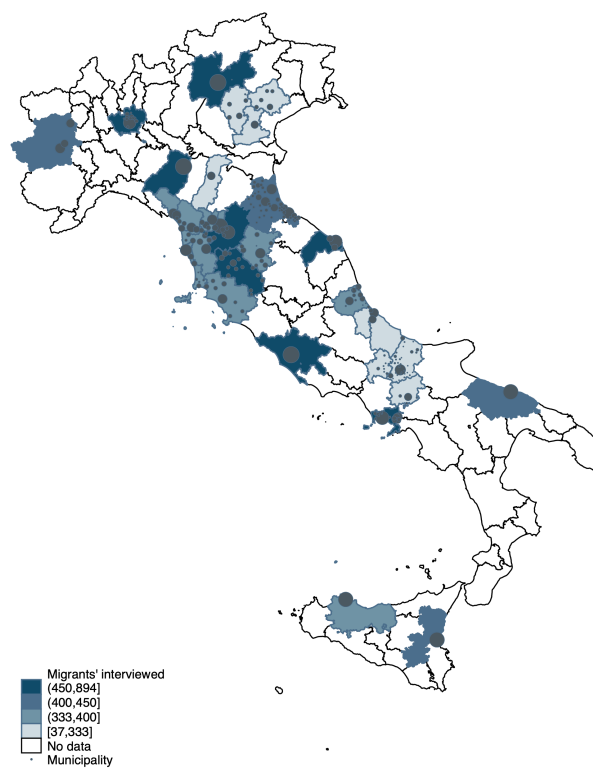
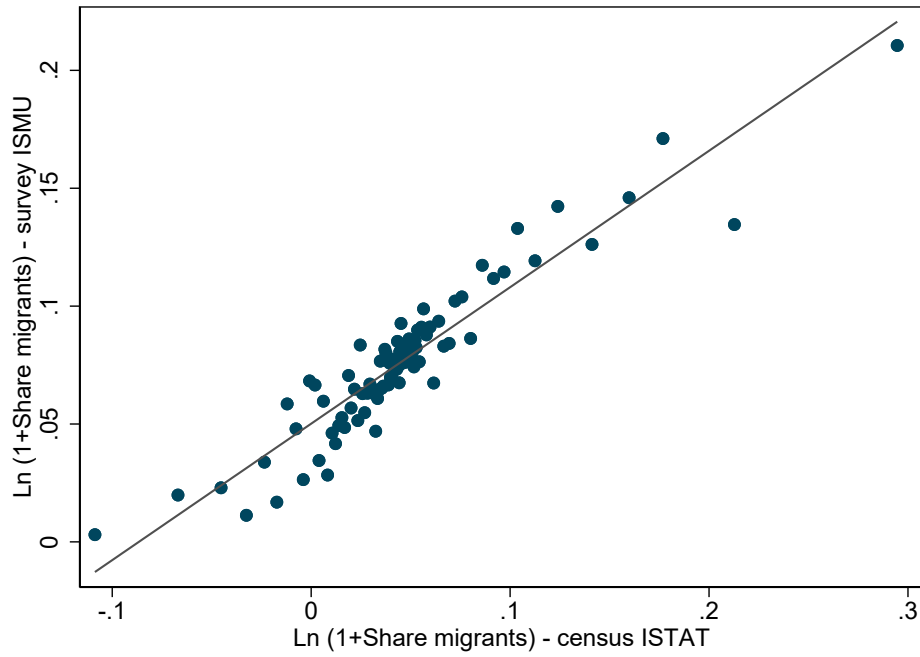


FIGURE A.3: DATA VALIDATION



Notes. The figure displays the binned scatter of the log of (1+) the share of migrants from country of origins o in municipality m as computed from the survey ISMU on the log of (1+) the share of migrants from country of origins o resident in municipality m from the official registry (ISTAT).

FIGURE A.4: ETHNIC GROUPS' SHARES AND DIVERSITY

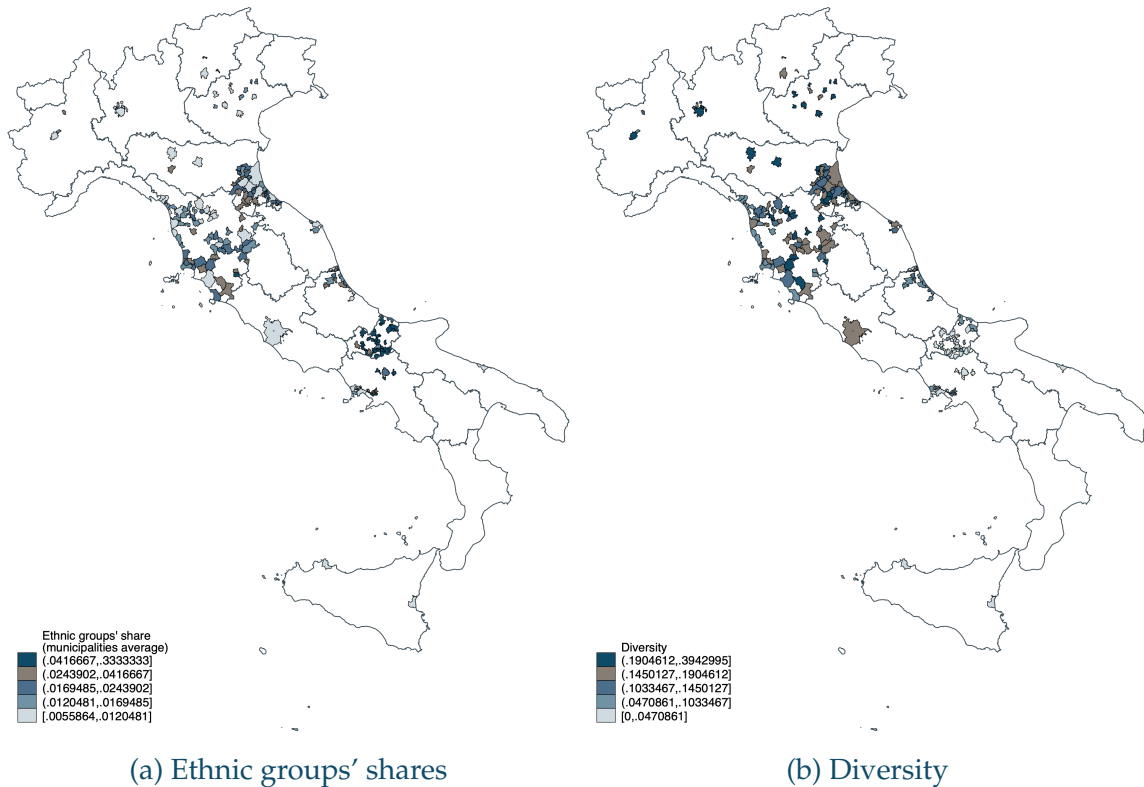
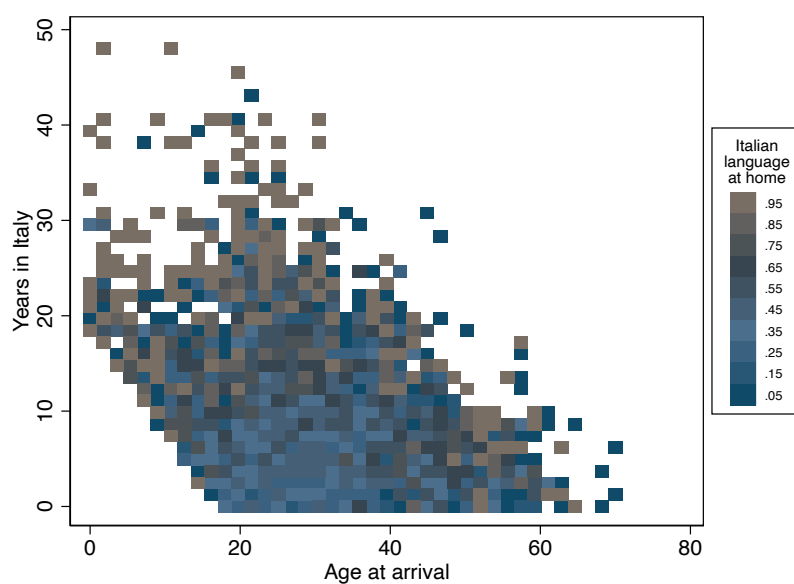


FIGURE A.5: AGE AT ARRIVAL AND YEARS IN ITALY INTERPLAY



Notes. The figure displays the distribution of the Italian language at home for different bins of age at arrival and years spent in Italy.

FIGURE A.6: INSTRUMENT ROBUSTNESS: COUNTRY EXCLUSION

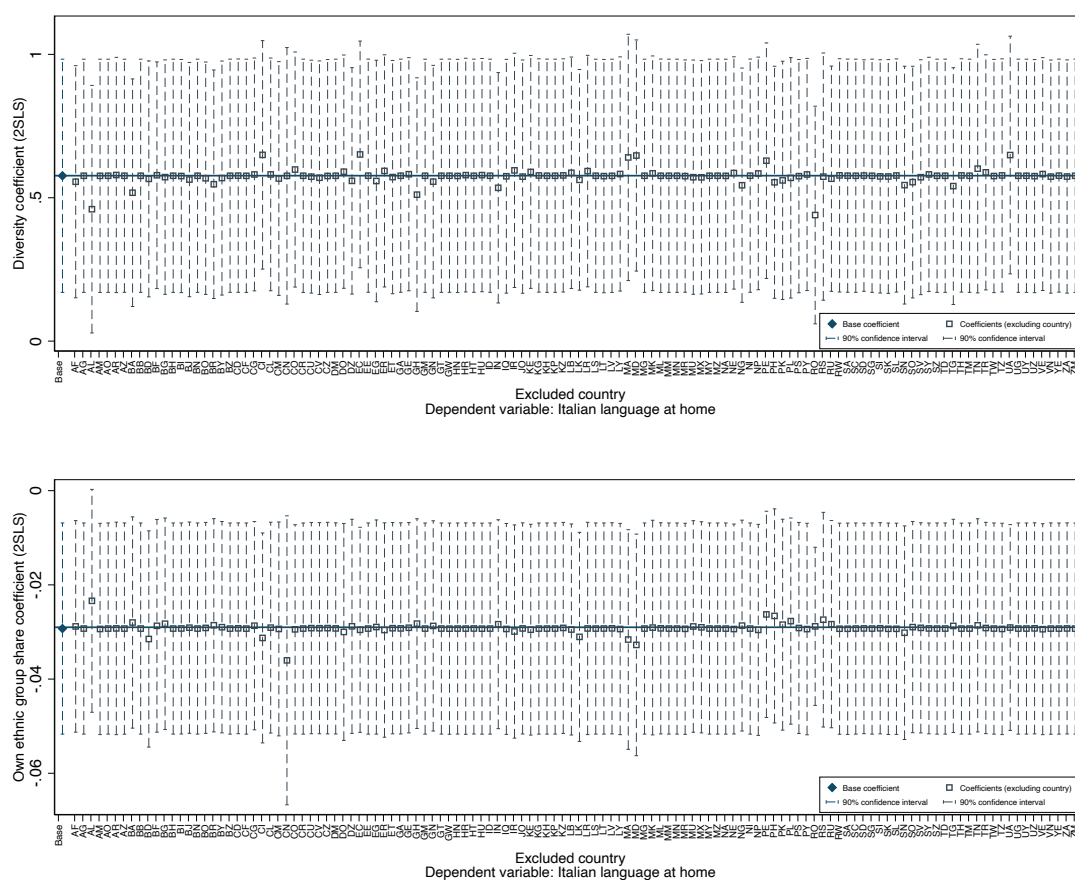


FIGURE A.7: INSTRUMENT ROBUSTNESS: PLACEBO FIRST-STAGE REGRESSIONS

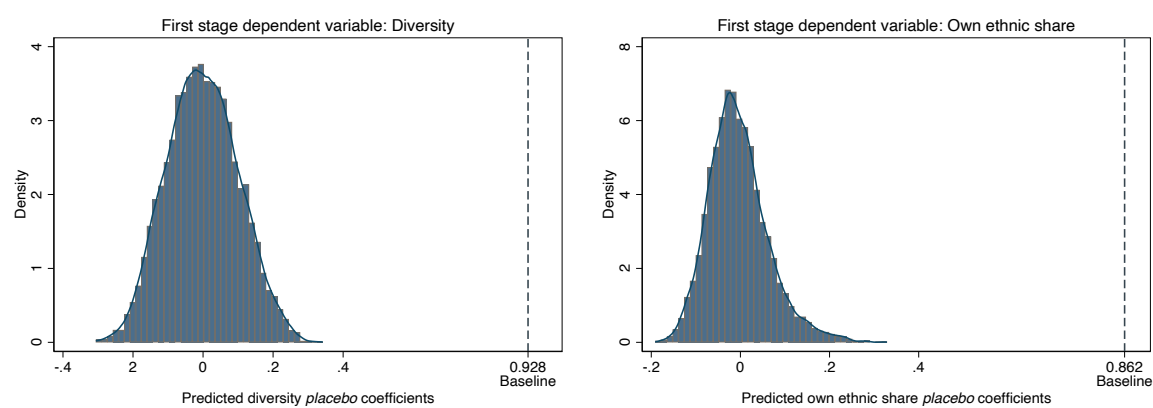


TABLE A.1: DATA VALIDATION

Source	The 10 most representative country in the survey ISMU (%)									
	Romania	Albania	Morocco	China	Philippines	Peru	Ukraine	Egypt	Bangladesh	Senegal
ISMU	(1st)	(2nd)	(3rd)	(4th)	(5th)	(6th)	(7th)	(8th)	(9th)	(10th)
	13.68	10.76	8.85	5.82	4.11	4.03	3.98	3.47	3.36	3.30
REGISTRY	1st	2nd	3rd	4th	6th	12th	5th	13th	18th	16th
	18.22	11.71	10.66	4.56	3.08	2.06	3.88	2.03	1.61	1.82

Notes. The Table reports the total population shares in p.p. and ranking of the ten most representative countries in the ISMU survey and the corresponding percentages and ranking of the countries from the official registry.

TABLE A.2: SUMMARY STATISTICS: INDIVIDUALS

	Observations	Mean	Std. Dev.	Min	Max
Italian language at home	11548	0.451	0.498	0	1
Host identity	11721	0.567	0.496	0	1
Glad to live in Italy	11949	0.689	0.463	0	1
Daughter intermarriage (approval)	10656	0.713	0.452	0	1
Host identity Index (PCA)	10030	0.500	0.500	0	1
Home identity	11870	0.908	0.289	0	1
<i>Places/Centers frequented by immigrants</i>					
Ethnic shops	12021	0.492	0.500	0	1
Service centers	12021	0.448	0.497	0	1
Assistance centers	12021	0.601	0.490	0	1
Local markets	12021	0.646	0.478	0	1
Shopping malls	12021	0.749	0.434	0	1
Workplaces	12021	0.331	0.471	0	1
Open places	12021	0.641	0.480	0	1
Private residence	12021	0.509	0.500	0	1
Training centers	12021	0.290	0.454	0	1
Worship centers	12021	0.481	0.500	0	1
Leisure	12021	0.564	0.496	0	1
Associations	12021	0.197	0.398	0	1
<i>Friends' type</i>					
Foreign	11883	0.168	0.374	0	1
Italian	11883	0.156	0.363	0	1
Both	11883	0.676	0.468	0	1
Male	11964	0.523	0.500	0	1
Single	11856	0.342	0.475	0	1
Married	11856	0.567	0.495	0	1
Widowed	11856	0.025	0.155	0	1
Divorced	11856	0.066	0.248	0	1
Having children	11919	0.567	0.495	0	1
Years in Italy	11905	7.788	6.191	0	60
Age at arrival	11862	28.53	9.752	0	82
Regular (with Legal permit)	11647	0.898	0.302	0	1
Born in Italy	11648	0.006	0.078	0	1
Proficiency in Italian	12001	3.503	1.094	1	5
No education	11675	0.062	0.242	0	1
Compulsory	11675	0.318	0.466	0	1
High school	11675	0.430	0.495	0	1
BA degree +	11675	0.190	0.393	0	1
Muslim	11596	0.339	0.474	0	1
Catholic	11596	0.266	0.442	0	1
Orthodox	11596	0.212	0.409	0	1
Coptic	11596	0.004	0.062	0	1
Evangelical	11596	0.022	0.147	0	1
Other christian	11596	0.019	0.137	0	1
Buddhist	11596	0.034	0.182	0	1
Hindu - Sikh	11596	0.021	0.145	0	1
Other	11596	0.009	0.094	0	1
No religion	11596	0.073	0.261	0	1
Employed	10191	0.817	0.387	0	1
Agriculture	8931	0.026	0.160	0	1
Industry	8931	0.190	0.392	0	1
Commerce	8931	0.203	0.402	0	1
Service firms	8931	0.119	0.324	0	1
Service people	8931	0.296	0.456	0	1
Other	8931	0.166	0.372	0	1
Income classes	10467	2.493	1.891	0	7
Spouse same country	6326	0.791	0.406	0	1
Spouse Italian	6326	0.162	0.368	0	1
Spouse Italian or other country	6326	0.209	0.406	0	1
Living alone	11857	0.197	0.398	0	1
Living with children	11857	0.358	0.479	0	1
Living with spouse same country	11857	0.449	0.497	0	1

TABLE A.3: SUMMARY STATISTICS: MUNICIPALITIES

	Observations	Mean	Std. Dev.	Min	Max
Diversity	233	0.126	0.075	0.000	0.394
Initial Diversity in 2003	232	0.076	0.051	0.000	0.267
Predicted diversity	232	0.117	0.073	0.000	0.356
Own ethnic group share	2664	0.345	0.714	0.000	8.475
Initial own ethnic group share in 2003	2655	0.191	0.433	0.000	6.545
Predicted own ethnic share	2655	0.315	0.718	0.000	10.629
City population	233	57435.811	212826.584	232	2604557
Initial city population in 2003	232	56863.078	210728.844	266	2552415
Ethnic population	2664	309.656	1679.849	0.000	41391
Initial ethnic shares in 2003	2655	0.721	2.956	0.000	45.556
City foreign population	233	4453.202	19752.468	0.000	218426
Initial city foreign population in 2003	232	2629.353	11521.667	0.000	122758
Foreign density	233	61.677	132.737	0.000	970.678
Initial foreign density in 2003	232	34.479	72.843	0.000	596.978
Income per-capita	233	12349.939	3238.266	5510.971	23745.592
Initial income per-capita in 2003	232	10692.935	3039.937	4417.265	20905.885
N. non-profit organizations per-capita (x1000)	231	4.785	2.126	0.672	12.177
Share high educated	232	30.822	7.143	12.800	49.000
Unemployment rate	232	9.718	8.054	2.600	47.900
Share workers in agriculture	232	8.193	7.522	0.600	34.100
Share workers in industry	232	34.898	9.817	16.200	64.400
Share workers in service	232	37.151	9.996	18.400	64.300
Share workers in commerce	232	19.760	5.043	2.700	40.000
Area (square km)	233	78.298	114.580	2.943	1286.601
Mean elevation	233	248.569	255.446	1.493	1554.165
Slope	233	454.712	411.161	6.000	2023.000
Coastal location	233	0.292	0.456	0.000	1.000

TABLE A.4: IV VALIDITY: LOCAL CHARACTERISTICS, INITIAL SHARES, AND DIVERSITY

	Dependent variables			
	Ethnic shares in 2003	Ethnic population in 2003	Diversity in 2003	Predicted Diversity
Initial municipality characteristics	(1)	(2)	(3)	(4)
Log Population	0.470*** (0.146)	0.905*** (0.063)	-0.002 (0.004)	-0.002 (0.005)
Log p.c. Income	1.025 (1.324)	-0.865 (0.607)	-0.004 (0.051)	-0.018 (0.070)
Share high educated	0.062 (0.041)	0.011 (0.017)	0.001 (0.001)	0.001 (0.001)
Unemployment rate	0.021 (0.024)	-0.015 (0.017)	0.000 (0.001)	0.000 (0.001)
Share workers in agriculture	0.747 (0.875)	-0.584 (0.607)	-0.050 (0.050)	-0.087 (0.070)
Share workers in industry	0.692 (0.879)	-0.600 (0.607)	-0.051 (0.050)	-0.087 (0.070)
Share workers in service	0.635 (0.886)	-0.607 (0.607)	-0.052 (0.051)	-0.089 (0.070)
Share workers in commerce	0.682 (0.881)	-0.597 (0.606)	-0.050 (0.050)	-0.086 (0.070)
P.c. non-profit organizations (x 1000)	-0.025 (0.039)	0.015 (0.030)	0.004 (0.002)	0.004 (0.003)
Mean elevation	0.001 (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Slope	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Coastal location	-0.094 (0.131)	0.077 (0.117)	-0.005 (0.007)	-0.009 (0.010)
Country of origin FE	yes	yes	–	–
Province FE	yes	yes	yes	yes
R^2	0.522	0.704	0.621	0.647
Observations	2654	2654	231	231
Sample	Municipality–Country of origin		Municipality	

Notes. In columns (1)–(2), the sample is composed by municipality (231)–country of origin (125) observations. In columns (1), the dependent variable is $100 \times N_{om}^{2003} / N_{o(IT)}^{2003}$. In columns (2), the dependent variable is $\log(1 + N_{om}^{2003})$, where N_{om}^{2003} is the level of the municipality–country of origin population in 2003. In columns (3)–(4) the sample is composed by the municipality observations only. All the controls refer to year 2001, but the population and per-capita income that refer to the year 2003. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.5: IV VALIDITY: LOCAL CHARACTERISTICS AND INITIAL SHARES OF TOP 10 COUNTRIES

	Dependent variables: 2003 ethnic shares of top 10 countries									
	Romania	Albania	Morocco	China	Philippines	Peru	Ukraine	Egypt	Bangladesh	Senegal
Initial municipality characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log Population	0.192* (0.100)	0.096*** (0.021)	0.120** (0.053)	0.299* (0.168)	0.549 (0.428)	0.587* (0.340)	0.105*** (0.031)	0.628 (0.504)	0.138* (0.075)	0.161*** (0.048)
Log p.c. Income	-0.254 (0.399)	0.014 (0.141)	-0.070 (0.257)	0.970 (1.193)	2.956 (2.980)	1.720 (2.255)	0.686 (0.490)	3.157 (3.516)	0.963 (0.650)	0.236 (0.322)
Share high educated	-0.002 (0.006)	-0.004 (0.003)	-0.001 (0.004)	-0.004 (0.013)	-0.004 (0.028)	-0.012 (0.024)	-0.014 (0.011)	-0.010 (0.035)	-0.015 (0.010)	-0.004 (0.006)
Unemployment rate	0.006 (0.006)	0.000 (0.002)	0.003 (0.004)	0.007 (0.013)	0.026 (0.027)	0.021 (0.022)	0.002 (0.003)	0.028 (0.031)	0.003 (0.007)	0.004 (0.005)
Share workers in agriculture	0.917 (0.714)	0.233 (0.159)	0.526 (0.344)	0.136 (0.648)	0.450 (1.138)	1.335 (1.225)	0.070 (0.132)	0.499 (1.391)	-0.720 (0.490)	0.087 (0.309)
Share workers in industry	0.910 (0.711)	0.230 (0.158)	0.521 (0.343)	0.114 (0.648)	0.402 (1.147)	1.294 (1.228)	0.066 (0.132)	0.447 (1.401)	-0.730 (0.491)	0.083 (0.309)
Share workers in service	0.908 (0.710)	0.231 (0.158)	0.519 (0.343)	0.104 (0.649)	0.373 (1.148)	1.277 (1.226)	0.066 (0.132)	0.414 (1.402)	-0.734 (0.492)	0.079 (0.309)
Share workers in commerce	0.914 (0.712)	0.229 (0.158)	0.521 (0.344)	0.119 (0.647)	0.395 (1.144)	1.292 (1.227)	0.068 (0.132)	0.443 (1.398)	-0.727 (0.488)	0.083 (0.309)
P.c. non-profit organizations (x1000)	0.001 (0.022)	0.014** (0.006)	0.005 (0.012)	0.052 (0.037)	0.088 (0.085)	0.061 (0.069)	0.012* (0.006)	0.097 (0.098)	0.032* (0.019)	0.014 (0.012)
Mean elevation	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.001 (0.001)	0.002 (0.002)	0.002 (0.002)	0.000* (0.000)	0.003 (0.002)	0.000 (0.000)	0.000 (0.000)
Slope	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.000)	-0.001 (0.001)	-0.000 (0.000)	-0.000* (0.000)
Coastal location	-0.072 (0.057)	-0.019 (0.021)	-0.052* (0.030)	-0.051 (0.074)	-0.064 (0.145)	-0.112 (0.143)	0.024 (0.029)	-0.109 (0.169)	-0.063 (0.052)	-0.051 (0.052)
Observations	231	231	231	231	231	231	231	231	231	231
R ²	0.719	0.656	0.455	0.465	0.478	0.445	0.812	0.257	0.898	0.410
Mean dependent	0.13	0.08	0.07	0.19	0.27	0.25	0.08	0.22	0.20	0.10

Notes. The dependent variables are the municipality shares, $100 \times N_{pm}^{2003} / N_{o(itr)}^{2003}$, of the ten most representative ethnic groups at national level. All regressions include province fixed effects. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.6: DIVERSITY AND ASSIMILATION: INITIAL CITY CONTROLS

	Municipality controls		
	Base	Extended	Initial
	(1)	(2)	(3)
Diversity	0.577** (0.246)	0.492* (0.270)	0.442* (0.262)
Own ethnic group share	-0.029** (0.014)	-0.031** (0.014)	-0.031** (0.014)
Log Population	0.009 (0.013)	0.018 (0.013)	0.016 (0.013)
Log Foreign density	-0.014 (0.019)	-0.009 (0.021)	-0.005 (0.020)
Log p.c. Income	0.116 (0.105)	0.087 (0.172)	0.165 (0.159)
P.c. non-profit organizations (x1000)	-0.009 (0.009)	-0.003 (0.009)	-0.004 (0.010)
Mean elevation	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Slope	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Coastal location	0.012 (0.028)	0.009 (0.032)	0.008 (0.032)
Share high educated		0.002 (0.005)	-0.000 (0.005)
Share workers in agriculture		-0.147 (0.161)	-0.124 (0.159)
Share workers in industry		-0.143 (0.160)	-0.120 (0.159)
Share workers in service		-0.146 (0.160)	-0.123 (0.159)
Share workers in commerce		-0.142 (0.161)	-0.119 (0.159)
Unemployment rate		-0.006 (0.007)	-0.006 (0.007)
Observations	9288	9288	9288
Kleibergen-Paap F-statistic	446.57	271.88	179.89

Notes. The dependent variables is a dummy for the use of Italian language at home. The specifications of all regressions are based on the base specification of column (6) of Table 2 presented in column (1), and include country of origin, province, day of week $\times$ place of interview and age at arrival $\times$ years in Italy fixed effects, as well as individual controls. Individual controls are: *proficiency in Italian language*, *gender (male dummy)*, *having children*, *marital status*, *compulsory school*, *high school*, *BA degree +*, *religion and income dummies*. Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.7: IDENTIFICATION STRATEGY ROBUSTNESS: OTHER INSTRUMENTS

	Dependent variable: Italian language at home								
	No leave-out shift	Countries aggregation	Diversity w/out EU entry in			+ Excluding EU entry in			Push instrument
	(1)	(2)	2004	2007	2004-2007	2004	2007	2004-2007	(9)
Diversity	0.577** (0.251)	0.616** (0.248)	0.578** (0.251)	0.608** (0.256)	0.608** (0.261)	0.565** (0.259)	0.480* (0.246)	0.489* (0.254)	0.562* (0.298)
Own ethnic group share	-0.030** (0.014)	-0.017** (0.008)	-0.029** (0.014)	-0.029** (0.014)	-0.029** (0.014)	-0.028** (0.013)	-0.028*** (0.010)	-0.024** (0.010)	-0.029* (0.018)
Observations	9288	9288	9288	9288	9288	8998	8212	7922	8941
Kleibergen-Paap F-statistic	519.30	444.54	446.51	575.62	571.45	442.00	570.02	555.24	47.36

Notes. Two-stage least square (2SLS) regressions. The dependent variables is a dummy for the use of Italian language at home. The specifications of all regressions are based on column 6 of Table 2 and include country of origin, province, day of week $\times$ place of interview and age at arrival $\times$ years in Italy fixed effects, as well as individual and municipality controls. Individual controls are: *proficiency in Italian language*, *gender (male dummy)*, *having children*, *marital status*, *compulsory school*, *high school*, *BA degree +*, *religion and income dummies*. Municipality controls are: *population (log)*, *foreign density (log)*, *per-capita income (log)*, *per-capita number of no-profit associations (x1000)*, *mean elevation*, *slope*, *coastal location*. Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.8: ROBUSTNESS OF MAIN RESULTS TO INDIVIDUALS' CHARACTERISTICS

	Dependent variable: Italian language at home			
	Base	Dropping		
		Overall Proficiency	Irregular immigrants	Second generations
		(2)	(3)	(4)
Diversity	0.577** (0.246)	0.688** (0.278)	0.610*** (0.233)	0.588** (0.249)
Own ethnic group share	-0.029** (0.014)	-0.037** (0.015)	-0.027** (0.013)	-0.030** (0.014)
Observations	9288	9297	8293	9237
Kleibergen-Paap F-statistic	446.57	444.49	442.94	447.61
Mean dependent	0.445	0.445	0.460	0.445

Notes. Two-stages least square estimates (2SLS). The dependent variables is the dummy equal to one for the use of Italian language at home. The specifications of all regressions are based on column 6 of Table 2 and include country of origin, province, day of week $\times$ place of interview and age at arrival $\times$ years in Italy fixed effects, as well as individual and municipality controls. Individual controls are: *proficiency in Italian language*, *gender (male dummy)*, *having children*, *marital status*, *compulsory school*, *high school*, *BA degree +*, *religion and income dummies*. Municipality controls are: *population (log)*, *foreign density (log)*, *per-capita income (log)*, *per-capita number of no-profit associations (x1000)*, *mean elevation*, *slope*, *coastal location*. Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.9: ROBUSTNESS: TRIMMING AND SAMPLING

	Dependent variable: Italian language at home			
	Dropping			Provinces with less than 3 municipalities
	Countries	Municipalities	Age at arrival & Years in Italy	
		with less than 10 obs.		
	(1)	(2)	(3)	(4)
Diversity	0.593** (0.244)	0.694** (0.278)	0.453* (0.236)	0.530** (0.220)
Own ethnic group share	-0.030** (0.014)	-0.029** (0.014)	-0.028* (0.014)	-0.029* (0.016)
Observations (9288)	9127	9092	7357	7052
Countries of origin (118)	70	118	113	111
Municipalities (224)	224	157	218	212
Provinces (33)	33	33	33	23
Kleibergen-Paap F-statistic	448.18	420.28	434.80	435.21
Mean dependent	0.445	0.446	0.425	0.432

Notes. The dependent variables is the dummy equal to one for the use of Italian language at home. The specifications of all regressions are based on column 6 of Table 2 and include country of origin, province, day of week $\times$ place of interview and age at arrival $\times$ years in Italy fixed effects, as well as individual and municipality controls. Individual controls are: *proficiency in Italian language*, *gender (male dummy)*, *having children*, *marital status*, *compulsory school*, *high school*, *BA degree +*, *religion* and *income dummies*. Municipality controls are: *population (log)*, *foreign density (log)*, *per-capita income (log)*, *per-capita number of no-profit associations (x1000)*, *mean elevation*, *slope*, *coastal location*. Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.10: ROBUSTNESS: DIVERSITY MEASURES ACROSS OTHER AGGREGATIONS

	Dependent variable: Italian language at home					
	Administrative aggregations					
	Municipality		SLL		Province	Region
	(1)	(2)	(3)	(4)	(5)	(6)
Diversity	0.577 [0.038] (0.246)** {0.167}***	0.542 [0.035] (0.198)*** {0.188}***	1.833 [0.077] (0.569)*** {0.588}***	1.324 [0.056] (0.343)*** {0.355}***	1.480 [0.053] (0.514)*** {0.701}**	1.848 [0.052] (1.075)* {1.242}+
Own ethnic group share	-0.029 [-0.033] (0.014)** {0.012}**	-0.036 [-0.040] (0.012)*** {0.010}***	-0.041 [-0.037] (0.022)* {0.022}*	-0.049 [-0.044] (0.019)*** {0.018}***	-0.082 [-0.068] (0.020)*** {0.018}***	-0.117 [-0.071] (0.034)*** {0.035}***
Province FE	yes	no	yes	no	no	no
Municipality controls	yes	no	yes	no	no	no
Province controls	no	yes	no	yes	yes	yes
Observations	9288	9289	9288	9289	9303	9303
Kleibergen-Paap F-statistic	446.57	320.55	48.17	46.18	564.23	349.94

Notes. Two-stage least square (2SLS) regressions. The dependent variables is the dummy equal to one for the use of Italian language at home. All regressions include country of origin, day of week $\times$ place of interview and age at arrival $\times$ years in Italy fixed effects, as well as individual covariates. The individual controls are: *proficiency in Italian language, gender (male dummy), having children, marital status, compulsory school, high school, BA degree +, religion and income dummies*. Municipality controls are: *population (log), foreign density (log), per-capita income (log), per-capita number of no-profit associations (x1000), mean elevation, slope, coastal location*. Province controls are: *population (log), foreign density (log), per-capita income (log), per-capita number of no-profit associations (x1000), mean elevation*. Standardized coefficients in brackets report the change in the dependent variable for one standard deviation increase in the independent variable ($= \beta * s.d.$). Sampling weights used. Robust standard errors in parentheses are clustered at municipality level, while those in curly braces are clustered at the level of the administrative aggregation reported on the top of the table, except for columns (6). Specifically, in columns (1) and (2), robust standard errors are clustered only at the municipality level. In columns (3) and (4), robust standard errors in curly braces are clustered at the SLL administrative level (91 clusters). Finally, in columns (5) and (6), robust standard errors in curly braces are clustered at the province administrative level (33), due to the small number of regions (13). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, + $p < 0.15$.

