Title

Essays on Migration and Taxation

Author
Ippedico, Giuseppe

Publication Date
2022

Peer reviewed|Thesis/dissertation

Essays on Migration and Taxation

By

Giuseppe Ippedico

Dissertation

Submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

ECONOMICS

in the

Office of Graduate Studies

of the

University of California

Davis

Approved:

Giovanni Peri, Chair

Monica Singhal

Santiago Pérez

Committee in Charge

2022

Contents

List of Figures

List of Tables

Abstract

Essays on Migration and Taxation

Migration flows of individuals across borders can have large economic consequences both on receiving and sending regions. The Schengen area, where individuals are free to move between European countries, constitutes a unique lab for economists to study the determinants and the effects of unrestricted migration flows across borders. This dissertation focuses on Italy, an excellent setting to study these topics because of i) the availability of administrative data on international migration of Italian citizens, ii) a context of "brain drain", with large emigration rates of young college graduates towards other European countries since the 2000s, and iii) the introduction of tax schemes to attract back high-skilled expatriates.

The first chapter, co-authored with Massimo Anelli, Gaetano Basso and Giovanni Peri, and conditionally accepted for publication in the *American Economic Journal: Applied Economics* at the time of writing, investigates the effect of the recent surge in emigration from Italy on firm-creation and innovation on the areas-of-origin, and its implications for local labor demand. Entrepreneurship requires a high degree of creativity, initiative, risk-taking and adaptability, and research has shown that some of these traits also increase the propensity to migrate. Hence, locations experiencing large emigration rates may be at risk of losing their entrepreneurial capital, with negative consequences on firm-creation. Using a rich set of administrative data on emigration and firm-creation and employment at the local level, and leveraging an exogenous source of variation which combines past emigration networks across origin locations with pull factors from destination countries, we find that emigration had a detrimental effect on firm-creation, especially by young entrepreneurs and of start-ups with high-growth potential. We also document negative effects on local employment, wage bill and employment-to-population ratios, consistent with a reduction in local labor demand stemming from the loss of potential job-creators entrepreneurs.

Among the policy responses that countries have introduced to mitigate the effects of brain drain, one of the most widespread is a preferential tax scheme for high-skilled expatriates who return to their origin country. Are tax incentives an effective policy to attract back individuals who migrated abroad? While the public finance literature has found that top earners are highly responsive to fiscal incentives, it is unclear whether tax discounts attract young college graduates, who are often at the beginning of their careers and not necessarily high earners. In the second chapter, co-authored with Jacopo Bassetto, we tackle this question by studying the effects of the Italian 2010 tax scheme, which drastically lowered income tax rates for expatriates who move back to Italy, conditionally on holding a college degree and to be born after January 1, 1969. Exploiting these two eligibility requirements in a difference-in-differences strategy, we find that return migration of eligible individuals increased substantially after the introduction of the scheme relative to non-eligible groups, while being on similar trends before 2010. We then replicate the analysis using social security data from Germany – the main destination country of Italian emigrants – which allow us to study the return-migration response of different subgroups of Italian expatriates to the tax scheme. Specifically, we find a large migration response throughout the wage distribution, suggesting that tax-schemes-induced mobility is a broader phenomenon than the relocation of top earners and results in a substantial reallocation of human capital across sending and receiving regions.

Last, the third chapter asks whether tax incentives also influence migration decisions within a country. In a context of fiscal decentralization, local income and property tax differentials between locations imply that individuals' tax liability can vary substantially across jurisdictions. By exploiting a series of reforms in the 2000s that increased local tax differentials in Italy, the main finding is that internal migration – as measured by transfers of residence for fiscal purposes – is influenced by both income and property tax differentials, with important implications for local economies.

Acknowledgments

A PhD is a long journey, and the key ingredient for getting to the finish line sane (and with some research done) is certainly the people you are surrounded by. I feel immensely grateful for the support and the inspiration I received from so many of you, way too many to fit in this short acknowledgments page, but I will try my best.

Thanks to my main advisor Giovanni, as well as to Monica, Santi and Diana for being a continuous source of guidance, feedback, inspiration and support, and for believing in me since the beginning. Advising is arguably the most delicate and crucial role of faculty, and I really hope to be able to give back when it will be my turn.

To my co-authors Massimo, Gaetano, Giovanni and Jacopo, for the great teamwork despite many challenges and for sharing all the stages of a research project. I learned so much working with all of you!

To the grad students and faculty in the applied micro group, the dissertation lab, the Global Migration Center, and to many conference participants for your feedback.

To my econ-friends: Derek, for introducing me to the Californian outdoors, a great source of inspiration for research; Thiago, for cooking great pastas that made me feel at home since day one; and the many other grad students with whom I shared incredibly fruitful moments, from brown bags and research chats, to dinners, beers, camping and skiing trips.

To my family in Italy, way too far from the West Coast, but always close enough to make me feel supported.

It is hard to find words to express gratitude for someone whose support and love went above and beyond throughout this entire journey. To the most special person I ever met and had the privilege to share everything with, from the most challenging to the most rewarding moments of these five years. This dissertation is dedicated to you, Iulia.

Chapter 1

Emigration and Entrepreneurial Drain (with Massimo Anelli, Gaetano Basso and Giovanni Peri)

1.1 Introduction

Entrepreneurship[1] requires a high degree of creativity, initiative, risk-taking and adaptability to new situations. Interestingly, research has shown that some of these non-cognitive traits also increase the propensity to emigrate. Jaeger et al. (2010) show that migrants have less aversion to risk than non-migrants, and Bütikofer and Peri (2021) show that individuals with a higher level of adaptability and cognitive ability are more likely to emigrate. Hence, countries and regions experiencing significant emigration rates may be at risk of losing a substantial amount of entrepreneurial potential, with negative consequences on firm-creation. This issue has long been a concern in developing countries. More recently, it has become a salient concern in Southern Europe, where the young cohorts have become significantly smaller in size due to the demographic transition and a substantially increased propensity to move to Central and Northern Europe. This was encouraged by free labor mobility within the EU and has been accelerated by the great recession of 2010, which hit Southern Europe much harder than

[1]The views expressed in this chapter are those of the authors and do not necessarily reflect the position of the Bank of Italy.

Northern Europe (Schivardi and Schmitz 2020).

In this paper, we investigate the causal effect of emigration on firm creation in the area of origin, and its potential implications for local labor demand. We then try to understand how much emigrants' selection on age, education and other features contribute to the loss in entrepreneurship, and how much this loss can negatively affect entrepreneurship rates among remaining individuals. Our empirical analysis focuses on Italy, an excellent setting to study this phenomenon due to the substantial surge in emigration rates during and after the Great Recession of 2008-10. Figure 1.1 shows the sharp increase in emigration flows which began in 2010 and increased three-fold by 2015. From 2008 to 2015, the cumulative emigration flows recorded by administrative data amount to a loss of almost 1% of the Italian population.[2] While emigration was occurring across all age groups, Figure 1.2 shows that its rate was especially high among young individuals (aged 25-44) and among college graduates.

Estimating the causal effect of emigration on local economic outcomes is challenging. The main threats to consistent estimation are reverse causality, as people may be more likely to leave regions with low firm-creation, and omitted variable bias, as several unobserved factors that push people to emigrate may also affect firm creation. Moreover, measurement error in recording emigration flows—resulting from delays and under-reporting in changes of residence—could attenuate the relationship between emigration and firm creation, and short-run measures of mobility can be especially noisy. To overcome these issues, we adopt an instrumental variable strategy in the spirit of Anelli and Peri (2017) and Fouka, Mazumder, and Tabellini (2020), and construct a measure of "network-driven" emigration. Such instrument captures the strength of existing networks of Italians abroad from specific local labor markets in the main destination countries (or sub-national regions, in a more detailed version of the instrument), measured in 2000, well before the Great Recession-era emigration wave.

[2] Comparable statistics on emigration flows across countries are hard to obtain. We were only able to locate a report from the Portuguese Observatory of Emigration (2015) indicating that the cumulative outflows of Portuguese citizens between 2011 and 2014 reached about 485,000 people or about 1.2% of the Portuguese population.

In our main instrument, we weight this measure of network intensity with the economic performance of destination countries from 2008 to 2015. While these destination country weights are likely independent of economic and labor market conditions at the origin, our instrumental variable, ultimately, leverages cross-sectional variation of network intensity across labor markets of origin. Specifically, we show that most of our identifying variation is driven by the pre-existing networks of Italians from each local labor market of origin towards Germany and Switzerland as a share of the labor market population. Those are two countries whose average income is higher than Italy's and performed better economically during the Great Recession.[3] The common push in the post-2008 period, generated by the recession, interacted with the pre-existing networks in economically more successful countries, generates the post-2008 variation of predicted emigration in our IV.[4] Our identification approach is supported by the fact that the IV passes the validity tests proposed by Goldsmith-Pinkham, Sorkin, and Swift (2020). The networks towards Germany and Switzerland are uncorrelated with pre-2008 trends and each of them is a good IV, passing an over-identification test when used jointly and producing similar 2SLS estimates when used one at a time.

Our results show that emigration—when instrumented with the network-driven IV—produced a decline in the number of existing firms, due to a lower birth rate and an unchanged death rate of firms. This is consistent with a significant loss of entrepreneurial capital. Namely, an increase in the emigration rate by one standard deviation (about 1.7 percent of the working age population) reduced the number of firms created in the average local labor market over the post-2008 period by 194 firms, corresponding to 4.76% of firm creation in the baseline period. This effect is significantly larger than what is implied by the simple mechanical subtraction of individuals with an average entrepreneurial ability, which only explains 36% of the decline. Using our data, we also show that an additional 17% of the effect is due to over-representation among emigrants of highly-educated and young individuals, who are

[3]We test the independence between those networks (shares) and pre-2008 economic trends.

[4]Before 2008, emigration was smaller and stable, and so we use that period for validity tests.

characterized by higher than average entrepreneurship rates (Liang, Wang, and Lazear 2018). Additionally, borrowing some estimates from the existing literature on human capital externalities, we estimate that the lost entrepreneurs exerted negative spillover effects on firm-creation among those who remain in the labor markets of origin. Our calculations suggest that between 36% and 47% of the effect may be due to these spillovers. The remaining 0-11% of the effect is left to the selection of (unobservable) high entrepreneurial types for given age and education (observable characteristics).

As firm creation, especially in highly-innovative sectors and by younger individuals, is an engine for the introduction of new technologies (Acemoglu, Akcigit, and Celik 2017), the loss of entrepreneurial capital due to emigration may be particularly damaging for economic growth. Our analysis shows a strong negative effect of emigration on the creation of firms whose owners and executives were younger than 45 years old, and a significant decline in the number of innovative start-ups operating in technology-intensive sectors. Finally, we study the potential effects of the emigration wave on overall employment and its composition. The departure of a portion of the labor force should create job opportunities for those left behind (as it represents a drop in labor supply) and, all else equal, should increase the employment-population ratio in the labor-markets-of-origin. Instead, we find that local labor markets with higher emigration rates experience unchanged employment-population ratio, in spite of the negative labor supply shock accompanied by no significant change in wages. These effects are consistent with a drop in labor demand accompanying the loss in labor supply, due to the loss of potential firm-creating entrepreneurs, who are more concentrated among migrants than among stayers.

There are three main areas in which this paper contributes to the literature. First, this paper extends the literature on the effects of emigration on country of origin outcomes, and is the first to focus on the effects on entrepreneurship. It provides reasonably strong identification and uses high quality administrative firm and emigration data. While the shift-share IV is not new, we innovate by exploiting a sudden emigration episode driven by a large recession in Italy combined with variation in the intensity of pre-existing networks of emigrants across local labor markets. The sudden change in emigration provides an event-style identification with a pre- and a post- period that we validate showing the low correlation of our IV with pre-recession emigration variables. Following the recent contributions of Goldsmith-Pinkham, Sorkin, and Swift (2020) and Borusyak, Hull, and Jaravel (2021), we scrutinize our identification assumptions, and we find that the network intensity –"share"– variation provides identification that is especially driven by the top two networks – in Germany and Switzerland. Hence, we test the latter's correlation with pre-trends and, alternatively, we use them directly as an IV. The results of these checks strengthen our confidence in this IV strategy.

Related papers analyzing the impact of high-skilled emigration, often referred to as *brain drain*, on developing economies are Mayr and Peri (2009a), Waldinger (2010), Docquier and Rapoport (2012a), Docquier, Özden, and Peri (2014) and Di Giovanni, Levchenko, and Ortega (2015). Less is known about the effects of brain drain on developed economies. One paper related to ours is Giesing and Laurentsyeva (2017a), which finds that high-skilled emigration from Eastern Europe after the EU enlargement of the 2000s had negative effects on firms' TFP in the countries-of-origin. Anelli and Peri (2017) and Ippedico (2017) looked at the relationship between emigration, political outcomes, and local firms but without a thorough investigations of the mechanisms, the instrument, or the identification strategy. Most of the previous brain drain/emigration literature considered this phenomenon as a decline in the country of origin labor supply. Hence, researchers studied the short-run impact on wages and employment opportunities of those who remained (Mishra 2007; Elsner 2013a; Elsner 2013b; Dustmann, Frattini, and Rosso 2015). These papers focus on wage effects of emigration episodes largely driven by emigrants of intermediate educational attainment (skill) level from Poland and Mexico attracted by strong pull factors in Europe (Elsner 2013b; Dustmann, Frattini, and Rosso 2015) or the US (Mishra 2007) rather than emigration of highly educated due to a recession. These papers find small positive effects on wages of intermediate skilled workers (Dustmann, Frattini, and Rosso 2015), on young workers (Elsner 2013b) or on the average worker (Mishra 2007) in countries of origin. In our case, there was stronger positive selection of emigrants possibly driven by the strong recession in the country of origin. The fact that people with higher entrepreneurial abilities left might have crucially weakened labor demand. Indeed, in this context we do not find a positive significant effect on average wages of stayers, nor on the employment/population ratio. We are the first to focus on the firm-creation effect of emigration and its implications for local employment/population ratios. Our paper shows that emigration can reduce labor demand which, as far as we know, is an unexplored economic effect among countries of origin.

Second, as the emigration wave we analyzed was mainly fueled by the mobility of young people, our paper has some bearing on the literature on the role of young individuals in starting up new firms (Barker and Mueller 2002; MacDonald and Weisbach 2004; Acemoglu, Akcigit, and Celik 2017). We find that emigration is a relevant force in reducing the number of young people and their innovative, entrepreneurial role. Related literature shows a positive relationship between the share of young people in a country (or region) and rates of entrepreneurship (Liang, Wang, and Lazear 2018), productivity (Ciccarelli, Gomellini, and Sestito 2019), growth (Engbom 2019) and start-ups (Karahan, Pugsley, and Sahin 2019). If innovative entrepreneurship is higher at a young age, as suggested by Kopecky (2017), the loss of young people may be associated with a loss of growth and innovative ideas.[5]

Finally, our paper complements the studies which find that immigrants, especially in the US, have a special propensity to innovate and to be entrepreneurs. Hunt and Gauthier-Loiselle (2010), Kerr and Lincoln (2010), Moser, Voena, and Waldinger (2014), and Burchardi et al. (2020) show that immigrants in the US are more likely to be active in patenting and innovation than comparable natives. Similarly, as reviewed in Fairlie and Lofstrom (2015), a significant number of studies finds that immigrants in the US have a higher probability of being self-employed and starting firms relative to natives. A recent paper by Azoulay et al. (2021) shows that immigrants act more

[5] This is, however, not yet fully established, as Azoulay et al. (2020) show that high-growth entrepreneurship peaks later in life.

as job-creators than job-takers by starting high-growth enterprises. This evidence, analyzed from the receiving country perspective, suggests a positive selection of immigrants among innovators and entrepreneurs. Our study complements that evidence from the sending country perspective.

A qualification of our findings is also needed. We are analyzing emigration in a relatively developed country during a deep recession. This context was characterized by strong positive selection of emigrants, strong negative effect on entrepreneurship and null effect on wages. This evidence is somewhat different from what is found in other studies (Mishra 2007; Elsner 2013a; Elsner 2013b; Dustmann, Frattini, and Rosso 2015) and can be due to the larger propensity of high skilled to leave under those circumstances. One should be cautious in generalizing these results. However, our results are informative about the emigration effects and of the selection of migrants in countries experiencing deep recessions, a relevant scenario in which the "brain drain" can exacerbate the negative effects of recessions. Thus, our findings add a very important aspect to the analysis of the long-run consequences of deep local recessions.

The rest of the paper is organized as follows. Section 1.2 describes the main data and trends for emigration and firm creation in Italian local labor markets. Section 1.3 introduces the empirical specification, then describes the 2SLS identification strategy and discusses its validity. Section 1.4 presents the main results, and Section 1.5 discusses several additional results. Section 1.6 reports the main robustness checks. Section 1.7 concludes the paper.

1.2 Data

1.2.1 Emigration flows and network

We obtained administrative data on emigration flows of each municipality from the Italian National Statistical Institute (Istat 2016a). The data are aggregated into year of emigration by municipality of origin by country of destination by age-group by education-level cells and cover the period 2002-2015. We also obtained data on the stock of emigrants directly from the Registry of Italians Residing Abroad (AIRE 2015;

Anelli and Peri 2017), which includes all individuals who permanently emigrated between 1990 and 2014 and were still abroad as of 2015, and which includes precise information about the destination country (and region), the municipality-of-origin and the year-of-emigration. These features allow us to construct the historic networks of emigrants from those individuals who emigrated before 2000.

Table 1.1 shows the stock of emigrants from Italy by country-of-destination as of 2000 in Panel A and the cumulative emigration flows between 2008 and 2015 by age group in Panel B. The table reveals two important facts. First, the top destination countries have been quite stable over time. While in recent years the economically successful UK and US have replaced some more historical destinations, such as Argentina and Belgium, we can see that Germany, Switzerland and France were among the most common destinations for Italian emigrants in both periods. Germany and Switzerland emerge as crucial in identifying the pull-driven migration from 2008 to 2015, as we discuss in Section 1.3.3. Second, confirming the trends in the aggregate data, the table shows that young people (25 to 44 years old) represent a very large share of migrants from 2008 to 2015 (column 2 of Panel B in Table 1.1).

A limitation of the administrative data described above is that despite the fact that Italian emigrants are required by law to register as living abroad within six months of emigration and have significant financial incentives to do so[6], there is anecdotal evidence of under-registration, especially in the early years after emigration, as not all changes of residence may be recorded in a timely manner by the Italian authorities. Figure 1.3 compares the outflows of Italians to the UK in the AIRE-Istat data and the registration of Italian immigrants recorded in the UK social security registry (NINo 2018). The UK data indicate that Italian migrants are underestimated by about two thirds (Panel (a)), and that annual immigrant changes from Istat data closely follow those from the UK social security registrations with one year of lag (Panel (b)). This lag is consistent with the six-month window allowed to migrants by Italian authorities to communicate their new residence abroad and with bureaucratic delays characteriz-

[6]Namely, registered emigrants do not pay income tax in Italy on income earned abroad.

ing the formal registration process. An analysis based on data from the Switzerland Statistical Office (BFS 2018) show similar patterns (Figure 1.4).[7]

In Appendix A.II, using these destination-country sources, we estimate that actual emigration flows of Italians are plausibly about 2.6 times larger than those registered in the AIRE-Istat records. Such measurement error, due mainly to delays and imperfect registration of temporary migrants, is an additional reason to use IV estimation. It is important to notice that measurement error is likely to be much smaller for the measure of pre-existing networks of Italians abroad (those who emigrated before 2000), as those numbers are not affected by delays or by the presence of temporary migrants. Hence, the cross-sectional distribution of historical Italian emigrants used to measure network intensity across municipalities, and at the core of the instrument construction, is likely a precise measure of the Italian diaspora, while the recent flows may be underestimated significantly. Finally we account for such under-counts when interpreting the magnitude of the effects relative to the size of the emigration rate, as we do in Section 1.4 below.

1.2.2 Firms, employment and local labor markets

We obtained firm-level data that cover the universe of Italian firms from the Chambers of Commerce (Infocamere 2017). We merged them with data from the social security administration (INPS 2017) on employment and wages. Data from the Chambers of Commerce include information on the stocks, births and deaths[8] of firms and demographic characteristics of owners, shareholders and executives of each firm over the period 2005-2015.[9] The latter is used to identify firms with a majority of owners and executives under 45 years old, to which we refer as "young-owned firms." Our data in-

[7]We performed a similar analysis for the US using data from the American Community Survey (ACS), which we show in Figure A1 in the Appendix. Despite the fact that the survey nature of the data does not allow us to precisely estimate the immigration of Italians, the analysis based on the US qualitatively confirms the evidence found using the UK and Swiss administrative data.

[8]As deaths are often recorded with delay in the Chambers of Commerce data, we estimate deaths as $deaths_t = -stock_t + stock_{t-1} + births_t$, which is standard practice in the literature.

[9]We consider a birth to be the appearance of a new firm in any given year, provided it survives at least through the end of the year.

clude all firms, some of which may be multi-plant (though the vast majority have only one establishment). The INPS data cover the period 2005-2015 and include information on the yearly number of employees (broken down by broad occupation categories, i.e., apprentices; production workers, often referred to as "blue collar" workers; non-production workers, often referred to as "white collar" workers, and managers), their average monthly wage, industry, and the geographic location of the employer.[10]

Our unit of analysis is the local labor market (LLM), defined using the Istat 2001 definition (Istat 2016b; Istat 2018). According to Istat, LLMs are geographic clusters of municipalities with commuting patterns mainly internal to the cluster, an analogue definition to that of Commuting Zones (CZ) for the US.[11] There are 686 LLMs in Italy covering the whole national territory. We focus our analysis on the period from 2005 to 2015, considering from 2008 to 2015 as the "treatment" period, as emigration increased suddenly and substantially in those years.

1.3 Empirical Specification and Identification

In our empirical model, the main outcome is the change in the stock of firms from 2008 to 2015 (equal to the difference between entries and exits in the period) in local labor markets, indexed by l. This variable is indicated as Δy_l in equation (1.1). The main explanatory variable is the cumulative outflow of Italians who are 25 to 64 years old from 2008 to 2015, indicated as $\sum_{t=2008}^{2015} m_{l,t}$.[12] Both variables are divided by the average pre-treatment LLM population aged 25 to 64, $pop_{l,pre}$. This normalization produces the emigration rate in the area-of-origin, l, relative to the initial population. In the baseline specification, we control for 2005 GDP per capita and unemployment

[10] Both the Chambers of Commerce and INPS data identify the location of a firm with its headquarters. The vast majority of Italian firms have only one establishment, so the headquarter address corresponds to the whole firm in most cases.

[11] Following the US literature on CZs, in the case where a LLM crosses provincial boundaries, we assign it to the province where most of the population resides. Such assignment is relevant when we include province fixed effects in the main empirical specification.