TABLE A.11: MECHANISMS: ETHNIC COMPOSITION AND FRIENDS' NATIONALITY

	Dependent variable: Friends' nationality		
	Foreign	Italian	Both
	(1)	(2)	(3)
Diversity	-0.364** [-0.024] (0.168)	-0.083 [-0.005] (0.172)	0.447* [0.029] (0.258)
Own ethnic group share	0.018** [0.020] (0.008)	-0.021*** [-0.023] (0.008)	0.003 [0.003] (0.009)
Observations		9588	
Kleibergen-Paap F-statistic		446.90	
Mean dependent	0.170	0.151	0.679

Notes. Two-stages least square estimates (2SLS). The dependent variables are a dummy equal to one if immigrants have friends of nationality, respectively, exclusively foreigner, mainly Italian nationalities, or of both nationalities and zero otherwise. The specifications of all regressions are based on column 6 of Table 2 and include country of origin, province, day of week $\times$ place of interview and age at arrival $\times$ years in Italy fixed effects, as well as individual and municipality controls. Individual controls are: *proficiency in Italian language*, *gender (male dummy)*, *having children*, *marital status*, *compulsory school*, *high school*, *BA degree +*, *religion* and *income dummies*. Municipality controls are: *population (log)*, *foreign density (log)*, *per-capita income (log)*, *per-capita number of no-profit associations (x1000)*, *mean elevation*, *slope*, *coastal location*. Standardized coefficients in brackets report the change in the dependent variable for one standard deviation increase in the independent variable ($= \beta * s.d.$). Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.12: MECHANISMS: ETHNIC COMPOSITION AND PLACES POPULAR AMONG IMMIGRANTS

Dependent variable: Popular places/centers												
	Ethnic shops	Service centers	Assistance centers	Local markets	Shopping malls	Work places	Open places	Private residence	Training centers	Worship centers	Leisure	Associations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Diversity	-0.866** [-0.056] (0.373)	-0.655* [-0.042] (0.378)	-0.728** [-0.047] (0.307)	-0.714** [-0.046] (0.361)	-0.462 [-0.030] (0.316)	-0.085 [-0.006] (0.378)	-0.184 [-0.012] (0.321)	-0.191 [-0.012] (0.519)	-0.103 [-0.007] (0.232)	0.207 [0.013] (0.248)	0.149 [0.010] (0.344)	-0.248 [-0.016] (0.237)
Own ethnic group share	0.030** [0.034] (0.013)	-0.002 [-0.002] (0.010)	-0.008 [-0.009] (0.011)	0.017* [0.019] (0.009)	0.024** [0.027] (0.010)	0.031*** [0.035] (0.011)	0.017* [0.019] (0.009)	0.020 [0.023] (0.013)	-0.001 [-0.001] (0.008)	-0.011 [-0.012] (0.011)	0.012 [0.013] (0.012)	0.003 [0.004] (0.008)
Observations	9675											
Kleibergen-Paap F-statistic	447.21											
Mean dependent	0.494	0.454	0.606	0.650	0.756	0.337	0.652	0.507	0.284	0.493	0.569	0.195

Notes. Two-stages least square estimates (2SLS). The dependent variables are a dummy equal to one for each type of place usually frequented. The specifications of all regressions are based on column 6 of Table 2 and include country of origin, province, day of week $\times$ place of interview and age at arrival $\times$ years in Italy fixed effects, as well as individual and municipality controls. Individual controls are: *proficiency in Italian language, gender (male dummy), having children, marital status, compulsory school, high school, BA degree +, religion and income dummies*. Municipality controls are: *population (log), foreign density (log), per-capita income (log), per-capita number of no-profit associations (x1000), mean elevation, slope, coastal location*. Standardized coefficients in brackets report the change in the dependent variable for one standard deviation increase in the independent variable ($= \beta \times s.d.$). Sampling weights used. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Identification strategy and instrument validity

In this Appendix, I describe how I build the set of instruments based only on push factors in the country of origin of the immigrants (Alesina et al., 2016; Bove et al., 2023; Fouka et al., 2022). I begin to predict the total number of immigrants arrived in Italy in the period 2003-2008 by estimating a (“zeroth stage”) gravity model of migration

$$\ln I_{ot} = \alpha_o + \alpha_t + \beta \ln GDP_{pc_{ot}} + \gamma \ln Pop_{ot} + \delta \mathbf{Distance}_o \alpha_t + \varepsilon_{ot}, \quad (\text{B.1})$$

where $\ln I_{ot}$ is the log of the inflows of immigrants to Italy from each countries of origin $o = 1, \dots, 182$ in the years $t = 2003, \dots, 2008$.¹² The model in equation (B.1) takes into account several exogeneous factors that capture only shocks in the country of origin and are thus unlikely correlated with unobserved local and national factors in Italy. In particular, the specification exploits the panel structure of the data controlling for countries of origin (α_o) and years fixed effects (α_t), the (log) GDP per-capita and population in the countries of origin. I additionally control for two distance measures interacted with year dummies, the geographical distance from the capitals of each countries of origin to Italy, and the linguistic distance between each country of origin and Italy;¹³ these variables capture the changes in connections costs and cultural frictions over time induced also by the EU enlargement. After computing the predicted number of immigrants entering in Italy from each country of origin o in the year 2008, $\hat{I}_o^{2008}$, I apportion these predicted inflows across the municipalities using the initial local settlements shares ($N_{om}^{2003} / N_{o(IT)}^{2003}$), as in eq. (3). Then, using the municipalities’ population P_m^{2008} , I construct the predicted shares of inflows from country o in municipality m as follows:

$$\widehat{SI}_{om} = \frac{N_{om}^{2003}}{N_{o(IT)}^{2003}} \frac{\hat{I}_o^{2008}}{P_m^{2008}} \quad (\text{B.2})$$

The shares $\widehat{SI}_{om}$ are theoretically and mechanically different from those used in the main analysis because they are derived only from push factors that predicted movements of the immigrants out of their country of origin towards Italy. Finally, I define the predicted diversity as the fractionalization index applied to these predicted shares as follows:

$$\widehat{PushDiversity}_m = \sum_o \widehat{SI}_{om} (1 - \widehat{SI}_{om}) \quad (\text{B.3})$$

Column (9) of Table A.7 presents the results when substituting the baseline instruments in equations (??)-(4) with the push instruments in equations (B.2)-(B.3). Despite the differences in the construction of the instruments, the coefficients on *Diversity* and *Own ethnic group share* are statistically significant and remain remarkably close in magnitude to the main findings.

¹²Data on the inflows are from the OECD International Migration Database.

¹³Following the literature, I define the linguistic distance between each country of origin o and Italy as $\delta_{oIT} = 1 - (branch_{oIT} / \max(branch_o, branch_{IT}))^\kappa$, with $\kappa = (0.05, 0.5)$, using the the fraction of shared branches on linguistic classification trees from the Ethnologue database 20 (Simons and Fennig, 2017).

The Effect of the Tariff War between Italy and France on the Big Surge of Mass Migration

Maria Rosaria Carillo, Viviana De Falco, and Luca Pennacchio

Abstract

This paper investigates the role of Italian protectionist policies in the late nineteenth century, and in particular the tariff war with France (1888–1898), as a key driver of the Great Italian mass migration to the New World. Using a novel dataset that combines regional migration and agricultural production data (1881–1913), the authors show that regions most affected by tariffs—primarily in Southern Italy and specialized in export-oriented crops such as wine and olive oil—experienced significantly higher emigration rates. The empirical analysis, based on pre-post and difference-in-differences models, reveals that the impact of protectionism on migration was substantial but delayed, reflecting adjustment costs faced by agricultural producers. The loss of foreign markets, especially France, triggered a severe agricultural crisis that was not offset by domestic industrialization, pushing surplus labor to emigrate. The paper contributes to the literature on the causes of Italian mass migration by providing quantitative evidence of the link between trade policy and migration flows, and suggests that protectionism exacerbated the North–South divide while shaping Italy’s structural economic transformation.

1 Introduction

Mass migration stands as one of the most significant episodes in modern history, yet its causes remain understudied. Between 1850 and 1914, fifty-five million Europeans emigrated to the New World, profoundly reshaping global population dynamics and economic activity (Hatton and Williamson, 1998). A defining feature of this phenomenon is the dramatic rise in long-distance migration rates compared to earlier periods.

Beyond the global factors that facilitated mass migration — such as falling transportation costs, higher returns to migration especially in the United States, and open-border policies designed to attract immigrants¹ (Eriksson, 2018) — a crucial question remains: were there country-specific determinants that further accelerated the Great Migration? Italy represents an exemplary case, having been one of the countries most deeply affected by this phenomenon and one whose diaspora has attracted sustained scholarly attention (Petraglia and Vecchione, 2020; Prinz, 2019; Hatton, 2004; Reis, 2004; Vallette, 2018).

Nevertheless, the factors driving emigration from Italy have not yet been fully analyzed. This paper deepens the investigation of both the causes that triggered emigration and the government policies implemented at the end of the nineteenth century.

Drawing on a unique dataset combining regional-level data on migration and agricultural production, we provide evidence that the Italian protectionist policy introduced to shield infant industry from 1887 onward had a significant effect on regional migration. Specifically, our results indicate that the regions most severely affected by protectionism — concentrated almost entirely in Southern Italy — subsequently experienced higher rates of emigration. These effects become particularly pronounced from the second half of the 1890s onward, suggesting their long-lasting character and their connection to the great surge of migration flows observable from 1900 through World War I.

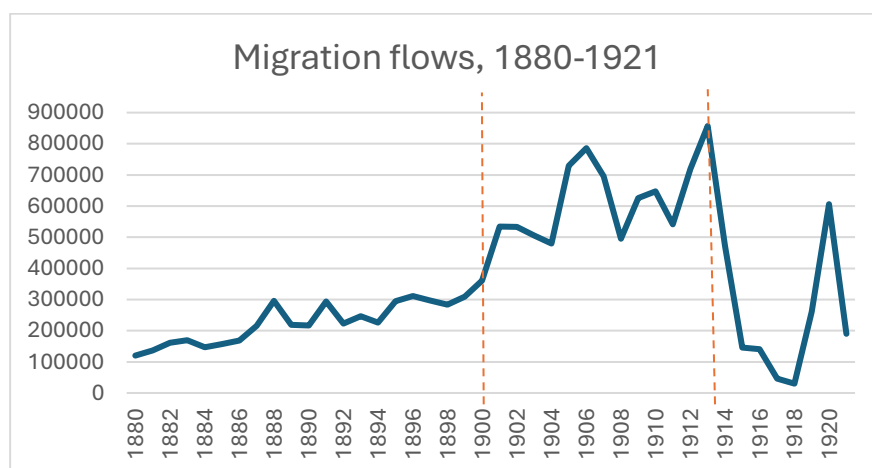
The literature on the long-run effects of mass migration spans two broad strands: impacts on sending countries and impacts on receiving countries. The effects on countries of origin primarily concern changes in native population composition, particularly questions of self-selection and brain drain, as well as the role of remittances. Evidence on self-selection has proven ambiguous (Abramitzky and Boustan, 2012; Haque and Kim, 1995): some studies find positive self-selection among migrants (Spitzer and Zimran, 2018; Beck-Knudsen, 2019), while

¹ South American countries — particularly Argentina and Brazil — adopted policies to attract European labor for agricultural development. Both countries subsidized passage tickets for migrants, primarily from Italy and Spain. While working conditions in Brazil were found to be harsh, Argentina offered greater social mobility and fiscal advantages such as exemption from land taxes. The United States imposed no immigration restrictions until the 1920s (Ferrie and Hatton, 2013; De Clementi, 2009).

others document negative self-selection (Beine et al., 2011; Fernández-Huertas Moraga, 2013; McKenzie and Rapoport, 2010). Research on brain drain — the loss of high-skilled individuals from origin countries — suggests that this process can be either a threat or an opportunity, depending on the strategies adopted by sending countries (Mountford, 1995; Haque and Kim, 1995; Bhagwati and Hamada, 1974). The literature on receiving countries predominantly examines the effects of mass migration on economic growth and cultural assimilation: the impact on wage convergence remains contested (Borjas, 2013; Ager and Hansen, 2017; Abramitzky et al., 2014), while cultural diversity appears to generate beneficial effects in labor markets (Sun et al., 2019; Ager and Brückner, 2013; Audetsch et al., 2009).

Despite this rich body of work, few studies focus specifically on Italian mass migration (among them, Petraglia and Vecchione, 2020; Hatton and Williamson, 1998; Moretti, 1999). To our knowledge, no empirical study has yet examined the motivations to migrate for Italians specifically during the period 1900–1913, when flows from Southern Italy reached a peak never seen before or since (Figure 1) — even though this constitutes one of the most debated topics in Italian economic history. The main objective of this study is accordingly to shed light on one possible driver of this migration wave: the industrial policy adopted during the 1880s to support Italy’s still-fragile infant industries. The Italian government introduced protectionism to shield industrial production, at the cost of sacrificing the agricultural exports of the country’s most innovative crops.²

Figure 1. Italian migration flows, 1880-1921



Source: Annuario dell’Emigrazione Italiana 1876-1925

² Italy’s principal agricultural exports at the time were high-value products such as citrus fruits, vegetables, olive oil, and wine, predominantly grown in the South.

Building on this premise, we empirically investigate the impact of France's 1889 retaliatory tariffs — a direct response to Italy's Tariff Law of 1887, which aimed to protect rising domestic industry — on Italian migration. We employ agricultural data from the *Annuario Statistico Italiano* covering 1881 to 1913, which reports regional-level information on agricultural production across the sixteen historical Italian regions. In particular, we focus on the crops most severely harmed by protectionism, namely wine and olive oil, both predominantly produced in Southern Italy. We further draw on historical migration data from the *Annuario Statistico dell'emigrazione Italiana*, also at the regional level, for the same period (1881–1913). To assess whether the regions most damaged by protectionism subsequently experienced higher migration growth, we estimate two different models — one in the spirit of a before-and-after approach and one following a difference-in-differences (DiD) framework — treating 1889 (the year of France's retaliatory tariffs and the ensuing mutual escalation) as the watershed. In both cases, the outcome variable is the regional migration growth rate. It is important to note that the protectionist reform constitutes an exogenous shock to migration: it was designed to protect Italian infant industry, and the subsequent increase in emigration flows was an unintended consequence. The mechanism we exploit is the heterogeneous exposure of Italian regions to the tariff shock: regions whose economies depended on exports of affected products suffered the most, and their inhabitants consequently sought better opportunities abroad. Crucially, the decision to migrate did not occur simultaneously with the adoption of the mutual tariffs; it was delayed by the time needed to liquidate land holdings or accumulate the capital necessary for departure.

Our empirical evidence shows that regions most specialized in olive oil and wine production experienced higher migration growth rates, particularly in the medium term. These results point to a structural transformation involving the exit of labor from the agricultural sector — surplus that was not fully absorbed by Italy's nascent industrial sector — which in turn pushed many agricultural workers to migrate, especially to the New World.

At the same time, the tariff policy likely foreclosed pathways to Southern growth not only through emigration, but also through the erosion of capital accumulation from the most innovative and profitable crops, as many farmers abandoned or sold their landholdings.

This investigation yields at least two major contributions. First, it provides quantitative support for a strand of literature that had previously addressed the relationship between protectionism and Italian mass migration only in qualitative terms. Second, it introduces a unique dataset combining historical agricultural and migration data that had not previously been digitized.

The paper is organized as follows. Section 2 is divided into three subsections reconstructing the historical context of late nineteenth-century Italy: subsection 2.1 introduces Italian mass migration, subsection 2.2 reviews the historical debate on tariff policy, and subsection 2.3 describes the events leading to the Tariff Law of 14 July 1884 and the ensuing Tariff War. Section 3 describes the data sources and provides descriptive evidence on the relationship between migration and productive specialization. Section 4 presents the empirical strategies used to investigate the central hypothesis: the pre-post analysis (section 4.1) and the difference-in-differences approach (section 4.2), with robustness checks in section 4.2.2. Section 5 concludes.

2. Historical background

2.1. *The characteristics of Italian mass migration*

Italian mass migration unfolded from the late nineteenth century (1880s) through World War I. The period from 1900 to 1913 is characterized by a steep upward trend in gross migration, marked by increasing fluctuations (Hatton and Williamson, 1998).

Italian mass migration between 1880 and 1925 can be divided into three distinct phases, during which the total number of Italians living abroad reached a peak of 14 million (Sanfilippo, 2009). During the first phase (1880–1900), approximately 14,399 people left Italy each year. In the second phase (1901–1913), this figure grew dramatically, nearly tripling to an annual average of 39,150 emigrants. The third phase (1914–1925), coinciding with World War I and its aftermath, saw a decline, with an average of 16,536 migrants per year (Figure 1). The trend was not steady, but rather characterized by alternating periods of growth and stagnation.

The central phase of Italian migration has been widely recognized as the *peak* or *big surge* of mass migration (Hatton and Williamson, 1998; Petraglia and Vecchione, 2021; Sanchez-Alonso, 2000; Ager, 2017; Spitzer and Zimran, 2018; Del Boca and Venturini, 2003). Southern regions dominated this movement, both in absolute levels and as a share of total migration.

Table 1 presents the regional distribution of annual migration rates (calculated by dividing regional emigrant counts by regional population). In the first phase, 0.8% of the population left Northern and Central Italy, with Veneto (3.45%) and Piedmont (1.2%) displaying the highest rates, while the Southern average was slightly lower (0.7%), though Calabria and Basilicata were exceptions. In the second phase, Southern regions drove overall migration (2.27% of the Southern population emigrated), with nearly every region at least doubling its emigration rate. In the final phase, emigration rates stood at 0.77% for the South and 0.68% for the North, with

Piedmont (1.23%), Veneto (1.17%), Abruzzi and Molise (1.04%), and Calabria (1.15%) posting the highest shares.

Table 1. Migration share on regional population, 1880-1925, average per year

Region	First period (1880-1900)	Second period (1901-1913)	Third period (1914-1925)
Piedmont	1.2%	2.14%	1.23%
Liguria	0.54%	0.65%	0.56%
Lombardy	0.5%	1.13%	0.56%
Veneto	3.45%	3.86%	1.17%
Emilia-Romagna	0.45%	1.36%	0.42%
Tuscany	0.57%	1.30%	0.56%
The Marches	0.35%	2.18%	0.74%
Umbria	0.07%	1.65%	0.49%
Lazio	0.08%	1.11%	0.28%
Abruzzi and Molise	0.87%	3.34%	1.04%
Campania	0.83%	2.16%	0.69%
Apulia	0.15%	1.19%	0.49%
Basilicata	1.65%	2.93%	0.96%
Calabria	1.02%	3.18%	1.15%
Sicily	0.37%	2.30%	0.80%
Sardinia	0.06%	0.80%	0.31%
Northern and Central Italy	0.80%	1.70%	0.68%
Southern Italy and Island	0.70%	2.27%	0.77%

Notes: The statistics involve 9 regions for Northern and Central Italy (Liguria, Lombardy, Veneto, Emilia-Romagna, Tuscany, the Marches, Umbria, Lazio) and 7 regions for Southern Italy (Abruzzo and Molise, Campania, Apulia, Basilicata, Calabria, Sicily and Sardinia.).

Examining each phase more closely reveals several noteworthy characteristics. Table 2 reports migration features with respect to gender, age, family status, and occupation, as well as region of origin and the share of migrants who traveled overseas. To test for differences across the three periods, we use the chi-squared test, given that the data are expressed as percentages.

Table 2. Italian Migration - Migrants' Individual Characteristics

	First period (1880-1900)	Second period (1901-1913)	Third period (1914-1925)
Characteristics	Average (%)	Average (%)	Average (%)
Male	88.30***	83.73***	73.17***
Over 15	86.32***	89.23***	87.08***
Leaving alone	63.40***	79.07***	74.02***
Employees in agriculture	49.26***	35.81***	24.13***
Daily workers	23.89***	29.68***	20.76***
Employees in secondary sector	21.63***	24.67***	30.94***
Share of migrants from Northern Italy	62.30***	41.07***	48.95***
Share of migrants from Southern Italy	30.36***	45.93***	39.32***
Transoceanic Destinations Share	62.99***	60.08***	49.28***

Notes: chi-squared (mean difference between the first and the second period). ***, **, * statistically significant at 1%, 5% and 10%.

The tests confirm that the three phases are statistically distinct from one another.

During the first phase, migration flows grew steadily. Liberal migration policy facilitated both spontaneous and irregular emigration (the 1888 Crispi Law formally guaranteed freedom of migration). Approximately 5,300,000 people left the country, with the sharpest growth occurring between 1885 and 1889 (Bevilacqua et al., 2009). The typical migrant was male (88.3%), over 15 years of age (86.32%), and emigrated alone (63.64%); nearly half were employed in agriculture (49.6%). Northern Italian regions³ accounted for 62.30% of migrants. Transoceanic destinations were preferred over European and Mediterranean ones (62.99% crossed the ocean), though many migrants chose countries engaged in large public works

³ Here Northern and Central regions refer to Piedmont, Liguria, Lombardy, Veneto, Emilia-Romagna, Tuscany, the Marches, Umbria, and Lazio; Southern regions are Abruzzo and Molise, Campania, Apulia, Basilicata, Calabria, Sicily, and Sardinia.

projects, where employment as unskilled labor was readily available, as Golini and Amato (2001) note.

The second phase (1901 to the eve of World War I) brought a slight decline in the share of male emigrants (83.73%), but an increase in those migrating alone (79.07%). Sanfilippo (2011) attributes this to migrant networks that helped compatriots join established communities abroad.⁴ The overall share of agricultural workers among emigrants fell to 35.81%, though Southern agricultural migrants outnumbered their Northern counterparts in absolute terms. Workers from the secondary sector (24.67%) continued to be attracted by large public works, including in North Africa, where Italian emigration had previously been limited (Cresti, 2008). The share preferring transoceanic destinations remained broadly stable but grew substantially in absolute terms; approximately 9,000,000 Italians emigrated during this phase, with 1913 marking the single highest year on record (870,000 departures). The surge in emigration prompted the enactment of Law 1901/23, which established the *Commissariato Generale dell'Emigrazione*, a government agency tasked with regulating and protecting emigrants — both from exploitative travel agents and couriers, and to a limited degree in cases of hardship abroad.

The third phase (1914–1925) opened with a near-complete halt to emigration during World War I. Migration continued to decline in the 1920s owing to restrictions imposed by receiving countries — including Argentina, Brazil, and above all the United States Quota Acts of 1921 and 1924 — as well as Fascist-era policies. Temporary migration was tolerated, but permanent emigration was initially frowned upon (colonial emigration excepted) and later prohibited outright. The *Commissariato Generale dell'Emigrazione* was abolished, and migration affairs were transferred to the Ministry of Foreign Affairs through its *Direzione Generale per gli Italiani all'Esterro*.

The share of female migrants rose to approximately 27%, and the proportion of under-15 migrants also increased, likely reflecting a process of family reunification. The share of emigrating families rose slightly compared to the previous phase (around 26%), while the proportion of agricultural workers fell to 24.13% — a trend perhaps partly attributable to the Fascist regime's promotion of land reclamation (*bonifica integrale*) (Golini and Amato, 2009). The share of Southern migrants on total emigration fell to 39.32%, and flows were increasingly directed toward European and Mediterranean destinations (50.72%).

⁴ The role of migrant networks will be examined further in the fourth essay of this volume.

Table 3 examines geographical heterogeneity across the three phases, comparing Northern-Central and Southern Italy. Several features stand out. The South consistently supplied the highest proportion of agricultural workers, particularly through 1913 (43.86% and 41.42% respectively). In the second phase, the share of Southern migrants choosing transoceanic destinations nearly doubled (82.15%) relative to their Northern counterparts (42.91%). Geographic proximity to Europe led Northern migrants to favor European destinations, while Southern migrants benefited from cheaper fares at Southern ports (Sanfilippo, 2009). Southern regions also consistently recorded higher shares of under-15 migrants (16.35% – 12.17% – 16.93%) and family emigration (38.89% – 15.84% – 25.84%) compared to the North.

Table 3. Migrants' Characteristics – Northern-Southern Comparison

	1880-1900		1901-1913		1914-1925	
	Northern and Central Italy	Southern Italy	Northern and Central Italy	Southern Italy	Northern and Central Italy	Southern Italy
Characteristics	Average (%)	Average (%)	Average (%)	Average (%)	Average (%)	Average (%)
Male	89.18	86.74	84.37	73.9	73.11	73.75
Over15	87.51	83.65	91.63	87.83	87.31	83.07
Leaving alone	64.66	61.11	81.56	74.16	76.47	69.26
Employees in agriculture	37.16	43.86	29.36	41.42	15.89	34.29
Daily workers	18.8	11.76	32.75	20.18	11.46	6.77
Employees in secondary sector	16.25	16.57	19.64	17.33	15.99	6.77
Transoceanic Destinations Share	54.37	74.06	42.91	82.15	31.19	74.2

Notes: The columns considering Northern and Central Italy include Liguria, Lombardy, Veneto, Emilia-Romagna, Tuscany, the Marches, Umbria, Lazio, while the columns concerning Southern Italy and the Islands include Abruzzo and Molise, Campania, Apulia, Basilicata, Calabria, Sicily and Sardinia.

What drove such a massive transatlantic movement of people? This essay offers a new explanation, rooted in the effects of industrial policy, which is examined in the sections that follow.

2.2. The debate on protectionism

The role and effects of trade policy on structural change in the Italian economy represent one of the most contested issues in Italian economic history (Gomellini, 2020). Barone (2007) argued that Italian mass migration followed two reinforcing trends, both of which intensified after 1887, in parallel with the agrarian crisis and the tariff war with France: migration flows surged sharply, and transoceanic migration came to dominate over European migration. Both trends persisted until World War I. The author attributes these shifts to agrarian backwardness and demographic transition, which together pushed emigration and put upward pressure on agricultural wages in the South.

Musella (1981) documented the divide between liberals and protectionists over the grain tariff. Among the tariff's main critics was the Socialist Party: Loria (1892) foresaw that the tariff would, in the medium term, become a battleground between industrialists and landowners, driving up the price of staple goods, pushing for higher wages, and curtailing consumer spending on non-essential items.

Giretti (1898), Einaudi (1960), and De Viti De Marco (Mosca, 2014) contended that the grain tariff, presented as a measure of general interest, was in reality designed to benefit industrialists and large landowners, to the detriment of the majority of the population — particularly specialized producers and consumers. Moreover, the tariff damaged Italy's most valuable agricultural exports (olive oil, wine, and fruit) and failed to generate the expected fiscal revenue. De Viti De Marco (2014) argued that the tariff was a failure on its own terms, failing to stimulate internal competition between North and South while actively worsening economic conditions in Southern regions.⁵

Einaudi, De Viti De Marco, and Salvemini more broadly championed a program of major reform — including liberalism, tax reduction, and fiscal redistribution — and identified the protectionism of 1887 as the principal driver of the North-South divide. Giretti (1912) argued that the Agrarian League (*Lega Agraria*) had successfully lobbied for the tariff in Parliament; the League was an alliance of landowners and industrialists formed with the specific aim of

⁵ De Viti De Marco (1929) wrote: "It was predictable that Northern landowners would demand the grain tariff, but the fact that Southern representatives agreed to hamper the most innovative crops and lock the Mezzogiorno into cereal farming is a case worthy of study by Doctor Colajanni" (author's translation).

securing duties on wheat, steel, iron, and textiles. Landowners supported the tariff to sustain high rental incomes. Einaudi in particular opposed the tariff on the grounds that it benefited grain cultivation alone, rather than agriculture as a whole. In his view, the Italian Mezzogiorno was doubly penalized: first by the loss of its export markets, and second by higher prices for domestically produced industrial goods.

By contrast, Luzzatti (1913) interpreted the demand for protectionism as an unavoidable response to the international situation, dominated by cheap American grain. Colajanni, a protectionism supporter, claimed that the 1887 tariff had in fact been driven by industrialists, with large landowners playing a more passive role. He agreed that protectionism damaged Southern Italy, but attributed the harm to a strategic political miscalculation by Prime Minister Depretis: as a member of the *Triplice*, Italy was looking for a pretext to sever its preferential trade relationship with France — its leading commercial partner — in order to align more closely with Germany and Austria.

More recently, Gerschenkron (1962) argued that the 1887 tariff failed to protect genuinely innovative industries such as chemicals, favoring instead traditional sectors such as textiles, iron, and steel. Fenoaltea (1993) criticized government intervention in agriculture specifically: grain tariffs raised nominal wages, but because the real grain wage was exogenously determined by migration flows, higher nominal wages reduced manufacturing employment and drove surplus labor to emigrate. Federico and Tena (1998) showed that the tariff primarily benefited landowners — particularly wheat and sugar producers — along with the nascent military-industrial complex (steel, iron, and, to a lesser extent, textiles). Barbagallo (2017) contends that during the 1890s, Italy effectively chose to export labor — and with it, income and consumers. In this way, agriculture, the sector from which most migrants came, was once again made to serve industrialization. In 1910, Coletti famously observed: “*What the great industry was for the North, emigration was for the South.*”⁶ Pescosolido (2011) argued that the North-South divide was not especially sharp between 1861 and 1887, since the North was building its industrial base while the South was investing in specialized agriculture. The Mezzogiorno was also closing its infrastructural gap: in 1861, Sicily and Sardinia had no railways, but by 1886 they boasted 393 kilometers of track; over the same period, the broader South expanded from 184 to 2,698 kilometers. Consequently, Southern Italy constituted a profitable market for Northern infant industry in the 1880s, and could also draw on emigrants’ remittances and the proceeds of specialized agricultural exports. According to this view, Northern industrialists,

⁶ Author’s translation.

focused on the short run and national markets, pushed for protectionism without thinking through its broader consequences. Toniolo (2014) reaches a similar conclusion at the aggregate level, arguing that protectionism was one of the underlying causes of Italy's relatively slow growth compared to other European economies. With respect to foreign investment, the tariff war with France damaged Italy's international reputation and deterred investors from expanding their activities in the country.

The combined burden on the most innovative agricultural productions — olive oil and wine — triggered a deep agricultural depression in the South. The domestic market did not compensate for lost exports, and Southern economies fell into crisis (Iuzzolino et al., 2011). Davis (1999) similarly argues that Southern regions lost the opportunity to accumulate capital and invest in growth, and that these dynamics opened the door to mass migration.

To investigate the effect of protectionism on migration empirically, we exploit the heterogeneous regional exposure to the tariff shock: regions whose economies depended on the export of affected products were hit hardest, and their residents reacted by seeking better opportunities abroad. Yet migration did not occur simultaneously with the imposition of tariffs; a lag was required while farmers sold land or accumulated capital sufficient for departure.

We adopt two econometric specifications: a lagged pre-and-post analysis and a lagged DiD, both using 1889 as the temporal threshold.⁷ The results suggest that regions specialized in the production of high-value exportable crops — olive oil and wine — experienced, on average, higher migration rates in the subsequent years.

2.3. The Tariff War of the 1880s and Its Effects on Agriculture

Italian trade policy at the time of Unification was shaped by the liberal tradition of the Kingdom of Sardinia. On 1 January 1861, Piedmont's low tariff schedule was extended to the newly unified country.

Initially, the leading industrialists — including producers of silk, wool, and cotton — were largely indifferent to the new tariff framework. Agriculture then accounted for approximately 60 percent of gross national product, while industry — still organized primarily along artisanal lines — contributed only about 20 percent. The infant industries of Piedmont and Lombardy were not yet in a position to exploit the national market and did not seek tariff protection. Italian producers did benefit from a form of indirect protection through the adoption of the forced

⁷ 1889 is the year in which both parties adopted mutual war tariffs.

currency (*corso forzoso*) in 1866,⁸ while the underdevelopment of transportation and communications still afforded local markets a degree of natural protection (Menghini and Travaglia, 2010). Some protectionist pressure from Northern capitalists emerged in the 1870s (Milza, 1978), but it was only in the 1880s that industry entered a period of deeper transformation. The cotton industry came under attack from English manufacturers offering high quality at low prices, who were able to penetrate European markets thanks to the expansion of transport networks. At the same time, Italy's dependence on iron imports — in 1880, the 25,000 tons imported already exceeded the 16,000 tons produced domestically (Coppa, 1970) — combined with the expansion and modernization of the Italian navy and the desire to strengthen military capacity, gave new impetus to heavy industry: the Terni steelworks were founded in 1884 to supply the Italian military navy. The *Banca Generale*, the *Credito Mobiliare*, and the *Società di Costruzione* participated in the establishment of Terni with an initial capital of 3 million lire, which soon grew to 16 million (Romeo, 1988).⁹ In 1881, the forced currency was abolished, ending the era of indirect protection.

Falling transportation costs also facilitated transoceanic imports: American grain flooded European markets, triggering tariff responses across the continent (Germany: 1879; France: 1881; Russia and Austria: 1882). In Italy, the crisis in cereal production was starkly evident: grain imports jumped from 1,474 million quintals in 1881 to 10,159 million in 1887, while domestic prices collapsed and production contracted sharply. Romeo (1988) documents a price fall from 33.11 lire per quintal in 1878–1880 to 22.80 lire in 1887. Some commentators had argued that overseas competition was ultimately beneficial. The government eventually reached the same conclusion and raised import duties on grain to 5.00 lire per quintal in 1888. When the cereal crisis persisted, the duty was increased to 7.00 lire by royal decree in February 1894 and rose further to 7.50 lire by year-end (Coppa, 1970). Italy's first tariff schedule of 1878 had already signaled the end of the free-trade era, ushering in a phase of depression driven by the agrarian crisis; in Parliament, Ruggero Bonghi countered that the real culprit was high land taxation, not foreign competition.

France had been one of Italy's closest trading partners for agricultural products since Unification. In 1862, the two countries signed a navigation convention, and in 1863 they

⁸ The *corso forzoso* (forced currency) was adopted by the Banca Nazionale in 1866 to manage Italy's debts, which had grown to fund military campaigns in Veneto (1866) and Lazio (1870) and to finance infrastructural investment. In this context, the Banca Nazionale gained economic and political prominence among the six banks then licensed to issue currency (Banca Nazionale Toscana, Banca di Credito, Banca Romana, Banco di Napoli, and Banco di Sicilia) (Coppa, 1970).

⁹ The iron industry was also expanding. The Boselli Law of 1885 encouraged shipbuilding; the chemical sector and the cotton industry also made significant gains.

concluded the first trade agreement, which ran until 1876 and was mutually beneficial. As a result, Italy joined the emerging *European treaty network* based on the *most-favoured-nation* clause (James and O'Rourke, 2011), and Italian tariffs fell to the second-lowest level in Europe. Prior to 1876, the renegotiation of the new trade agreement had proceeded with France voting for a 24% duty increase on manufactures that did not harm Italian exports, while Italy agreed to major reductions on textiles and metals. The treaty was not signed by the Italian government until 1882, however, because the Tunisian crisis of 1881 stalled negotiations.¹⁰ The agreement was set to expire on 1 February 1892, with an option to withdraw effective from 1888 (Milza, 1978).

At the end of 1882, Finance Minister Magliani introduced a new, moderate tariff project aimed at renegotiating trade agreements with Austria and France. The proposal fueled the lobbying efforts of large Southern landowners and representatives of Northern infant industry, both pushing for higher protection.

In 1886, the French Parliament failed to ratify the navigation treaty with Italy that the Italian minister Robilant had already signed. Italian protectionists seized on this setback to press for new duties on wheat, grain, steel, iron, and clothing (Romeo, 1988).

The Italian government formally denounced the 1881 Franco-Italian trade agreement in 1888. General Menabrea, Italy's chief negotiator, declared that Italy wished to open new talks with France, as the existing treaty no longer reflected Italian interests. Parliamentary debate on tariffs intensified. Northern industrialists — particularly in heavy industry — lobbied through their representative Luigi Luzzati (of the historical Right Party), who championed resistance to French tariffs on Italian exports and opposed compromise. The official sticking point was livestock: the French Parliament refused to reduce tariffs on livestock, and Prime Minister Crispi broke off negotiations on 28 February 1888. From 1 March onward, both countries reverted to their general tariff schedules. France responded by increasing duties on several key Italian agricultural exports, including olive oil, wine, citrus fruits, wool, cotton, and silk, to offset Italy's higher industrial tariffs (Figure 2). Italy retaliated in kind (Arnauné, 1911). Events proved that the Italian economy was far more vulnerable than the French: silk and wine producers across the peninsula were devastated by the loss of their primary market. The tariff war, which began effectively in June 1888, lasted until 1898, when a new trade agreement was finally signed, sacrificing the interests of French small industry and Italian innovative

¹⁰ France's occupation of Tunisia in 1881 deepened the breach between the two countries, prompting Italy to join Germany and Austria in the Triple Alliance. The two countries now found themselves in opposing diplomatic camps, which cast a shadow over subsequent trade negotiations (Coppa, 1970).

agriculture alike (including the Rhenish silk industry and the wine, oil, and fruit cultivations of Southern Italy).

Romeo (1988) notes that at the same time, Italy's textile industry was consolidating its grip on the domestic market following the 1887 tariff. Agriculture in the Po Valley was protected, but Southern Italy bore the heaviest costs. The silk industry was also severely affected, with exports falling from 200 million to 60 million lire. These economic dislocations contributed to the unrest of the *Fasci siciliani* in Liguria and Sicily (1892–1893) and the Milan riots of 1898.¹¹

The breakdown in Franco-Italian trade relations and the introduction of the 1887–1888 tariffs constituted a genuine seismic shock to Italian trade, given that France had absorbed 44% of Italian exports and supplied 23% of its imports (Barone, 2007). The value of exports to France plummeted from 497 million lire in 1887 to 187 million lire in 1892 (Table 4). James and O'Rourke (2011) report that between 1887 and 1897, Italian exports to France fell by 57%, while French exports to Italy declined by 21%. Only in 1898 did France restore minimum-tariff status to Italy. In an effort to offset these losses, Italy expanded exports to Austria (from 95 to 109 million lire), Germany (from 115 to 148 million lire), Switzerland (from 100 to 178 million lire), the United Kingdom (from 79 to 114 million lire), and Belgium (from 19 to 25 million lire), but the combined gains of 166 million lire fell far short of recovering the 310 million lire lost from France.

The new tariff burden hit the North disproportionately through its effects on silk producers and livestock breeders with French markets, but it was Southern Italy that suffered most (D'Angiolini, 1969). French retaliatory tariffs struck hardest at fruits and vegetables, citrus fruits, wine, and olive oil. Between 1881 and 1886, wine exports had reached approximately 2 million hectoliters (three-quarters of which were destined for France, Table 4), with a peak of 3.6 million hectoliters exported in 1887. Following the outbreak of the tariff war, wine exports fell sharply: to 1.8 million hectoliters in 1888, 1.4 million in 1889, and 0.9 million in 1890 (Table 5).

New trade agreements in the 1890s brought some recovery in wine exports, though pre-war levels were never regained.¹² Apulia and Sicily paid the steepest price, though the French wine

¹¹ The *Banco di Napoli* had extended credit to farmers and small landowners to fund the conversion of traditional farming to higher-value crops (olives, fruit, vines). The tariff war with France drove many of these borrowers into bankruptcy (Barone, 2007).

¹² The Vice-President of the Trapani Chamber of Commerce wrote: "We mourn the loss of the French market, which was the primary destination for Southern wines. The agrarian crisis — and the wine crisis in particular — has destroyed our properties and collapsed the value of our land. The banks that financed land acquisition and conversion now demand repayment" (Presidenza del Consiglio di Gabinetto, 1888; author's translation).

tariff affected Southern Europe broadly, given France's role as a major wine importer. Once France overcame its own difficulties — first through the eradication of phylloxera and then through the development of Algerian vineyards — it moved toward national self-sufficiency in wine production. After the war tariff came into effect, Italian wine barrel exports fell by roughly two-thirds by 1890; until then, Italian bulk wine had been used to supplement depleted French production during the phylloxera crisis (Ciuffoletti, 2017). As a reaction to the trade shock, Italy might have pivoted toward Northern European markets — particularly the United Kingdom — but political and economic constraints prevented this.¹³

As a consequence of the war, trade costs with France rose by 3–4 percent over the following decade, while the gains from trade fell from 4 to 2 percent (Gomellini, 2020). During this period, Northern Italy increasingly specialized in industrial crops (cocoons, hemp, and sugar beet), while Southern Italy deepened its specialization in the affected export crops. By 1910, gross product per hectare stood at 420 lire in the North and only 240 lire in the South (Barone, 2007). Among the further indirect costs of the grain tariff was a rise in the price of the staples of the Italian diet — bread, pasta, and polenta — which in turn forced upward adjustments in industrial wages, eroding one of Italian industry's principal competitive advantages: cheap labor.

From the 1890s, Italian trade policy began to adjust. In 1891 and 1892, Italy concluded two trade agreements with the Central Powers. Giolitti declared in 1892 that the government's primary objective in commercial treaties was to secure export markets for Italy's agricultural produce; accordingly, he negotiated comparable agreements with Switzerland, Germany, and Austria. Italian tariffs were lowered to allow increased imports of German industrial products — especially chemicals, textiles, machinery, and motors — whose combined value rose from 134 million lire in 1887 to 203 million in 1900. In exchange, Italy retained privileges inherited from earlier treaties and secured new concessions for Southern products such as wine and citrus fruits. Austria opened its borders to Italian wines, giving new impetus to that export sector (Coppa, 1970).

Barone (2007) further notes that as emigration to Latin America and the United States accelerated, new commercial treaties were concluded with those countries, with imports rising from 16% to 35% of the relevant total and exports from 13% to 34%.

¹³ Barone (2007) reports that the Secretary of the Italian Chamber of Commerce proposed a partnership with the United Kingdom through an Italian-British shipping company, with the *Banco di Napoli* and the *Banco di Sicilia* to participate on the condition that the ports of Naples and Palermo would be used exclusively. The terms were judged too restrictive, and the company closed after three years.

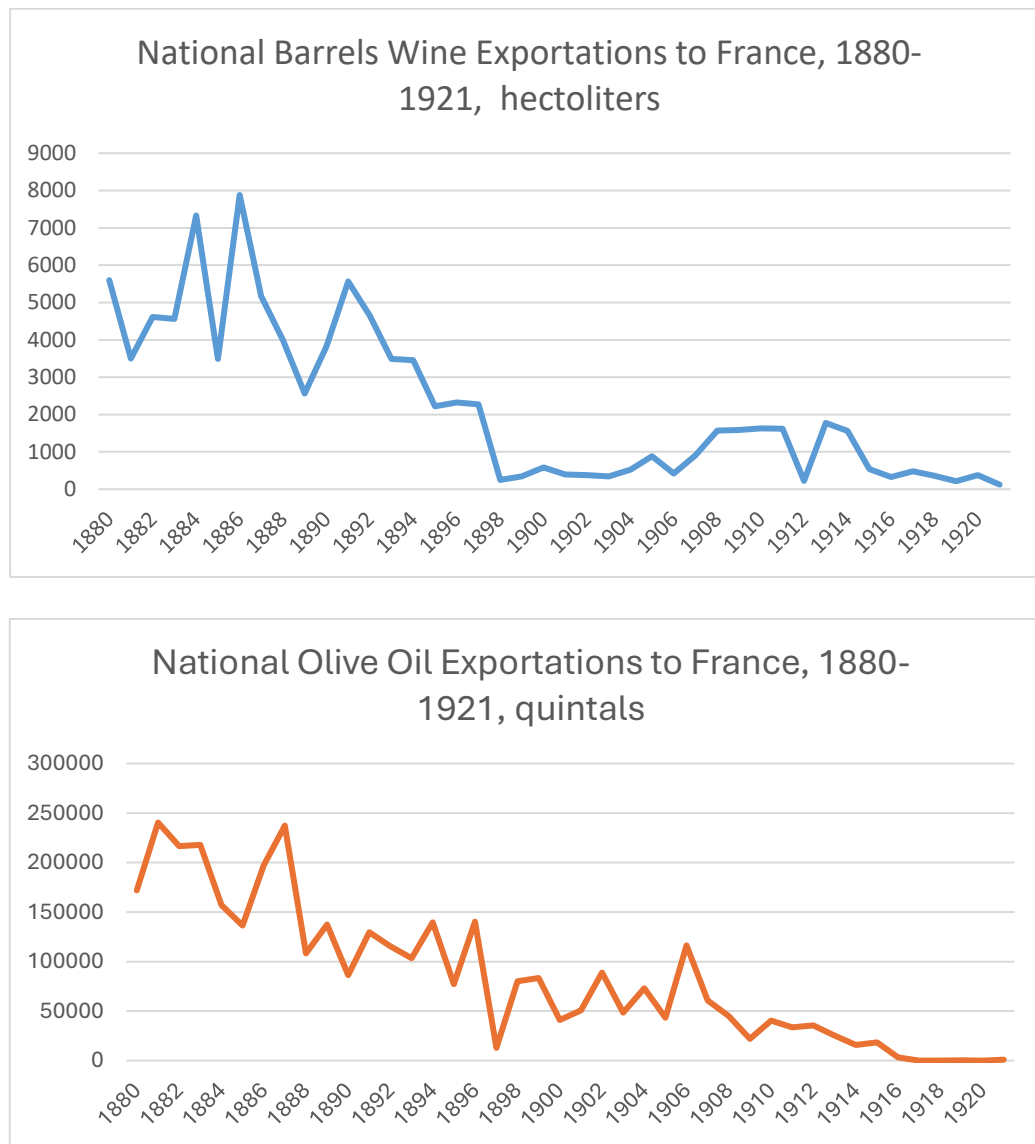
Notwithstanding these adjustments, the decade spanning the late 1880s to the late 1890s was a period of profound transformation and dislocation for the Italian economy.

Table 4. Italian exports to France, Olive Oil and Barrels of wine, 1880-1898 (until the end of the Tariff war)

Year	Olive Oil (quintals)	Wine in barrels (hectoliters)
1880	171945	5596
1881	240357	3501
1882	216532	4617
1883	217808	4561
1884	157106	7336
1885	136504	3492
1886	197279	7881
1887	237186	5168
1888	108318	4000
1889	137370	2563
1890	86188	3826
1891	129590	5569
1892	115235	4640
1893	103294	3494
1894	139547	3456
1895	77204	2220
1896	140385	2326
1897	12854	2272
1898	80160	248

Source: Bankit -FTV

Figure 2. National Exports to France, 1880-1921



Source: Bankit – FTV

Table 5. Italian exports (principal products), thousands of quintals

	Cheese	Wheat Pasta	Citrus Fruits	Dried Fruits	Wines and Vermouth*	Olive Oil	Raw Hemp
1861	14	13	51	65	255	372	94
1862	14	21	464	136	217	468	106
1863	19	35	692	134	468	363	96
1864	35	31	662	146	236	440	131
1865	32	18	695	129	275	657	193
1866	34	35	908	183	359	648	157
1867	23	60	672	137	299	378	247
1868	20	56	715	158	241	523	253
1869	22	64	881	147	287	776	261
1870	18	67	770	264	240	578	213
1871	18	73	878	226	243	841	437
1872	20	77	876	259	609	674	240
1873	23	79	836	209	309	603	249
1874	22	68	717	228	272	477	250
1875	20	93	960	194	363	927	288
1876	22	48	933	274	507	813	230
1877	22	57	1008	176	363	602	243
1878	23	54	978	240	537	514	335
1879	27	56	995	192	1077	887	364
1880	24	68	932	273	2206	577	282
1881	28	65	1287	204	1760	677	281
1882	33	71	1197	292	1332	814	281
1883	38	71	1587	201	2629	806	273
1884	38	57	1734	250	2381	539	357
1885	35	59	1523	289	1481	360	316
1886	39	42	1249	265	2354	648	301
1887	50	35	2298	331	3603	641	310
1888	54	30	1653	251	1829	524	384
1889	63	77	1943	228	1439	553	350

1890	57	73	1906	291	936	378	369
------	----	----	------	-----	-----	-----	-----

*Wine and Vermouth are in thousands of hectoliters.

Source: *Sommario di Statistiche Storiche Italiane, 1861-1955*.

3. Data description and descriptive analysis

The dataset used in this analysis is the product of a substantial digitization effort.¹⁴ It combines annual data on Italian international migration by region with regional agricultural production and specialization measures, along with regional population controls.

The analysis covers the sixteen regions of historical Italy: Piedmont, Liguria, Lombardy, Veneto, Emilia-Romagna, Tuscany, The Marches, Umbria, Lazio, Abruzzo and Molise, Campania, Apulia, Basilicata, Calabria, Sicily, and Sardinia.

Regional annual emigration data from 1876 to 1925 are available only in PDF-OCR format, digitized by the ISTAT Digital Library. The source is the *Annuario Statistico dell'emigrazione Italiana dal 1876 al 1925*, a volume containing annual statistics on Italian emigration collected by two principal government agencies: the *Direzione Generale di Statistica* and the *Commissariato Generale dell'Emigrazione* (active from 1901). Under the law, emigrants were required to follow a two-step procedure: first, they applied for a *nulla osta*, a document certifying eligibility for a passport, after which they collected the passport itself, valid for three years. The statistical series of the *Direzione Generale di Statistica* are based on the *nulla osta*. Two potential measurement issues arise, though neither gives reason to expect systematic regional bias: not all *nulla ostas* would be converted into passports, for a variety of reasons, and not all destination countries required passports.

This study shares a limitation common to all research using these data: it is impossible to account for return migration or seasonal migration. As in Hatton and Williamson (1998) and Gray et al. (2019), however, this constraint should not invalidate the analysis, since the focus is on how policy choices affected migration in a broad sense.

The data used here cover the period 1882–1913 in absolute terms, with migration growth rates calculated at the regional level. Destination countries are not considered, as the analytical focus is on the domestic context and on identifying one key push factor.

Regional population data are taken from the Italian Census for the years 1881 to 1921. Population figures by region have been digitized and are entered in natural logarithms. Regional

¹⁴ The digitization was performed using Able2Extract Professional (available at <https://go.investintech.com/able2extract-pdf-software>), which was used to convert OCR-formatted PDF files into Excel spreadsheets.

GDP is estimated by Brunetti et al. in Vecchi (2011), expressed as an index number with Italian GDP normalized to 100; like the Census data, this variable is available at decennial intervals. Agricultural data are digitized from the volumes of the *Annuario Statistico Italiano*, comprehensive publications reporting socio-economic conditions in Italy. The agricultural section provides regional data on olive oil and wine production (in hectoliters) from 1881 to 1913. Since the *Annuario* is not available for every year of interest, missing observations have been linearly interpolated. Unfortunately, data on citrus fruits — another high-value crop of interest — could not be used because the gaps were too extensive to bridge through interpolation. Regional olive oil and wine production are used to construct two specialization indices, defined as the share of regional production in national production. A joint wine-and-olive oil specialization measure is also constructed as the average of the two individual indices. Finally, data on Italian exports to France (wine and olive oil) referenced in the introduction come from the BANKIT-FTV database developed by Federico et al. (2012), with the support of Banca d'Italia. This database covers Italian trade from 1862 to 1939 and is based on the official publication *Movimento Commerciale del Regno d'Italia*. Products are classified according to SITC Revision 2.0 at the four-digit level. Wine data are reported in hectoliters; olive oil data in quintals. While olive oil is captured in a single category, wine data are distributed across several product categories. As Romeo (1988) notes, Southern Italy primarily exported bulk wine (*vino da taglio*) to France, with the exception of some Chianti; accordingly, the analysis draws on data for that specific category.

Table 6. Variable Description and Source

Variable name	Variable description	Source
Migrant growth rate	Growth rate of migration, by Region	Annuario dell'Emigrazione Italiana 1876-1925
Historical GDP per capita	Mean of Index numbers (Italy=100) by region	A. Brunetti et al. (2011)
Historical regional population	Mean of population by region, log	Censimenti italiani 1871,1881,1901,1911,1921
Export of oil and wine	Flows, levels, by Region	FTV – Banca d'Italia Italian Trade
Wine production	Log of wine hectoliters, by Region	Annuario Statistico Italiano 1881-1921

Wine productive specialization	Hectoliters of regional production/national total	wine	Annuario Statistico Italiano 1881-1921
Olive oil production	Log of oil hectoliters, by Region		Annuario Statistico Italiano 1881-1921
Olive oil productive specialization	Hectoliters of regional production/national total	oil	Annuario Statistico Italiano 1881-1921
Olive oil and wine production	Log of the average of wine hectoliters and oil hectoliters, by Region		Annuario Statistico Italiano 1881-1921
Olive oil and wine productive specialization	Average of hectoliters of regional oil and production/average national total		Annuario Statistico Italiano 1881-1921

Table 7 shows that the migration growth rate exhibits considerable variation, across both time and regions. Some regions recorded extreme values: for example, Abruzzi and Molise went from just a few emigrants in 1892 to several hundred the following year. Wine and olive oil specialization also show substantial regional variation: Piedmont, for instance, recorded zero olive oil production.

Table 7. Summary Statistics on variables of interests

Variable	Observations	Mean	St. dev.	Min	Max
Migration growth rate	496	25.79	122.1	-97.89	1581.25
Wine productive specialization	496	0.625	0.053	0.001	0.51
Olive oil productive specialization	496	0.625	0.077	0	0.505
Wine and oil specialization	496	0.625	0.577	0.004	0.507

Notes: Migration growth rate is expressed as percentage; productive specialization is expressed as regional share on national production concerning that production.

Table 8 reports pairwise correlations among the main variables. At first inspection, migration growth rate is negatively and weakly correlated with all productive specialization measures (-0.067 for wine, -0.027 for oil, and -0.049 for the joint specialization), while wine and olive oil specialization are positively and strongly correlated with each other, consistent with the

observation by Barone (2007) and Coppa (1970) that regions investing in innovative export crops tended to diversify across them.

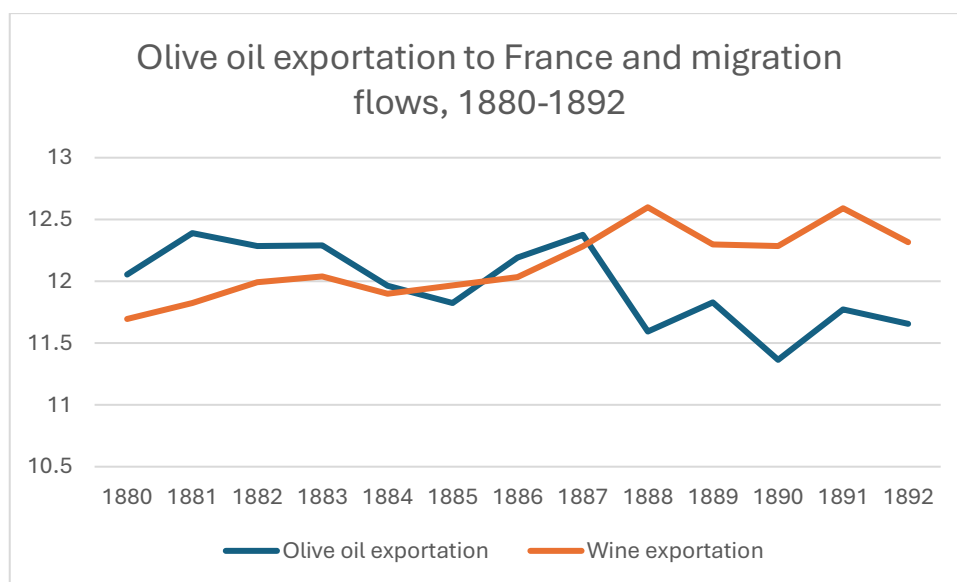
Table 8. Correlations matrix

Variable	(1)	(2)	(3)	(4)
(1) Migration growth rate	1.00			
(2) Wine productive specialization	-0.067	1.00		
(3) Olive oil productive specialization	-0.027	0.537*	1.00	
(4) Wine and oil specialization	-0.049	0.823*	0.92*	1.00

Notes: * Statistically significant at 1%.

A visual inspection of the relationship between olive oil production and migration growth rate — up to the ratification of the new trade agreement — confirms opposite trends in the two series (Figure 3).

Figure 3. Trends of olive oil exportations to France and migration flows before and after the Tariff War, 1880-1892, natural logarithm



Source: Bankit – FT

4. Empirical strategy

The empirical strategy rests on the assumption that the protectionist tariff first inflicted direct damage on regions with higher productive specialization in wine, olive oil, and citrus fruits — crops predominantly exported to France. This economic depression would subsequently trigger emigration.

Since data on citrus fruits are incomplete, the analysis focuses on olive oil and barrels of wine. Ideally, to rule out endogeneity, the analysis would use FAO land suitability data as an instrument; however, these data are unavailable for the crops of interest. We therefore rely on regional productivity shares as a second-best proxy. Data on historical trade flows at the regional level are also unavailable, as ISTAT began conducting regional trade surveys only in 1990.

To assess the effect of the tariffs, we estimate two models — one in the spirit of a before-and-after approach and one following a difference-in-differences approach — treating 1889 (the year of France’s retaliatory tariffs and the mutual escalation) as the watershed year. In both cases, the dependent variable is the regional migration growth rate. The analysis covers all sixteen Italian regions.

The protectionist reform is treated as an exogenous shock to migration: it was designed to protect Italian infant industry, and the subsequent acceleration of emigration was an unintended consequence. The Italian government could not have anticipated that the loss of the French market would be so persistent, particularly given that Italy was soon able to conclude new trade agreements with the Central Powers (Coppa, 1970). Reverse causality is further ruled out by assuming a lagged response: the collapse of exports would have hurt migration outcomes only after a delay, as the most affected regions lost the opportunity to sustain their development trajectory.

The decision to use two complementary methods reflects the fact that the DiD approach additionally controls for any pre-existing trend, providing a more stringent test.

4.1. A Pre-and-post analysis

To identify the effects on migration following the tariff war, we estimate a model in the spirit of the *before-and-after* (or pre-and-post) approach. We use regional productive specialization in olive oil and wine — indirectly affected by protectionism — as a proxy for the severity of the damage the policy inflicted on regions that were leading producers of these export crops.

Pre-post designs are a form of quasi-experimental research that allow for causal inference by testing a dependent variable before and after an intervention (Cameron and Trivedi, 2005). The

independent variable in this case — regional productive specialization — is intrinsically exogenous with respect to the migration growth rate, for the reasons outlined above. Moreover, farmers affected by the tariff shock could not immediately repay loans advanced by banks to fund crop conversion, nor could they quickly sell or reconvert their land (Barone, 2007). We therefore expect a delay between the shock and emigration, and allow for this by introducing lags of up to five years, following the approach of Carillo and Jappelli (2021). The empirical specification is:

$$y_{it} = \alpha_i + \delta_t + \sum_{s=0}^5 \beta_s \text{specialization1889}_{t-s} * \text{Post} + X\varphi + \varepsilon_{it} \quad (1)$$

where Y_{it} is the migration growth rate in a region i at time t , α_i is a set of regional fixed effects, δ_t is a set of time fixed effects, i are Italian regions (1,...,16), t are the years from 1883 to 1913, s are the years-lag introduced in the model (from contemporary to 5 years later), *specialization1889* is, according to the case, regional productive specialization on oil, on wine or on both of them in 1889, the year of the French reprisal on Italian duties. Productive specialization is measured as the regional share of that production on national production, as described in section 3. X_i is matrix of regional controls (population and GDP per capita) interacted by trend, φ is the vector of estimated coefficients. *Post* are years after the 1889 event and ε_{it} is the error term, clustered at region level. β_s represent the association between productive specialization and migration growth rate at the same years.

Table 9 presents the results. Panel A refers to olive oil specialization. From the second-year lag onward, the coefficient is statistically significant at the 1% level: regions more specialized in olive oil production experienced a 3.5 percentage point higher migration growth rate two years after the French reprisal. The estimated effect is similar at the fourth (3.54 p.p., significant at 5%) and fifth (3.42 p.p., significant at 10%) year lags.

Panel B reports results for wine specialization. Here the coefficients are uniformly significant, and their magnitude increases over time — confirming the long-lasting character of the policy effect. At lags four and five, the estimated impact on migration growth rate is approximately 4.8 percentage points.

Panel C combines wine and olive oil specialization. Again, coefficients are consistently significant, growing in magnitude with time. At lags four and five, the joint specialization measure is associated with migration growth of approximately 4.7 percentage points.

These findings are broadly consistent with Gray et al. (2019), who estimate comparable effects (ranging from 1.5% to 5.8% depending on specification) of globalization-induced agricultural price shocks on Italian mass migration decisions.

Table 9. Protectionism, Productive Specialization and Migration Growth Rate – Pre/Post Analysis

		Migration growth rate					
		t	Lag1	Lag2	Lag3	Lag4	Lag5
Panel A	Specialization1889*post	2.19	2.362	3.521* **	2.882	3.54**	3.422*
		(1.67)	(1.663)	(1.08)	(1.713)	(1.76)	(1.864)
	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed Effects	yes	yes	yes	yes	yes	yes
	Obs	496	496	496	496	496	496
	R ²	0.2	0.2	0.2	0.2	0.2	0.2
		t	Lag1	Lag2	Lag3	Lag4	Lag5
Panel B	Specialization1889*post	3.137**	3.418***	3.521* **	3.76***	4.8***	4.812***
		(1.10)	(1.285)	(1.08)	(1.38)	(1.11)	(1.143)
	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed Effects	yes	yes	yes	yes	yes	yes
	Obs	496	496	496	496	496	496
	R ²	0.2	0.2	0.2	0.2	0.2	0.2
		t	Lag1	Lag2	Lag3	Lag4	Lag5
	Specialization1889*post	3.157**	3.378**	3.568**	3.814**	4.787***	4.697***
		(1.359)	(1.285)	(1.297)	(1.38)	(1.388)	(1.471)

Panel C	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed Effects	yes	yes	yes	yes	yes	yes
	Obs	496	496	496	496	496	496
	R ²	0.2	0.2	0.2	0.2	0.2	0.2

Notes: Panel A refers to regional productive specialization on olive oil, Panel B refers to regional productive specialization on wine and Panel C refers to regional productive specialization on both wine and olive oil. Explanatory variables are defined in the text. All specifications include a constant term. Clustered Errors in parenthesis. ***, **, * statistically significant at 1%, 5% and 10.

4.2. Difference-in-Differences (DID) Analysis

To provide more robust evidence of the impact of protectionism on mass migration, we estimate a model in the spirit of difference-in-differences (DiD). This approach allows the treated group to be compared with the control group both before and after the event of interest (Bertrand et al., 2004).

The event in question is the Italy-France tariff war, and its effect on migration is expected to be strongest for regions that had the most to lose — that is, the leading producers of wine and olive oil. The treatment is defined by productive specialization. However, since nearly all regions produced some wine and olive oil (with Piedmont as a partial exception for olive oil), it is not possible to define treated and untreated groups with a clean binary cutoff. To partly address this limitation, we use the national average production level as a threshold: regions whose specialization exceeds the national average are classified as treated, and the remaining regions serve as controls.

As in the pre-post analysis, we test the model for three definitions of productive specialization (olive oil, wine, and both combined). For oil specialization, the treated group consists of Abruzzi and Molise, Calabria, Campania, Lazio, Apulia, Sardinia, Sicily, and Tuscany. For wine specialization, the treated group comprises Piedmont, Apulia, Sardinia, Sicily, and Tuscany. For the joint specialization, the treated group includes Calabria, Campania, Apulia, Sicily, Tuscany, and Sardinia. Given the expected delayed migration response, We estimate six specifications corresponding to lags from t (1889) through lag 5 (1894).

The model is:

$$y_{it} = \alpha_i + \delta_t + \gamma T_{it} + \rho t + \beta_1 T_{it}t + \beta_7 X_{it}\varphi + \varepsilon_{it} \quad (2)$$

where Y_{it} is the outcome variable, namely the regional migration growth rate; T_i takes value 1 if the region has a productive specialization higher than the national average for that production takes value 0 otherwise. The time period t takes value 1 if it considers the years from 1889 to 1913 and 0 otherwise in the first specification. Also this model show up to 5-year lags in the following specifications. X_i is matrix of regional controls (population and GDP per capita) interacted by trend, φ is the vector of estimated coefficients. ε_{it} , the error term, is clustered at regional level. Finally, time (δt) and region (αi) fixed effects are included. Table 7 shows the results.

Table 10 presents the results. Panel A (olive oil specialization) shows that the coefficient β_1 at lag 5 (1894) estimates an approximately 45 percentage point higher migration growth rate for treated regions relative to controls (significant at 10%). In Panel B (wine specialization), treated regions show significantly higher migration growth from the fourth-year lag: about 43 percentage points higher at lag 4 and 44.55 percentage points higher at lag 5 (both significant at 10%). Panel C (joint specialization) yields a 35-percentage-point differential at both lags 4 and 5 (significant at 10%).

Table 10. Protectionism, Productive Specialization and Migration Growth Rate – Difference-in-Differences Analysis (treated: productive specialization: > average threshold)

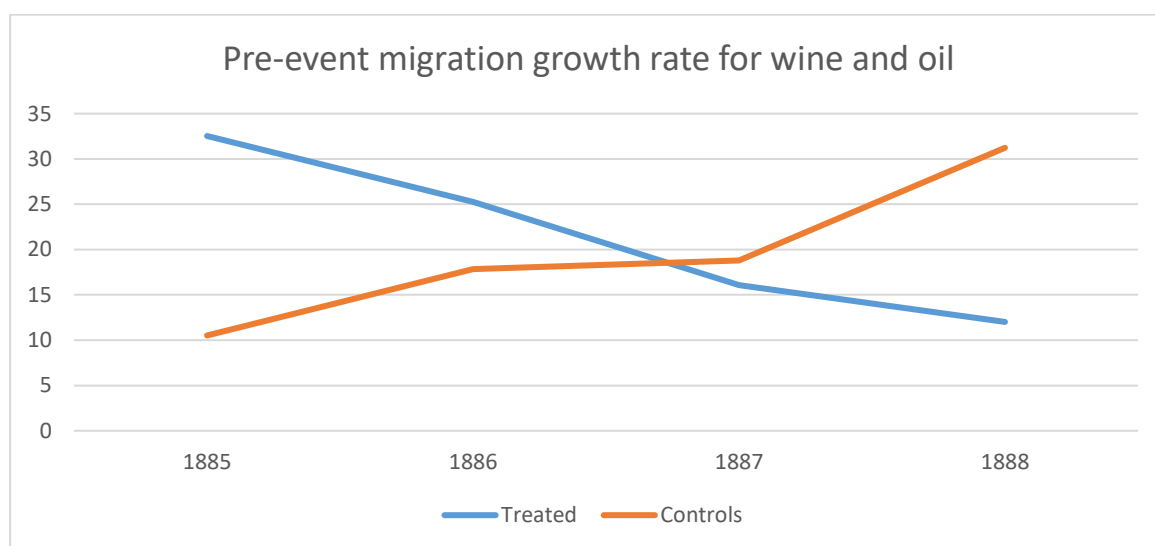
		Migration growth rate					
		t	Lag1	Lag2	Lag3	Lag4	Lag5
Panel A	Oil Treated*Post	4.803	9.794	12.748	7.562	34.982	45.193*
		(14.965)	(15.873)	(16.003)	(16.384)	(25.306)	(25.796)
	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed effects	yes	yes	yes	yes	yes	yes
	Obs	512	512	512	512	512	512
	R ²	0.07	0.07	0.07	0.07	0.07	0.10
		t	Lag1	Lag2	Lag3	Lag4	Lag5
Panel B	Wine Treated*post	12.902	18.285	19.125	19.772	42.925*	44.549*
		(18.706)	(19.103)	(19.429)	(20.083)	(23.456)	(24.282)
	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed effects	yes	yes	yes	yes	yes	yes
	Obs.	512	512	512	512	512	512
	R ²	0.2	0.2	0.2	0.2	0.2	0.2
		t	Lag1	Lag2	Lag3	Lag4	Lag5
	Wine-and-Oil-Treated*post	8.480	11.019	10.855	14.267	35.619*	35.402*
		(15.645)	(16.093)	(16.173)	(16.267)	(22.175)	(22.499)

Panel C	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed Effects	yes	yes	yes	yes	yes	yes
	Obs	512	512	512	512	512	512
	R ²	0.2	0.2	0.2	0.2	0.2	0.2

Notes: Panel A refers to regional productive specialization on olive oil, Panel B refers to regional productive specialization on wine and Panel C refers to regional productive specialization on both wine and olive oil. The explanatory variables are defined in the text. All specifications include a constant term. Robust errors in parenthesis. ***, **, * statistically significant at 1%, 5% and 10%.