[12] Data on emigration flows from Istat are divided into four age groups, 0-25, 25-44, 45-64 and 65+. We exclude people under 25 and over 64, as their contribution to firm creation and employment is marginal.

rate to account for the economic performance of the LLM before treatment, denoted by $X_{l,2005}$. We also include either twenty regions or 110 provincial fixed effects (ϕ_p) that capture time-invariant, unobserved geographic and institutional factors common to all LLMs within a region or a province, and we cluster standard errors at the province-level. We thus estimate the following equation:

$$\frac{\Delta y_l}{pop_{l,pre}} = \alpha + \beta \frac{\sum_{t=2008}^{2015} m_{l,t}}{pop_{l,pre}} + \phi_p + \gamma X_{l,2005} + \varepsilon_l \tag{1.1}$$

If the size of migration outflows were distributed randomly across LLM, the OLS estimate of equation (1.1) would deliver the causal effect of emigration on the number of firms. However, this is unlikely because such outflows are likely correlated with local economic and social conditions, which in turn might affect our outcomes of interest. On one hand, if LLMs with more intense entrepreneurial and economic activity tend to have a stronger connection with foreign economies and possibly more migrants as a consequence of this (notice in Figure 1.5 how many LLMs in Northern Italy—the more economically entrepreneurial part of the country—have large emigration rates), the OLS estimates would be biased upwards, possibly enough to find a positive or zero correlation between emigration and entrepreneurial intensity. On the other hand, if individuals are more likely to leave LLMs when labor demand declines and economic activity slows, then there would be a negative correlation between emigration and entrepreneurship and, thus, a downward bias, towards a negative effect. Moreover, because of delays and missing reports of short-term migration, as discussed in Section 1.2, the measures of emigration rates from 2008 to 2015 could suffer from measurement error, biasing the estimated coefficient toward zero. All these reasons suggest the existence of potential bias in the OLS estimates, although its direction is *a priori* unclear. Hence, while the OLS estimates of the β coefficients in Table 1.2 indicate no significant correlation between the LLM emigration rate and changes in firm stock, entry or exit, we should be aware of the significant potential bias. To correct the omitted variable and measurement error biases of OLS estimates, we exploit variation in migration flows driven by historical networks (which are measured more precisely) and weighted by economic pull factors, both of which are only very weakly correlated with local economic conditions in the place-of-origin.

1.3.1 Identification: The IV approach

The basic intuition for our main instrumental variable, a version of a shift-share/Bartik IV, is that LLMs have connections with specific foreign countries through previously established networks of emigrants. These pre-existing networks may share information or even job referrals to individuals living in the LLM-of-origin. Such networks exert a stronger pull the larger they are relative to the LLM population and if they connect to countries with strong economic opportunities. Building on this intuition, we interact the intensity of pre-existing networks with the economic growth of destination countries from 2008 to 2015. We use the number of people who emigrated from each LLM l to each foreign country c before 2000, as a percentage of the LLM population in 2000 (in a robustness check we also consider 1992 diaspora networks), as a measure of the network. We then weight these shares with the growth rate of GDP per capita in destination countries during the treatment period.[13] Summing across destination countries for each LLM produces an economic weighted, network driven factor, for the 2008 to 2015 period, specific to the LLM. The variable is defined as follows:

$$Pull_l = \sum_c NTWK_{l,c} * G_c \tag{1.2}$$

In expression (1.2), the first term, $NTWK_{l,c}$, is the number of Italians who moved from LLM l to country c before year 2000 (or 1992) and are still residents of c as of 2015, as a share of the LLM population in year 2000 (or 1992).[14] It captures the size of the

[13]GDP data are obtained from the World Bank national account database (World Bank 2019). We were able to match 184 destination countries.

[14]To maximize precision, we use the LLM population as of the 2001 Census to proxy population in 2000.

historic diaspora from LLM l in country c, which affects the potential for subsequent emigration outflows from l to c. The second term, $G_c = GDP_c^{2015}/GDP_c^{2008}$, is the growth rate of GDP per capita of country c during the treatment period, which includes the Great Recession and the sovereign debt crisis disproportionately hitting Southern European countries. Table 1.1 summarizes the variation in GDP growth from 2008 to 2015 for the main countries-of-destination. The variable defined in equation (1.2) is used as an instrument for the actual emigration rate, $(\sum_{t=2008}^{2015} m_{l,t})/pop_{l,pre}$, which is the main explanatory variable in the estimating equation (1.1). Let us emphasize, however, that the identifying variation generated by the IV, as we show below, depends primarily on the variation in network size, especially in Germany and Switzerland, across LLMs, much more than on the economic weights given to these networks. We therefore also use the network size in main destination countries (as share of population) directly as an IV in additional estimation results.

1.3.2 Instrument validity: Pre-trends

Our key identifying assumption is that the strength of the pre-existing diaspora networks weighted by the economic pull of destination countries from 2008 to 2015 is uncorrelated with unobserved factors specific to an LLM that may affect firm creation in the same period. However, our identification strategy is threatened if past economic shocks in an LLM persist over time and affect emigration before 2000 and firm creation in the treatment period. To increase confidence in the assumption underlying our IV, we perform several checks.

We first note that we include province fixed effects in our preferred specification to control for economic, institutional and policy trends (as the specification is in differences), which may vary substantially across locations in Italy. Province-specific trends capture the potential impact of policies common to these areas of about 500,000 people on average. Most importantly, the inclusion of fixed effects implies that the identifying variation of the IV is across labor markets that are geographically close to each other and have similar economic and social conditions, but can still be quite different in their diaspora network due to historical events and the highly-localized nature of migrant networks.

Figure 1.6 shows a quite informative representation of the raw data on the stock of firms per capita, our main outcome variable. We plot the average number of firms over time among LLMs above (below) the median value of the "instrumented emigration" (predicted by the first stage of the IV) with a solid (dashed) line (both values are standardized to one in 2005). First, we notice that the two groups have parallel trends up to 2009, which marks the onset of the emigration surge. Second, after 2009, the lines start progressively diverging, and they show a substantial difference by the end of the treatment period, in 2015. Firms per capita were fewer by the end of the period in local labor markets with "instrumented emigration" above median than in those below median.[15]

To confirm the "event" nature of the migration surge starting in 2009 and the independence of the IV from pre-existing economic trends, we check the within-province correlation of our instrument with the 2005-2008 trends of our key outcomes, as well as other economic and demographic variables[16]. In Table 1.3, we regress the 2005-2008 change in the stock, cumulative births, and cumulative deaths of firms on the post-2008 IV-predicted emigration. The estimated coefficients are very small in magnitude and not statistically significant.[17] Thus, the estimates of Table 1.3 are consistent with our identifying assumption, that the IV is not correlated with pre-2008 firm creation and destruction rates. We also estimate similar regressions on the other outcomes of interest, namely the firms owned by young entrepreneurs, total employment, employment-population ratio, total wage bill, and the number of blue and white collar workers and

[15]Appendix Figure A2 shows not just raw data, but the differences in an event-study graph. Each dot represents the estimate of an interaction between our IV and a year dummy. The pattern confirms that while our IV is not correlated to the stock of firms before the emigration episode, we estimate increasingly negative and statistically significant coefficients in the treatment years.

[16]2005 is the earliest year for which our firm data are available. We therefore cannot extend the analysis of pre-trends to years before 2005.

[17]In comparison, Appendix Table A18 shows the reduced form estimates obtained by regressing the 2008 to 2015 change in the same outcomes on our IV. The reduced form effect on firm creation in this case is statistically significant and 20 times larger than the coefficient in our validity check.

of managers. When we consider changes in those outcomes between 2005 and 2008, we never find a significant correlation with the IV, as shown in Appendix Tables A2, A3 and A4. These validity checks are consistent with interpreting our identification as hinging on the sudden post-2008 emigration surge instrumented by the different LLM intensity of local networks.

An additional concern is that the instrument may be correlated with other dimensions of local mobility. If the IV predicts internal migration[18] or inflows of immigrants into the local labor markets, then the causal interpretation of IV estimates would be problematic. In Appendix Table A7, we show that there is neither significant correlation of the IV with 2008 to 2015 internal migration or with the immigration rate of foreign-born individuals. This is not surprising, as the countries-of-origin of immigrants to Italy (mainly from Eastern Europe and North Africa) are different from those where Italian emigrants reside.

1.3.3 Shift-share diagnostics

The IV we construct has the structure of a Bartik/shift-share instrument. Specifically, it combines the variation in the past cross-sectional distribution of emigrants' population shares by destination country (the share component) with the destination countries' aggregate economic growth post-2008 (the shift component). Goldsmith-Pinkham, Sorkin, and Swift (2020) show that a sufficient condition for identification in this setting is that the past population shares of emigrants across LLMs are uncorrelated with the error term.[19] To test whether this is the case in our setting, we scrutinize the cross-sectional components of the IV. We first calculate the weights that the instrument attributes to each share (the so-called "Rotemberg weights"). Higher weights

[18]Internal migration flows are also from Istat, and they are based on transfers of residence between municipalities.

[19]Borusyak, Hull, and Jaravel (2021) show that a necessary and sufficient condition for identification is that the interaction between the shares and the shift components is asymptotically uncorrelated with the error term. This can be satisfied by a large number of uncorrelated shift terms, which is unlikely in our setting, as there are only a dozen important destination countries, and their growth rates are likely correlated. Borusyak, Hull, and Jaravel (2021) point out that the condition they propose is also satisfied by the exogeneity of shares proposed by Goldsmith-Pinkham, Sorkin, and Swift (2020).

correspond to greater relevance in the identifying variation. We then test whether the population share of emigrants receiving higher weights are correlated with pre-2008 observable characteristics of the LLM-of-origin.

Tables 1.4 and 1.5 report the main results of diagnostic tests as suggested in Goldsmith-Pinkham, Sorkin, and Swift (2020). Table 1.4 shows four sets of tests. First, in Panel A, we show the share of Rotemberg weights ($\hat{\alpha}_c$) that are positive and negative. Almost all of them are positive, indicating that the separate shares are positively correlated with the IV, thus suggesting that our instrument is a convex combination of the country-specific estimated $\hat{\beta}_c$ coefficients and does not show signs of mis-specification. Panel B reports correlations among the components of the IV (G_c and $NTWK_c$), the Rotemberg weights ($\hat{\alpha}_c$), the power of the IV ($\hat{F}_c$) and the estimated coefficients of equation (1.1) with per-capita stock of firms as the dependent variable ($\hat{\beta}_c$).[20] An informative statistic is the correlation between each component of the IV (G_c and $NTWK_c$) and the Rotemberg weights ($\hat{\alpha}_c$). A larger correlation implies higher relevance of that component of the IV in generating the identifying variation. We see that while the share component ($NTWK_c$) has a correlation of 0.84 with the weights, the shift component (G_c) has very low and even negative correlation (-0.05). This confirms that the share variation is what generates most of the identifying variation in our setting. Therefore, it is important to check that those emigration shares receiving the highest weights are associated with estimates of β similar to our main estimate and that they are not correlated with pre-2008 local characteristics.

Panel C of Table 1.4 reports the five destination country shares receiving the highest weight and, hence, driving most of the identifying variation. The share of emigrants to

[20] In all specifications we cluster the standard errors at the province level to capture potential error correlations of geographically close labor markets. This is consistent with what discussed in Adão, Kolesár, and Morales (2019) and Goldsmith-Pinkham, Sorkin, and Swift (2020). Following an exercise proposed by Adão, Kolesár, and Morales (2019) and implemented in Fouka, Mazumder, and Tabellini (2020), in Appendix Figures A4 and A5 we perform two placebo exercises where we substitute the shifters and the shares, respectively, with random numbers extracted from $N(0,5)$. The two exercises confirm that the clustered standard errors are valid and, if anything, too conservative: in the case of shifters, only 0.4 percent wrongly reject the null hypothesis of $\beta = 0$ at the 10 percent level, and never at the 5 percent level (0 out of 500 replications in the case of shares randomization).