Prior to the tariff war, there is no a priori reason to expect that more specialized regions would have been systematically exposed to higher migration growth rates. To assess the plausibility of the parallel trends assumption, Figure 4 plots the migration growth trends for treated and control groups before and after the reform.

Figure 4. Trend of migration growth rate among the different treated and untreated groups for wine and olive oil specialization, 1885-1888 and 1894-1896



Source: Annuario dell'Emigrazione Italiana 1876-1925

Although the parallel trend assumption is violated in the pre-reform period, the divergence is not at odds with the central argument of this essay. Before 1889, the migration growth rate for the untreated group follows an increasing trend while the treated group's trend declines — which is, if anything, consistent with the idea that the tariff shock reversed the relative fortunes of these regions.

4.2.1. Robustness Checks

This section presents robustness checks based on alternative DiD thresholds. As noted, the treated and control groups cannot be sharply identified; we therefore experiment with different productivity thresholds. The baseline analysis used the national average production as the threshold, with above-average regions classified as treated. Here we use the national median as an alternative threshold. The median allows all sixteen regions to contribute to the identification

of both groups, whereas using the first or fourth quartile would result in groups too small for adequate analysis. The time span and lag structure remain unchanged.

Using the national median as threshold, the treated group for oil specialization now comprises Abruzzi and Molise, Calabria, Campania, Lazio, Liguria, Apulia, Sardinia, Sicily, Tuscany, and Umbria. For wine specialization, it includes Abruzzi and Molise, Campania, Emilia-Romagna, Lazio, Piedmont, Apulia, Sardinia, Sicily, and Tuscany. For the joint oil-and-wine specialization, the treated group consists of Abruzzi and Molise, Campania, Lazio, Apulia, Sardinia, Sicily, and Tuscany. All remaining regions serve as controls. The empirical model is identical to model (2), differing only in group composition.

Table 11 presents the results. For wine productive specialization (Panel B) and the joint wine-and-oil specialization (Panel C), the effect on regions above the median is significant at 10% at lag 5, confirming the expected delay. In particular, the migration growth rate for the most wine-specialized regions is approximately 44 percentage points higher than for controls after five years; regions jointly specialized in both crops show a differential of around 51 percentage points in 1894. These findings corroborate the baseline estimates, providing consistent results across both threshold definitions and across both empirical approaches.

Table 11. Protectionism, Productive Specialization and Migration Growth Rate – Difference-in-Differences Analysis (treated: productive specialization: >median threshold)

		Migration growth rate					
		t	Lag1	Lag2	Lag3	Lag4	Lag5
Panel A	Oil Treated*Post	22.036	12.599	21.345	19.792	-4.062	11.943
		(15.529)	(16.167)	(16.474)	(16.784)	(19.624)	(19.874)
	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed effects	yes	yes	yes	yes	yes	yes
	Obs	512	512	512	512	512	512
	R ²	0.2	0.07	0.2	0.2	0.2	0.2
		t	Lag1	Lag2	Lag3	Lag4	Lag5
Panel B	Wine Treated*post	3.365	10.548	7.498	2.479	35.180	43.852*
		(15.529)	(16.625)	(16.783)	(17.049)	(25.073)	(25.61)
	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed effects	yes	yes	yes	yes	yes	yes
	Obs.	512	512	512	512	512	512
	R ²	0.2	0.2	0.2	0.2	0.2	0.2
		t	Lag1	Lag2	Lag3	Lag4	Lag5
	Wine-and-Oil -Treated*post	8.367	12.625	16.019	10.431	37.833	50.918**

Panel C		(15.672)	(16.489)	(16.097)	(17.203)	(25.109)	(25.788)
	Region Fixed effects	yes	yes	yes	yes	yes	yes
	Time Fixed Effects	yes	yes	yes	yes	yes	yes
	Obs	512	512	512	512	512	512
	R ²	0.2	0.2	0.2	0.2	0.2	0.2

Notes: Panel A refers to regional productive specialization on olive oil, Panel B refers to regional productive specialization on wine and Panel C refers to regional productive specialization on both wine and olive oil. Explanatory variables are defined in the text. All specifications include a constant term. Robust errors in parenthesis. ***, **, * statistically significant at 1%, 5% and 10%.

5. Discussion and conclusion

Quantitative analyses of the determinants of Italian mass migration have remained relatively rare in the literature. This essay contributes to filling that gap by investigating the relationship between protectionism — and the ensuing tariff war with France, Italy's principal trading partner — and the great surge of emigration that peaked from 1900 through World War I.

The results provide evidence that protectionism was a significant driver of migration: regions that suffered the greatest export losses display migration growth rates roughly twice as high as the others. Across all specifications, the estimated increase in migration growth rate attributable to productive specialization in the affected crops ranges from 3.5 to 4.8 percentage points. The further novelty of this study lies in the construction of an original dataset from historical sources that were previously available only in analog format and required substantial digitization effort. The main limitation of the identification strategy concerns the DiD approach, where the boundary between treated and untreated groups is necessarily blurred; future work could explore alternative methodologies or robustness checks to further validate these findings and probe the heterogeneity of the effects.

Deepening our understanding of Italian mass migration and its causes is relevant for at least three broader reasons. First, protectionist policies may have played a dual role in shaping the North-South divide: by sheltering Northern infant industry on one hand, and by diverting resources away from the most innovative Southern crops — olive oil and wine, potential engines of capital accumulation — on the other. This hypothesis, not pursued here, merits further investigation.

Second, this chapter sheds light on the process of structural transformation in late nineteenth-century Italy: the transition to industrialization was expected to displace workers from agriculture, who would then be absorbed by the industrial sector. This reallocation was incomplete, however, as industrialization had heterogeneous regional effects. In regions where industry failed to absorb surplus agricultural labor, emigration served as the adjustment mechanism.

Third, this study offers a new perspective on the mass migration literature and on the long-run consequences of government policies (Adelman, 1999). The results show that the effects of protectionism on migration did not materialize immediately but unfolded over the medium term, pointing to the probable emergence of further long-run consequences for divergent regional growth trajectories in Italy.

References

- Abramitzky, R., Boustan, L.P., Eriksson, K. (2012). Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review*, 102 (5): 1832-56.
- Abramitzky, R., Sin, I. (2014). Book Translation as Idea Flows: the Effects of the Collapse of Communism on the Diffusion of Knowledge. *Journal of the European Economic Association*, 12 (6): 1453-1520.
- Adelman I., (1999), The Role of Government in Economic Development, *WP 890*, University of California at Berkley.
- Ager, P., Brückner, M. (2013). Cultural diversity and economic growth: Evidence from US during the age of mass migration. *European Economic Review*, 64: 76-97.
- Ager, P., Hansen, C.W. (2017), Closing Heaven's Door: Evidence From the 1920s U.S. Immigration Quota Acts. *Discussion Paper*, 17 (22): 1-35.
- Arnauné A. (1911), *Le Commerce extérieur et les tarifs de douane*, Bibliothèque Nationale de France.
- Bhagwathi, J. N., Hamada, K. (1974). The Brain Drain, International Integration of Markets for Professional and Unemployment. *Journal of Development Economics*, 1: 19-42.
- Barbagallo F. (2017), *La Questione Italiana. Il Nord e il Sud dal 1860 a oggi*, Editori Laterza.
- Barone G. (2007), La modernizzazione italiana dalla crisi allo sviluppo, in Sabbatucci G. e Vidotto V. (a cura di), (2007), *Storia d'Italia, Volume 3, Liberalismo e Democrazia*, Editori Laterza.
- Beck-Knudsen, A.S. (2019). Those Who Stayed: Individualism, Self-Selection and Cultural Change during the Age of Mass migration. *Discussion Paper*, 19 (01): 1-71.
- Beine M., Docquier F., Ozden C., (2011), Diasporas, *Journal of Development Economics*, 95: 30-41.
- Bertrand M., Duflo E., Mullainathan S. (2004), How much should we trust difference-in-differences estimates), *The Quarterly Journal of Economics*, Volume 19, Issue 1: 249-275.
- Bevilacqua P., De Clementi A., Franzina E. (2009), *Storia dell'Emigrazione Italiana, Partenze*, Donzelli.
- Borjas, G. J. (2013). The Analytics of Wage Effect of Immigration. *IZA Journal of Migration*, 2 (22): 1-25.

- Cameron A.C., Trivedi P.K. (2005), *Microeconometrics, method and applications*, Cambridge University Press.
- Carillo M., Jappelli T. (2021), Pandemics and regional economic growth: evidence from the Great Influenza in Italy, *European Review of Economic History*, 00: 1-29.
- Ciccarelli C., Dalmazzo A., Vuri D. (2018), Home Sweet Home: the Effect of Sugar Protectionism on Emigration in Italy, 1876-1913, *CEIS*, Vol. 16, Issue 5, no. 437.
- Ciuffoletti Z. (2017), La svolta protezionistica e le sue conseguenze nell'agricoltura italiana, Globalizzazione e modernizzazione diseguale, *Georgofili, Quaderni*, I: 61-73.
- Coppa F.J. (1970), The Italian Tariff and the Conflict Between Agriculture and Industry: The Commercial Policy of Liberal Italy, 1860–1922, *The Journal of Economic History*, Volume 30, issue (4).
- D'Angiolini P. (1969), L'Italia al termine della crisi agraria alla fine del XIX secolo, in Altieri Magliozzi E., (2002), *Scritti archivistici e storici*, Ministero per i Beni e le Attività Culturali.
- Daniele V., Malanima P. (2011), *Il divario Nord-Sud Italia, 1861-2011*, Rubbettino.
- Davis J.A. (1999), Mutamenti di prospettiva sul cammino dell'Italia verso il XX secolo”, Ciocca P., Toniolo G. (a cura di), *Storia economica d'Italia. I. Interpretazioni*, Laterza
- Del Boca D., Venturini A. (2003), Italian Migration, *IZA DP* No.938.
- De Viti De Marco (1929), *Un trentennio di lotte politiche (1894-1922)*, Giannini.
- Einaudi L. (1960), *Cronache economiche e politiche di un trentennio I*, Einaudi.
- Fenoaltea S. (1993), Politica doganale, sviluppo industriale, emigrazione, verso una riconsiderazione del dazio sul grano, *Rivista di Storia Economica* (10): 65-77.
- Fernández-Huertas Moraga, J. (2013), Understanding different migrants selection pattern in urban and rural Mexico, *Journal of Development Economics*, 103:182-201.
- Gerschenkron A. (1962), *Economic Backwardness in Historical Perspective*, Belknap Press of Harvard University.
- Giretti E. (1912), La nuova fase doganale e gli interessi del Mezzogiorno d'Italia, *L'Unità*, 17 febbraio.

- Giretti E. (1898), Le illusioni e i danni del protezionismo. Considerazioni offerte agli “agrari” italiani, *Giornale degli economisti*, settembre.
- Gomellini M. (2020), Breve storia delle Barriere agli Scambi in Italia, *Quaderni di Storia Economica*, n. 46, Banca d’Italia.
- Gray R., Narciso G., Tortorici G. (2019), Globalization, agricultural markets and mass migration: Italy, 1881-1912, *Explorations in Economic History*, 74 101276.
- Haque, N., Kim, S. J. (1995). Human Capital Flight: Impact of Migration on Income and Growth. *IMF Staff Papers*, 42: 577-607.
- Hatton T., Williamson J.G. (1998), *The Age of Mass Migration: Causes and Economic Impact*, Oxford University Press.
- Iuzzolino G., Pellegrini G., Viesti G., (2011), Convergence among Italian Regions, 1861-2011, *Quaderni di Storia Economica*, n. 22, Banca d’Italia.
- James H., O’Rourke K. (2011), Italy and the First Age of Globalization, 1861-1940, *Quaderni di Storia Economica*, n.16, Banca d’Italia.
- Loria A. (1892), Il dazio sui cereali, *Critica Sociale*.
- Luzzatto G. (1913), Prezzi, salari e dazi, *L’Unità*, 19 dicembre.
- Lupo S. (1990), *Il giardino degli aranci. Il mondo degli aranci nella storia del Mezzogiorno*, Marsilio.
- Menghini M., Travaglia M.L., (2010), L’evoluzione dell’industria italiana: peculiarità territoriali, *Dossier Tagliacarne Web*, 1/2010.
- Milza P. (1978), *Les origines de la guerre douanière franco-italienne de 1888-1898*, Relations internationales, n°15, Protectionnisme et Relations Internationales, 1: 235-254.
- Mosca M. (2014), De Viti de Marco e il Mezzogiorno: Nuova Luce dai Quotidiani, *Core*, AC.UK.
- Mountford, A. (1997). Can a Brain Drain Be Good for Growth in the Source Economy?. *Journal of Development Economics*, 53: 287-303.
- Musella L. (1981), La valutazione storica del dazio sul grano nel dibattito tra liberisti e protezionisti a cavallo fra Otto e Novecento, *Archivio Storico Italiano*, 1981, vol. 139, no. 3 (509).
- Nitti S., (1958), *Il Bilancio dello Stato dal 1862 al 1896-97*, Saitta A. (a cura di), Laterza Editore.

- Pescosolido G. (2011), Una società immobile? Sviluppo pre-unitario e questione meridionale, in Galasso G. (2011), *Mezzogiorno, Risorgimento e Unità d'Italia, Atti del Convegno, 18,19 e 20 Maggio, 2011*, Presidenza del Consiglio dei Ministri, Dipartimento per l'informazione e l'editoria: 73-82.
- Petraglia C., Vecchione G. (2020), Long-run pro-trade effects of diasporas: evidence on Italian regions, *Spatial Economic Analysis*, Volume 16, Issue 1: 47-72.
- Presidenza del Consiglio di Gabinetto (1888), Considerazioni e proposte di S. Giacomazzi-Favara, Vicepresidente della Camera di Commercio di Trapani, *Archivio Centrale dello Stato*, fasc. 390, V.
- Romeo R., (1988), Breve Storia delle Grande Industria in Italia, 1861-1961, Il Saggiatore Arnoldo Mondadori Editore.
- Sabatucci G., Vidotto V. (a cura di), (2007), *Storia d'Italia, Vol. 3, Liberalismo e Democrazia*, editori Laterza.
- Sanchez-Alonso, B. (2000). Those Who Left and Those Who Stayed Behind: Explaining Emigration from the Regions of Spain, 1880-1914. *The Journal of Economic History*, 60 (3): 730-755.
- Sanfilippo M. (2009), Tipologie dell'emigrazione di massa, *Storia dell'Emigrazione Italiana*, Donzelli.
- Spitzer, Y., Zimran, A. (2018). Migrant self-selection: Anthropometric evidence from the mass migration of Italians to United States. *Journal of Development Economics*, 134: 226-247.
- Sun, B., Zhu, P., Li, W. (2019). Cultural diversity and new firm formation in China. *Regional Studies*, 53 (5): 1-14.
- Tabellini, M. (2018). Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration. *Harvard Business School BGIE Unit Working Paper*, 19-005.
- Toniolo G. (2014), La crescita economica italiana, 1861-2011, in Toniolo G. (a cura di) (2014), *L'Italia e l'economia mondiale dall'Unità a oggi*, Il Mulino: 6-51.
- Vecchi G. (2011), *In ricchezza e in povertà. Il benessere degli italiani dall'Unità a oggi*, Il Mulino.

The Long-Lasting Effects of Italian Mass Migration on Foreign Direct Investment Inflows

Maria Rosaria Carillo, Viviana De Falco and Luca Pennacchio

Abstract

This paper examines whether Italian mass migration generated long-lasting cross-border ties that shape the geographical allocation of inward foreign direct investment (FDI) across Italian regions. We construct a novel dataset by digitizing historical emigration flows from Italian regions to destination countries between 1880 and 1925 and combining them with current data on the regional stock of foreign-participated firms. Using an extended gravity framework, the results reveal a positive and statistically significant association between past emigration flows and present-day inward FDI across all three phases of Italian mass migration. The relationship is stronger for the 1901–1913 period, which coincided with the peak of Italian emigration. Italian institutions abroad in 1927 may serve as a channel of cultural persistence. Their interaction with historical migration is positively associated with inward FDI, with particularly pronounced effects for economic and commercial institutions. These findings are consistent with the hypothesis that diaspora networks foster persistent informational, economic, and cultural ties between migrants' places of origin and destination.

1. Introduction

The age of mass migration produced one of the largest population movements in modern history, as millions of Europeans migrated to the New World between the second half of the nineteenth century and the early twentieth century (Hatton and Williamson, 1998).

Although scholars have extensively studied the immediate demographic and economic consequences of these movements, a growing literature shows that migration can also produce effects that persist across generations. Historical migration has influenced economic development, human capital accumulation, trade, political preferences, and cultural attitudes in both origin and destination countries (Parsons and Vézina, 2016; Sequeira et al., 2018; Beck-Knudsen, 2019). These findings raise a broader question: can population movements that occurred more than a century ago continue to shape contemporary economic relationships between migrants' places of origin and destination?

This paper builds on this line of research by examining the relationship between Italian mass emigration and the geographic distribution of inward FDI across Italian regions. More specifically, it investigates whether countries that received larger numbers of migrants from a particular Italian region between 1880 and 1925 subsequently developed stronger investment ties with that region. The analysis therefore considers whether historical migration created persistent bilateral relationships that continue to influence the location of foreign economic activity within Italy. It connects two different strands of literature, namely one related to persistence of historical events (among the others, Abramitzky, 2014, Petraglia and Vecchione, 2020, Sequeira et al., 2018, Tabellini, 2010) and the other connected to the impact of migration and network on foreign direct investments (Kugler and Rapoport, 2007; Malan, 2015; Javorick, 2011; Tong, 2005; Mayda, 2021; Buchardi, 2018).

The Italian case provides a particularly useful setting for investigating these long-term relationships. Italian mass emigration was both large and geographically heterogeneous. Italian regions differed considerably in the timing, intensity, and destinations of their emigration flows, while migrant communities abroad frequently maintained strong connections with their places of origin. Moreover, Italian mass migration was not homogeneous. This paper distinguishes among three phases: the initial expansion of emigration between 1880 and 1900, the peak period between 1901 and 1913, and the subsequent phase between 1914 and 1925. Examining these periods separately allows us to assess whether the relationship with contemporary FDI varies with the scale and characteristics of historical migration.

The empirical analysis draws on a novel dataset constructed by digitizing historical records of emigration from 16 Italian regions to 26 foreign destinations between 1880 and 1925. These

data are obtained from the *Annuario Statistico dell'Emigrazione Italiana dal 1876 al 1925* and are combined with information from the Italian population censuses and the *Censimento degli Italiani all'Estero*. The resulting dataset contains 416 region–country dyads. We link historical migration data to the regional inward foreign direct investments, measured by the regional stock of foreign-participated firms in 2017 disaggregated by investors' country of origin, as well as to a range of historical, contemporary, regional, and bilateral controls.

The empirical strategy follows the spirit of the gravity approach applied to two model specifications. The first model investigates the relationship between past migration (either in terms of mass migration flows or former ancient Italian communities settled abroad). The second model, by contrast, also includes Italian institutions settled abroad before 1927. We consider institutions as a moderating factor in explaining peer cultural transmission, which strengthened the sense of belonging to the homeland.

The results show a positive and statistically significant association between historical emigration and contemporary inward FDI across all three phases of Italian mass migration. The relationship is strongest for emigration during 1901–1913, when Italian migration reached its highest intensity. This pattern is consistent with the idea that larger migrant communities established more extensive and durable networks connecting destination countries to their regions of origin. The positive association remains when the three phases are combined, although its magnitude is smaller, underscoring the importance of accounting for the heterogeneity of Italian mass migration over time.

The paper also investigates a potential mechanism underlying the persistence of these relationships. Beyond family-based intergenerational transmission, migrant communities may maintain connections with their country of origin through associations and other collective institutions. As a community grows, it becomes more institutionalized, and Italian institutions abroad can serve as an additional channel for cultural transmission.

To examine this channel, we use newly digitized information on Italian institutions operating abroad in 1927, including educational, charitable, political, social, and economic organizations. The interaction between historical migration flows and the presence of these institutions is positively associated with contemporary inward FDI. The association is particularly pronounced for economic and commercial institutions. These findings align with the view that formal migrant organizations helped preserve information, social capital, and economic connections within the Italian diaspora.

The paper contributes to the literature in three respects. First, it extends research on the long-term consequences of mass migration by examining an outcome—contemporary inward FDI—

that has received comparatively limited attention in studies of Italian emigration. Second, it introduces a subnational perspective by exploiting variation across Italian regions and destination countries. This level of analysis matters in the Italian context, given pronounced regional differences in historical migration patterns and economic development. Third, the paper examines migrant institutions as a possible channel through which diaspora networks and bilateral economic ties persist across generations. In doing so, it connects the literature on migration and international investment with research on cultural persistence and the long-run influence of historical population movements.

The remainder of the paper is organized as follows. Section 2 reviews the literature on migration, FDI, and cultural persistence. Section 3 describes the principal phases and regional characteristics of Italian mass migration. Section 4 presents the data and empirical strategy. Section 5 reports the main results. Section 6 examines the role of Italian institutions abroad as a potential mechanism of persistence. Section 7 concludes.

2. Literature review

Two main strands of literature relate to this analysis: one concerning the relationship between migration and FDI, and another analyzing cultural persistence.

2.1. Migration and FDI

Foreign direct investment should be considered a driver of the simultaneous migration effects on sending and receiving countries. Migrants support FDI by spreading information through spatial networks (acting as a bridge between their origin and host country) and by fostering demand for foreign goods. Several authors examine the relationship between FDI and migration and whether they are complementary or substitutes. Malan (2015) analyses the impact of West-East FDI flows on migration flows and, conversely, the impact of East-West migration flows on FDI. The fall of the Berlin Wall is used as a natural experiment to test the impact of migration. Using a two-gravity-equation model, the study shows a significant relationship between FDI and migration flows: a 1% increase in the natural logarithm of east-west migration leads to a 63% increase in the natural logarithm of west-to-east FDI. The conclusion is that investments and migrations are complements. Kugler and Rapoport (2007) investigate complementarity or substitution between FDI and migration by decoupling short- and long-run effects. The authors account for the different effects of skilled versus unskilled migration on FDI, finding that FDI and migration are contemporaneous substitutes and dynamic complements. Their estimates imply that a 1% increase in the stock of migrants with tertiary

education residing in the US in 1990 is associated with a 1% rise in the annual growth rate of FDI inflows throughout the period covered (from 2.6% to 3.6%). Javorick et al. (2006) investigate the relationship between migrants in the US and outward FDI from the US, providing evidence of a positive correlation between migration and FDI, mainly driven by migrants with college education. As the authors claim, FDI generally require long-term focus and interactions with different groups of economic agents. The authors suggest that, beyond remittances, the migrant diaspora may serve as a channel for information transfer across international borders and may contribute to greater integration.

Also, binomial trade and FDI, and binomial trade and migration, are complements (Rauch and Trindade, 2002; Parsons and Vézina, 2014). Helpman et al. (2004) consider a model involving FDI, migration, and exportation; migrants can supply the information needed to foster FDI and bridge gaps between the host and sending countries. Also, Aubry et al. (2019), building on this model, find that skilled migration positively affects trade and FDI; they find complementarity among FDI, trade, and migration through migrants' international information transmission about business opportunities in their country of origin. In their study, FDI are more sensitive to migration than trade.

Moreover, FDI are essential for economic development and maintains relationships between origin and destination countries. Rauch and Trindade (2002) and Tong (2005), by analyzing the Chinese diaspora, provide evidence that migrant networks encourage FDI.

In addition, the literature finds that migrants reduce transaction costs through demand and information channels; migrants can also foster trust by reducing information frictions. According to Javorick et al. (2011) and Tong (2005), migrants' networks seem to matter more for FDI than for trade. Cuadros et al. (2019) quantify migrants' job positions in fostering FDI: higher shares of migrants with managerial skills in a given country promote investments from that country, while migrants who do not occupy a qualified position negatively affect investments. Foley and Kerr (2011) also provide evidence of a relationship between skilled migration and FDI using American patent data at the firm-level.

Another strand of the literature has focused on the effect of historical migration events on FDI and trade, though this issue remains insufficiently addressed. Buchardi et al. (2018) analyze the causal effect of ancestry composition on U.S. counties' foreign direct investment, both sent and received by local firms. They find persistent migration effects: the presence of a defined ethnic group in a US County increases the likelihood of FDI sent and received and appears to operate through reduced information frictions. The magnitude of the effect varies by instrument; however, a 1% increase in the historical presence of an ethnic group leads to a 0.2% increase

in FDI from the country of origin in 2014. Mayda et al. (2021) exploit US policies to provide evidence of the pro-investment effects of refugees; a 10% increase in refugees increases FDI outflows by 0.54%, FDI projects by 0.24%, and FDI jobs by 0.72%. Parsons and Vézina (2016), in their study on Vietnamese emigration to the USA after the Vietnam War of the '70s, find that migration has a long-term impact on economic growth, reflected in USA exports to Vietnam at the beginning of the New Millennium. Their findings show exports increasing by 4.5% to 13.8% for a 10% increase in migration flows. Javorick et al. (2011) and El Yaman et al. (2007) use qualified migrant networks established during the Sixties to assess the effect on current FDI and find positive links.

Regional analysis of FDI have been rarely explored; among the studies conducted at the regional level, Bajo-Rubio and Diaz-More (2010), who try to assess the effect of FDI on Spanish regions, find that FDI plays an important role in promoting productivity growth. Furthermore, Bronzini (2004) finds that specialized regions – namely regions that host a concentration of firms within the same sector – attract FDI, while diversified regions – namely regions with a concentration of firms within different sectors – only attract FDI in the industrial sector.

Manfredonia and Florio (2021) find a link between intergenerational propensity to migrate and CEOs' attitudes towards risk: a firm with a CEO who has at least one ancestor who emigrated is found to have, on average, productivity more than 4% higher than the average firm.

2.2. *Cultural persistence and migration*

The first question one could pose is why an event that occurred more than one century ago should affect current economic choices. In fact, many scholars have considered history one of the main determinants of current economic development (North, 1981; Hall and Jones, 1999; Acemoglu et al., 2001).

Putterman and Weil (2010) find that past migration – and overall people's movement – has clear implications for the course of economic development: people who moved across borders brought with them human culture, institutions, languages, and social relationships. In particular, it has become almost commonplace to refer to *cultural persistence*, a concept analyzed across a wide strand of literature. This kind of persistence can have long-run effects through two channels: *intergenerational transmission* and *peer transmission*. Giavazzi et al. (2019) focus on these channels by investigating the speed of evolution of values and beliefs representing culture among different generations of immigrants in the USA. The authors present three main

findings: cultural persistence reduces but does not disappear over generations (they consider up to the forth and higher generations); cultural convergence is faster for some values and attitudes and slower for others (attitudes towards politics, government, sexual morality, abortion and religion persist longer); cultural persistence also depends on ancestors' country of origin in terms of strength of family concept, distance of the mother tongue from destination language and residential segregation. Values and beliefs are affected by both vertical and horizontal transmission. Vertical transmission develops within the family; the authors define it as parents shaping their children's preferences, balancing the desire to share common values *with the concern for teaching traits that will make it easier for their children to function in the social environment in which they live*¹. Horizontal transmission, instead, refers to children's exposure to the world outside the family and their subsequent social imitation and learning outside the family, which usually happens through institutions.

Giuliano and Nunn (2020) deepen the cause of cultural persistence, that is, the similarity of the environment across generations. The scholars find that the more ancestors lived in an unstable cross-generational environment, the less likely cultural persistence is. This pattern suggests that, in ethnic aggregations, enclaves find it easier to maintain existing customs. Among the other strategies, the authors test the behavior of the descendants of immigrants of different origin living in the same city and their results show that if ancestors come from more unstable countries, children are more likely to deviate from cultural persistence (they measure cultural persistence using two practices, namely marrying someone from the same ancestry and speaking the ancestral language at home, but they also take into account the intergenerational transmission of education and occupation, reaching the same results).

Diensen and Sørderkson (2018) use social trust as a proxy for culture. They define social trust *as the inclination to trust unknown others in situation, where little information exists about their trustworthiness*. To better understand trust, the authors review studies on immigrants and show that in Europe and the USA, two sources feed social trust: experience and cultural persistence, especially for first-generation migrants². Outside the USA, Helliwell, Wang, and Xu (2016) conduct a worldwide analysis of migrants (130 destination countries) and find that the destination country has a three times stronger effect on trust than ancestral heritage. The fact that research conducted in the USA shows more results on the side of cultural persistence

¹ Among the others, Gruber and Hungerman (2008), Alesina and Fuchs-Schuendeln (2007), Giuliano and Spilimbergo (2014).

² Dienesen (2013) finds that, for instance, that Italian migrants, regardless to the grade of descendants they represent, show a low grade of social trust, as Italians in Italy.

can be due to the creation of more ethnic enclaves in this country, which pushes toward a strong cultural persistence (Luttmer and Singhal, 2011).

Costa-Font and Sato (2016) find that time in the country of residence strengthens cultural association with the country of ancestry. They study health self-assessments of first- and second-generation migrants and test the association between migrants' health assessment and that of their ancestral country. Health assessment is a proxy for cultural persistence because evaluations of health status depend on how the family of origin thinks about it, so it can be considered a tradition. Alba and Nee (1997), Antecol (2000), and Borjas (1992) find cultural persistence effects affecting the second generation of immigrants, and so do Costa-Font and Sato.

Zucker (1977) deepens the mechanisms underlying cultural persistence through an ethnomethodological approach to institutionalization³. He finds that cultural persistence – decomposed into its three main aspects, namely generational uniformity of cultural understanding, maintenance of these understandings, and resistance to change of cultural understandings – is stronger in the case of institutionalization.

Grompone and Sessa (2014) study the trade-off between persistence and dynamics in social capital and culture in Southern Italy and find that, at a certain point, dynamic aspects of culture overcome cultural persistence. The studies by Guiso et al. (2011), Alesina et al. (2013), and Voigtlander and Voth (2012) go in the same direction, supporting cultural and social capital persistence. The authors find that culture has a certain degree of persistence, but it can also be recently determined.

Another part of the literature considers either cultural persistence or cultural convergence, reflecting the economic long-run effects of mass migration. Beck Knudsen (2019) deepens the individualism attribute among Northern European migrants during the mass migration age; she finds that people who left home came from individualistic households. In the short term, individualists migrate; in the long term, the home-country composition of individualists decreases (she finds a loss of individualism of -3.7% in Denmark, -9.4% in Sweden, and -13.6% in Norway). Sequeira et al. (2018) conduct a classical growth analysis that accounts for the effects of mass migration. The results are divided into short-term and long-run effects. In the short term, counties with more immigrants see growth in the share of industrialization and innovation, and an increase in agriculture. In the long run, they report higher incomes, less

³ This approach considers reality experienced as an intersubjective world known-or-know-able-in-common-with-others, which exists historically prior to the actors and furnishes the resistant objective structures, which constrain action (Zucker, 1977, pp. 727-728).

poverty, less unemployment, and higher rates of urbanization and education. These benefits are not associated with extra social costs. The authors find a positive effect of persistence on urbanization, education, and income.

Abramitzky et al. (2019) use foreign names as a proxy for cultural persistence and assess the speed of migrant cultural assimilation today compared with the past (the age of mass migration). The authors report that the speed of assimilation does not vary over time. Also, Fouka et al. (2019) investigate the cultural dynamics of inclusion and exclusion. Their work aims to understand whether the arrival of new ethnic groups in a country accelerates assimilation among existing groups. The analysis simultaneously assesses migrants' assimilation effort and assimilation equilibrium, defined as migrants' acceptance by the receiving country. The historical case of African arrival in the USA shows that, with new ethnic arrivals, previous migrant groups increase their assimilation effort, but assimilation equilibrium is reached only by migrants from Northern and Western Europe. Stereotypes concerning Southern and Eastern European migrants, however, decrease.

Also, the use of native language can be considered a legacy of cultural persistence. Fouka (2019) reports on the impact of restrictive language policies on German migrants after World War I in some US States, namely Ohio and Canada. The author collects different census data and identifies three proxies to test the backlash of language prohibition in schools. The proxies used to measure a sense of cultural identity are the number of volunteers among American German natives and, in World War II, the number of interethnic marriages and the German name diffusion among the immigrants' children. The results show a decrease in the number of volunteers and an increase in the other two parameters. Petraglia and Vecchione (2020) argue that the cultural phenomenon of Italian mass migration has affected current trade. Their regional study shows heterogeneous effects on both imports and exports in the long run.

Regarding the effects of communities abroad and their ties to the country of origin, social network theory is well established in the migration literature (e.g. Massey et al. 1993; Massey et al. 1998; Gurak and Caces 1992; Fawcett 1989; Tsuda 1999; Pellegrino 2004). Networks are defined by Granovetter (1973) as “*sets of actors who know each other relevant characteristics or can learn them through referral*”, while Guiso et al (2010), define social capital as “*those shared customary beliefs and values that ethnic, religious and social groups transmit from generation to generation and enable their members to act together more effectively to pursue shared objectives*”.

Migrant pioneers move to a country, and therefore, non-migrants in the origin areas also connect to migration destinations, where better employment opportunities and increased chances for

socio-economic upgrading are thought to be found. Therefore, the flow continues beyond the pioneers, as the latter mediate the migration of friends and family. Migrant networks help newcomers by providing information about the new society and assistance, and they reduce migration costs over time. Massey (1990), with his theory of cumulative causation, defines social capital as “*sets of interpersonal ties that link migrants, former migrants and non-migrants in origin and destination areas by ties of kinship, friendship and community origin*”. Family can also help reduce financial and psychological costs and risks of moving.

Eurenius (2020) investigates the importance of family ties in an individual’s propensity to migrate. The author examines regional mass migration from Sweden to the USA in the late nineteenth century; her results show that chain migration has a role in increasing the individual’s propensity to migrate among both men and women. Migrant networks are a form of social structure that generate and accumulate social capital. Moretti (1999) studies Italian mass migration from 1876 to 1913 and finds that the probability of migrating is affected by the social network that links the migrant to the country. Eriksson (2018) examines the role of ethnic enclaves during Norwegian mass migration to the USA and finds a strong (negative) impact of larger ethnic networks on occupational outcomes. Moreover, Hatton and Williamson (2005) found that the stock of previous migrants eases emigration (network effects). Other studies, however, consider not only family ties but also other sources of aggregation. As mentioned earlier, peer cultural transmission plays a role in shaping a sense of belonging to a community. For instance, Van Meeteren and Pereira (2013) examine Brazilian migration to Portugal and the Netherlands. They find a complementarity between social networks based on community or kin relationships and multiple configurations involving different agents in both the origin and destination countries (such as institutions or web communities).

3. Historical Background: the mass migration in Italy

Economic and demographic determinants played a paramount role in 19th-century migration, with richer countries attracting larger inflows and poorer countries with younger populations and larger shares in agriculture experiencing heavier emigration, further reinforced by network effects. In this context, European countries reached the peak of their mass migration in different decades, from 1860 to 1910.

In particular, Italian mass migration can be settled between 1876 and 1928, when about 17 million Italians emigrated (Del Boca and Venturini, 2003); 14 million left the country before World War I, opening the way to what has been defined as *Italian diaspora* (Ben-Ghiat et al., 2016), one of the greater episodes of European mass migration.

However, from the 1880s to the 1920s, Italian migration is not considered a steady and flat movement; at the turn of the century, one witnesses a sharp rise in migration, which is qualitatively and quantitatively different, and then, during and after World War I, flows change their features again.

During the first twenty years after the Unification, Italian migration was small, seasonal, and mainly directed towards Europe. Inhabitants of the Po Valley migrated to France, Belgium, Switzerland, and Germany for temporary work. Venetians moved to South America to cultivate land. Southern small land tenants moved to North America, intending to return home as soon as their economic conditions improved (Bevilacqua, 2009). Pre-migration attitudes also shaped how people understood the social impact of family members' absence and their economic consequences.

From the mid-1880s onward, migration grew, and transoceanic destinations entered the migration landscape. This movement continued until World War I, and the Great Migration can be divided into three periods.

The first phase, from 1880 to 1900, shows around 14,399 people leaving Italy every year; during the second period (1901-1913) this number dramatically grows, becoming almost three times bigger (every year, on average, 39,150 people emigrate from the Country) and then, in the phase that starts with World War I (1914-1925), levels of migration decrease, touching the average of 16,536 migrants per year. The trend is not steady, but it is characterized by periods of growth and stagnation (Figure 1).

Therefore, the beginning of the XX century represents the first turning point of migration; in 13 years, almost 9 million people leave Italy. World War I marks another step change, ending in the interruption of migration declared by fascism. The years from 1901 to 1913 are crucial, as the huge movement of people can be defined as a real diaspora, bringing social and economic changes.

As Figure 2 shows, in the first period (from 1880 to 1900) migration levels increase not steeply, and Northern Italian regions lead the flows; the gap in the selection of European and Transoceanic destinations is not so large (Figure 4), and those employed in agriculture mainly emigrate from Northern Italy (Figure 3). In the second period (from 1901 to 1913), migration levels increase steeply; furthermore, in this phase, Southern Italian regions lead the flows. The gap between European and Transoceanic destinations widens, and those employed in agriculture mainly emigrate from Southern Italy. During the third period (from 1914 to 1925), World War I stops migration flows; after the war, European and Mediterranean destinations are selected over transoceanic destinations. Migration flows from Northern Regions are higher than

flows from the South; agriculturists from Southern Italy still emigrate more than agriculturists from Northern Italy. Again, the biggest effect is in size and the shift in destination choice (transoceanic versus European countries). These three periods will be separately employed in the next estimates.

Figure 1: Italian Migration Flows, 1880-1921

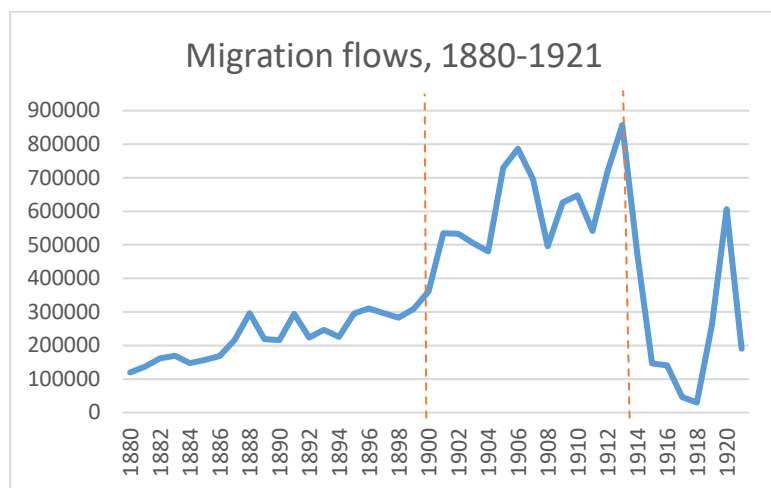


Figure 2: Migration flows by geographical aggregation, 1880-1925

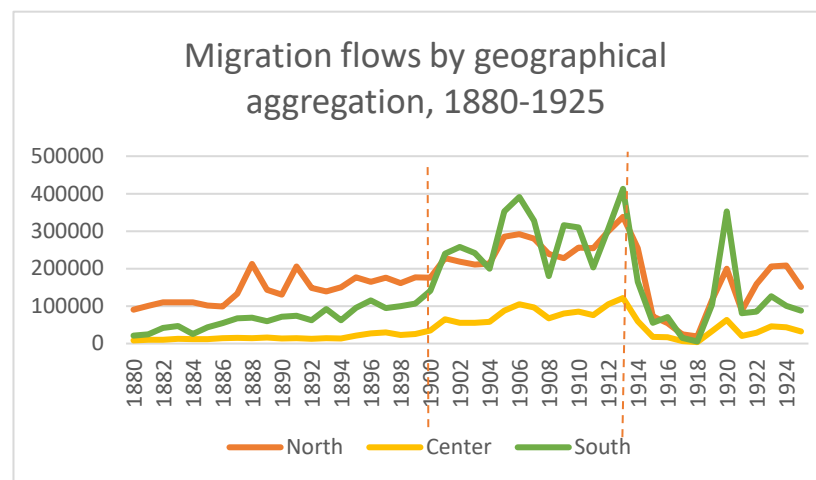


Figure 3: Agricultural Migrants, geographical distribution 1880-1925

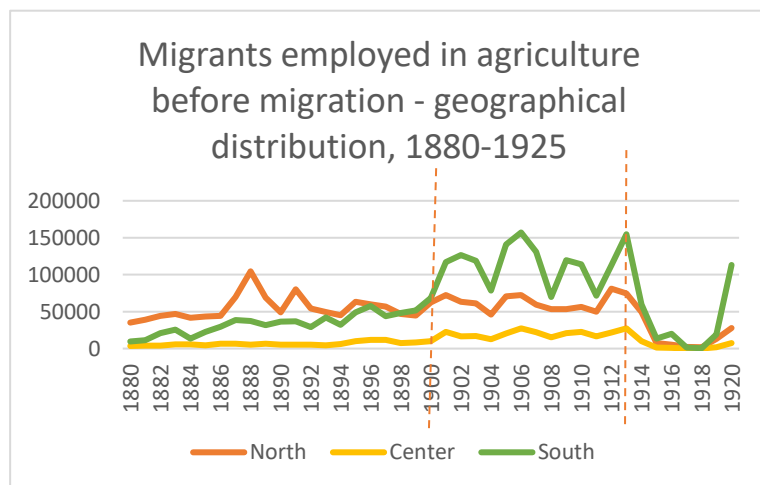
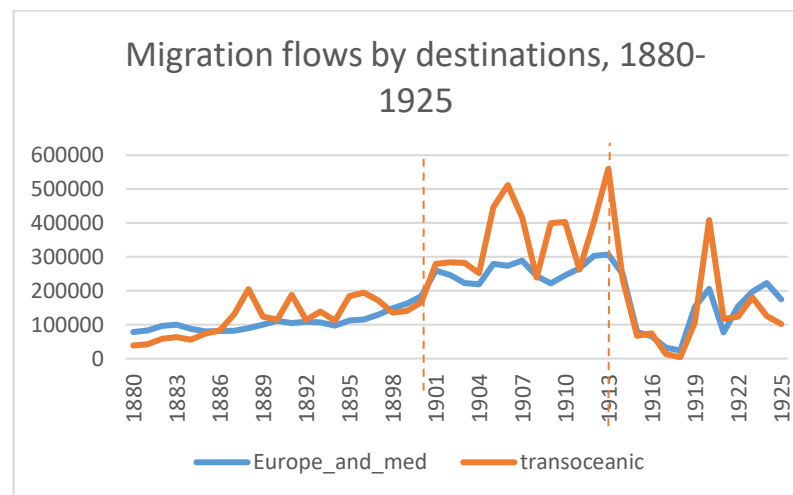


Figure 4: Migration flows by destinations, 1880-1925



4. Empirical strategy

4.1. Data and descriptive statistics

The dataset results from a significant digitization effort⁴. In this case, we created a unique dataset by combining annual regional data on Italian international migration with historical and current data on regional conditions, as well as regional data on FDI and trade.

The sixteen regions involved in the analysis are Piedmont, Liguria, Lombardy, Veneto, Emilia-Romagna, Tuscany, The Marches, Umbria, Lazio, Abruzzo and Molise, Campania, Apulia, Basilicata, Calabria, Sicily and Sardinia, namely the regions existing at the time of mass migration. For reasons of consistency, the current variables also refer to these 16 Regions⁵. Foreign countries considered are Russia, Sweden, Norway, Denmark, UK, Ireland, Belgium, Luxembourg, Netherlands, France, Switzerland, Germany, Austria, Hungary, Spain, Portugal, Gibraltar, Malta, Greece, Romania, Bulgaria, Former Yugoslavia, Canada, USA, Mexico, Colombia, Venezuela, Guyana, Ecuador, Peru, Bolivia, Chile, Brazil, Argentina, Uruguay, Paraguay, Algeria, Tunisia, Morocco, Egypt, Turkey and Oceania. The choice of foreign countries depends on available historical data⁶. For the Former Yugoslavia, which no longer exists, we use current data for the successor countries: Bosnia, Serbia, Kosovo, Montenegro, and Macedonia. Table 1 reports the units we consider.

Table 1: Units of analysis: Italian Regions and Foreign Countries

Italian Regions	Foreign Countries
Piedmont	Russia
Liguria	Sweden, Norway, Denmark
Lombardy	UK, Ireland
Veneto	Belgium, Luxembourg, Netherlands
Emilia-Romagna	France
Tuscany	Switzerland
The Marches	Germany
Umbria	Austria
Lazio	Hungary
Abruzzo and Molise	Spain, Portugal
Campania	Gibraltar, Malta

⁴ We employed the software Able2Extract – Professional version, purchasable at the web site: https://go.investintech.com/able2extract-pdf-software/?gclid=CjwKCAjwu_mSBhAYEiwA5BBmf6yMrPxO5D-MCEhsQJaz9yR2ypAlk9wKe305H5iUQxtXdYIB9VzqXhoCLRoQAvD_BwE, to convert all the OCR files in Excel Format.

⁵ Trento and Bolzano will be separated by Veneto in 1927; Valle d'Aosta will be established in 1946, while Molise will be separated from Abruzzo in 1963.

⁶ We avoided taking the countries defined in the *Annuario* as “other” or “undefined”.

Apulia	Greece, Bulgaria, Romania, Former Yugoslavia
Basilicata	Canada
Calabria	USA
Sicily	Mexico
Sardinia	Colombia, Venezuela, Guyana, Ecuador
	Peru, Bolivia, Chile
	Brazil
	Argentina
	Uruguay, Paraguay
	Algeria
	Tunisia
	Morocco
	Egypt
	Turkey
	Oceania

Current data

Data on FDI are taken from the Reprint Database, provided by the Italian Trade and Investment Agency (ICE), in collaboration with the Politecnico of Milan. We collect data at the provincial level and aggregate them at the regional level. We use the stock of inward FDI (the number of enterprises with foreign participation, transformed into logarithms) in 2017, as data are available only for this year.

Data on regional GDP and regional GDP per capita are from ISTAT (expressed in euros) for the period 2015-2017 (average, logarithm). ISTAT also provides the average number of regional firms from 2015 to 2017.

Data on regional imports and exports come from Coeweb, a specialized online system implemented by ISTAT that has collected monthly trade data since 1991. In this case, we use figures from 2015 to 2017, averaged, expressed in euros and referred to Ateco 2007, 1-letter classification (in this case, we also use the variable in logarithms).

Unfortunately, another limitation concerns the measurement of residents abroad: because figures on current regional emigration flows to the foreign country of destination are not available for this research, we have used data on residents abroad, using their average number from 2015 to 2017. This information is taken from the Report of Migrantes 2015, 2016 and 2017, based on data by AIRE (*Anagrafe Italiana dei Residenti all'Esterro*).

In order to be consistent with the FDI stock (2017), as in other studies (i.e. Cuadros et al., 2019), we apply the average value of the last three years (2015-2017) for the variables *Residents abroad 2015-2017*, *Regional GDP 2015-2017*, *Regional GDP per capita 2015-2017*, *Import 2015-2017*, *Export 2015-2017* and *Regional Firms 2015-2017*.

Finally, we calculated the distance in kilometers from each Italian regional capital to each foreign country's capital city using Google Maps.

Historical data

Regional annual emigration data from 1876 to 1925 are only available in a PDF OCR format, edited by ISTAT Digital Library. *Annuario Statistico dell'emigrazione Italiana dal 1876 al 1925*, indeed, is the volume that contains annual figures and universal statistics on Italian emigration for the cited years. Figures were collected by the two main governmental migration agencies of the time: *Direzione Generale di Statistica* and *Commissariato Generale dell'Emigrazione* (the latter active since 1901). According to the law, migrants had to follow a two-steps procedure: first, they had to apply for a *nulla osta*, a document proving their eligibility for the passport and then they collected their passports, valid for three years. Statistical series of the *Direzione Generale di Statistica* are based on the *nulla osta*. The data have two caveats, but neither suggests systematic error across regions: first, not all *nulla osta* are converted into passports for various reasons; second, not all foreign countries request passports.

Moreover, this paper shares a limitation with studies employing the same data, as it is not possible to account either for return or for seasonal migration, but as for Hatton and Williamson (1998) and Gray et al. (2019), such constraints should not invalidate the analysis, as we are interested in investigating how political choices might influence migration at large.

We digitized data on migration flows on absolute terms, by region of origin and country of destination from 1880 to 1925. The units of analysis for the outcome variable (FDI) and the variables of interest (migration flows) are therefore 416 region-country dyads.

For historical controls, we collect regional population data from the Italian Census, from 1881 to 1921. We digitized figures on population size and the number of literates by region and then took their natural logarithms. A special section of the Census covers inhabitants' professions, and we use it to select data on Italy's agricultural structure at the time as a proxy for occupational structure, following Petraglia and Vecchione (2020) and Hatton and Williamson (1998). Hence, we introduce into the analysis the number of landlords, sharecroppers, and land employees, as we expect them to affect the region's entrepreneurship differently. Finally, data on regional agricultural wages are taken from a reconstruction by Daniele and Malanima (2011) and are expressed in thousands of lira, while data on regional GDP are estimated by Brunetti et al., in Vecchi (2011); this variable is expressed as Index number, compared to Italian GDP that takes value 100 and it is estimate every decade as well as the other Census data.

From *Censimento degli Italiani all'estero, 1871*, we digitized data on the stock of Italian migrants abroad, collected by Italian Embassies in foreign countries, in 1871. In this case, again, we use the natural logarithm of the stock of migrants from Italian regions in each foreign country considered for the analysis.

Finally, we discuss data on Italian institutions abroad, taken from *Censimento degli italiani all'Estero 1927*, in Section 6.1.

Table 2: Variable Description and Source

Variable name	Variable description	Source
<i>Contemporaneous Variables</i>		
FDI (2017)	Log of the number of inward FDI by country, in 16 Italian Regions (stock at 2017)	Italian Trade and Investment Agency (ICE), Reprint Database
Import 2015-2017	Log of total import flows, by region and country, from 2015 to 2017	Coeweb Istat
Export 2015-2017	Log of total import flows, by region and country, from 2015 to 2017	Coeweb Istat
Residents abroad 2015-2017	Mean of resident abroad, by origin region and country of residence, from 2015 to 2017	Migrantes Reports on AIRE data
Distance	Log of geographical distance between the regional and the historical country capital city	Google Instruments
Regional GDP 2015-2017	Mean of Italian regional GDP from 2015 to 2017, log	ISTAT
Regional GDP per capita 2015-2017	Mean of Italian regional GDP from 2015 to 2017, log	ISTAT
Regional firms 2015-2017	Mean of Italian firms by region from 2015 to 2017, log	ISTAT
<i>Historical variables</i>		
Migration flow	Log of migration flows from 1880 to 1900, from sending region to receiving Country;	Annuario dell'Emigrazione Italiana 1876-1925

	Log of migration flows from 1901 to 1913, from sending region to receiving Country; Log of migration flows from 1914 to 1925, from sending region to receiving Country	
Migrants 1871	Log of the stock of migrants abroad, from sending region to receiving Country	Censimento degli Italiani all'Estero 1871
Italian Institutions 1927	Log of the number of Italian institutions abroad in 1927	Censimento degli Italiani all'Estero 1927
Historical Landlords	Mean number of landlords, by region, log	Censimenti italiani 1871,1881,1901,1911,1921
Historical Sharecroppers	Mean number of sharecroppers, by region, log	Censimenti italiani 1871,1881,1901,1911,1921
Historical Land Employees	Mean number of land employees, by region, log	Censimenti italiani 1871,1881,1901,1911,1921
Regional agricultural wages	Mean of agricultural wages, in thousandth of lira, log	Daniele and Malanima (2011)
Historical Regional Literacy	Mean number of literates by region, log	Censimenti italiani 1871,1881,1901,1911,1921
Historical regional population	Mean of population by region	Censimenti italiani 1871,1881,1901,1911,1921

Tables 3 and 4 report the descriptive statistics for the main variables of interest. FDI, the stock of migrants in 1871, and recent imports show high variability. These statistics therefore suggest heterogeneity among the units considered.

Table 3: Descriptive statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
FDI (2017)	416	28.23	106.54	0	1200
Migration flows	416	38475.36	126338.7	2	1173330
Institutions (1927)	416	4251.15	1169.80	2	6105
Residents abroad (15-17)	416	9983.74	21040.92	64	225762
Migrants 1871	416	574.55	2916.59	0	40010
Distance	416	4456.22	4680.56	209	34395
Export (2015-17)	416	2210	4760	6030	46100
Import (2015-17)	416	1800	5420	0	74300

Regional firms (2015-17)	416	262267.9	189218.2	34884	811771
Regional GDP (2015-17)	416	11000	90,2	12,2	375

Notes: FDI are expressed as the number of enterprises with foreign participations; migration flows are expressed in levels; institutions are expressed in levels; residents abroad (2015-2017) and migrants 1871 are expressed in number; distance is expressed in km; import (2015-17), export (2015-17) and regional GDP (2015-17) are expressed in thousand euros; regional firms are expressed in levels.

The correlation matrix shows positive relationships between FDI and mass migration flows, between institutions and FDI, and between the 1871 migrant stock and FDI. The current regional economic fabric (trade, firms, and GDP) positively correlates with FDI, while distance is negatively related to investments. Finally, the number of current residents abroad is positively related to contemporaneous FDI.

Table 4: Correlations matrix for the main variables (n=416)

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) FDI (2017)	1.00									
(2) Migration flow	0.415*	1.00								
(3) Institutions (1927)	0.165*	0.634*	1.00							
(4) Migrants 1871	0.337*	0.414*	0.207*	1.00						
(5) Residents abroad (2015-17)	0.463*	0.547*	0.544*	0.477*	1.00					
(6) Distance	-0.329*	0.071	0.243*	-0.074	0.165*	1.00				
(7) Export (2015-17)	0.682*	0.337*	0.118*	0.202*	0.228*	-0.357*	1.00			
(8) Import (2015-17)	0.585*	0.314*	0.117*	0.254*	0.277*	-0.302*	0.757*	1.00		
(9) Regional firms (2015-17)	0.384*	0.296*	0.000	0.166*	0.186*	-0.041	0.510*	0.539*	1.00	
(10) Regional GDP (2015-17)	0.398*	0.270*	0.000	0.223*	0.169*	-0.036	0.505*	0.525*	0.902*	1.00

Notes: *Statistically significant at 5%.

4.2. Empirical model

The main issue we examine is whether inward FDI volume is related to past mass migration through cultural persistence.

Different approaches have been used for FDI, trade, and migration. The most commonly used one is the gravity model. The classical gravity model assumes that flows between two countries are directly proportional to their size (GDP or population) and are inversely proportional to the physical distance between them. The simplest version of gravity models involves bilateral migration from the origin and destination countries and the distance between them. However, in recent studies (Ramons, 2016), researchers have expanded basic gravity models by adding different push and pull factors that affect the outcome variables (trade, FDI, or migration). Therefore, the enlarged form of the gravity model is considered more suitable for trade and FDI. The enlarged form of the gravity model is known as the extended or modified gravity model (Greenwood, 1997); it adds variables beyond distance and GDP to explain the effect on the outcome variable. Mele and Quarto (2017) extend the standard gravity model by adding variables such as political risk, the degree of openness to international trade, a proxy for natural resources, and cultural proximity, while Wajdi et al. (2017) use an extended gravity model for interregional migration. Finally, Petraglia and Vecchione (2020) use a gravity model that considers both contemporaneous and historical variables to estimate the effect of mass migration on trade.

In this analysis, we use a version of the extended gravity model that, beyond the effects of migration, also accounts for possible regional pulling factors for FDI. Moreover, we introduce historical regional controls to capture regional heterogeneity at the time of mass migration. Indeed, the choice to use dyadic data reflects the fact that, between the end of the XIX century and the beginning of the XX century, not only country differences mattered but regional cultural differences also counted. Italian regions were marked by their own peculiar linguistic, social, and economic contexts, and we suppose those peculiarities also carried over into migrants' communities abroad.

The gravity model used in the analysis is as follows:⁷

$$y_{i,j} = \alpha_j + \beta_1 m_{i,j} + X_{i,j} \theta + C_i \phi + \varepsilon_{i,j} \quad (1)$$

⁷ In contrast to the classical gravity model, here we do not consider bilateral GDP (in levels or per capita), bilateral firms number or bilateral population, but we consider Italian regional GDP (in levels and per capita), regional firms number and regional population (see also Petraglia and Vecchione, 2020).

where i stands for the 16 Italian regions, j indicates the 26 foreign destination countries, $y_{i,j}$ is the stock of inward FDI from country j to region i in 2017, $m_{i,j}$ is the main explanatory variable of interests measuring mass migration flows from region i to country j , β_1 is the estimated coefficient; $X_{i,j}$ is a matrix considering dyadic control variables, that capture factors influencing the current trade relationships and are also related with the level of development of a region in its turn correlated with the past migration flows, C_i is a matrix of current and past regional controls, θ and ϕ are the vectors of estimated coefficients for the controls; α_j are foreign country fixed effects and ε_{ij} is the error term. The dependent variable and all explanatory variables are expressed in logarithms. Equation (1) is estimated by OLS. Standard errors are clustered by region of origin to account for arbitrary correlations among error terms across observations.

The matrix of dyadic control variables is composed of:

- *Italian Residents abroad 2015-2017*, coming from region i to country j , as the current migration can either influence FDI or can be explained by the old migration;
- *Distance* from region i to country j , hence it could affect the choice of the area of investment. Gravity-model literature finds that investment costs rise with distance (Mele and Quarto, 2017); moreover, distance can proxy not only geographical separation among countries or regions but also history, culture, and social relations (Portes and Rey, 1999);
- *Import 2015-2017* from country j to region i and *Export 2015-2017* from region i to country j , are relevant for trade relationships between the considered region/country. Trade literature (among others, Wajdi et al., 2017) recognizes that the latter can stimulate economic growth once international trade has been liberalized. While imports are expected to be negatively related to FDI (as the number of foreign firms in an area reduces the need to import foreign production), export volume is positively related to FDI, as foreign firms can increase local production;
- *Migrants 1871*, namely the stock of migrants from region i who settled in country j in 1871, is a proxy for a rooting effect to the homeland, transferring through the networks. It persists over time and can affect recent economic events, such as investment choices. Indeed, Wajdi et al. (2017) report that migrant stock is a popular proxy to measure the network effects of migration. Migrant stock is defined as the accumulated number of previous immigrants to the destination who have migrated from a particular region of origin, and it describes the linkages between potential migrants in the region of origin

and their relatives and friends who have already settled as migrants in the destination areas. Relatives and friends are supposed to facilitate migration for their beloved.

The current regional controls present in the model are:

- *GDP (2015-2017)*, a measure of the wealth produced by the region. The higher the destination's GDP, the more attractive it is. It can act as a pull factor for inward FDI (Wajdi et al., 2017);
- *GDP per capita (2015-2017)*: it is a key determinant factor, as it indicates the expected earnings of an individual. This variable does not always positively affect FDI, as it can reflect regional labor costs (Mele and Quarto, 2017);
- *Number of firms (2015-2017)* is a proxy for the entrepreneurial ability of the region, and its magnitude can signal good firm policy in the area.

Regional GDP (2015-2017), *Regional GDP per capita (2015-2017)* and *Number of Regional firms (2015-2017)* proxies the current economic regional structure that can be considered as a pull factor in the case of FDI, namely a characteristic of the destination area that – when healthy – can encourage inward FDI and discouraged outward FDI (Bodvarsson and Van den Berg, 2013, Brunea, 2012, Greenwood, 1997).

Past historical regional controls capture features that affect both migration and the attractiveness of foreign direct investment in each region. We include *Historical regional literacy*, a proxy for regional education, and historical regional population, which captures regional size and remains related to both current regional population and economic scale. Heterogeneity among regions was quite strong over the period considered (1880-1925). Finally, we consider historical regional agricultural wages, as low wages can push migration. Indeed, Hatton and Williamson (1998) argue that migration was a crucial factor in reducing the real wage gap with the rest of Europe before World War I. Finally, the agricultural occupational structure (here represented by *historical landlords*, *historical sharecroppers*, and *historical land employees*) reflects that most of the Italian population was employed in agriculture and captures differences in propensity to migrate.

As mentioned before, we expect heterogeneity across the three phases of mass migration, which have distinct qualitative and quantitative characteristics that may affect future FDI differently. For these reasons, we run three separate regressions for the first period (1800-1900), the second period (1901-1913), and the third period (1914-1925). In any case, we also run the regression for the full period (1880-1925), even though, because of heterogeneity, we find it more

informative to keep the three periods separate to show the effects of past migration on future events.

In addition, to partly control for potential endogeneity from unobserved variables, we include regional and country fixed effects in the baseline model and the second specification; in the third and fourth specifications, we replace regional fixed effects with regional controls. However, we acknowledge that endogeneity issues cannot be completely ruled out, especially given the potential problem of time-variant omitted variables.

5. Results

Table 5 reports the results concerning the first period (1880-1900). Column (1) shows the baseline specification, which assesses the relationship between migration flows and FDI. Here, a 1% increase in migration flows is correlated with an 11% increase in FDI. Column (2) shows the results when the model includes all the dyadic variables, both contemporaneous and historical. Column (3) presents the results when the model only includes contemporaneous variables, either dyadic or regional, and Column (4) shows the results for the full model, including regional historical and contemporaneous variables.

The coefficient on migration flows remains positive and highly significant across all specifications, ranging from 0.072 to 0.112. The variable capturing the stock of migrants established abroad before the mass migration, when present (Migrants 1871), is significant and positively related to FDI (0.07). In fact, migrants abroad preferred to concentrate in the same cities, even if they arrived in different eras; as a result, entire neighborhoods in destination towns replicated sending cities. Indeed, Munshi (2016) suggests that the networks variable is typically measured by the stock of migrants from the individual's origin community at each destination. We use this measure because we expect ancient communities established abroad to be positively linked to the outcome variable, suggesting a rooting effect for ethnic groups. The literature finds the coefficient on migrant stock to be positive, as it reduces migration costs through information provided by relatives and friends who previously migrated.

Imports are negatively correlated with FDI, consistent with FDI and trade models, as the presence of foreign business reduces the need for imports. Regional firms are also positively correlated with foreign investment, confirming that the regional economic fabric plays a crucial role in investment decisions.

Table 5: The effect of mass migration on FDI - First period (1880-1900)

	FDI (2017)			
	(1)	(2)	(3)	(4)
Migration flow	0.112*** (0.026)	0.079*** (0.030)	0.072** (0.026)	0.082*** (0.023)
Migrants 1871	-	0.075** (0.038)	-	0.070* (0.035)
Import (2015-2017)	-	- 0.051*** (0.018)	- 0.510** (0.019)	-0.065*** (0.019)
Export (2015-2017)	-	0.49 (0.033)	0.045 (0.031)	0.09* (0.045)
Distance	-	-0.188 (0.173)	-0.213 (0.169)	-0.181 (0.165)
Regional Firms (2015-2017)	-	-	0.184 (0.322)	1.26*** (0.22)
Regional GDP (2015-2017)	-	-	0.228 (0.15)	-0.636* (0.333)
Region fixed effects	Yes	Yes	No	No
Country fixed effects	Yes	Yes	Yes	Yes
Observations	416	416	416	416
R ²	0.84	0.84	0.84	0.83

Notes: The dependent variable is the FDI (logarithm of the number of foreign enterprises). The explanatory variables are defined in the text. All specifications include a constant term. (1) considers migration flows with region and country FE; (2) includes all dyadic variables, with regional and country FE; (3) includes dyadic variables and regional controls with country FE. It does not include historical variables. (4) includes all dyadic variables and all regional variables, with country fixed effects. ***, **, * statistically significant at 1%, 5% and 10%.

Table 6 reports results for the period 1901-1913. Here we find larger coefficients for migration flows, ranging from 0.105 to 0.146. The highest coefficient magnitude may reflect the peak of mass migration during this period. As shown earlier, migration flows increase, on average, by 300% over the period considered. This result suggests that the correlation between FDI and migration grows especially in large-scale people movement, namely in the case of diaspora, where almost an entire community moves and brings with it its own habits and customs, strengthening cultural persistence.

The correlation between imports and FDI, as well as the correlation between the stock of migrants in 1871 and FDI, does not change much in the first period. The coefficient of regional firms decreases in magnitude but remains highly significant (ranging from 0.442 to 0.757).

Results for the third period (1913-1925) are consistent with the first period; as reported in Section 3, the number of migrants decreases in this phase. The coefficient of migration flows ranges from 0.064 in the fourth specification to 0.113 in the first specification. The coefficient on migrants' network in 1871 remains positive and significant. The correlations between regional firms and FDI range from 0.706 to 1.19, while the coefficient on imports ranges from -0.054 to -0.064, both confirming the trends from earlier periods.

Table 8 presents results when heterogeneity across the three phases of mass migration is not accounted for. Mass migration flows are significant and positive in all four specifications, but their magnitude, on average, tends to be smaller than before (from 0.037 to 0.052). The variable migrants 1871 is still positive, but it is not significant. By contrast, the regional economic variables (import and firms) remain stable across periods and model specifications.

To conclude, the results suggest that past matters in the sense that cultural persistence helps sharing the knowledge of their own origin, searching a way to come back home: foreign direct investments, mostly coming from countries that experienced in higher degree Italian mass migration, may symbolize the sense of attachment to the ancestors' country or the intention to know more about the land that one's heard in parents and grandparents tales. Moreover, it could be a way to keep in touch with relatives and start business with them, as family trust is found to be stronger than other types of trust (Eurenius, 2019).

Indeed, mass migration flows and FDI are positively correlated in every period; however, we found a stronger effect in the second period, with higher flows, consistent with literature on the *diaspora effect*. On average, the results show a positive correlation between the presence of older communities in the past and current FDI. Imports and FDI show a negative correlation, consistent with literature on FDI and trade models. Finally, inward FDI are positively linked to regional firm composition.