Germany explains about 45 percent of the total instrument variation, and the share of emigrants to Switzerland generates an additional 28 percent. Hence it is important that we test their correlation with other variables, as we do in Table 1.5 discussed below. Panel C also shows that emigrant shares to France, Australia and Belgium receive non-negligible weights as well; however, when used individually, the F-statistic shows they are very weak instruments. A reassuring feature of our IV is that all estimates of the main coefficient of interest (β in equation 1.1), obtained using any of the top five shares as unique, just-identified, instruments, are all negative and similar in magnitude to the main estimate (-0.414). Estimates obtained using the German or Swiss share only, each of which exhibit a reasonably high F-statistic above 10, are both negative and significant (-0.388 and -0.202 respectively). The 95% confidence interval for the German and Swiss shares are in the negative range.[21] Panel D of Table 1.4 shows in columns (2)-(5) the estimates of the main coefficient using as instrument the share of past migrants to Germany, to Switzerland, the two shares jointly and the shares to the top 5 destinations jointly. We also report the test of over-identification, which never rejects the null that all the instruments produce the same coefficient estimate. In column (6), we show the Limited Information Maximum Likelihood (LIML) estimate obtained using shares of all networks as instruments. This method is more robust to weak IVs bias. Even in this case, the over-identifying restrictions are not rejected. Moreover, the estimates of the coefficient of interest are always negative and significant, consistently with the estimate obtained with our main specification IV (reported for convenience in column 7).

Table 1.5 shows the correlation of the emigrant shares to the top 5 destination countries (according to their Rotemberg weights) with the observable characteristics of the origin LLMs measured in the pre-period, from 2005 to 2008. Germany and Switzerland are particularly important, and a strong correlation of those shares with pre-existing economic trends would cast doubts on the validity of our instrument. From

[21] Following Goldsmith-Pinkham, Sorkin, and Swift (2020), we construct weak instrument robust confidence intervals using the Chernozhukov and Hansen (2008) method.

the regressions, however, we see no systematic correlation between the population share of emigrants to each of the main destination countries and the LLM-of-origin growth in the number of firms, firm birth or death rates, the unemployment rate, and GDP per capita before 2008.

To further validate our identifying variation, we perform two additional exercises. First, we hold the emigration networks (the shares) fixed and randomly assign the GDP changes (the shifts). Considering that the shares drive most of the identifying variation, the random permutation of the shifts should still allow us to identify our results. Indeed, Appendix Figure A6 shows that our main effect (see Section 1.4) is replicated under this permutation. To the contrary, when we randomized the main identifying variation (shares) we are not able to identify any effect (Appendix Figure A7).

As an additional exercise that can potentially increase the power of the instrument, we split the emigrants' destinations into smaller geographical units (consular areas) corresponding to European regions (Eurostat NUTS-2 classification) rather than countries, whenever this information is available in our data (i.e., for Germany, Switzerland, Belgium and the UK). The instrument constructed with this richer set of destinations—otherwise identical to the one used so far—does not show significantly higher power and has similar properties when subject to Goldsmith-Pinkham, Sorkin, and Swift (2020)'s tests (reported in Appendix A.IV). A large share of the variation is driven by two German regions (Stuttgart/Friburg and Dortmund/Koln) and two Swiss regions (Zurich and Lugano). Similar to what we find using Germany and Switzerland, we obtain values that are extremely close to our main estimate and to each other when we use only these most important regions of destination to estimate β_c.

Overall, these diagnostic tests indicate a prominent role of networks in Germany and Switzerland driving most of the variation in emigration, and therefore the IV variation. Most importantly they confirm that there is no systematic reason to believe those shares violate the identifying assumptions. Rather, the sufficient conditions for identification outlined by Goldsmith-Pinkham, Sorkin, and Swift (2020) are satisfied.

1.3.4 First stage and compliers characterization

In Table 1.6, we report the first stage results where we predict the emigration rate with the instrument, $Pull_l$. In the regressions, we control for GDP per capita and the unemployment rate in 2005, and we include region fixed effects in column (2) and province fixed effects in column (3). These controls capture pre-determined economic conditions in the LLM-of-origin. The estimates in the first row of Table 1.6 show that the $Pull_l$ has a significant predictive power for actual emigration. The size of the coefficient is stable across specifications.[22] The first stage F-statistics lie between 14.9 and 29.6, well above the standard rule of thumb value of 10, below which weak instrument concerns would arise.[23]

Among the three specifications, the one including province fixed effects is the most restrictive, as it leverages variation only within provinces (smaller than regions). The fixed effects account for unobservable characteristics generating common trends to LLMs within the same province. In the rest of the paper, we use this more demanding specification.

Figure 1.5 shows the geographic variation of the emigration rates (the endogenous variable) in Panel (a) and of their predicted values (IV) in Panel (b). The provincial boundaries are marked in bold. Based on historical emigration patterns, the IV predicts more emigration from the South, while the actual emigration in the treatment period was predominantly from Central and Northern regions, which are also richer and more dynamic in terms of business creation. The broad North-South variation, however, is not used in identifying our effects, as we include the fixed effects. This evidence will help us interpret the main results on firm creation we find below.

Finally, our IV strategy delivers treatment effects that can be interpreted analo-

[22]In Table A9 in the Appendix, we show the corresponding first stage estimates using the 1992 emigration shares. While the estimates are similar to those of Table 1.6, the instrument power is slightly lower, consistent with an older expatriate network being less relevant for emigrants in 2008-2015.

[23]For transparency, in the Appendix Figure A3, we also show scatter-plot correlation of the IV and endogenous variable after cleaning for the partial correlation with controls and province fixed effects. The Figure shows that our first-stage variation is not driven by outlier LLMs.

gously to Local Average Treatment Effects (LATE-like).[24] Thus, we attempt to characterize the local labor markets that comply in a LATE-like sense, i.e., local labor markets that experienced larger emigration rates because they happened to have a sizeable network of expatriates abroad, and that would not have experienced large emigration rates absent such a network. To characterize these LLMs, in Figure 1.7 we show the first stage coefficients and F-statistics obtained by splitting the sample along several dimensions. The plot shows that LLMs with a younger baseline population, a larger share of college graduates and higher firm creation rates before the onset of the Great Recession have a relatively larger first stage coefficient (as well as higher F-stats). This suggests that LLMs with a more dynamic, younger and more highly-educated labor force are more likely to be the "compliers", i.e. those locations where emigration responds more strongly to emigration opportunities as proxied by the IV. Our main IV estimates are, therefore, likely to reflect the effect of emigration on these dynamic LLMs and to be an upper bound of the average treatment effect.

1.4 Main results

1.4.1 Effects on firm creation

Table 1.7 shows the main results of the paper. The coefficients reported are from 2SLS regressions where we instrument the emigration rate with the pull factor IV. We also include pre-shock economic controls and province fixed effects.[25] The dependent variables are the change in the stock of firms in column (1), cumulative firm births in column (2), and cumulative firm deaths in column (3), all measured over the period 2008-2015. All the outcomes are standardized by the LLM population 25-64 years old before the emigration episode (average over 2005 to 2008) and expressed in percentage

[24] As shown by Goldsmith-Pinkham, Sorkin, and Swift (2020), such interpretation is possible due to the assumption of constant linear effects within a location over the whole support of the covariates. Moreover, as shown in Section 1.3.3, the Rotemberg weights in our setting are non-negative, thus limiting the extent of non-convex weights in our Bartik-style estimate.

[25] In the Appendix Tables A15, A16 and A17, we show robustness of these results to the exclusion of fixed effects or the inclusion of region- instead of province-level FEs. The underlying logic of our instrument should hold even without fixed effects if past emigration networks are not correlated with current economic trends. The estimates are similar to the main ones.

points. The emigration rate is normalized by subtracting its mean and dividing by its standard deviation, so that the coefficient can be interpreted as the change in the number of firms per one hundred people (25-64 years old) in response to an increase of the emigration rate by one standard deviation (which corresponds to about 1.7 percent of the average LLM population 25-64 years old). Standard errors are clustered at the province level to account for correlation of unobserved local factors.

The estimates indicate that after 2008, in areas with larger emigration flows, the total stock of firms declined. This effect is fully driven by fewer firm births (that is, lower firm creation) rather than higher firm deaths: on average, a one standard deviation increase in the emigration rate is associated with a decline of about 0.43 firms created per one hundred persons in the LLM. As shown at the bottom of Table 1.7, the average pre-2008 firm creation across districts was 9.08 firms per 100 people; hence, the loss that we attribute to one standard deviation of emigration is about 4.76 percent of the total firm creation in the pre-treatment period.[26] The predominant effect on firm creation suggests that emigration deprives the area of individuals with high entrepreneurship propensity and that potential negative externality/spillovers may be at work. In fact, the loss of firm creation due to emigration is 2 to 3 times larger than what the simple subtraction of people with average propensity to be entrepreneurs would have produced. In Section 1.4.2 below, we show how the selection of emigrants on age and education, plus plausible negative externalities of emigrant entrepreneurs on other LLM residents' entrepreneurial success, would generate firm creation losses consistent with our estimates.

The small and non-significant coefficient of emigration on the number of firm deaths is also reassuring: significant correlation between emigration and firm deaths could suggest a reverse channel of causation, as firm losses—and related job losses—due to the recession may encourage people to emigrate. We check the robustness of these results by controlling for lagged values of the outcomes in Appendix Table A11. Consistent with

[26] Alternatively, a one percentage point increase in the emigration rate generates a 2.8% decrease in the number of firms created.

the checks in Section 1.3.2 showing no significant pre-2008 trends, this specification in Table A11, aimed at purging the dependent variable from potential correlation with persistent pre-2008 shocks, shows coefficients which are not significantly different from those in Table 1.7.

1.4.2 Average subtraction, selection and spillover effects

A fraction of the estimated loss in firm creation is simply a *subtraction of people* effect. Namely, if emigration drained people at random (i.e., with an entrepreneurship rate equal to the population average), fewer firms would be created. Three additional and more interesting effects are present, however. First, emigrants are potentially more likely to start a firm than stayers because of their distribution of age and schooling (education and age selection). Second, emigrants differ in other less observable characteristics (e.g. risk-aversion, adaptability) from non-migrants, and this may affect their entrepreneurship– a *residual selection* effect. Third, other people in the LLM-of-origin may be less inclined to start firms without the inspiration, learning potential, peer effects, local demand—*spillovers*, in short—of the departed entrepreneurs.

As *residual selection* and *spillovers* are hard to identify without strong assumptions, we first leave them in a residual term. Thus, we decompose the loss in firm births due to emigrants as follows:

$$\begin{aligned} \Delta \text{Firms Birth} &= \text{Subtraction of People} + \text{Education\&Age Selection} + \\ &+ \text{Residual Selection and Spillover} \end{aligned}$$

We first translate the estimated effect from Table 1.7 into total firm creation lost in the treatment period in the hypothetical average-size Italian LLM (whose 25-64 years old population was 44,805 between 2005 and 2008) as the emigration rate increases by one standard deviation. The estimated loss is equal to 194 fewer new firms over the seven years from 2008 to 2015.[27]

[27] This is because one standard deviation of the emigration rate which generates a 0.432 decline–the coefficient in Table 1.7– of firms per 100 people, which multiplied by (44,805/100) equals 194.

The *subtraction of people* effect is simply obtained by multiplying the pre-2008 firm-creation rate (new firms as share of the working age population), which was 9.1% (corresponding to a 1.3% annually, added over the seven years considered), by the number of people who left the country in the average LLM (760).[28] The *subtraction of people* effect, due simply to subtracting people with average entrepreneurial ability, is equal to 69 firms and accounts for about one third of the total estimated firm loss.