Table 6: The effect of mass migration on FDI - Second period (1901-1913)

	FDI (2017)			
	(1)	(2)	(3)	(4)
Migration flow	0.146*** (0.028)	0.116*** (0.032)	0.110* (0.053)	0.105** (0.052)
Migrants 1871	-	0.068* (0.036)	-	0.064 (0.04)
Import (2015-2017)	-	- 0.052*** (0.017)	- 0.051** (0.008)	-0.06*** (0.017)
Export (2015-2017)	-	0.030 (0.035)	0.025 (0.035)	0.0282 (0.159)
Distance	-	-0.185 (0.158)	-0.189 (0.158)	-0.186 (0.159)
Regional Firms (2015-2017)	-	-	0.442** (0.207)	0.757** (0.031)
Regional GDP (2015-2017)	-	-	0.046 (0.092)	0.033 (0.39)
Region fixed effects	Yes	Yes	No	No
Country fixed effects	Yes	Yes	Yes	Yes
Observations	416	416	416	416
R ²	0.84	0.85	0.85	0.84

*Notes: The dependent variable is the FDI (logarithm of the number of foreign enterprises). The explanatory variables are defined in the text. All specifications include a constant term. (1) considers migration flows with region and country FE. (2) includes all dyadic variables, with regional and country FE. (3) includes dyadic variables and regional controls with country FE. It does not include historical variables. (4) includes all dyadic variables and all regional variables, with country fixed effects. ***, **, * statistically significant at 1%, 5% and 10%.*

Table 7: The effect of mass migration on FDI – Third period (1913-1925)

	FDI (2017)			
	(1)	(2)	(3)	(4)
Migration flow	0.113*** (0.025)	0.07*** (0.027)	0.076* (0.037)	0.064* (0.04)
Migrants 1871	-	0.089** (0.035)	-	0.083** (0.39)
Import (2015-2017)	-	- 0.054** (0.017)	- 0.054*** (0.017)	-0.064*** (0.016)
Export (2015-2017)	-	0.047 (0.033)	0.04 (0.03)	0.058* (0.029)
Distance	-	-0.188 (0.159)	-0.22 (0.15)	-0.202 (0.16)
Regional Firms (2015-2017)	-	-	0.706*** (0.214)	1.19*** (0.21)
Regional GDP (2015-2017)	-	-	-0.064 (0.115)	0.668** (0.30)
Region fixed effects	Yes	Yes	No	No
Country fixed effects	Yes	Yes	Yes	Yes
Observations	416	416	416	416
R ²	0.84	0.84	0.84	0.83

*Notes: The dependent variable is the FDI (logarithm of the number of foreign enterprises). The explanatory variables are defined in the text. All specifications include a constant term. (1) considers migration flows with region and country FE. (2) includes all dyadic variables, with regional and country FE. (3) includes dyadic variables and regional controls with country FE. It does not include historical variables. (4) includes all dyadic variables and all regional variables, with country fixed effects. ***, **, * statistically significant at 1%, 5% and 10%.*

Table 8: The effect of mass migration on FDI – Overall period (1880-1925)

	FDI (2017)			
	(1)	(2)	(3)	(4)
Migration flow	0.052*** (0.013)	0.04*** (0.014)	0.045* (0.014)	0.037* (0.015)
Migrants 1871	-	0.066 (0.04)	-	0.018 (0.036)
Import (2015-2017)	-	- 0.052*** (0.017)	- 0.05*** (0.017)	-0.061*** (0.017)
Export (2015-2017)	-	0.036 (0.033)	0.031 (0.033)	0.102** (0.03)
Regional GDP (2015-2017)	-	-	0.082 (0.102)	0.62 (0.424)
Regional Firms (2015-2017)	-	-	0.365 (0.23)	1.19*** (0.21)
Region fixed effects	Yes	Yes	No	No
Country fixed effects	Yes	Yes	Yes	Yes
Observations	416	416	416	416
R ²	0.84	0.84	0.84	0.83

*Notes: The dependent variable is the FDI (logarithm of the number of foreign enterprises). The explanatory variables are defined in the text. All specifications include a constant term. (1) considers migration flows with region and country FE. (2) includes all dyadic variables, with regional and country FE. (3) includes dyadic variables and regional controls with country FE. It does not include historical variables. (4) includes all dyadic variables and all regional variables, with country fixed effects. ***, **, * statistically significant at 1%, 5% and 10%.*

6. A possible channel of cultural transmission: Italian institutions abroad

How does cultural transmission strengthen and survive? Can we consider institutions as a channel for fostering current FDI through peer cultural transmission, in addition to the intergenerational transmission channel? In this section, we answer this question by considering Italian institutions settled abroad, as recorded in the *Censimento degli Italiani all'Estero* of 1927. As mentioned before, we expect to find a possible cultural rooting effect, with institutions as a potential transmission channel. To explore eventual heterogeneity, we isolate economic institutions from other kinds of associations, and our results show a stronger effect of the latter on current FDI.

Italian institutions abroad served different functions: they represent places where Italians take care of Italians, operate to preserve customs and traditions, and provide value through community identification. The history of the associations – and then also institutions – set up by Italians abroad is tangled with the social dynamics of integration (or segregation) in destination countries and includes the relationship with the home country. In the destination country, belonging to an Italian institution could be both a source of isolation and a limit on segregation. In destination countries, associations often played political and economic roles alongside a strong cultural role. Cultural institutions abroad fostered religious celebrations and the publication of newspapers and journals. For instance, according to a publication by the Italian Minister of Foreign Affairs, *Associazioni Italiane nel mondo* (2000), in Rosario, Argentina, the institutions born during the age of mass migration have the highest number of enrolled members today (Prencipe, 2019).

Among the first Italian associations abroad, there was the political group of the *Risorgimentali*, Italian political refugees who fled after the failure of the riots of 1821-1861. These associations were open to Italian emigrants willing to join, without restriction. Murat et al. (2011) claim that, by the end of the nineteenth century, associations had created strong links with the home government (which recognized and sometimes subsidized them). When the middle class began emigrating, some associations offered different kinds of services to new migrants, such as job placement (known as the USA Padrone System); however, opportunistic interests drove the actions of certain associations, hiding their purposes behind offers of assistance.

The American Catholic Church tried to overcome the closure of locally based associations to foster national identity and a sense of belonging after Italian Unification. Their efforts included offering hospitality to migrants, exerting political pressure on local and national governments, providing education, raising public awareness of emigrants, and offering health assistance. Moreover, as diplomacy worked to change stereotypes about Italians, commercial associations

(Chambers of Commerce) grew in almost every foreign country. Additionally, Unions started supporting labor, especially issues concerning Italian migrants, who were rarely less protected than natives (Prencipe, 2019).

At the eve of the Great War, Italian associations abroad collected money and resources to support Italian contribution to World War I. Regional associations gave way to national ones, and Fascism entered the world of associations, but Union trades and surviving regional associations opposed it.

Along with the crisis of 1929, migrant associations stopped attracting members, and many associations disbanded. At the beginning of the '50s, Italian associations faced a crisis because many Italians who embraced fascism left the peninsula and moved, especially to South America. During the regime, Fascism encouraged the establishment of Italian political fascist associations abroad to strengthen ties with the home country and gain easier consensus from emigrants. For this reason, many associations dissolved, and new forms emerged that favored cultural and social ties (festivals and religious celebrations, regional provenience, gastronomic traditions, local music). After World War II, moreover, Italian national institutions lowered their attention to Italian institutions abroad, and localism – regionalism and provincialism – became the main feature of this phase; new generations of migrants felt the distance from the previous and suffered difficulty in assimilation. Religious associations, however, continued to feel the gap left by institutions. During the '60s, migrant associations strengthened their ties with the Italian government. In contrast, the decade of the '70s was marked by the establishment of Italian Regional Governments, which became the privileged interlocutor for Italian institutions abroad. At the same time, English-speaking countries recognized the role of cultural diversity and took steps to value migrant communities, such as establishing community centers, opening public spaces for debate, and planning local government-supported cultural events. During the '80s and the '90s, economic associations played a major role in fostering imports and exports and fought to protect and promote Italian brands worldwide (Colucci, 2019).

Among studies on the relationship between migration and institutions, Nifo and Petraglia (2013) deepen the role of institutions in attracting human capital; they find that the most skilled migrants seek a better institutional context, along with better employment opportunities and higher incomes. North (1990) defines institutions as the rules of the game in a society that affect behavior and individual choices. Hall and Jones (1999) and Acemoglu et al. (2001) highlight the impact of history and link institutions to a country's economic performance. Institutions generate mutual trust, foster innovation, expand common knowledge, and enhance local economic activity, making them attractive for human capital migration.

Baudassé et al. (2016) reflect on the fact that migration can both be caused by the quality of institutions (either as a pull factor with respect to institutions in destination countries or as a push factor, with regard to institutions in the origin country), and it can also affect the quality of institutions in sending (through temporary migration) and receiving countries. Durkheim (1982) defines institutions as “*all the beliefs and models of behavior instituted by collectivity*”. Institutions are classified as formal or informal (North, 1990), in institutions providing resource allocation, in governance level (the way in which contractual relations are ruled), in institutional environment (the rules of the game), and in social embeddedness level, which includes norms, customs, habits, and traditions (Williamson, 2000).

Rodrik and Subramanian (2003) divide institutions in *marketing creating institutions* (institutions creating property rights and enforcing contracts), *market regulating institutions* (dealing with market failures), *market stabilizing institutions* (preventing crisis) and *market legitimizing institutions* (providing social protections, insurance, redistributions). In addition, economic institutions result from the interaction between political power and resource redistribution (Acemoglu et al., 2005). Bertocchi and Strozzi (2008) conduct their analysis on data of the 19th century and find that, beyond economic and demographic factors, institutions (political and migration institutions) also played a key role among the determinants of international migration. Finally, Acemoglu and Johnson (2005) compare the relative strength of different sets of institutions, while Persson and Tabellini (2006) decompose the impact of different forms of political institutions (democracy). Engerman and Sokoloff (2002, 2005) conduct a broad comparative analysis of institutional evolution in relation to America's growth. However, for the case of Italy, institutions do not seem to have received adequate treatment in the literature; a study on this issue has been conducted by Murat et al. (2011); they find a positive effect of emigrant associations abroad on bilateral FDI, especially for the inward FDI coming from the countries where these associations settled before. According to the authors, associations represent transnational social networks that bring social capital. They also show that the stock of Italian institutions abroad is independent of host countries' institutions. Their results suggest that the stock of associations captures the transnational social capital accumulated by the Italian diaspora up to that date and is more related to culture than to economic characteristics. Their findings show that the effect of associations on inward FDI ranges from an elasticity of 0.32 to 0.41, while the effect on outward FDI ranges from 0.39 to 0.76.

Considering the long-term effects of migration, we hypothesized that institutions can positively mediate peer cultural transmission for migrants' communities, and that their effect is stronger where these communities are better established.

We consider FDI at the regional level by introducing an interaction term that accounts for the size of the flows, supporting the idea that the larger the Italian presence, the stronger the cultural transmission. In Section 6.1. we introduce the new dataset. In Section 6.2. we present the OLS estimation, and in Section 6.3. we discuss the main findings.

6.1. Data

The *Censimento degli Italiani all'Estero 1927* was a publication that collected data from Italian embassies worldwide. The last census of Italians abroad dated back to 1881, before the 1927 census, but the section on institutions was included for the first time in the 1927 collection. In the *Censimento*, institutions are divided into multiple categories:

- Health institutions: hospitals and clinics, hosting patients from every nationality;
- Educational Institutions: they included schools, research centers, and great international associations aimed at promoting Italian culture all over the world;
- Charity and mutual associations: the former were the first established, and took care of all aspects of migrants' lives, acting as a bridge between Italy and the receiving Country. The latter were mainly established to solve disputes among compatriots;
- Economic associations: this category especially involved Chambers of Commerce;
- Political associations: they were linked to existing Italian political parties or might represent the interests of migrants in that Country;
- Community centers: sports associations, retired clubs, recreational groups (*dopolavoro*): they were less rooted in the social fabric of the country and maintained inconsistent ties with Italian formal institutions.

A foreign country collected institutional data. Unfortunately, the data do not allow us to allocate the institutions to the different Italian regions (as they concern *Italian* associations). Furthermore, this is the stock of associations existing in 1927, but there is no clue about their date of foundation. Up to 1927, the *Censimento* measures 7044 Italian institutions abroad and 4009 different types of press (quarterly journals, daily journals, bulletins, etc.).

Table 9: Number of Italian institutions abroad sorted by continent, 1927

Continent	Europe	America	Africa	Oceania
Type of institution				
Health	2	15	11	2
Education	493	1442	290	0
Charity	488	3232	44	2
Economics	31	29	6	1
Politics	193	227	45	4
Communities centers	112	341	27	7
Total	1319	5286	423	16

Source: Censimento degli Italiani all'Estero 1927

Because we introduce this new variable in the model, Section 3 provides the reference description of the other variables used.

6.2. Empirical specification

The model relies on the idea that cultural persistence transfers via intergenerational and peer channels, and when both factors act, cultural transmission is stronger, so the effect on current FDI is eventually higher. Intergenerational transmission is represented by migrants' families and subsequent generations, while peer transmission is represented by different networks (associations, friends, institutions), as stated before.

As in the previous analysis, we implement a gravity model that extends equation (1) to include two additional variables: $inst_j$, representing Italian institutions in destination countries (in logarithm) and $m_{ij} * inst_i$ that is the interaction term between Italian institutions and migration flows. This interaction allows us to investigate whether the impact of mass migration on current FDI is strengthened by cultural transmission, captured by the presence of Italian institutions in foreign countries. In other words, we test whether Italian institutions in destination countries positively moderate the relationship between mass migration and current FDI inflows. Hence, the second model of the analysis is as follows:

$$y_{i,j} = \alpha_j + \beta_1 m_{i,j} + \beta_2 inst_j + \beta_3 m_{i,j} * inst_i + X_{i,j} \theta + C_i \phi + \varepsilon_{i,j} \quad (2)$$

where i stands for the 16 Italian regions, j indicates the 26 foreign destination countries, $y_{i,j}$ is the stock of inward FDI from country j to region i in 2017, $m_{i,j}$ is the main explanatory variable of interests measuring mass migration flows from region i to country j , β_1 is the estimated coefficient; $instj$ represents the stock of Italian Institutions abroad in 1927 in country j , $X_{i,j}$ is a matrix considering dyadic variables (namely *Italian Residents abroad 2015-2017*, coming from region i to country j , *Distance* from region i to country j , *Import 2015-2017* from country j to region i , *Export 2015-2017* from region i to country j , *Migrants 1871*, namely the stock of migrants from region i settled in country j in 1871), C_i is a matrix of current and past regional control, θ and ϕ are the vectors of estimated coefficients for the controls; α_j are foreign country fixed effects and ε_{ij} is the error term adjusted for clusters in regions.. The dependent variable and all explanatory variables are expressed in logarithms.

Again, the advantage of using logged coefficients is that β s can be interpreted as elasticity. In this case, another limitation arises: the variable of interest, namely institutions, varies across countries but not across regions, so it should be treated as a kind of fixed effect. Therefore, this specification does not allow adding country-fixed effects. Equation (2) is estimated by OLS with standard errors clustered at the regional level.

We also run four regressions, covering different time spans (1880-1900, 1901-1913, 1914-1925, and the full period). Additionally, we run the same regressions by exclusively considering economic institutions, as we expect them to have a different – positive, higher in magnitude – impact on our outcome.

6.3. Results

Table 10 reports the results. The coefficient on migration flows is significant and positive in all periods considered: a 1% increase in migration flows increases FDI by 0.040% in the overall period to 0.12% in the second period (note that the second period has the largest migration flow). The interaction between migration flows and institutions shows a positive link with FDI, even though, on average, it seems to reduce the magnitude of the flows' effect. The negative coefficient on institutions may reflect the absence of country fixed effects, so it may not capture the correlation between this variable and FDI. Moreover, the negative sign could reflect cultural distance between Italy and the destination country; in this case, institutions were established where cultural, geographical, and political distance was higher to protect migrants' habits. However, where Italian institutions were established, migrants could preserve tighter cultural ties with Italy and, over time, transmit a sense of belonging to the home country; this cultural

transmission is reflected in the positive interaction between institutions and migration flows. It is worth remembering that the variable Migrants 1871 represents the ancient Italian regional community settled abroad, and it is no longer significant. One possible explanation is the rooting effect of Italian migrant communities: ancient settlements lose strength in cultural transmission as they transfer into institutions.

Distance remains negative and highly significant, consistent with gravity-model literature: a 1% increase in distance reduces inward FDI by around 3.3% across all periods considered.

The coefficient of exports ranges from 0.318 to 0.328; its positive sign aligns with literature finding complementarity between FDI, migration, and trade (Kugler and Rapoport, 2007; Aubry and Rapoport, 2019). Current residents abroad keep a positive and significant link with FDI, confirming migration relevance for foreign direct investments; for an increase of 1% in the number of residents abroad, FDI rises from 3.1% to 3.4% (Tong, 2005, Parsons and Vézina, 2014, Mayda et al., 2011).

Table 10: FDI, migration and institutions

	FDI (2017)			
	1880-1900	1901-1913	1914-1925	1880-1925
Migration flow	0.068*** (0.021)	0.116*** (0.028)	0.101* (0.037)	0.040*** (0.009)
Institutions	-0.165** (0.58)	- 0.250** (0.062)	- 0.220** (0.063)	-0.234*** (0.063)
Migration flow*Institutions	0.019* (0.10)	0.055** (0.009)	0.052** (0.009)	0.014*** (0.003)
Migrants 1871	-0.003 (0.049)	0.019 (0.041)	0.014 (0.04)	0.002 (0.020)
Residents abroad (2015-2017)	0.342*** (0.044)	0.305*** (0.052)	0.314*** (0.102)	0.321*** 0.029
Distance	-0.342*** (0.044)	0.329*** (0.108)	0.319*** (0.102)	0.328*** (0.062)
Export (2015-2017)	0.3839 (0.072)	0.318*** (0.069)	0.322*** (0.06)	0.328* (0.115)
Import (2015-2017)	0.01 (0.04)	0.026 (0.05)	0.028 (0.049)	0.125 (0.054)
Regional Firms (2015-2017)	0.109 (0.284)	0.077 (0.347)	0.261 (0.252)	0.379 (1.005)
Regional GDP (2015-2017)	-0.722 (0.209)	0.161 (0.404)	0.109 (0.312)	-0.118 (0.370)
Historical regional controls	Yes	Yes	Yes	Yes
Observations	416	416	416	416
R ²	0.63	0.65	0.65	0.64

*Notes: The dependent variable is the FDI (logarithm of the number of foreign enterprises). The explanatory variables are defined in the text. All specifications include a constant term. Errors clustered by Regions. ***, **, * statistically significant at 1%, 5% and 10%.*

In Table 11, we address institutional heterogeneity and account for a special type of institution that should be strongly related to FDI, namely commercial and economic institutions (mostly Chambers of Commerce). As expected, the interaction between migration flows and institutions is larger than before, suggesting a stronger moderating role for economic associations. Coefficients are consistently positive and significant across the entire period analyzed. The stronger effect, suggesting the possible explanation we provided before, is during the second period, when an increase of 1% in the interaction entails FDI of 15%; for the other two periods, the effect of associations on inward FDI ranges from an elasticity of 0.11 to 0.14, a bit lower than the findings reported by Murat et al. (2011), who also consider recent institutions.

Table 11: FDI, migration and economic institutions

	FDI (2017)			
	1880-1900	1901-1913	1914-1925	1880-1925
Migration flow	0.0465*** (0.015)	0.082*** (0.023)	0.071*** (0.032)	0.025*** (0.008)
Economic Institutions	-0.307** (0.120)	-0.442*** (0.102)	-0.391*** (0.107)	-0.411*** (0.112)
Migration flow*Institutions	0.106*** (0.024)	0.157*** (0.019)	0.143*** (0.016)	0.050*** (0.006)
Migrants 1871	0.018 (0.018)	0.035 (0.018)	0.028 (0.018)	0.025 (0.019)
Residents abroad (2015-2017)	0.295*** (0.029)	0.259*** (0.028)	0.265*** (0.027)	0.270*** (0.025)
Distance	-0.357*** (0.056)	-0.372*** (0.056)	-0.356*** (0.055)	-0.375*** (0.056)
Export (2015-2017)	0.376** (0.123)	0.328** (0.118)	0.323 (0.118)	0.328** (0.120)
Import (2015-2017)	0.008 (0.052)	0.017 (0.05)	0.019 (0.048)	0.015 (0.049)
Regional Firms 2015-2017)	0.177 (0.239)	0.072 (0.315)	0.217 (0.237)	0.193 1.021
Regional GDP (2015-2017)	-0.845 (0.302)	0.171 (0.399)	0.084 (0.311)	-0.110 (0.363)
Historical regional controls	Yes	Yes	Yes	Yes
Observations	416	416	416	416
R ²	0.64	0.65	0.65	0.65

*Notes: The dependent variable is the FDI (logarithm of the number of foreign enterprises). The explanatory variables are defined in the text. All specifications include a constant term. Errors clustered by Regions. ***, **, * statistically significant at 1%, 5% and 10%.*

The role of institutions in the destination country seems to encourage integration, rather than assimilation, where Italian and local cultures coexist. Following Lockwood (1964), sociological theory of social systems has developed the concepts of system integration and social integration: social integration refers to the inclusion of individuals in a system, the creation of relationships among individuals, and their attitudes towards society. It results from the conscious, motivated interaction and cooperation of individuals and groups. Assimilation, by contrast, is understood as a one-sided process in which immigrants and their descendants give up their culture and adapt completely to the society they have migrated to. From this perspective, Italian communities abroad, especially those with a localist character, appear to help spread Italian culture, either among Italians or throughout destination countries, through the interaction and cooperation Lockwood describes.

7. Conclusion

Cultural persistence has recently been considered a crucial factor in current choices. Research across multiple fields has examined the role of history in explaining today's economic conditions (Alesina, 2007, 2013; Acemoglu, 2005). Migration has also been used in studies of long-run cultural effects (Abramitzky et al., 2020; Beck-Knudsen, 2019; Haque and Kim, 2019; Abramitzky et al., 2014; Sequeira et al., 2018), and this paper addresses these issues. This work aims to deepen understanding of one of the long-run effects of Italian mass migration: the development of regional FDI inflows.

Using a relatively unexplored database, through a gravity model and by separately considering the three different waves of Italian mass migration from 1880 to 1925, we observe that there is a positive correlation between mass migration flows and FDI in every period, even though we found a greater effect in the period with higher migration flows, as predicted by the literature on the *diaspora effect*. On average, the results also show a positive correlation between the presence of older communities in the past and current FDI.

Moreover, by enriching the model with an interaction between past Italian institutions abroad and migration flows, a possible channel through which cultural persistence is transmitted - along with intergenerational transmission - appears, namely the peer channel represented by those associations. Results suggest that institutions in the destination country seem to encourage a pattern of integration, rather than assimilation, in which Italian and local cultures coexist.

Therefore, the main innovative aspects of this paper are two-fold. First, using an almost unique dataset, we study FDI and Italian mass migration at the regional level. Little literature considers

FDI at the regional level, and, as far as we know, no other studies connect regional FDI with past regional migration, at least for Italy.

Second, we propose an alternative channel for cultural transmission by explaining how Italian institutions abroad established before 1927 can moderate the relationship between migration flows and FDI.

The main limitation of this study is that the empirical strategy is a gravity model estimated with OLS, so the results should be taken with a grain of salt, as they show correlations between the variables of interest rather than causal relationships. While one should be cautious about interpreting these findings as a causal relationship between past mass migration and current FDI because endogeneity is not fully addressed, we believe this study can answer questions relevant today.

The degree of historical background that every individual brings with him or herself, the diffusion of cultural belonging through structured communities - one may say the proudness of being Italian, in this case- may be treated as a universal value that can be useful in implementing integration policies not only for the people who migrate for economic reasons, but also for refugees, political or climate migrants.

References

- Abramitzky, R., Boustan, L., Eriksson, K., (2020), Do Immigrants assimilate more slowly today than in the past?. *American Economic Review Insight*, 2 (1): 125-141.
- Abramitzky, R., Boustan, L.P., Eriksson, K. (2014). A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration. *Journal of Political Economy*, 122 (3): 467-506.
- Acemoglu D., Johnson S. (2005), Unbundling Institutions, *Journal of Political Economy*, 113: 949-995.
- Acemoglu D., Johnson S. and Robinson J.A. (2005), Institutions as a fundamental cause of long-run growth, *Handbook of economic growth*, 1: 385-472.
- Acemoglu D., Johnson S. and Robinson J.A. (2001), The colonial origin of comparative development: an empirical investigation, *American Economic Review*, 91: 1369-1401.
- Alba R., Nee V. (1997), Rethinking Assimilation Theory for A New Era of Immigration, *International Migration Review*, 31(4): 826-874.
- Alesina and Fuchs-Schuendeln (2007), Good bye Lenin (or not?) The Effect of Communism on People's Preferences, *The American Economic Review* 97(4): 1507-1528.
- Alesina A., Giuliano P., Nunn. N. (2013), On the Origins of Gender Roles: Women and the Plough, *The Quarterly Journal of Economics*, Vol. 128: 469-530.
- Anchol H. (2000), An Examination of Cross-Country Differences in the Gender Gap in Labor Force Participation Rates, *Labor Economics* 7(4): 709-426.
- Aubry A., Rapoport H. (2019), Migration, FDI and the Margin of Trade, *Discussion Paper LEM* 2019-22.
- Bajo-Rubio O., Diaz-Mora C. (2010), Foreign Direct Investment and Regional Growth: An Analysis of the Spanish Case, *Regional Studies*.
- Beck-Knudsen, A.S., (2019), Those Who Stayed: Individualism, Self-Selection and Cultural Change during the Age of Mass migration. *Discussion Paper*, 19 (01): 1-71.
- Ben-Ghiat R., Malia Hom S., (2016), Beyond the Italies? Italy as a Mobile Subject, in *Italian Mobilities*, Routledge University Press: 20-45.
- Bertocchi G., Strozzi C. (2008), International Migration and the Role of Institutions, *Public Choice* (2008) 137: 81-102.

- Beverelli C., Keck A., Larch M., Yotov Y. (2018), Institutions, trade and development: A quantitative Analysis. *Technical Report, LeBow College of business*. Drexel University.
- Bevilacqua P., De Clementi A., Franzina E. (2009), *Storia dell'Emigrazione Italiana, Partenze*, Donzelli.
- Boyd R. and Richerson P.J. (2005), *Culture and the Evolutionary Process*, University of Chicago Press.
- Bodvarsson O., Van den Berg H., 2013, *The Economics of Immigration, Theory and Policy*. 2nd Eds, New York Springer.
- Borjas G.J. (1992), Ethnical Capital and International Mobility, *Quarterly Journal of Economics*, 107(1): 123-150.
- Bronzini R., (2004), Foreign direct investment and agglomeration: Evidence from Italy, *Temi di discussion del Servizio Studi*, N. 526, Banca d'Italia.
- Buchardi, K., Chaney, T., Hassan, T., (2018), Migrants, Ancestors and Foreign Investments. *Review of Economic Studies*, 0: 1-39.
- Bunea D. (2012), Model Gravity Model in International Migration: The case of Romania, *Theoretical and Applied Economics*, 19(4): 127-144.
- Cavalli-Sforza L.L., (2001), *Genes, Peoples, and Languages*, University of California Press.
- Chiswick, B. R., & Hatton, T. J. (2003). International migration and the integration of labor markets. In M. Bordo, A. Taylor, & J. A. Williamson (Eds.), *Globalization in historical perspective*: 65-119. Cambridge, National Bureau of Economic Research.
- Colucci M. (2009), L'associazionismo di emigrazione nell'Italia repubblicana, in Bevilacqua P., De Clementi A., Franzina E. (2009), *Storia dell'Emigrazione Italiana*, Donzelli.
- Costa-Font J., Sato A. (2016), Cultural Persistence of Health Capital: Evidence from European Migrants, Health, *Econometrics and Data Group, The University of York*, WP 16/09.
- Cuadros A., Martín-Montaner J., Paniagua J. (2019), Migration and FDI; the role of job skill, *International Review of Economics and Finance*, 59: 318-332.
- Daniele V., Malanima P. (2017), Regional wages and the North-South disparity in Italy after Unification, *Rivista di Storia Economica*, Issue 2: 117-158.

- Dinesen P.T (2013), Where you come from or where you live? Examining the cultural and institutional explanation of generalized trust using migration as a natural experiment. *European Sociological Review* 29: 114-128.
- Dinesen P. T., Sønderskov K. M. (2018), Cultural Persistence or Experiential Adaptation? A Review of Studies using Immigrants to Examine the Roots of Trust, *The Oxford Handbook on Social and Political Trust*, Oxford University Press.
- De Clementi A., (2009), La grande migrazione: dalle origini alla chiusura degli sbocchi americani, *Storia dell'Emigrazione Italiana, Partenze*, Donzelli.
- Del Boca D., Venturini A., (2003), Italian Migration, *IZA DP* No.938.
- Durkheim E. (1982), The rules of sociological method and selected texts on sociology and its method, *The Free Press*, New York.
- El Yaman S., Kugler M., Rapoport H. (2007), Migrations et investissements directs étrangers dans l'espace européen, *Revue économique*, vol, 58(3): 725-733.
- Engerman, S., & Sokoloff, K. L. (2002), Factor endowment, inequality, and paths of development among New World economies, *Economía*, 3: 41-109.
- Engerman, S., & Sokoloff, K. L. (2005), Institutional and non-institutional explanations of economic differences, in C. Ménard & M. M. Shirley, *Handbook of new institutional economics*: 639-665, Springer
- Eriksson K. (2018), Ethnic Enclaves and Immigrants outcomes: Norwegian Immigrants During the Age of Mass Migration, *NBER Working Paper* n. 24763.
- Eurenius A. (2019), A family Affair: Evidence of Chain Migration during the Mass Emigration from the county of Halland in Sweden to The United States in the 1890s, *Population Studies*, Vol. 74, No.1: 103-118
- Ferrie J.P., Hatton T.J., (2013), Two century of International Migration, *IZA PD* No.7866.
- Florio E., Stefano M. (2021), Ancestors, inter-generational transmission of attitudes and corporate performance: Evidence from the Italian Mass Migration, *GLO Discussion Paper* no. 830.
- Foley C.F., Kerr W.R. (2011), Ethnic Innovation and U.S. Multinational Firm Activity, *NBER Working Paper* no. w17336.

- Fouka, V., (2019), Backlash: the Unintended Effects of Language Prohibition in US Schools after World War I, *The Review of Economic Studies*, RDZ024: 1-52.
- Fouka, V., Mazumder, S., Tabellini, M., (2019). From Immigrants to Americans: Race and Assimilation During the Great Migration. *Harvard Business School BGIE Unit Working Paper* 19-018: 1-86.
- Gheasi, M., Nijkamp, P. (2017). A Brief Overview of International Migration Motives and Impacts, with Specific References to FDI. *Economies*, 11 (3): 1-11.
- Giavazzi F., Petkov I., Schiantarelli F. (2019), Culture: Persistence and Evolution, *Journal of Economic Growth*, vol. 24 (2): 117-154.
- Giuliano P., Nunn N., (2021), Understanding Cultural Persistence and Change, *Review of Economic Studies*, 88: 1541-1581.
- Giuliano P., Spilimbergo A., (2014), Growing Up in a Recession, *The Review of Economic Studies* 81(2).
- Granovetter M.S. (1973), The Strength of Weak Ties, *American Journal of Sociology*, 78(6): 1360-1380.
- Greenwood M.J., (1997), International Migration in Developed Countries, in *Handbook of Population and Family Economics*, Vol. 1B:647-720, Elsevier B.V.
- Grompone A., Sessa L. (2015), Cultural Persistence? Evidence from an Administrative Reform on Borders of Southern Italy, Bank of Italy, Mimeo.
- Gruber J. and Hungerman D.M., (2008), The Church Versus the Mall: What Happens When Religion Faces Increased Secular Competition?, *Quarterly Journal of Economics*, 123(2): 831-862.
- Guiso L., Sapienza P. and Zingales L., (2011), Civic Capital as the Missing Link, *EIEF Working Papers Series* 1005, Einaudi Institute for Economics and Finance (EIEF)
- Hall R., Jones C.I., (1999), Why do some countries produce so much output per worker?, *Quarterly Journal of Economics*, 114: 83-116.
- Haque, N., Kim, S. J. (1995). Human Capital Flight: Impact of Migration on Income and Growth. *IMF Staff Papers*, 42: 577-607.
- Harach M., Rodriguez-Crespo E., (2014). Foreign Direct Investment and trade: A bi-directional gravity approach. Kiel Advanced Studies Working Papers, n. 467.

- Hatton T.J., Williamson J.G. (2005), *Global migration and the world economy: two centuries of policy and performance*, Cambridge, Mit Press.
- Hatton T., Williamson J.G., (1998), *The Age of Mass Migration: Causes and Economic Impact*, Oxford University Press.
- Helliwell, J.F., S. Wang and J. Xu (2016), How durable are social norms? Immigrant trust and generosity in 132 countries. *Social indicators Research* 128: 201-219.
- Helpman E., Melitz M.J., Yeaple S.R. (2004), Export versus FDI with Heterogeneous Firms, *American Economic Review*, vol. 94, n. 1: 300-316.
- Javorck B.S., Neagu C., Ozden C. (2011), Migrant Networks and Foreign Direct Investment, *Journal of development Economics*, vol. 94, n.2: 231-241.
- Javorick B. S., Ozden Ç., Spatareanu M., Neagu C. (2006), Migrant Networks and Foreign Direct Investment, *World Bank Policy Research Working Paper* 4046.
- Kogut B., Singh H. (1988), The effect of national culture on the choice of entry mode, *Journal of International Business Studies*, 19(3): 411-432.
- Kugler, M., Rapoport, H., (2007), International Labor and Capital Flows: Complements or Substitute?, *Economic Letters*, 94: 155-62.
- Lockwood D. (1964), Social Integration and System Integration, in Zolischank K. and Kirsh W., *Explorations in Social Change*, London Routledge and Krgan.
- Luttmer E.F. and Singhal M. (2011), Culture, Context and the Taste for Redistribution. *American Economic journal: Economic Policy* 3: 157-159.
- MacDonald J. S. , MacDonald L.D. (1964), Chain migration, ethnic neighborhood formation and social networks, *Milbank Memorial Fund Quarterly* 42 (1): 82-97.
- Malan, F.A. (2015), Complementarity between FDI and Migration: Using the Fall of the Berlin Wall as a Natural Experiment, *IFAC Papers On Line*, 48 (24): 201-206.
- Massey Douglas S. (1990), Social structure, household strategies and the cumulative causation of migration, *Population Index*, 56(1): 3-26.
- Mayda A.M., Parsons C., Pham H., Vézina P.L. (2022), Refugees and Foreign Direct Investment: Quasi-Experimental Evidence from US Resettlements, *Journal of Development Economics*, 156.

Mele M. (2009), L'impatto della cultura sul commercio internazionale nel Sud Est asiatico, *Mercati, Valori, Regole*.

Mele M., Quarto A. (2017), A Gravitational Model for Estimate the Determinants of Outward Foreign Direct Investment of China, *International Journal of Economics and Financial Issues*, 7(1): 1-5.

Moretti E. (1999), Social networks and migration: Italy 1876-1913, *The International Migration Review*, 33(3): 640-657.

Murat M., Pistoiesi B., Rinaldi A. (2011), Transnational Social Capital and FDI. Evidence from Italian Associations Worldwide, *Working Paper Series*, Recent, Center for Economic Research.

Nifo A., Petraglia G. (2013), Do Institutions Play a Role in Skilled Migration? The Case of Italy, *Regional Studies*, 2013.

North D.C. (1990), *Institutions, Institutional Changes and Economic Performance*, Cambridge University Press, Cambridge.

Ottaviano G.I.P., Peri G. (2006), The Economic Value of Cultural Diversity: Evidence from US Cities, *Journal of Economic Geography*, 6: 9-44.

Parsons C., Vézina, P.L. (2016), Migrants Networks and Trade: The Vietnamese Boat People as a Natural Experiment, *IZA DP* no. 10112.

Persson, T., & Tabellini, G. (2006). Democracy and development: The devil in the details. NBER Working paper No. 11993.

Petraglia C., Vecchione G. (2020), Long-run pro trade effects of diaspora: evidence on Italian regions, *Spatial Economic Analysis*, vol.16 (1): 47-72.

Portes R, Rey H. (1999), The Determinants of Cross-Borders Equity Flows, *Working Paper* 7336, NBER.

Prencipe L. (2019), L'associazionismo italiano all'estero: una continua storia di "relazioni". Il contributo al processo di unificazione, *Working Paper*, available at <https://www.migrantes.it/wp-content/uploads/sites/50/2019/03/associazionismo-estero.pdf>.

Putterman L., Weil D.N., (2010), Post-1500 Population Flows and The Long-Run Determinants of Economic Growth and Inequality, *The Quarterly Journal of Economics*.

Ramons R. (2016), Gravity model: A tool for migration analysis, *IZA World of Labor*, 239.

- Rapoport H. (2016), Migrations et flux financiers internationaux, *Revue d'économie financière*, no. 122: 69-79.
- Rauch J.E. (2002), Ethnic Chinese Network in International Trade, *The Review of Economic and Statistics*, vol. 84, no.1: 116-130.
- Rodrik D., Subramanian A. (2003), The primacy of institutions, *Finance and Development*, 40(2): 31-34.
- Rose A.K. (2018), Currency wars? Unconventional monetary policy does not stimulate exports. *Working paper 24817*, NBER.
- Sequeira, S., Nunn, N., Qian, N., (2018), Immigrants and the Making of America. *Review of Economic Studies*, 2020, 87(1): 382-419.
- Tabellini G., (2010), Culture and Institutions: economic Development in the Regions of Europe, *Journal of the European Economic Association, Papers and Proceedings*, Vol.8: 677-716.
- Thierry B., Baziellier R., Issifou I. (2016), Migration and Institutions: Exit and Voice (from Abroad)?, *hal-02484265*.
- Tong S. Y. (2005). Ethnic networks in FDI and the impact of institutional development, *Review of Development Economics*, 9(4): 563-580.
- Van Meetren M., Pereira S. (2013), The Differential Role of Social Networks: Strategies and Routes in Brazilian Migration to Portugal and the Netherlands, *Norface Migration, Working Paper* no. 2013-10.
- Vecchi G. (2011), *In ricchezza e in povertà. Il benessere degli italiani dall'Unità a oggi*, Il Mulino.
- Voigtlaender V., Voth J. (2012), Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany, *Quarterly Journal of Economics*, Vol. 127: 1339-1392.
- Wajdi N., Adioetomo S.M., Mulder C.H. (2017), Gravity Models of Interregional Migration in Indonesia, *Bullettin of Indonesian Economic Studies*, 5(3): 309-332.
- Wickramasinghe A.A.I.N., Wimalaratana W. (2016), International Migration and Migration Theories, *Social Affair*, Vol.1 (5): 13-32.
- Williamson O.E. (2000), The new institutional economics: taking stocks, looking ahead, *Journal of Economic Literature*: 595-613.

Zucker L.G. (1977), The Role of Institutionalization in Cultural Persistence, *American Sociological Review*, Vol. 42, No.5: 726-743.

Depopulation of Rural and Inner Areas in Southern Italy

Dina Moawad^{1,2}

Abstract

Depopulation due to the demographic changes in Italy is an urgent policy challenge that requires immediate attention. Drawing from 678 questionnaires administered as part of the "COVID-19 and Migration: A Local Development Perspective for Inner Areas" research project, this Policy Brief aims to highlight the depopulation problem in rural and inner areas of southern Italy, especially after COVID-19, outline the significant development challenges and opportunities, evaluate different policy responses, and propose emerging policy recommendations at the national level.

¹ Department of Business and Economics, University of Naples Parthenope, Naples, 80132, Italy; ² Department of Economic Policy University of Gastronomic Sciences, Pollenzo – 12042, Italy; d.moawad@unisg.it

1- Introduction

The demographic change in Italy is not a concern for the distant future but a pressing policy challenge that demands immediate attention. Over the past few decades, many rural, inner areas have experienced a decline in population as agriculture has been restructured, leading to a concentration of population and employment in urban centers. By 2050, the population of urban regions in Europe is projected to increase by 24.1 million people, providing homes to almost half of the EU-28 population. In contrast, the population of rural areas is projected to decrease by 7.9 million (Eurostat, 2016).

The COVID-19 pandemic brought significant changes in social and demographic dynamics, leading to considerable relocation from urban centers to rural and inner areas. The adoption of remote work and distance learning has facilitated this shift. University students and workers who had previously migrated to major urban centers might choose to return to their home regions due to economic uncertainty, concerns about contagion, and the desire to reduce the high living costs associated with studying and working in large cities. These migratory movements may have a significant impact on the development of inner areas, taking various factors into account. Hence, the COVID-19 pandemic has prompted a critical examination of the shortcomings of the dominant urbanization model (Lucatelli, 2020; Sharifi & Khavarian-Garmsir, 2020) and emphasized the importance of prioritizing inner areas in new sustainable development approaches.

Based on the findings of a research project titled "**COVID-19 and Migration: A Local Development Perspective for Inner Areas**," which involved 678 questionnaires to evaluate the impact of COVID-19 on migrants from rural and inner areas, this Policy Brief aims to highlight the issue of depopulation in these areas of southern Italy following the COVID-19 pandemic. It aims to present the significant development challenges and opportunities, evaluate different policy responses, and propose emerging policy recommendations at the national level.

2- The Depopulation in Rural and Urban Areas, Its Challenges, Opportunities, And Current Policies in Italy

Depopulation in rural and inner areas can be interpreted differently based on national, regional, and local contexts. This topic has become the focus of increasing research and policy attention, particularly following the significant impact of COVID-19 on Italy and its demographic characteristics. The depopulation manifests as a declining population, leading to a growing imbalance between the supply and demand of services. This imbalance creates challenges for both the public and private sectors. Weak local markets result in underutilized and poorly maintained services that often become unviable and have to be discontinued. Consequently, local living conditions and quality of life deteriorate, unemployment increases, and skilled labor becomes scarce, leading to abandonment and obsolescence. Therefore, it is crucial to take a broader perspective and use an analytical framework to understand depopulation and propose mitigating or coping strategies.

In rural and inner areas, there are significant social, economic, and environmental challenges such as:

- 1- The high concentration of **elderly individuals** which is often linked to depopulation. The proportion of people over 65 is exceptionally high in rural and inner areas and has increased over time. Due to the aging population and depopulation, ensuring the availability and accessibility of public services in rural and inner areas becomes complex.
- 2- The limited availability of **broadband Internet** in rural and inner areas is a significant concern. Italian telecom companies currently only offer broadband services in municipalities with a minimum of 1,000 customers, leaving approximately 5,600 smaller

municipalities with limited access to broadband. This impacts technology adoption in these areas.

- 3- Depopulation significantly impacts labor productivity and undermines the sustainability of education and healthcare services due to a decrease in the number of workers and students and increased demand for healthcare.
- 4- Depopulation, with all its consequent impacts, exerts substantial pressure on public budgets, which are burdened with reduced fiscal income, escalating costs, and increased social expenditures. This precarious situation places governance capacity in a precarious position while simultaneously hindering proactive entrepreneurship policies. There is a strong correlation between underdeveloped regions, particularly in eastern and southern Europe, and poor governance quality, which is identified as the main obstacle to development in low-growth areas ([European Commission, 2017](#)).

While the flow of foreign workers could offset the depopulation of rural and inner areas, even integrating them poses a challenge, as rural regions are not currently equipped for this process. Instead of considering it an opportunity to revitalize and diversify rural areas, immigrants are often viewed primarily as a source of cheap labor for agriculture and other sectors.

In Italy, the rural policy approach is complex and diverse, with several innovative projects and strategies contributing to rural development. The National Strategy for Inner Areas (SNAI) is an innovative public policy addressing depopulation in designated inner areas, determined by their distance from essential service centers. SNAI relies on an ambitious place-based policy that utilizes new multilevel local governance for integrated local promotion and development. It aims to address demographic challenges and respond to the needs of territories penalized by significant geographical and demographic handicaps. The success of this policy is essential in responding to the new planning challenges brought on by different events, such as the COVID-19 pandemic.

The inner areas, as defined by SNAI, are firstly significantly distant from the main centers that provide essential welfare services such as education, healthcare, and mobility. Secondly, these areas are rich in environmental resources, including water resources, agricultural systems, natural and human-made environments, and cultural resources such as historical villages and craft centers. Lastly, the territory is diverse due to the different dynamics of natural systems and human activity ([Barca et al., 2014](#)). As a result, inner areas make up about 60% of the national territory, including over 4000 municipalities and more than 20% of Italian residents ([De Matteis, 2013](#)). Among them, 72 project areas, grouping several municipalities classified as inner areas, have been selected for the SNAI's implementation (Figure 1).

Figure 1. Seventy-two Italian project areas were selected with the National Strategy for Inner Areas.



Source: Agency for Territorial Cohesion

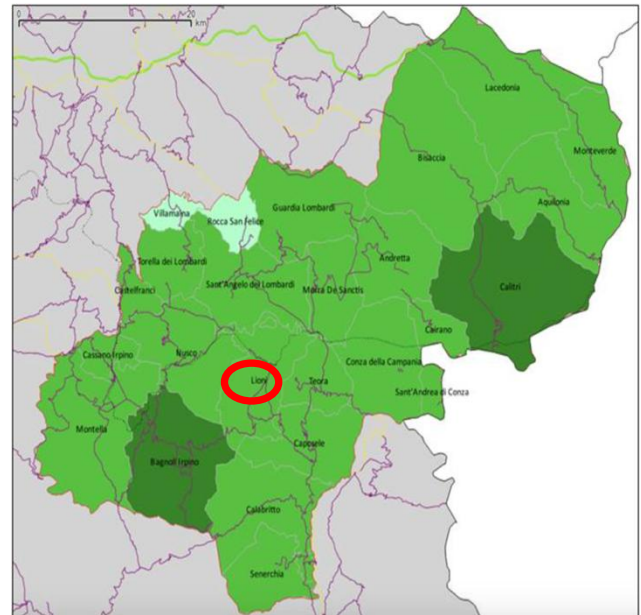
3- Project of “COVID-19 and Migration: A Local Development Perspective for Inner Areas”

The analysis and policy recommendations in this brief are based on the research project “COVID-19 e Migrazione: Una Prospettiva di Sviluppo Locale per le Aree Interne” which was conducted by Università degli Studi di Napoli Parthenope in collaboration with Ministero dell'Università e della Ricerca (MUR) to evaluate whether the migratory flows from major cities to the inner areas of Italy due to COVID-19 which facilitated by remote work and distance learning, can help alleviate territorial disparities. This potential reversal of the migration trend may contribute to the repopulation and economic development of

these areas if these migratory flows become a long-term phenomenon. The research aligns with the goals of the National Strategy for Inner Areas (SNAI) in addressing Italy's inner areas, particularly in the southern part of Italy.

Therefore, this research selected Lioni, a municipality in southern Italy designated by the SNAI and located in the Alta Irpinia region of Campania (Figure 2).

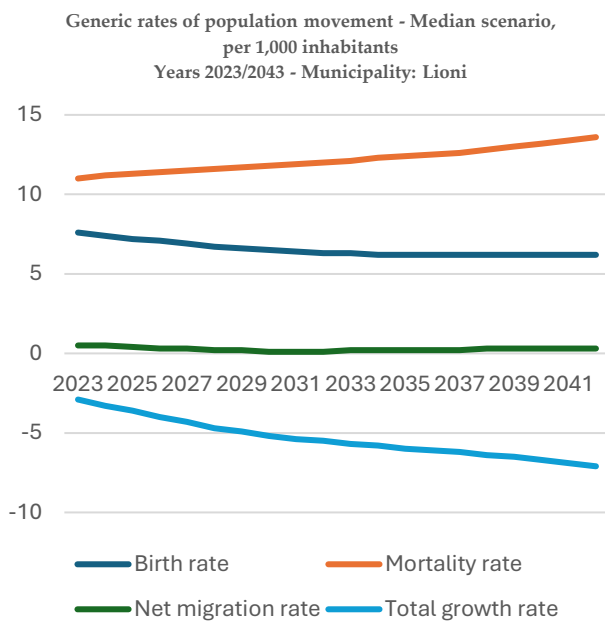
Figure 2: Lioni located in the Alta Irpinia region as chosen by SNAI



Source: Agency for Territorial Cohesion

Lioni is also an interesting case as it is experiencing a significant demographic shift, with clear signs of depopulation. According to ISTAT, projections for the total population indicate a decrease from 5942 inhabitants in 2023 to 5339 inhabitants in 2042, reflecting a negative growth rate of -7.1%, as shown in (Figure 3).

Figure 3: Expected Demographic Trends in Lioni 2023-2043



Source: ISTAT

Moreover, based on the analysis of the conducted questionnaire in Lioni, the study findings confirmed that individuals who migrated during the COVID-19 pandemic, whether internally or internationally for work, education, or other reasons, faced significant vulnerabilities compared to non-migrants, which led them to choose to return to their small origins. Most of those who returned during the pandemic to Lioni subsequently remigrated again, primarily to resume in-person work or studies that could not be conducted remotely or to pursue new job opportunities after facing challenges in finding opportunities in their municipality. Therefore, it is premature to conclude that the return of migrants due to COVID-19 was a short-term occurrence. Even though they have the potential for remote work or distance learning but, they may lack the necessary resources.

Another question was added at the end of the questionnaire by the Lioni municipality governing body's inquiry on preventing depopulation of rural areas like Lioni. Most of the population has emphasized the importance of protecting and enhancing the cultural, historical, and artistic heritage of the region. This sentiment stems from the profound significance of preserving their history and heritage, as they are integral to people's identities and are seen as crucial for the well-being of future generations. Others have suggested that safeguarding and enhancing the landscape heritage, biological diversity, and natural resources would be more beneficial. They argue that improving existing infrastructure is a more cost-effective approach than starting from scratch. Furthermore, they believe that leveraging the unique qualities of Lioni municipality's nature would contribute to the area's strength and resilience. Lastly, some individuals advocate for considering new building areas to revitalize the territory economically, viewing investment in the economic aspect as crucial for creating job opportunities and preventing outward migration.

The previous evidence has emphasized the importance of including small and rural municipalities, such as Lioni, in efforts to address depopulation and its resulting socio-economic impact, like the approach taken by SNAI.

4- Policy recommendations:

4.1. Immigration as a solution to depopulation and labor shortages in rural and inner areas.

With the rise in immigration to smaller cities and rural areas, scholarly focus is broadening beyond the traditional emphasis on larger cities and metropolitan areas. The increasing political significance of immigration and integration is being felt globally, particularly in the EU, which experiences high influxes of migrants each year. Meanwhile, with the increasing trend of depopulation, several scholars characterized immigration as the solution to labor shortages and depopulation in rural and inner areas.

The issue in Italy is the growing flow of mixed immigrants arriving by its borders each year, especially the sea. A record 150,000 migrants came by sea to Italy in 2023, higher than the 94,000 who arrived in 2022 (ISTAT 2023). To deal with this issue, in November 2023, the Italian government signed an agreement with the Albanian authorities that aims to move immigration and asylum detention outside Italian territory at 40,000 per annum. Italy will pay €16.5 million to inland areas in Albania in a five-year plan as well as cover all running costs of operating the centres during this time – although the figure is likely now to be anything between €300 million and €900 million.

Some studies argued that there is concern about the Protocol's impact on search and rescue operations and asylum in Italy and the EU. It raises legal, practical, and constitutional issues under Italian, EU, and international law. It poses a significant risk of lowering and systematically violating common standards for asylum procedures, reception conditions, expulsions, and detentions (Carrera et al., 2023). Other practical and ethical problems are raised (Juss, 2024). Firstly, there is a track record of such arrangements being unsuccessful. In April 2023, a comparable agreement between the EU and Tunisia fell through. Also, in June 2023, a British court suspended the refugee agreement between the UK and Rwanda due to concerns that Rwanda was not considered a “safe third country.”

Furthermore, there are some questions without answers: What happens to asylum-seekers rejected in Albania if their country does not take them back? Will Italy accept them? Italy has limited repatriation agreements, so how will it handle the return of failed asylum-seekers? Out of 150,000 migrants who entered Italy in 2023, only 4,000 were repatriated. When transferred to Italy, both successful and unsuccessful applicants become Italy's responsibility, burdening the country. If unsuccessful asylum-seekers in Albania travel to another EU country like Croatia, how will other EU countries perceive this? Could these migrants become vulnerable to human trafficking? Albania remains a significant source of vulnerable individuals seeking refuge in Europe. During the 2015 migration crisis, Albanian refugees ranked second only to Syrians, with 54,762 asylum claims lodged in Germany alone (Eurostat, 2016).

Before this, the Italian government has taken steps to regulate the employment of foreign migrants in certain sectors, notably agriculture and tourism. In 2017, the government issued the 'Decreto Flussi,' which allowed for a maximum of 17,000 non-EC foreign seasonal workers in these sectors. However, this number is widely considered to be an underestimation of the actual demand for foreign labor, as previous estimates suggested a much more significant presence of temporary workers without Italian citizenship, potentially numbering around 320,000 (Osti, 2018). Foggia province tops the list with approximately 20,000 seasonal migrants, followed by South Tyrol (Alto Adige) with around 18,000 (Eurostat, 2016). Overall, foreigners constitute over 10% of the total workforce in Italy, with approximately half of the labor force in the primary sector being comprised of foreign workers. This concentration of migrants is particularly pronounced in the agriculture sector, where they are involved in both permanent jobs, such as dairy and livestock activities, as well as temporary jobs in fruit and horticulture production throughout the country. **This is further evidence that immigrants are needed not only to repopulate rural and inner areas but also to help the agricultural sector flourish and absorb all the labor shortage.**

Subsequently, For the social side, Italy has implemented the "**Italy 2020 Integration Plan: Security, Identity, and Engagement**," which applies to all migrants and focuses on integrating them into society. Regional authorities play a significant role in integration, particularly in areas such as labor, housing, social welfare, vocational training, and health. The success of this policy was evident in certain inner areas, such as in (Case Study 1).

Case Study 1: Riace Municipality in Reggio Calabria, Southern Italy **(European Committee of the Regions, 2020)**

Riace is a small town in Reggio Calabria with about 2000 inhabitants and about 470 migrants.¹⁸² It is an interesting example of the agility of small cities to respond promptly and creatively to new migration realities and to take advantage of migration to regenerate the town. It is also, however, an example of the sustainability risks that come with measures implemented on the significant push of a small group of individuals in the town. Since the early twentieth century, the very low employment rate in Riace pushed local inhabitants to relocate to Northern Italy and other European countries (Sky tg24, 2018). However, since the late 1990s, Riace has started to receive inward migration, with the arrival of two hundred refugees from Kurdistan to Riace Marina (the maritime part of Riace). The town, first through NGOs and then through the mayor, interpreted the arrival of refugees as an opportunity to revive the territory from the progressive depopulation it had suffered until then (La Repubblica, 2019). Reggio Calabria has other successful examples of municipalities recognized for having “Best Practices” in managing international migrants, such as Camini (Guillory, 2023).

Hence, rather than attempting to address these challenges by, for example, relocating immigrants to inland areas such as Albania at a substantial cost, Italy could leverage the influx of human capital to repopulate its own inland regions, which are experiencing a concerning trend of depopulation and its associated socioeconomic consequences.

Case Study 2: The Atlantic Immigration Program in Canada

Targeted government initiatives, such as the Atlantic Immigration Program and various Provincial Nominee Program pathways, position immigration as a solution to labor shortages and depopulation in smaller communities. The Atlantic Immigration Program (AIP) is a critical initiative by the Canadian government that offers a pathway to permanent residency for foreign workers and international graduates seeking to work and settle in one of Canada's four Atlantic provinces, which are considered remote areas. This program addresses the labor market needs and demographic challenges in these provinces. As a pilot of the program, in December 2019, 5,590 newcomers arrived in the Atlantic Provinces. Almost three-quarters (73%) of newcomers arrived in Canada in 2019 and then initiated steps to be involved in the AIP program. Overall, the Pilot met the expected outcomes and helped Atlantic Canada fill labor market needs ([Government of New Brunswick, Canada, 2025](#)).

These immigrants could help repopulate and fill the labor shortages with lower salaries in these areas. At the same time, the funds that would have been spent on relocation could instead be invested in developing the infrastructure and overall community to accommodate them. The government could also attract public and private investments in rural and inner areas to enhance agricultural capacity by utilizing the available labor force of immigrants at a lower cost.

As we have seen, Italy has recently established some efficient policies about immigrants and foreign workers, but they are separate. At this stage, it is essential to establish a comprehensive and systematic framework that integrates current immigrant and foreign worker policies to effectively regulate their presence and leverage their contribution as human capital. By implementing it strategically, it could serve as a best practice example for other European countries to emulate as an Italian model. Canada provides an example of such strategies, as illustrated in (Case Study 2).

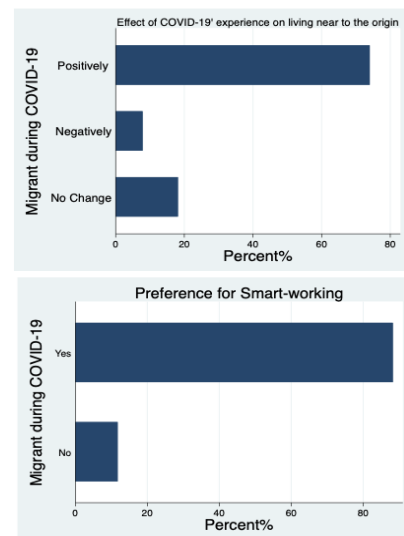
4.2. Solutions to address the resource gap for individuals working or studying remotely in small towns: Co-working spaces (CSs)Model.

Coworking spaces (CSs) are considered spaces for "joining a community" ([Akhavan, 2021](#); [Gandini, 2015](#); [Spinuzzi, 2012](#)). They function as community-based organizations built on principles of mutual assistance and collaboration, which play a pivotal role in fostering innovation. They also provide adequate spaces not only for freelancers but also for individuals who work or study remotely and cannot work from home or prefer to have an outside space ([Capdevila, 2022](#)). Previous studies have emphasized that establishing CSs in rural and inner areas can have several benefits. These include reducing traffic congestion and pollution, improving the well-being of workers by cutting down on commuting time, enhancing work-life balance, and contributing to the local community by retaining local knowledge workers and attracting new ones. This, in turn, can lead to the socioeconomic development of the area by utilizing previously vacant spaces ([Avdikos & Merkel, 2020](#); [Akhavan et al., 2019](#); [Boutillier et al., 2020](#)). Due to COVID-19, the rural and inner areas of southern Italy have witnessed a phenomenon where there was returned migration, which led to the establishment of "**community garrisons**." These garrisons are public and aim to accommodate remote workers and retain young people by providing them with opportunities ([Mariotti, 2022](#)). The shift to remote work and distance learning during the pandemic has created an opportunity for the South and Italian inner regions to be repopulated, attracting skilled professionals and retaining young and qualified talent. It also contributes to the long-term development and well-being of these areas.

According to the questionnaire conducted in Lioni, 74% of individuals who were migrants during the pandemic reported having a more positive experience about their place of origin due to COVID-19. Additionally, 88% of them expressed a desire to permanently reside in Lioni if they had the opportunity to work or study remotely. Hence, they identified critical infrastructure needs for remote work, including a reliable internet connection (84%) and access to public workspaces (53%), as shown

in (Figure 4). However, it is crucial to recognize that while the nature of inner areas can improve the quality of life, the lack of essential infrastructure, such as broadband Internet, can hinder various types of work (Mariotti & Di Matteo, 2020). The existing evidence demonstrates the significance of having public coworking spaces (CSs) in inner areas like Lioni.

Figure (4): COVID-19's effect on migrants living near the



Hence, policymakers should encourage the establishment of CSs in inner areas by accommodating them in public libraries or other public spaces (Di Marino & Lapintie, 2018). CSs have a profitable nature, which means it can be an investment or at least can be self-financing. In that case, it is essential to assess (i) the potential demand from coworking and evaluate the long-term sustainability of the project; (ii) the technical feasibility and initial costs; and (iii) the risk of diluting the coworking concept, potentially leading to a less dynamic and engaging environment in favor of a more static and utilitarian use of the space (Di Marino & Lapintie, 2018). Therefore, it is anticipated that targeted policies, combined with grassroots initiatives, will foster the proliferation of new workspaces in rural and inner areas, promoting entrepreneurship and collaborative work culture while maintaining the spontaneous and flexible nature of coworking (Fuzi, 2015). These spaces can help retain migrants who are interested in working or pursuing academic opportunities remotely for the long term. Subsequently, this is also beneficial for the repopulation and overall development of these areas. This approach also aims to reduce the disparity between main and inner areas, explicitly addressing the issue of "places that do not matter" (Rodríguez-Pose, 2018).

Conclusion

In the upcoming decades, many rural areas are expected to continue depopulating, which will require a carefully managed economic transition. Lioni, like other inner areas in southern Italy, is facing a clear risk of depopulation due to expected demographic changes in the future. The questionnaire conducted in the joint project "COVID-19 and Migration: A Local Development Perspective for Inner Areas" in Lioni confirmed that COVID-19 prompted migrants to return to their hometowns, using the advantage of work or studying remotely. The analysis indicated that this was a short-term trend, as these individuals migrated again after the end of the pandemic. However, the analysis revealed that the pandemic altered the potential for migrants to return to and work from small towns if they have the necessary resources and services. As it is difficult to accept the depopulation socially and politically or rely solely on local development strategies to combat it, this policy brief offers two recommendations to mitigate this phenomenon. Firstly, to address depopulation, it is vital to consider foreign migrants and refugees as a potential solution to depopulation and labor shortages in rural and inner areas such as Lioni. Secondly, providing the necessary services to support the return of migrants in a long-term phenomenon, utilizing the availability of working or studying remotely. One crucial service is the coworking spaces (CSs), which allows individuals to work remotely from their small towns, such as Lioni.

References

- Akhavan M (2021) Third places for work: a multidisciplinary review of the literature on coworking spaces and maker spaces. In: Mariotti I, Di Vita S, Akhavan M (eds) *New workplaces: location patterns, urban effects and development trajectories. A worldwide investigation*. Springer, Cham, pp 13–32.
- Akhavan M., Mariotti I., Astolfi L., Canevari A., 2019, "Coworking Spaces and New Social Relations: A Focus on the Social Streets in Italy." *Urban Science*, 3, 2: 1-11. Doi:10.3390/urbansci3010002.
- Avdikos V., Merkel J., (2020), "Supporting open, shared and collaborative workspaces and hubs: recent transformations and policy implications." *Journal of Urban Research and Practice*, 13, 3: 348-357. Doi: 10.1080/17535069.2019.1674501
- Barca, F.; Casavola, P.; Lucatelli, S. *Strategia Nazionale Per le Aree Interne. Definizioni, Obiettivi e Strumenti di Governance*. Mater. UVAL 2014, 31, 10 Available at: https://www.agenziacoesione.gov.it/wp-content/uploads/2020/07/MUVAL_31_Aree_interne.pdf accessed on 14 November 2023.
- Boutillier S., Capdevila I., Dupont L., Morel L., (2020), "Collaborative Spaces Promoting Creativity and Innovation." *Journal of Innovation Economics & Management*, 31, 1: 1-9. Doi: 10.3917/jie.031.0001.
- Capdevila I (2022) *Building communities in rural coworking spaces*. In: Mérindol V, Versailles DW (eds) *Open labs and innovation management: the dynamics of communities and ecosystems*. Routledge, UK.
- Carrera, S., Campesi, G., & Colombi, D. (2023). THE 2023 ITALY-ALBANIA PROTOCOL ON EXTRATERRITORIAL MIGRATION MANAGEMENT The worst practice in migration and asylum policies. In <https://www.ceps.eu/ceps-publications/the-2023-italy-albania-protocol-on-extraterritorial-migration-management/>. PAPERS ON LIBERTY AND SECURITY IN EUROPE (CEPS). Retrieved February 10, 2024, from https://cdn.ceps.eu/wp-content/uploads/2023/12/ASILE_Italy-Albania-MoU-Extraterritorial-Migration-Management.pdf
- De Matteis, G. Montagne e aree interne nelle politiche di coesione territoriale italiane ed europee. *Territorio* 2013, 66, 7–15. Available at: <https://www.documentigeografici.it/index.php/docugeo/article/viewFile/64/56> accessed on 15 November 2023
- Di Marino M., Lapintie K., 2018, «Exploring multi-local working: challenges and opportunities for contemporary cities». *International Planning Studies*, 25, 2: 1-21. Doi: 10.1080/13563475.2018.1528865.
- European Commission., (2017), *Annual activity reports*. Retrieved from https://commission.europa.eu/publications/annual-activity-reports-2017_en accessed on 12 October 2023
- European Committee of the Regions. Commission for Citizenship, Governance, Institutional and External Affairs. & British Institute of International and Comparative Law. (2020). *Integration of migrants in middle and small cities and in rural areas in Europe*. Publications Office. <https://data.europa.eu/doi/10.2863/281960>.
- Eurostat., (2016) *Annual activity report 2016*. European Commission. Retrieved from https://commission.europa.eu/publications/annual-activity-report-2016-eurostat_en accessed on 12 October 2023
- Fuzi A., (2015), «Co-working spaces for promoting entrepreneurship in sparse regions: the case of South Wales». *Regional Studies, Regional Science*, 2, 1: 462-469. Doi: 10.1080/21681376.2015.1072053.
- Gandini A (2015) The rise of coworking spaces: a literature review. *Ephemera* 15(1):193–205
- Government of New Brunswick, Canada. (2025, February 19). *Atlantic Immigration Program*. <https://www2.gnb.ca/content/gnb/en/corporate/promo/immigration/immigrating-to-nb/atlantic-immigration-program.html>
- Guillory, S. (2023). Camini: a little Italian village that's making a large global impact. Su Guillory. <https://www.susanguillory.com/post/camini-a-little-italian-village-that-s-making-a-large-global-impact>
- ISTAT., 2023, *Migrazioni interne e internazionali della popolazione residente*. Retrieved from: <https://www.istat.it/it/files/2024/05/Migrazioni-interne-e-internazionali-della-popolazione-residente.pdf> accessed on 15 October 2023
- Juss, S. S. (2024). *How the EU Pact on Migration and Asylum Fails*. Retrieved October 10, 2024, from: https://www.questionegiustizia.it/data/speciale/articoli/1206/speciale_luglio2024_juss_eng.pdf
- La Repubblica, 'Nasce la nuova Riace: una fondazione per rilanciare quel modello di accoglienza', 12 January 2019. https://www.repubblica.it/cronaca/2019/01/12/news/nasce_la_nuova_riace_una_fondazione_per_rilanciare_quel_modello_di_accoglienza-216427574/
- Lucatelli, S. Riflessioni sulle Aree Interne, all'indomani del Covid-19. In *Aree Interne e Covid*; Fenu, N., Ed.; LetteraVentidue: Siracusa, Italy, 2020; pp. 12–23. Available at: https://iris.polito.it/retrieve/e384c434-8098-d4b2-e053-9f05fe0a1d67/2020_Aree%20interne%20e%20COVID.pdf accessed on 13 November 2023.
- Mariotti I, Pais I (2022) *Lavoro a distanza e nuovi luoghi del lavoro: il ruolo dei coworking pubblici in Italia*. In: Mirabile M, Militello E (ed) a cura di, *South Working. Per un futuro sostenibile del lavoro agile in Italia*. Donzelli Editore, Roma, pp 81–87
- Mariotti I., Di Matteo D., 2020, «Coworking in emergenza Covid-19: quali effetti per le aree periferiche?». *EyesReg*, 10, 2: 1-5.
- Osti, G. (2018). Migrants to rural areas as a social movement: Insights from Italy. *Bio-Based and Applied Economics*, 243-257 Pages. <https://doi.org/10.13128/BAE-23339>
- Rodríguez-Pose A., (2018), «The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society*, 11, 1: 189-209. Doi: 10.1093/cjres/rsx024.
- Sharifi, A., & Khavarian-Garmsir, A. R. (2020). The COVID-19 pandemic: Impacts on cities and major lessons for urban planning, design, and management. *The Science of the Total Environment*, 749, 142391. <https://doi.org/10.1016/j.scitotenv.2020.142391>
- Sky tg24, 'Riace, il borgo abbandonato divenuto modelli di integrazione', 2 October 2018, available at: <https://tg24.sky.it/cronaca/approfondimenti/riace-modello-integrazione.html>
- Spinuzzi C (2012) Working alone together: coworking as emergent collaborative activity. *J Bus Tech Commun* 26(4):399–441

Family Reunification, Integration and Well-being Outcomes of Immigrants

Maria Rosaria Carillo and Tiziana Venittelli

Abstract

This paper examines the relationship between family reunification and immigrants' subjective well-being, as well as the role played by the institutional conditions under which reunification takes place. Using data from the 2011 ISTAT Survey on the Social Integration of Immigrants in Italy, we first compare immigrants living in Italy with their spouses with those whose spouses remain abroad. Living with one's spouse is associated with significantly higher life satisfaction, even after controlling for a broad set of demographic, socioeconomic, migration, integration, and residential characteristics. We then shift the focus from the presence of the spouse in Italy to the institutional context and timing of the reunification process. Restricting the sample to immigrants who reunited with their spouses, we compare reunifications occurring before and after the introduction of the Bossi-Fini Law in 2002. Post-reform reunification is associated with substantially longer periods of spousal separation, lower life satisfaction, and a lower probability of feeling at home in Italy. The findings therefore suggest that, although living with one's spouse is positively associated with subjective well-being, restrictive reunification procedures and prolonged family separation may weaken these benefits and hinder immigrants' integration into the host society.