There are two demographic dimensions of selection—namely, age and education—for which we observe the aggregate distribution of Italian emigrants and for which we can assign group-specific firm-creation rates. As migrants are more concentrated among the groups with higher firm-creation rates relative to the overall Italian population, these calculations generate a measure of the additional firm loss due to selection along those two dimensions. We calculate such a loss for the average emigrant composition. Specifically, we split the population between young (25-44 years old) and old (45-64 years old), as 76% of migrants were young (versus only 51% of the population in 2008). Given that the pre-2008 firm-creation rate among the young was 1.8% (versus 0.8% among the old), selection on age further lowers firm creation by 14 firms.[29] Similarly, we split the population between college and non-college educated: as about 30% of migrants have tertiary education (versus only 10% in the population as of 2008), and the firm-creation rate of college-educated individuals was 2.7% (versus 1.2% for non-

[28]More precisely, the annual firm-creation rate pre-2008, r_{pre}, is defined as the average number of firms created in the 2005-08 period divided by the average 2005-08 25-64 years old population. The firm-creation rate in the Italian data is consistent with comparable estimates from other countries. For instance, based on the business applications data collected by the US Census Business Formation Statistics (BFS) program, there are between 0.9 (using high-propensity business applications) and 1.6 (total business applications) new firms per person, 25-64 years old, from 2005 to 2008.

[29]The age selection term is computed as follows:

$$\text{Age Selection} = \overbrace{\underbrace{\text{Emig}^{25-44}}_{-580} \times \underbrace{(r_{pre}^{25-44} - r_{pre})}_{(0.018-0.013)*7}}^{\text{correction for young}} + \overbrace{\underbrace{\text{Emig}^{45-64}}_{-180} \times \underbrace{(r_{pre}^{45-64} - r_{pre})}_{(0.008-0.013)*7}}^{\text{correction for old}} = -14$$

We calculated the age-specific firm-creation rates using our Chambers of Commerce data, as we explain in Section 1.4.3.

tertiary educated), selection on education explains the loss of 19 additional firms.[30,31]

The age and education selection suggests that emigrants are more likely than stayers to be entrepreneurs and start new firms. Likewise, they could also be selected on characteristics that we do not observe and that are positively correlated with entrepreneurship.[32] Existing studies show that emigrants have lower risk aversion (Jaeger et al. 2010) and higher "adaptability" to new people and situations (Bütikofer and Peri 2021), both of which may positively affect the probability of being an entrepreneur. Hence, there can be additional selection on pro-entrepreneurial characteristics that we do not observe.

As the effect of *unobserved selection* as well as the magnitude of spillover are harder to quantify, we first populate the decomposition shown at the beginning of the section with the estimated firm loss on the left hand side and and with the components estimated so far. We show in blue the average firm loss due to *Subtraction*, in green the one due to *Age selection* and *Education selection* of emigrants, and in red the firm loss

[30]The education selection is computed as follows:

$$\text{Edu Selection} = \overbrace{\underbrace{\text{Emig}^{LowEdu}}_{-532} \times \underbrace{(r_{pre}^{LowEdu} - r_{pre})}_{(0.012-0.013)*7}}^{\text{correction for Low Edu}} + \overbrace{\underbrace{\text{Emig}^{HighEdu}}_{-228} \times \underbrace{(r_{pre}^{HighEdu} - r_{pre})}_{(0.027-0.013)*7}}^{\text{correction for High Edu}} = -19$$

The education-specific firm-creation rates have been calculated based on Istat administrative data. Istat combines several administrative data sources to perform an individual-level linkage of firms' owners and managers from Chambers of Commerce data to their educational level, which they obtain from Ministry of Education data as well as from the 2011 Census (Istat 2014a). While we do not have access to these data, we use information from Istat (2014b) that reports the share of new entrepreneurs with a college degree (25.4%) in 2014 (the earliest year available). We adjust this share downward to the 2005-08 period by dividing it by the annual growth rate (about 4%) of the share of college graduates among the Italian population, and then applying the resulting shares to the firms created in the 2005-08 period. Reassuringly, if we perform the same procedure for entrepreneurs' age (which we observe in our data), we find remarkably similar estimates: the share of under-35 among new entrepreneurs in 2014 from the Istat (2014b) report is 34.4%, while in our data it is 36.2%.

[31]As an alternative, we calculated the age and education composition of "complier" LLMs, i.e., those with predicted emigration above median. As their average share of young is also 76% (equal to the average LLM) and their share of college educated is 31% relative to the 29% of the average LLM, this decomposition produces very similar results, with 7% of the effect due to age selection and 10% (-19.5 firms rather than -19) due to education selection.

[32]The direction and magnitude of the documented selection of Italian emigrants is consistent with those found in several other studies of emigrants. For instance, Grogger and Hanson (2011a) find positive selection on schooling, Parey et al. (2017) shows selection on pre-migration earnings, and Patt et al. (2021) on occupational skills.

due to *Residual Selection and Spillovers*:

$$-194 \quad = \quad -69-14-19-92$$

Expressing each component as a percentage of the total estimated firm-loss effect, we have:

$$100\% \quad = \quad 36\%+7\%+10\%+47\%$$

The decomposition indicates that one third of the loss is due to pure subtraction of people, one sixth to their selection on age and education and the residual one half to selection on non-observable variables plus spillovers. To make progress on decomposing this residual term, we consider estimates of human capital spillovers in the literature. Those are measured as the elasticity of productivity (wages) of the labor force to increases in the college share at the city or state level (as produced for instance in Moretti (2004), Iranzo and Peri (2009) and Winters (2014) for the US). We then translate those magnitudes into effects of the loss of entrepreneurs' share of the population (due to emigration) on firm-creation rate (as if it were productivity) of the rest of the population, and we calculate what decline in firm-creation over 7 years such spillovers would produce. We consider the decline in firm-creation rate from emigration, as measured by the subtraction of people plus age-and-education selection (equal to a loss of 102 firms), and we calculate the externality as a reduction that this loss of entrepreneurship has on the firms created by the rest of the population. Using an externality elasticity between 0.7 and 1, which is in the low-range of those estimated by Winters (2014) and in Moretti (2004) (Table 3), and compatible with those found in Iranzo and Peri (2009), we obtain a loss of 7-year firm-creation rate for the population left behind in the range of 0.16-0.2 percent. Applied to the population in working age in the average LLM, these figures generate a loss of 71 to 92 firms due to negative spillovers.

Specifically, this range is obtained by multiplying the percent decline in firm creation from the subtraction and selection effect (-102/4,068=-0.025) by the externality elasticity range (0.7-1.0) and by the average firm creation rate over 7 years (7*0.013), and finally scaling this rate for the average population aged 25-64 in a LLM (equal to 44,805).[33] Based on the strong assumptions of this exercise, the spillover effects equal between 77 and 100% of the residual effect, and hence, 36 to 47% of the total effect. This would leave selection on non-observed characteristics responsible for a loss between 0 and 21 firms.

Such an accounting exercise, albeit simple, provides some guidance to thinking about the channels of the estimated firm loss. First, the loss of firm creation would have been one third of the observed one if emigrants were randomly selected from the population, and there was no spillover from their departure. Second, we see that selection on age and education is substantial, explaining 17% of the effect, but leaves some role for non observable characteristics in explaining up to 11% of the effect. Finally, using externality/spillover elasticity estimates from the human capital literature, we calculate that 37-47% of the total effect may be due to the spillover effect on remaining individuals.[34]

1.4.3 The loss of young people and innovative start-ups

Entrepreneurship of young people is likely to introduce new and "creatively" disruptive technologies. Hence, the loss of entrepreneurial capital due to emigration may be particularly damaging for economic growth if it is also associated with a drain of young innovative entrepreneurs. We extend our previous analysis and focus on the possible impact on innovative potential by focusing on firms created by young entrepreneurs and on firms that operate in technology-intensive sectors. We call this latter group of

[33] 102 is the total from subtraction of people plus the education and age selection terms (-69, -14, -19). The average baseline number of new firms, 4,068, is obtained by multiplying the baseline 2005 to 2008 firm-creation rate, 9.08, by the average population size, 44,805. Considering that the upper bound of the spillover effects, -0.025∗1∗7∗0.013∗44,805=-102, slightly exceeds the residual unexplained loss after factoring out subtraction and education-age selection effects, we consider -92 as the plausible maximum spillover effect, corresponding to an externality elasticity of 0.9.

[34] This is consistent with evidence that entrepreneurs whose benefits are lost in the place-of-origin may help create agglomeration of innovation in destination areas (e.g. Kerr et al. 2017b), confirming that their departure exerts negative spillover effects on the local economies-of-origin.

firms "innovative start-ups," as they are those more likely to embody new technologies and ideas.[35]

In Table 1.8, we first look at the creation and destruction of firms whose owners and executives are younger than 45. The age of owners and executives is reported in the data from the Chambers of Commerce, and this information is used to construct a synthetic measure that identifies a firm as "owned and managed" by young people if the majority of owner-executives are under 45 years old. We then look at the effects of pull-driven emigration on the number (column 1), creation (column 2), and destruction (column 3) of this subset of firms. The results in Table 1.8, which mirror those of Table 1.7, indicate that a one standard deviation increase in emigration (as induced by our pull instrument) reduced the number of firms created by young individuals by 0.23 firms per 100 people, which is equivalent to a 3.6% decrease relative to baseline firm creation. More than half of the loss in new firms generated by emigration occurs because of fewer firms created by young individuals.

In column (4), we focus instead on the net cumulative entry of innovative start-ups in each LLM in the post-2008 period as a dependent variable.[36] The estimated coefficient is statistically significant and indicates that the larger the migrant outflows from Italian LLMs, the less likely those LLMs are to birth innovative start-ups. While on average there were 0.01 innovative start-ups per 100 people in the average LLM (or 1 per 10,000), a one standard deviation higher emigration rate induced a lower creation of about 0.004 start-ups per 100 people (or 0.4 per 10,000). Emigration is associated with a 40% decline in innovative start-up creation, a substantial and alarming decline in the creation of innovative firms which are likely responsible for job creation and growth.

[35] Data on start-ups come from the *Registry of Innovative Start-ups*, a special section of the Italian firms registry (Infocamere 2016). Newly born firms which develop, produce or sell highly innovative products or services can apply to this registry if they satisfy one of the following conditions: i) 1/3 of their workforce hold a PhD or 2/3 hold a graduate degree; ii) R&D expenditures amount to at least 15% of revenues (or costs, if higher); or iii) they hold at least one patent of an innovative nature. Firms can maintain this status up to 5 years after registration provided their revenues do not exceed 5 million euros. They cannot be spin-offs of larger established firms.

[36] The outcome is a *net* entry rate, as we observe only a 2015 snapshot of firms started since 2009, and we only capture those start-ups that were able to survive over the entire period. Moreover, since the registry starts in 2009, we are not able to test for pre-trends with this particular outcome.

Such a large effect can be explained by the fact that young start-up entrepreneurs are a rather small group in the population, and it is reasonable to expect that they are also concentrated in few LLMs, where the pull factors are stronger. Considering the well known tendency of STEM (Science, Technology, Engineering and Math) professionals to dominate the group of highly educated migrants to countries such as the US (see Peri, Shih, and Sparber 2015) or the UK, and considering their significant contribution to innovation in their destination countries (see Kerr and Lincoln 2010), there could be a corresponding decline of innovation in their countries-of-origin.