Keywords: Migration, Family Reunification, Life Satisfaction, Integration

JEL classification: F22, I31, J12

1. Introduction

Family reunification represents one of the main channels of legal migration to the European Union, accounting for approximately one third of residence permits issued to non-EU migrants in recent years. The European Union recognises family reunification as an important means of safeguarding family life and supporting immigrants' integration into the host society. Living with one's family can provide emotional support, stability, and a sense of security, helping immigrants cope with the challenges of migration and adaptation to a new social environment.

A growing interdisciplinary literature documents the adverse psychological consequences of prolonged family separation. Migrants who spend extended periods apart from their family members tend to experience greater emotional distress, anxiety, depressive symptoms, post-traumatic stress, and an increased risk of mental disorders (Choummanivong et al., 2014; Miller et al., 2018; Liddell et al., 2021; Hvidtfeldt et al., 2022). These consequences may be particularly severe for children and young adults, for whom separation from their parents can generate psychosocial difficulties that persist over time (Eremenko and Bennett, 2018; Cebolla Boado and González Ferrer, 2022).

Consistent with this evidence, empirical studies suggest that family reunification can improve immigrants' well-being, particularly among refugees who have experienced prolonged separation from their relatives (Löbel, 2020; Löbel and Jacobsen, 2021). By restoring emotional support, companionship, and family stability, reunification may reduce loneliness and alleviate the psychological distress associated with separation. Beyond its effects on well-being, the presence of family networks may also facilitate immigrants' economic and social integration by providing support with employment, housing, childcare, and adaptation to the destination country (Gubernskaya and Dreby, 2018).

Although the beneficial effects of family reunification are widely recognised, the literature also highlights that their intensity and degree depend on the conditions under which reunification occurs. The design of reunification policies can influence the duration of separation, the administrative and financial burdens faced by applicants, and the circumstances in which family members begin living together in the host country. More accessible procedures may reduce the emotional costs of separation, whereas restrictive requirements may generate uncertainty, financial pressure, and lengthy waiting periods (Nika, 2017; Phillimore et al., 2023). Moreover, even after reunification, inadequate economic and integration support may leave families exposed to employment insecurity, unsuitable housing, childcare difficulties, and problems of post-reunification adjustment (Bonizzoni, 2009). The relationship between family reunification and immigrants' well-being is therefore likely to be shaped

both by the restoration of family unity and by the institutional and socioeconomic conditions surrounding the reunification process and the policies governing it.

This paper contributes to this debate by examining the relationship between family reunification, subjective well-being, and social integration among married immigrants from non-EU countries living in Italy. We use microdata from the 2011 ISTAT Survey on the Social Integration of Immigrants in Italy, which provides detailed information on family arrangements, migration histories, socioeconomic conditions, and immigrants' attitudes towards the host country. The empirical analysis proceeds in two stages. First, we compare married immigrants living in Italy with their spouses with those whose spouses remain abroad. Second, among immigrants who were joined by their spouses after arriving in Italy, we compare reunifications occurring before and after the introduction of the Bossi-Fini Law in 2002. The reform tightened the Italian family reunification framework by strengthening income, housing, residence, and documentary requirements and by making the administrative procedure more demanding.

Our findings show that immigrants living with their spouses in Italy report significantly higher life satisfaction than those whose spouses remain abroad, even after controlling for a broad set of demographic, socioeconomic, migration-related, integration, and residential characteristics. However, the institutional conditions under which reunification occurs also appear to matter. Immigrants who reunited after the introduction of the Bossi-Fini Law experienced substantially longer periods of separation from their spouses and report lower life satisfaction than those who reunited before 2002. They are also less likely to report a strong attachment to Italy, as they are less likely to say that they feel at home in the country, suggesting a lower degree of social integration. These associations remain when the analysis is restricted to immigrants who had already arrived in Italy by 2002, thereby reducing concerns that the results are driven entirely by changes in the composition of migration flows induced by the legal reform.

These findings provide a more nuanced perspective on family reunification. Although living with one's spouse is positively associated with subjective well-being, the magnitude of this benefit appears to depend on how and when reunification takes place. Lengthy periods of separation, together with the administrative, financial, and emotional burdens associated with restrictive procedures, may erode part of the well-being gains generated by the restoration of family unity. Family reunification should therefore not be viewed simply as the admission of a family member, but as a process whose consequences are shaped by its duration and by the institutional and socioeconomic conditions under which it occurs.

The lower levels of attachment to Italy observed among immigrants who reunited after the Bossi-Fini reform also suggest that restrictive reunification policies may have consequences extending beyond individual well-being. More demanding requirements and longer waiting periods may weaken immigrants' identification with the host country and reduce their sense of belonging to the receiving society, by generating uncertainty, frustration, and sense of exclusion. Taken together, the results indicate that facilitating timely and less burdensome family reunification may support not only immigrants' subjective well-being, but also their longer-term social integration into the host country.

The rest of the paper is structured as follows. Section 2 presents the institutional framework governing family reunification in Italy and discusses the potential effects of the policy changes introduced in 2002. Section 3 describes the data and the empirical strategy. Section 4 reports the results. Finally, Section 5 concludes by discussing the policy implications and outlining directions for future research.

2. Family reunification in Italy

In Italy, the legal framework governing family reunification for non-EU immigrants was substantially reshaped and tightened in the early 2000s with the introduction of the Bossi-Fini Law (Law No. 189/2002). The reform adopted a more restrictive approach by strengthening minimum-income and housing requirements, placing greater emphasis on stable legal residence in Italy, limiting the categories of family members eligible for reunification, and expanding documentary obligations. It also reorganised the administrative procedure, which required the involvement of several domestic and consular authorities. Taken together, these provisions made family reunification more selective and administratively demanding, potentially increasing the time needed to complete the process.

The rationale underlying these restrictions was to ensure that immigrants had adequate economic and housing conditions before being joined by their family members. Family reunification may increase household expenditure, housing needs, childcare responsibilities, and the pressures associated with the reorganisation of family roles. Requiring applicants to demonstrate sufficient income, suitable accommodation, and a stable residence status may therefore reduce the risk that newly reunited households experience severe financial hardship, overcrowding, or immediate dependence on public assistance. From this perspective, a certain degree of selectivity may help provide more stable conditions for family settlement and facilitate the integration of newly arrived family members.

However, when legal requirements and administrative procedures become excessively burdensome, they may impose substantial costs on immigrant families. Applicants may need considerable time to secure suitable housing, meet income thresholds, obtain the required documentation, and complete procedures involving multiple administrative and consular authorities. These obstacles can prolong family separation, increase uncertainty and financial pressure, and generate additional emotional distress. The time and effort required to satisfy the conditions for reunification may therefore reduce immigrants' well-being even before family unity is restored.

Moreover, the consequences of prolonged separation may persist after reunification. Spouses who have lived apart for an extended period may have developed different routines, expectations, responsibilities, and social ties, making the reconstruction of family life more difficult. Lengthy separation may also delay the formation of shared social networks and stable household arrangements in the destination country. The burdens created by restrictive procedures may consequently affect not only individual well-being and family adjustment but also immigrants' integration into the host society.

Family reunification policy therefore involves a potential trade-off. On the one hand, income, housing, and residence requirements may help ensure that families reunite under sufficiently stable material conditions. On the other hand, excessive administrative burdens and prolonged waiting periods may create an additional source of insecurity, undermining the well-being and integration benefits that reunification is intended to promote. More accessible procedures may shorten the duration of separation and facilitate the restoration of family unity, although their effectiveness may still depend on the availability of adequate economic, housing and integration support.

The Bossi-Fini reform provides a particularly relevant setting in which to examine this trade-off. By tightening eligibility requirements and making the reunification procedure more demanding, the reform may have prolonged the period during which spouses remained separated and increased the uncertainty surrounding the process. The resulting delays may have weakened, without necessarily eliminating, the positive association between living with one's spouse and subjective well-being. In other words, immigrants reunited under the post-reform framework may still enjoy higher well-being than those whose spouses remain abroad, but lower well-being and greater difficulties in post-reunification family adjustment than immigrants who reunited under the less restrictive pre-reform regime. The consequences of family reunification therefore depend not only on whether family unity is eventually restored, but also on how long reunification takes and on the institutional conditions under which it occurs.

3. Data and Empirical analysis

For the empirical analysis, we use survey data collected by Italian National Institute of Statistics (ISTAT) on the Social Integration of Immigrants in Italy in 2011. The survey provides a rich set of individual and family-level data on foreign-born residents.

The questionnaire collects detailed information on family structure, including the spouse of the respondent, their children, and all the other relatives living in the household. It also records the year of arrival in Italy for each household member, which allows us to reconstruct the timing of reunification. Additional sections document immigrants' migration histories, including the number of times they had previously travelled to Italy, their mode of arrival, reasons for migration, country of departure, the individuals or organizations that assisted them, and who hosted them upon arrival. The survey also collects information on employment experience in both the country of origin and Italy, language proficiency, educational background, and vocational training undertaken either in Italy or abroad. Health status, perceived discrimination, and subjective integration are also measured.

The full survey includes data on around 9,500 immigrant families, from which about 6,000 coming from extra EU countries. represent Europeans the majority (60%), with Romanians alone accounting for 23.5% and Albanians for 13.7%. Other respondents come from Africa (20%, mostly Moroccan), Asia (13.3%, notably Filipinos and Chinese), and Latin America (6.5%). Only 0.2% of the sample come from North America and Oceania.

From the broader dataset, we select married immigrants from non-EU countries who were interviewed in Italy in 2011 and whose family migration decisions were therefore subject to Italian national immigration legislation. Within this sample, we distinguish between immigrants living alone in Italy, whose spouses remained abroad, and immigrants living with their spouses in Italy. The latter group includes both individuals who migrated together with their spouses and those who initially migrated alone and were subsequently joined through family reunification. This distinction allows us first to examine whether living with one's spouse in Italy is associated with greater life satisfaction than living apart. We then restrict the analysis to immigrants who experienced family reunification and for whom the timing of the spouse's arrival can be identified. Within this subgroup, we compare reunifications occurring before and after the introduction of the Bossi-Fini Law in 2002, which made the reunification process more restrictive and increased the time and administrative burden involved.

We measure immigrants' subjective well-being, our main outcome, using life satisfaction. Respondents were asked to assess their overall satisfaction with life on a scale ranging from 0, indicating the lowest possible level of satisfaction, to 10, indicating the highest.

We estimate the following model:

$$Y_{icr} = \alpha + \beta \text{Reunited vs. Alone}_{icr} + X'_{icr}\theta + \lambda_c + \gamma_r + \varepsilon_{icr}$$

where Y denotes the life satisfaction of individual (i), that belong to country of origin (c) and is located in a given macroarea (r); *Reunited vs. Alone*, is a dummy variable equal to one for married immigrants living in Italy with their spouse and equal to zero for married immigrants living alone in Italy while their spouse resides abroad. The coefficient of interest, β , captures the conditional difference in life satisfaction between immigrants living with their spouse in Italy and those living apart from their spouse. The vector X contains individual-level controls, including gender, age, education, religion, years in Italy, Italian proficiency, intention to stay in Italy, willingness to acquire Italian citizenship, reason for migration, being hosted upon arrival, attachment to home country.¹ Country-of-origin (λ) and macro-area of residence (γ) fixed effects are included to account for time-invariant differences across immigrants' origin countries and place of residence², respectively. Standard errors are clustered at the country-of-origin level.

In the second part of the analysis, we examine whether reunification under different policy regimes is associated with immigrants' well-being. To this end, we restrict the sample to immigrants who were joined by their family members after arriving in Italy and distinguish between those who experienced

¹ In particular, educational attainment is represented by four dummy variables—*Elementary*, *Middle*, *High school*, and *BA degree* +, with *No education* as the reference category. *Christian* is a dummy equal to one for respondents classified as belonging to a Christian religion and zero otherwise. *Years in Italy* measures the number of years the respondent has lived in Italy. The attachment to the origin country is captured by the variable “*Proud of being a foreigner*”, that equals one when respondents report being “very proud” of being a foreigner and zero when they report being “quite”, “slightly”, or “not at all” proud. *Hosted at Arrival* equals one if the respondent was hosted by someone upon first arriving in Italy and zero if the respondent arranged accommodation independently. The reasons for migration are grouped into four categories: *Economic* (that includes lack of employment opportunities, the desire to earn more, the pursuit of a better quality of life, the search for new experiences); *Family* (that includes affective and family-related reasons); *Education* (that identifies study-related migration); and *War* (that groups war and persecution). *Employment status* is a dummy equal to one for currently employed respondents and zero otherwise. *Intention to stay long-term* equals one for respondents who intend to remain permanently in Italy and zero for those planning either to return to their country of origin or to move elsewhere. *Italian Citizenship Intention* equals one if the respondent would like to acquire Italian citizenship and zero otherwise. *Italian Language* equals one if the respondent reports understanding Italian very well or fairly well and zero if their understanding is limited or absent.

² Place of residence is captured by the geographical indicators *North*, *Centre*, and *South*, with North serving as the reference category. We include also Municipality type, which is divided into *Big city*, covering municipalities located in the centre or periphery of a metropolitan area; *City (≤10,000 inhabitants)*; and *City (>10,000 inhabitants)*, with Big city as the reference category.

reunification before the introduction of the Bossi-Fini Law and those who reunited after its implementation, under a more restrictive policy regime. We estimate the following model:

$$Y_{icr} = \alpha + \beta \text{Reunification after 2002}_{icr} + X'_{icr}\theta + \lambda_c + \gamma_r + \varepsilon_{icr}$$

where the main explanatory variable, *Reunification after 2002*, is a dummy equal to one for immigrants who experienced family reunification under the more restrictive regime introduced by the Bossi-Fini Law and zero for those who reunited under the less restrictive pre-2002 regime. All other variables are defined as in the previous specification.

3.1. Descriptive statistics

Table A.1 reports descriptive statistics for the main variables used in the empirical analysis. The baseline sample includes 2,032 married immigrants from non-EU countries. Approximately 72.5% live in Italy with their spouses, while the remaining 27.5% live apart from spouses who reside abroad. Among immigrants who experienced reunification, 62.2% reunited after the introduction of the Bossi-Fini Law in 2002. The sample is predominantly male, with men accounting for almost 80% of respondents, and the average age is approximately 44 years. Educational attainment is concentrated at the middle- and high-school levels, which together account for about 69% of the sample, while around 13% have no formal education and fewer than 9% hold a university degree. Employment rates are relatively high, with 83% of respondents currently employed, and immigrants have lived in Italy for an average of almost 14 years.

Regarding migration experiences and integration, the descriptive statistics indicate that almost 79% were hosted by relatives, friends, or other individuals upon arrival, and economic considerations represent the main reason for migration for approximately 86% of the sample. Around 64% intend to remain in Italy permanently, while 74% report that they would like to acquire Italian citizenship. In terms of residential location, immigrants are distributed almost equally between Northern and Southern Italy, while a smaller proportion resides in the Centre. Slightly more than half live in municipalities with more than 10,000 inhabitants, excluding metropolitan areas.

Taken together, these statistics describe a sample composed mainly of working-age immigrants with relatively long migration histories and a generally strong intention to remain in the host country.

Table A.2 compares immigrants living with their spouses in Italy with those whose spouses remain abroad. The two groups differ along several observable dimensions. Immigrants living with their spouses are more likely to be male, are younger on average, and have spent more years in Italy. They

also display stronger attachment to the host country: approximately 74% intend to remain in Italy permanently, compared with 39% among those living apart, while 80% would like to acquire Italian citizenship, compared with 59% in the other group. Reunited immigrants are also more likely to reside in Northern Italy and less likely to live in the South or in large metropolitan areas. By contrast, employment status, having been hosted upon arrival, and most reasons for migration are broadly similar across the two groups. These statistically significant differences therefore underscore the importance of controlling for demographic, migration-related, integration, and residential characteristics in the regression analysis.

Finally, Table A.3 compares immigrants who reunited before and after 2002. Those who reunited after the reform are substantially younger and have spent fewer years in Italy than those who reunited before 2002. Post-2002 reunifications are also more frequently observed among immigrants who migrated for economic reasons, whereas education and war-related migration are relatively more common among the pre-2002 group. However, the two groups are broadly similar in terms of gender, educational attainment, religion, employment status, attachment to the country of origin, and intentions to remain in Italy or acquire Italian citizenship. The regression analysis therefore adjusts for observable differences between the two groups, particularly those related to age, migration history, reasons for migration, and residential location.

4. Results

4.1. Living with family and immigrants' well-being

Table 1 reports the association between living with one's spouse in Italy and immigrants' subjective well-being, measured by life satisfaction. In column (1), the coefficient on *Reunited* vs. *Alone* is positive and statistically significant at the 1 percent level. Married immigrants living with their spouses in Italy report, on average, a life satisfaction score 0.431 points higher than that of married immigrants whose spouses remain abroad, indicating a positive association between family presence and subjective well-being.

The inclusion of the full set of individual controls in column (2) reduces the estimated coefficient to 0.309 but leaves the main conclusion unchanged. The coefficient remains positive and statistically significant at the 1 percent level, suggesting that the observed difference in life satisfaction is not entirely accounted for by immigrants' demographic, socioeconomic, migration-related, and integration characteristics. Column (3) additionally controls for geographic macro-area, municipality

type, and country of origin. The estimated coefficient increases to 0.370 and remains statistically significant at the 1 percent level. Relative to the sample mean of approximately 7.5, this difference corresponds to about 5 percent of average life satisfaction. The positive association between living with one's spouse and subjective well-being is therefore robust across specifications.

Taken together, these results are consistent with the idea that family presence may provide an important source of emotional support, companionship, and stability in the destination country, while also allowing spouses to share household responsibilities, financial pressures, and the practical challenges associated with adapting to a new country. At the same time, reunification may reduce the emotional strain generated by prolonged separation, including loneliness, uncertainty about the future of the relationship, and the difficulty of maintaining closeness across countries.

Table 1. Reunited family and immigrants' life satisfaction

	Dependent variable: Life Satisfaction		
	(1)	(2)	(3)
Reunited vs. Alone	0.431*** (0.083)	0.309*** (0.071)	0.370*** (0.085)
Controls	no	yes	yes
Geographic macroarea	no	no	yes
Municipality type	no	no	yes
Country of origin	no	no	yes
Adjusted R^2	0.012	0.069	0.078
Observations	2124	2032	2032
Mean dependent	7.5099	7.4995	7.4995

Notes. OLS estimates in columns (1)-(3). Controls in column (2) include demographic and economic characteristics (gender, age, education, religion, employment status, reasons of migration), integration variables (years in Italy, Italian fluency, hosting upon arrival, long-term stay intentions, citizenship intentions, attachment to origin country). Column (3) includes country of origin and residential location. Standard errors in parentheses, clustered at the country level; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

However, the causal interpretation of these findings requires caution. The estimated coefficient may partly reflect unobserved differences between immigrants who live with their spouses and those who remain separated, including differences in financial resources, legal status, housing conditions, or migration plans. Given the cross-sectional nature of the data, the estimates should therefore be interpreted as conditional associations rather than as evidence of a causal effect of family reunification.

4.2. Family reunification policies and immigrants' well-being

Table 2 examines whether family reunifications occurring after the introduction of the Bossi-Fini Law in 2002 are associated with differences in immigrants' life satisfaction, time required to reunite with their families, and their integration into Italian society. According to column (1), the coefficient on *Reunification after 2002* is negative and statistically significant, indicating that immigrants who reunited under the post-reform policy framework report a 0.202-point lower level of life satisfaction than those who reunited before 2002. Column (2) restricts the sample to immigrants who had already arrived in Italy by 2002. This restriction helps reduce concerns that the results are driven by changes in the composition of immigrants entering Italy after the reform. The estimated coefficient remains unchanged at -0.202 and continues to be statistically significant, suggesting that the negative association is not simply explained by differences between immigrants arriving in Italy before and after the policy change.

Column (3) provides evidence on a potential mechanism underlying this result. Immigrants who reunited after 2002 experienced an increase of approximately 5.8 years in the interval between their own arrival in Italy and the subsequent arrival of their spouses³. The coefficient is large and statistically significant at the 1 percent level. This finding is consistent with the view that the more restrictive requirements and greater administrative burden associated with the Bossi-Fini framework substantially prolonged the reunification process. The policy change therefore appears to have affected not only access to family reunification but also the length of time during which immigrants remained separated from their spouses.

Furthermore, the results in column (4) suggest that the consequences of post-2002 reunification extend beyond life satisfaction and also concern immigrants' integration into the host country. *Reunification after 2002* is associated with a statistically significant 0.062-point reduction in the integration indicator. Since this outcome captures whether immigrants report feeling at home in Italy, the estimate corresponds to an approximately 6.2 percentage-point lower probability of expressing a sense of belonging to the host country. Thus, immigrants whose reunification took place under the more restrictive policy framework are not only less satisfied with their lives but are also less likely to perceive themselves as integrated into Italian society.

This pattern is consistent with the idea that prolonged family separation may weaken the integration process. Immigrants waiting for reunification may face emotional stress, uncertainty, and the financial burden of supporting family members living in different countries. Longer periods of separation may

³ This difference remains virtually unchanged when the analysis is conducted on the full sample.

also postpone the development of stable family routines and shared social networks in Italy. Moreover, when reunification eventually occurs, spouses may have accumulated different experiences and expectations, making post-reunification adjustment more difficult. These pressures may reduce the emotional benefits normally associated with family presence and weaken immigrants' attachment to the host country. More restrictive procedures may therefore generate costs that persist even after family unity has been restored. Nevertheless, given the cross-sectional nature of the data, these results should be interpreted as conditional associations consistent with this mechanism, rather than as causal effects of the Bossi-Fini reform.

Table 2. Family reunification policy and immigrants' well-being and integration

	Dependent variable:			
	Life Satisfaction		Time to Reunification	Integration
	(1)	(2)	(3)	(4)
Reunification after 2002	-0.202* (0.109)	-0.202* (0.114)	5.816*** (0.443)	-0.062*** (0.019)
Controls	yes	yes	yes	yes
Geographic macroarea	yes	yes	yes	yes
Municipality type	yes	yes	yes	yes
Country of origin	yes	yes	yes	yes
Adjusted R^2	0.085	0.079	0.608	0.035
Observations	1184	923	923	916

Notes. OLS estimates in columns (1)-(4). Controls in columns (1)-(4) include demographic and economic characteristics (gender, age, education, religion, employment status, reasons of migration), integration variables (years in Italy, Italian fluency, hosting upon arrival, long-term stay intentions, citizenship intentions, attachment to origin country), country of origin and residential location. Columns (2)-(4) include only immigrants who arrived in Italy by 2002. Standard errors in parentheses, clustered at the country level; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5. Conclusions

This paper examines the relationship between family reunification and immigrants' well-being. Using microdata from the 2011 ISTAT Survey on the Social Integration of Immigrants living in Italy and focusing on married individuals from non-EU countries, we first show that immigrants living in Italy with their spouses report significantly higher life satisfaction than those whose spouses remain abroad. This finding supports the view that family presence is importantly associated with immigrants' subjective well-being.

The analysis further indicates that well-being may be related not only to whether spouses live together, but also to the institutional conditions under which reunification takes place. Indeed, in the

second part of the analysis, we focus on immigrants who experienced family reunification and compare those who reunited before and after the introduction of the Bossi-Fini Law in 2002. The reform made the reunification framework more restrictive by narrowing eligibility and increasing the administrative and documentary burden associated with the procedure. We find that immigrants who reunited after the reform experienced substantially longer periods of separation from their spouses and reported lower life satisfaction. They were also less likely to feel at home in Italy, suggesting that the costs associated with a prolonged and more demanding reunification process may extend beyond individual well-being and affect immigrants' integration into the host society.

These findings challenge the view that family reunification yields comparable benefits irrespective of the policy framework under which it takes place. Although living with one's spouse is positively associated with life satisfaction, prolonged separation, administrative uncertainty, and the economic and emotional effort required to complete the reunification process may weaken these benefits. Family reunification should therefore be understood not merely as the final legal admission of a family member, but as a multidimensional process whose consequences are shaped by the length of separation and by the institutional, social, and economic conditions surrounding it.

The policy implications are consequently important. Policies that impose lengthy and burdensome procedures may generate costs that persist even after family unity has been restored. Reducing unnecessary delays and administrative complexity may therefore contribute not only to family cohesion but also to immigrants' subjective well-being and sense of belonging to the host country. At the same time, faster procedures should be accompanied by measures supporting access to employment, adequate housing, language training, social networks, and other integration services, so that reunification can translate into more favourable outcomes for both immigrants and their families.

The results should nevertheless be interpreted with caution. The analysis is based on a single cross-sectional survey and compares distinct groups of immigrants according to their living arrangements and the retrospective timing of reunification. It therefore identifies conditional associations rather than the causal effects of family reunification or of the Bossi-Fini reform. Although the inclusion of a broad set of observable characteristics and the restriction of the post-reform analysis to immigrants who had already arrived in Italy by 2002 reduce concerns about compositional differences, selection on unobservable characteristics cannot be ruled out. Further research using longitudinal data or stronger sources of exogenous variation is therefore needed to identify the causal effects of family reunification and to clarify the mechanisms through which reunification policies influence immigrants' well-being and integration.

References

- Bonizzoni, P. (2009). Living together again: families surviving Italian immigration policies. *International Review of Sociology*, 19(1), 83-101.
- Cebolla Boado, H., & González Ferrer, A. (2022). The impact of physical separation from parents on the mental wellbeing of the children of migrants. *Journal of Ethnic and Migration Studies*, 48(10), 2436-2454.
- Choummanivong, C., Poole, G. E., & Cooper, A. (2014). Refugee family reunification and mental health in resettlement. *Kotuitui: New Zealand Journal of Social Sciences*, 9(2), 89-100.
- Eremenko, T., & Bennett, R. (2018). Linking the family context of migration during childhood to the well-being of young adults: Evidence from the UK and France. *Population, Space and Place*, 24(7).
- Gubernskaya, Z., & Dreby, J. (2018). US immigration policy and the case for family unity. *Journal on Migration and Human Security*, 5(2), 417-430.
- Hvidtfeldt, C., Petersen, J. H., & Norredam, M. (2022). Waiting for family reunification and the risk of mental disorders among refugee fathers: a 24-year longitudinal cohort study from Denmark. *Social Psychiatry and Psychiatric Epidemiology*, 57, 1061-1072.
- Liddell, B. J., Byrow, Y., O'Donnell, M., Mau, V., Batch, N., McMahon, T., Bryant, R. A., & Nickerson, A. (2021). Mechanisms underlying the mental health impact of family separation on resettled refugees. *Australian & New Zealand Journal of Psychiatry*, 55(7), 699-710.
- Löbel, L. M. (2020). Family separation and refugee mental health—A network perspective. *Social Networks*, 61, 20-33.
- Löbel, L. M., & Jacobsen, J. (2021). Waiting for kin: a longitudinal study of family reunification and refugee mental health in Germany. *Journal of Ethnic and Migration Studies*, 47(13), 2916-2937.
- Miller, A., Hess, J. M., Bybee, D., & Goodkind, J. R. (2018). Understanding the mental health consequences of family separation for refugees: Implications for policy and practice. *American Journal of Orthopsychiatry*, 88(1), 26-37.
- Nika, A. (2017). *Dear family: How European migration policies are keeping families apart*. Oxfam Briefing Paper.
- Phillimore, J., D'Avino, G., Strain-Fajth, V., Papoutsis, A., & Ziss, P. (2023). Family reunion policy for resettled refugees: Governance, challenges and impacts. *Frontiers in Human Dynamics*, 5, 1075306.

Appendix

Table A.1 – Descriptive statistics

	count	mean	sd	min	max
Reunited vs. Alone	2032	0.725	0.447	0.000	1.000
Reunification after 2002	1184	0.622	0.485	0.000	1.000
Male	2032	0.796	0.403	0.000	1.000
Age	2032	43.664	9.440	22.000	88.000
No education	2032	0.133	0.340	0.000	1.000
Elementary	2032	0.086	0.281	0.000	1.000
Middle	2032	0.342	0.474	0.000	1.000
High school	2032	0.351	0.477	0.000	1.000
BA degree +	2032	0.088	0.283	0.000	1.000
Christian	2032	0.323	0.468	0.000	1.000
Employment status	2032	0.833	0.373	0.000	1.000
Years in Italy	2032	13.691	6.547	2.000	52.000
Proud of being a foreigner	2032	0.601	0.490	0.000	1.000
Hosted at Arrival	2032	0.789	0.408	0.000	1.000
Reason: Economic	2032	0.861	0.346	0.000	1.000
Reason: Family	2032	0.062	0.240	0.000	1.000
Reason: Education	2032	0.011	0.104	0.000	1.000
Reason: War	2032	0.067	0.250	0.000	1.000
Intention to stay long-term	2032	0.640	0.480	0.000	1.000
Italian Citizenship Intention	2032	0.742	0.438	0.000	1.000
Italian Language	2032	0.124	0.330	0.000	1.000
North	2032	0.434	0.496	0.000	1.000
Centre	2032	0.142	0.349	0.000	1.000
South	2032	0.424	0.494	0.000	1.000
Big city	2032	0.231	0.421	0.000	1.000
City ($\leq 10,000$ inhabitants)	2032	0.255	0.436	0.000	1.000
City ($> 10,000$ inhabitants)	2032	0.514	0.500	0.000	1.000

Table A.2 – Descriptive statistics - Reunited vs. Alone

	(1) Reunited	(2) Alone	(3) ttest
Male	0.856	0.639	-0.217***
Age	42.673	46.277	3.605***
No education	0.109	0.197	0.087***
Elementary	0.083	0.095	0.012
Middle	0.369	0.268	-0.101***
High school	0.361	0.324	-0.037
BA degree +	0.077	0.116	0.039**
Christian	0.293	0.403	0.110***
Employment status	0.823	0.857	0.033
Years in Italy	14.367	11.911	-2.457***
Proud of being a foreigner	0.597	0.612	0.015
Hosted at Arrival	0.783	0.807	0.024
Reason: Economic	0.849	0.891	0.042*
Reason: Family	0.068	0.045	-0.023
Reason: Education	0.012	0.009	-0.003
Reason: War	0.071	0.055	-0.016
Intention to stay long-term	0.737	0.385	-0.352***
Italian Citizenship Intention	0.799	0.590	-0.209***
Italian Language	0.104	0.177	0.073***
North	0.470	0.336	-0.134***
Centre	0.149	0.125	-0.023
South	0.381	0.538	0.158***
Big city	0.208	0.292	0.084***
City(<= 10000 inhabitants)	0.284	0.179	-0.105***
City(>10000 inhabitants)	0.508	0.530	0.021

Table A.3 – Descriptive statistics - Reunited before vs. reunited after 2002

	(1) After2002	(2) Before2002	(3) ttest
Male	0.875	0.886	0.011
Age	39.611	47.221	7.610***
No education	0.125	0.107	-0.018
Elementary	0.080	0.087	0.007
Middle	0.386	0.357	-0.029
High school	0.342	0.384	0.042
BA degree +	0.067	0.065	-0.002
Christian	0.254	0.268	0.014
Employment status	0.841	0.806	-0.035
Years in Italy	12.189	19.565	7.376***
Proud of being a foreigner	0.607	0.623	0.015
Hosted at Arrival	0.822	0.790	-0.032
Reason: Economic	0.898	0.819	-0.079***
Reason: Family	0.064	0.056	-0.008
Reason: Education	0.004	0.020	0.016**
Reason: War	0.034	0.105	0.071***
Intention to stay long-term	0.723	0.748	0.025
Italian Citizenship Intention	0.830	0.797	-0.033
Italian Language	0.107	0.069	-0.038*
North	0.515	0.482	-0.033
Centre	0.121	0.165	0.044*
South	0.364	0.353	-0.011
Big city	0.167	0.248	0.081***
City(<=10000 inhabitants)	0.295	0.299	0.004
City(>10000 inhabitants)	0.538	0.453	-0.085**