1.5 Labor demand effects, skill composition and wages

The evidence presented so far highlights two important facts. First, emigration produced a loss of entrepreneurship, reducing firm creation by a significant amount. Second, this loss was larger than what the simple "subtraction" of average individuals would imply, suggesting that emigrants were more likely to be entrepreneurs than the average individual. A mechanical consequence of this higher propensity to be entrepreneurs is a lower propensity to be employees. Emigration is traditionally exemplified as a loss of labor supply, and symmetrically immigration is modeled as an increase in labor supply. However, if emigrants are significantly more likely to be entrepreneurs (relative to non-migrants), and the firms they start create additional jobs, then emigration may actually reduce local labor demand together with labor supply. There is significant evidence that immigrants are more likely to be entrepreneurs relative to natives, especially in the US, as summarized by Fairlie and Lofstrom (2015). Our paper is the first, to our knowledge, to suggest that emigrants are selected among highly entrepreneurial individuals relative to non-migrants in the country-of-origin. If entrepreneurship (including human capital and know-how to start a firm) is a scarce factor complementary to labor, and it is needed in production, then the loss of one person can be thought of as a loss of a fraction of one worker and a fraction of one entrepreneur. If emigrants are more likely to be entrepreneurs, then their loss reduces firm creation and the demand for local workers more than it reduces the labor supply. If this is the case, then larger emigration would be associated with lower employment rates, weaker labor markets and lower wages. Such a crucial role of entrepreneurship as a scarce factor, generating labor demand in a local economy, is emphasized in Beaudry, Green, and Sand (2018). In that study, they show that an increase in local population due to internal migration does not depress local wages, rather it increases local entrepreneurship. In their model, a decrease in population with larger propensity to be entrepreneurs would decrease labor demand and lower employment rates.

Table 1.9 shows the correlation of employment outcomes with emigration, instrumented with the $Pull_l$ IV. First, we test the impact on employment in column (1). The estimate shows a negative and significant effect of emigration on employment. The magnitude of the coefficient is 4.6% fewer employees per one standard deviation of emigration, which corresponds to a 1.7 percent emigration rate. A back-of-the-envelope calculation shows that for the average LLM, a one standard deviation increase in the emigration rate, which corresponds to 760 emigrants, would imply a decrease of 769 employed LLM residents.[37] Such an impact is much larger than what obtained by subtracting the average number of employed people among the population that migrated. Based on the employment rate in Italy in 2005 (equal to 0.57), the number of employees lost because of additional 760 emigrants would have been only 438, rather than 769. Therefore this implies the loss of additional jobs on top of those subtracted by a simple loss in labor supply. Column (2) shows, consistently, that the employment-population ratio—a measure capturing the number of jobs per capita in a local economy—declines in response to emigration, albeit not significantly. Column (3) shows that the average firm size did not significantly change in response to emigration, again suggesting that it was not simply a subtraction of workers from a fixed number of existing firms, which would have implied a decline in average firm size. Finally, column (4) shows that the overall wage bill in the LLM experienced a non-significant negative change in response to emigration, signaling a decline in labor income in the local economy. Taken

[37]This is simply the product of 4.6% times the baseline employment in the average LLM, 16,709.

together, these results do not suggest that the departure of emigrants was associated with a tightening of the labor market.[38] Instead, the overall picture is more consistent with the idea that emigration reduced labor demand as much as labor supply.

Furthermore, in Table 1.10, we explore whether emigration has altered the relative skill composition of employment in the economy. In particular, we analyze whether emigration rates affected employment of specific skill groups more than others. We distinguish in increasing order of average wage, between blue collar, white collar, and managerial jobs (as defined by the INPS dataset). We find that, while there is a small, insignificant negative effect on the number of blue collar workers in the labor market, there is a larger, negative and significant effect on white collar workers. Emigration is also associated with a large, negative change in managers, although imprecisely estimated and statistically insignificant.

These findings are consistent both with the selection of emigrants among the high-skilled and with the notion that the loss of new firms depressed demand for skilled labor more than that for unskilled labor. Overall, a local economy that lost emigrants experienced lower firm creation, fewer innovative start-ups, a (non-significant) decline in employment-population ratio and a decline in skilled employment. Taken together, these effects appear consistent with a loss in local entrepreneurship generating a drop in labor demand together with a decline in labor supply.

1.6 Robustness checks: other forms of mobility and trade

Emigration abroad is only one of the potential flows of individuals to and from a local area. Local economies also experienced internal migration flows of Italian citizens who moved within the country, as well as inflows of foreign immigrants. Those flows may be correlated with local economic conditions and, hence, with the flows of Italians moving abroad. Moreover, they can partially compensate for the impact of emigration

[38]We do not show the effect on average wages because it is insignificant in most specifications and its interpretation is less clear; its effect combines a change in employment composition (as shown in Table 1.10 below) together with the relative demand and supply effects.

on firm creation. If the IV is not entirely uncorrelated with other migration flows into or out of the local area, their presence may generate spurious results. To address the potential confounding effect of other migration flows, we perform several robustness checks. First, in column (1) of Table 1.11, we augment the basic specification by adding, as a control, the immigration rate of foreigners to each LLM. The estimated effect of the emigration rate is negative and significant, but slightly attenuated relative to our main specification. This is not surprising, as immigrants to Italy are from countries in Eastern Europe and North Africa (while emigrants go to Germany and Switzerland) and settle in locations hardly correlated with those with large emigration networks.[39]

As a second robustness check of our results, we exclude those areas which are more likely to be strongly affected by international commuting and trade, which are also potentially correlated with emigration. The map in Figure 1.5 (a) shows that migration outflows are more intense in border regions, which are also strongly connected with foreign countries in terms of commuting patterns and local trade. Trade relations and migration flows may be correlated (Rauch 1999; Rauch 2001), and both are correlated with past economic conditions, so we exclude the Italian LLMs touching a border with other countries, for which this correlation may be stronger. The results of this exercise are presented in column (2) of Table 1.11. The point estimate of the effect on firm creation barely changes, offering reassurance that our main conclusions are not biased by the presence of international commuting or trade. A more direct way of controlling for potential trade flows is presented in column (3), where we control for the share of firms in the tradable sector as of 2005. Including this control is associated with a slightly larger coefficient on the emigration rate, suggesting that the effects are not due to a spurious correlation with trade.[40]

[39] We formally confirm this finding in Table A7 of Appendix A.V, where we show a placebo first stage regression of the *Pull IV* on immigration flows.

[40] In Table A8 in Appendix A.V, we perform additional checks to further prove that the effects we find are not driven by trade linkages correlated with emigration networks.

1.7 Conclusions

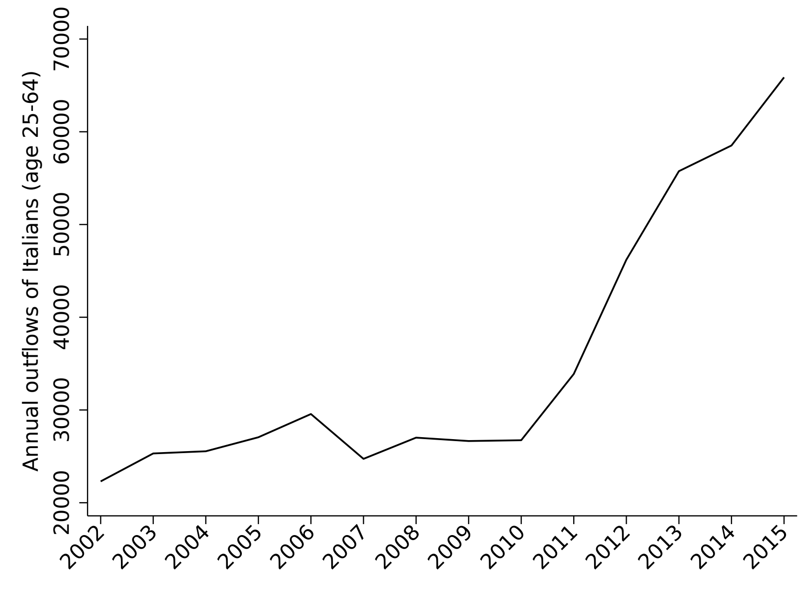
In this paper, we provide empirical evidence on an important question about which we know little: does emigration affect firm creation in the country-of-origin? We shed light on this question by taking advantage of a sudden and large emigration wave from Italy, that occurred between 2008 and 2015, and by using an instrumental variable strategy to isolate pull factors that are uncorrelated with local economic conditions. We combine data on emigration at the local labor market level with data on firm creation and start-ups to assess the impact of the former on the latter. We show that the IV-induced variation in emigration rates across local economies is independent of pre-2008 local trends in firm creation and economic outcomes. This is consistent with the validity of the exclusion restriction and with causal interpretations of our IV estimates.

Our results indicate that Italian local labor markets that lost more people due to emigration also experienced less firm creation. Moreover, we observe fewer births of innovative start-ups in those areas, as well as a decrease in employment and in the share of skilled workers workers. We find that a one standard deviation increase in the emigration rate over the considered period generated a loss of around 190 new firms in the average LLM. Put differently, a 1.7 percent increase in the local emigration rate decreased the local rate of firm-creation by about 4.8 percent of the average over the period. We then show that this effect is consistent with a decomposition into four parts. The first is a simple subtraction of people with average population characteristics, and accounts for about 36% of the total. The second component is associated with the selection of emigrants of younger age and higher education than the average—those who are more inclined to be entrepreneurs—and accounts for about 17% of the total effect. The third component captures entrepreneurship spillovers, which based on the magnitude of productivity spillovers estimated in the literature, can be as large as 36-47%. Finally, a residual component (0-11%) is the potential effect of the selection of emigrants on unobservable characteristics such as lower risk aversion and greater "adaptability" – characteristics that are also associated with entrepreneurship.

The findings in this paper have two main implications. First, international migration implies much more than "labor supply" changes once one considers the impact on firm-creation and consequently on job-creation. Migrants' roles as job-creators can be larger than their roles as employees; thus, traditional models that focus only on changes in labor supply may be missing a crucial part of the story. Second, our results suggest that emigrants are a highly-selected group with high entrepreneurial abilities, and that their loss can generate a significant shortage of the "entrepreneurship factor" crucial for job creation. This is in line with recent research showing that migrants have a higher propensity to take risks (Jaeger et al. 2010) and greater intensity of traits such as "adaptability to new circumstances" (Bütikofer and Peri 2021). This positive selection of migrants on non-cognitive traits may be very important for understanding their economic impacts and potential as workers, entrepreneurs and professionals in the receiving countries, and we hope to stimulate more research in this area.

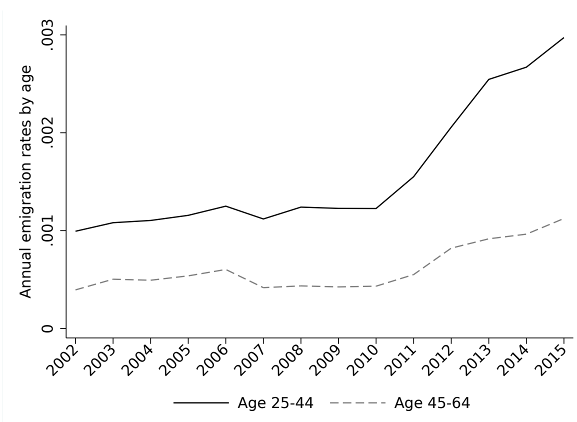
Figures

Figure 1.1: Emigration flows of Italians 25-64 years old

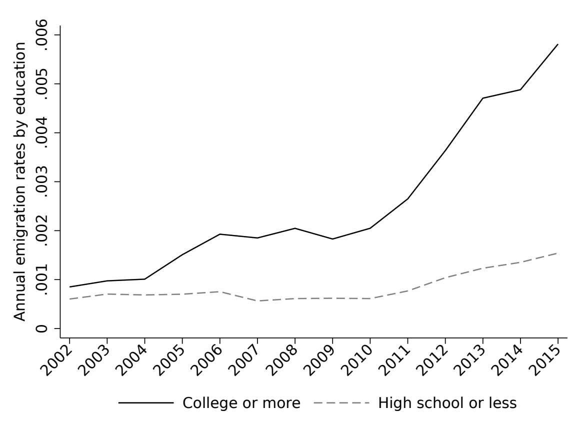


Notes: Annual outflows of Italian citizens 25-64 years old. Source: AIRE-Istat.

Figure 1.2: Emigration rates by age and education



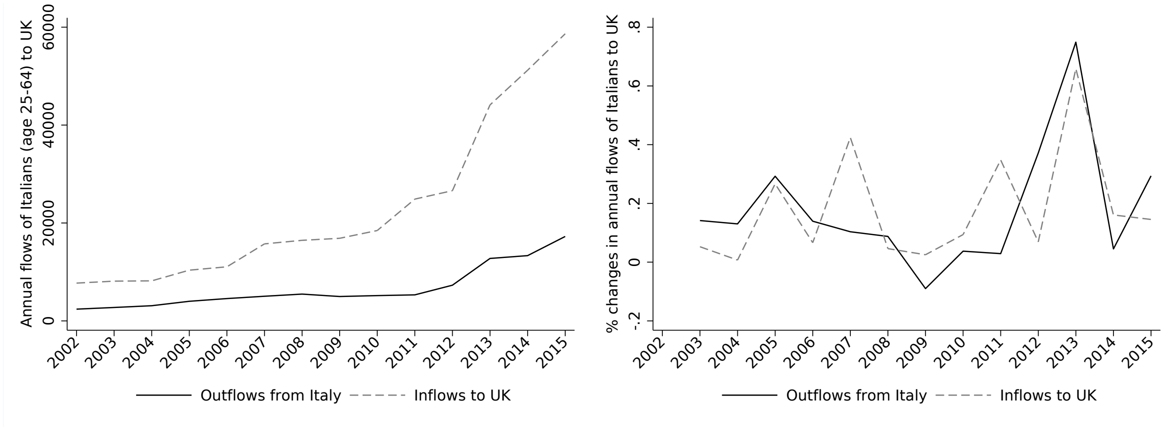
(a) Annual emigration rates, by age



(b) Annual emigration rates, by education

Notes: Annual outflows of Italians 25-64 years old. In Figure (a), emigration rates are as a fraction of the Italian resident population in 2002 by age group. In Figure (b), emigration rates are as a fraction of the Italian resident population by education group, as of the 2001 Census (Istat 2005). Source: AIRE-Istat.

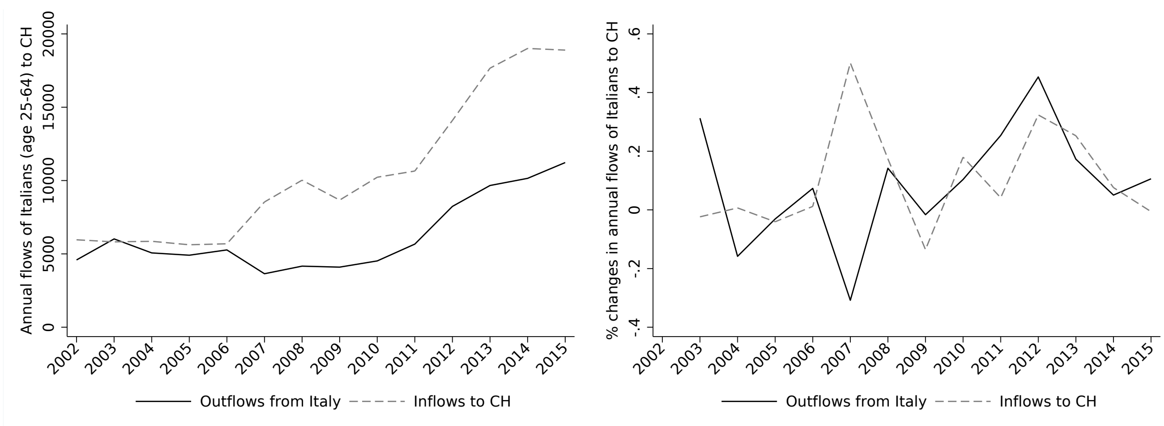
Figure 1.3: Recorded Emigration and Registered Inflows of Italians to the UK



(a) Annual emigration of Italians to UK (b) % changes in annual flows of Italians to UK

Notes: In Figure (a), the black solid line shows the annual outflows of Italians to the United Kingdom recorded in the AIRE-Istat data, while the grey dashed line shows the corresponding annual inflows of Italians to the UK according to the UK Social Security Registry data. Figure (b) shows the percentage changes in the annual flows from the two data sources.

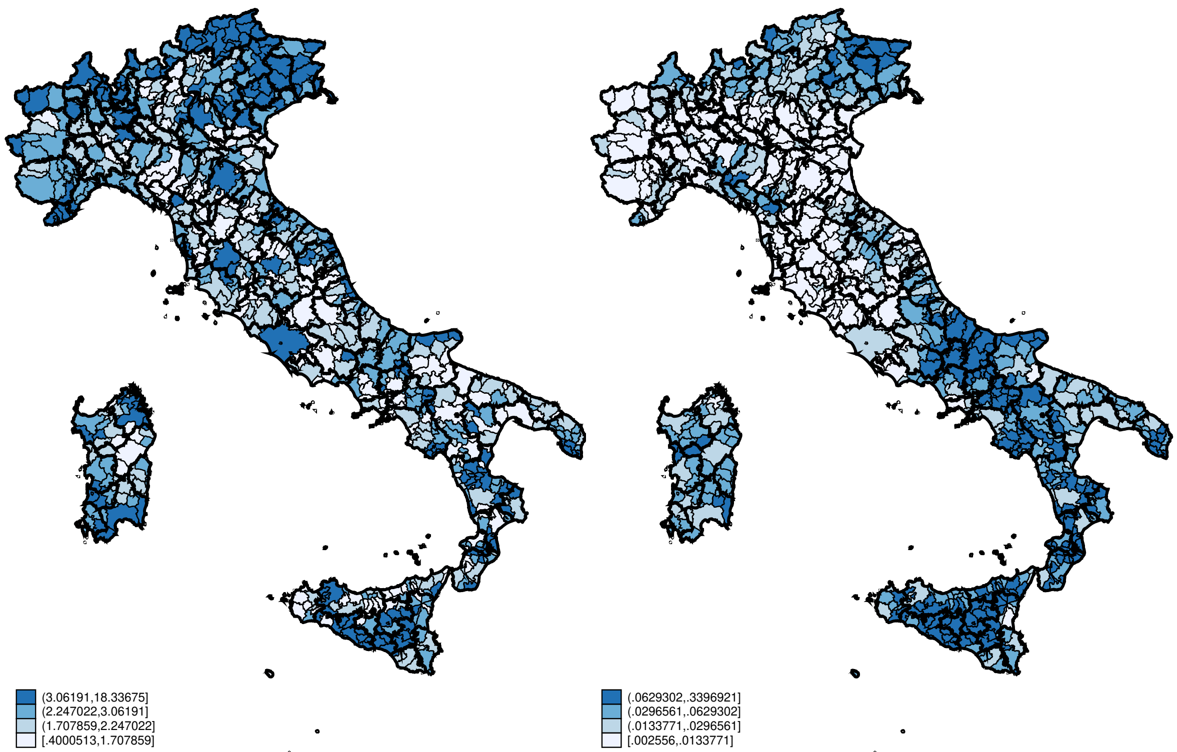
Figure 1.4: Recorded Emigration and Registered Inflows of Italians to Switzerland



(a) Annual emigration of Italians to CH (b) % changes in annual flows of Italians to CH

Notes: In Figure (a), the black solid line shows the annual outflows of Italians to Switzerland recorded in the AIRE-Istat data, while the grey dashed line shows the corresponding annual inflows of Italians to Switzerland according to the Swiss Bundesamt für Statistik (BFS) data. Figure (b) shows the percentage changes in the annual flows from the two data sources.

Figure 1.5: Actual and Predicted Emigration from Italian LLMs

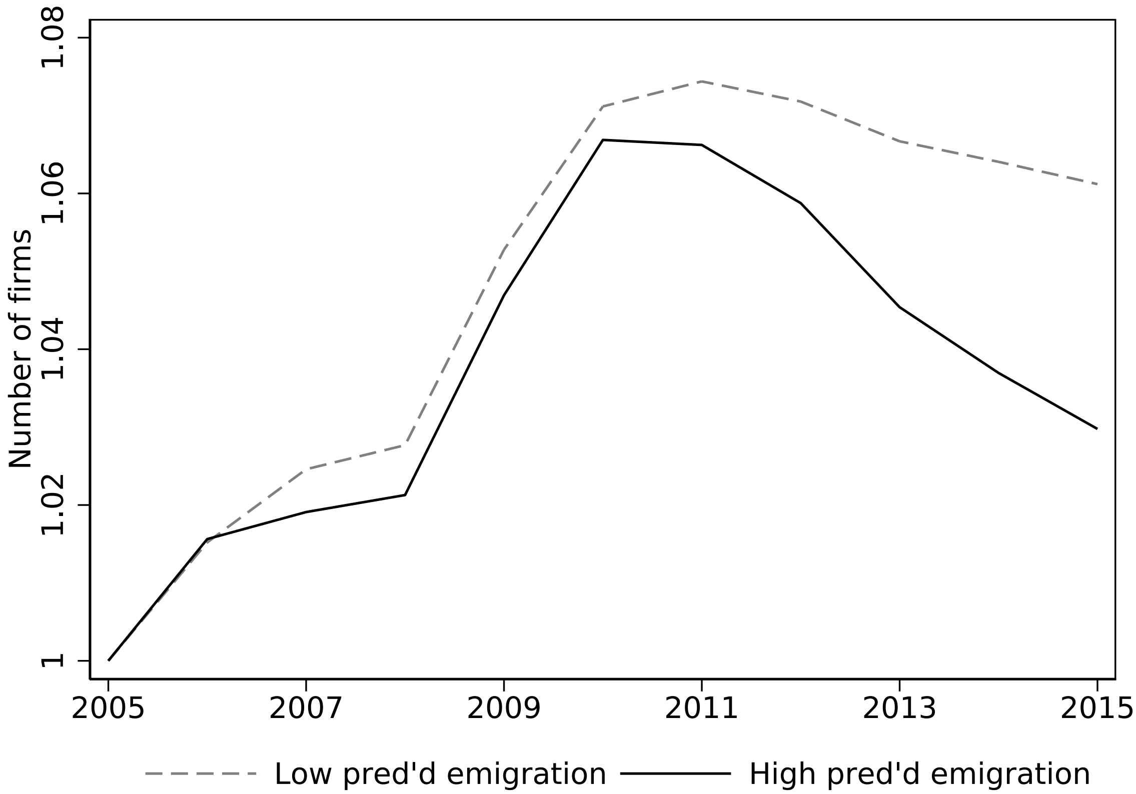


(a) Emigration

(b) Predicted Emigration

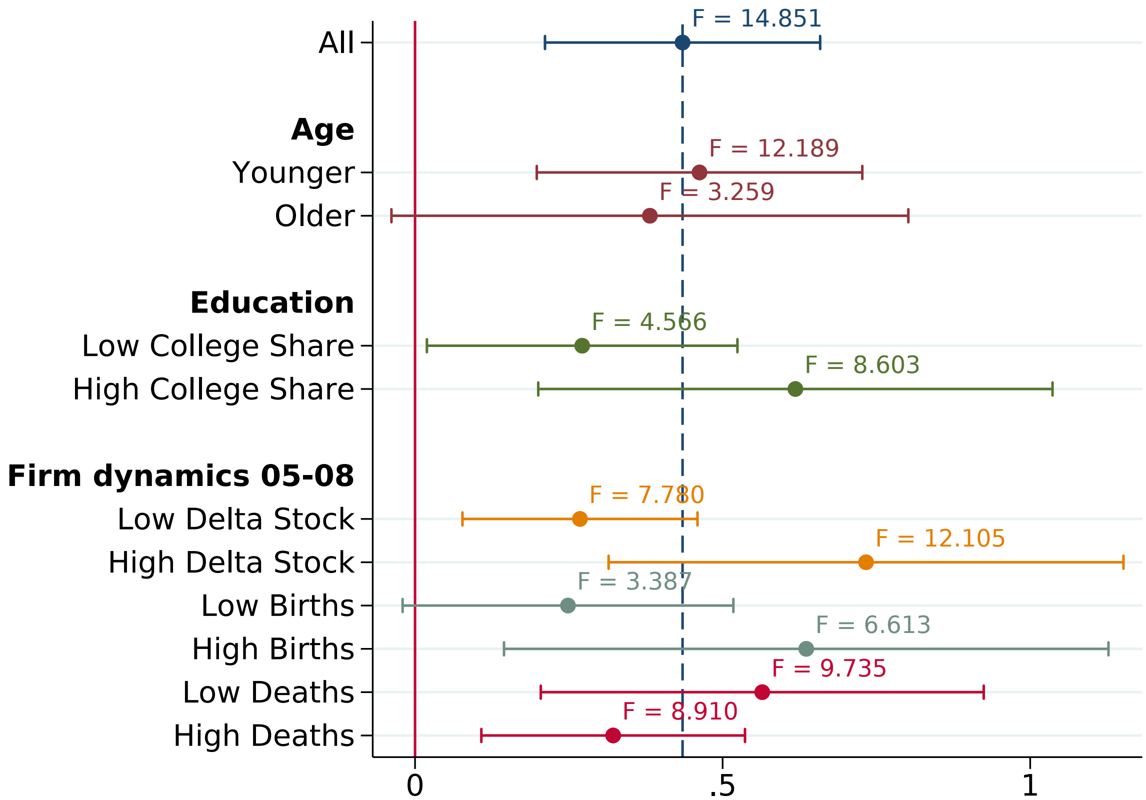
Notes: Figure (a) plots the cumulative emigration rate between 2008-2015 in percentage points, i.e. the number of Italian citizens 25-64 years old migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), multiplied by 100 and normalized to have mean zero and unit variance. Figure (b) plots the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$, normalized to have mean zero and unit variance. The black solid lines show province boundaries.

Figure 1.6: Firm stock in predicted high and low emigration LLMs, 2005-2015



Notes: The graph plots the stock of firms per person (25-64 years old) in LLMs predicted to have low and high emigration. The two series are normalized by their levels in 2005 (0.19 and 0.18 for low and high predicted emigration LLMs respectively).

Figure 1.7: Characterizing Complier LLMs in the IV



Notes: The graph plots the coefficients from separate first-stage regressions based on different breakdowns of the 686 Local Labor Markets (LLM), as well as the corresponding F-statistic on the excluded instrument. The first row simply reports the coefficient of Table 1.6, column (3), which includes all the 686 LLMs. In panel "Age" we split LLMs based on whether the average age of LLM population in 2005 is below/above median. In panel "Education" we split LLMs based on whether the LLM population share of college graduates is below/above median as of the 2001 Census (Istat 2005). In panel "Firm dynamics 05-08" we split LLMs along three different dimensions of baseline firm dynamism: first, by whether the change in stock of firms between 2005-2008 is below/above the national median; second, by whether the cumulated firm-creation between 2005-2008 is below/above median; third, by whether the cumulated firm exit between 2005-2008 is below/above median. Confidence intervals are at the 5-percent level.

Tables

Table 1.1: Emigration by country-of-destination, top 5 countries: 2000 stock, 2008-2015 flows and 2008-2015 GDP performance

Panel A Top countries in 2000	Stock of Emigrants	GDP 2015/2008
Germany	286,570	1.07
Switzerland	228,725	1.01
France	165,244	1.01
Belgium	117,935	1.00
Argentina	99,506	1.04
Panel B Top Countries in $2008-15$	Flows	$\%\ of\ 25-44-y.o.$
Germany	70,104	48.6
U.K.	66,094	61.2
Switzerland	53,567	52.3
France	45,046	46.8
United States	27,563	54.9

Notes: Panel A reports the top 5 countries in terms of size of the emigration network as of 2000 as measured in the AIRE data, and the GDP per capita growth between 2008-2015 based on World Bank data (out of a total of 184 countries considered). For reference, GDP per capita growth was 1.02 and 1.06 in UK and US respectively and 0.9 in Italy. Panel B reports the cumulative emigration flows to the top 5 destination countries in the period 2008-2015 and the share of 25-44 years old as measured in the Istat data. Stocks, flows, and the denominator of the share of young individuals include emigrants of all age groups.

Table 1.2: OLS regressions of LLM firms dynamics on observed emigration rates

VARIABLES	(1) All Firms Δ Stock 2008-15	(2) All Firms $\sum$ Births 2008-15	(3) All Firms $\sum$ Deaths 2008-15
Emig Rate	0.037	-0.022	-0.059
	(0.047)	(0.077)	(0.076)
Unemp Rate 2005	6.125	0.020	-6.105
	(3.960)	(3.911)	(4.680)
GDP PC 2005	7.302	4.089	-3.213
	(4.481)	(0.682)	(4.406)
Observations	686	686	686
R-squared	0.185	0.567	0.241
Avg. Baseline Outcome	0.790	9.078	8.288
Mean Emig Rate	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in firm stock (1), cumulative firm entry (2) and exit (3) between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008; source: Istat 2013), and normalized to have mean zero and unit variance. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level (source: Istat 2014c) as well as 110 province FEs. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Table 1.3: Instrument validity: effect of the instrument on pre-shock (2005-08) change in stock, cumulative entry and exit of firms

	(1)	(2)	(3)
	All Firms	All Firms	All Firms
	Δ Stock	$\sum$ Births	$\sum$ Deaths
VARIABLES	2005-08	2005-08	2005-08
Pull IV	-0.046	-0.008	0.038
	(0.056)	(0.054)	(0.046)
Observations	686	686	686
R-squared	0.161	0.627	0.181
Avg. Outcome	0.339	3.891	3.552
Mean Pull IV	0.046	0.046	0.046
S.d. Pull IV	0.049	0.049	0.049
Controls	X	X	X
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in firm stock (1), cumulative firm entry (2) and exit (3) between 2005-2008 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$, and normalized to have mean zero and unit variance. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Standard errors are clustered at the province level (110 clusters).

Table 1.4: Pull IV diagnostics

Panel A: Negative and positive weights

	Sum	Mean	Share
$\hat{\alpha}_c \leq 0$	-0.007	0	0.007
$\hat{\alpha}_c > 0$	1.007	0.009	0.993

Panel B: Correlations

	$\hat{\alpha}_c$	G_c	$\hat{\beta}_c$	$\hat{F}_c$	$Var(NTWK_c)$
$\hat{\alpha}_c$	1.0000				
G_c	-0.0476	1.0000			
$\hat{\beta}_c$	0.0043	0.0694	1.0000		
$\hat{F}_c$	0.0140	0.0299	0.0036	1.0000	
$Var(NTWK_c)$	0.8430	-0.0919	0.0034	0.0046	1.0000

Panel C: Top 5 destination countries

	$\hat{\alpha}_c$	G_c	$\hat{\beta}_c$	$\hat{F}_c$	95% C.I.
Germany	0.452	1.075	-0.388	12.98	(-2.10, -0.20)
Switzerland	0.277	1.01	-0.202	16.44	(-0.40, 0.00)
France	0.075	1.007	-0.364	3.56	($-\infty$, 1.10)
Australia	0.039	1.064	-0.197	0.60	($-\infty$, ∞)
Belgium	0.029	1.005	-0.081	0.84	($-\infty$, ∞)

Panel D: OLS and IV estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Share	Share	Shares	Shares	All	Pull
	OLS	DE	CH	DE & CH	Top 5	shares	IV
$\hat{\beta}$	0.037	-0.388	-0.202	-0.311	-0.311	-0.704	-0.414
	(0.047)	(0.231)	(0.111)	(0.139)	(0.142)	(0.476)	(0.155)
$\hat{F}$		12.976	16.442	16.448	6.868	411.213	14.851
Over ID				0.383	0.924	0.455	

Notes: The table reports the Pull IV diagnostics as suggested by Goldsmith-Pinkham, Sorkin, and Swift (2020). Panel A reports the sum, the mean and the share of negative and positive Rotemberg weights $\hat{\alpha}_c$. Panel B reports correlations between the weights ($\hat{\alpha}_c$), the 2008-2015 destination country GDP growth (G_c), the just-identified coefficients ($\hat{\beta}_c$), the first stage F-statistics for the just-identified instruments ($\hat{F}_c$) and the variance in the emigrant networks across destination countries ($Var(NTWK_c)$). Panel C reports the top five destination countries according to the Rotemberg weights. The 95% CI are the weak instrument robust confidence intervals obtained with the Chernozhukov and Hansen (2008) method with a range from -10 to 10 ($(-\infty, \infty)$ indicates that the CI is undefined). The coefficients $\hat{\beta}_c$ are based on the regression of Table 1.7, column (1), where the outcome is the change 2008-2015 in the stock of firms per capita, and control variables include LLM value added per capita and unemployment rate in 2005 as well as 110 province FEs. We computed the Rotemberg decomposition using Goldsmith-Pinkham, Sorkin, and Swift (2020)'s Stata package. Panel D shows our main coefficient of interest estimated using different IVs: 2SLS estimates using the share to Germany ("Sh DE") and to Switzerland ("Sh CH") as instruments, both separately and jointly ("Sh DE & CH"), as well as LIML estimates using the top-5 shares of Panel C jointly ("Top 5") and all shares jointly ("All shares"). We report the first stage F-statistic and the p-value of the Sargan over-identification statistic when appropriate. The "OLS" and "Pull IV" columns show the coefficients of Tables 1.2 and 1.7, column (1), for comparison.

Table 1.5: Relationship between destination countries' emigration networks and pre-period LLM characteristics

VARIABLES	(1) Share to Germany	(2) Share to Switzerland	(3) Share to France	(4) Share to Australia	(5) Share to Belgium
Δ Stock	-0.002	-0.005	0.002	-0.001	0.002
	(0.006)	(0.005)	(0.002)	(0.002)	(0.004)
$\sum$ Births	0.169	-0.094	-0.085	0.004	0.002
	(0.160)	(0.078)	(0.084)	(0.060)	(0.093)
$\sum$ Deaths	0.005	-0.001	-0.010	0.003	-0.002
	(0.007)	(0.005)	(0.008)	(0.005)	(0.006)
Unemp Rate 2005	0.059	-0.053	-0.006	-0.053	0.004
	(0.078)	(0.046)	(0.028)	(0.037)	(0.046)
GDP PC 2005	-0.014	-0.017	-0.012	-0.003	-0.005
	(0.010)	(0.012)	(0.009)	(0.002)	(0.004)
Observations	683	683	683	628	660
Avg. Outcome	0.010	0.008	0.006	0.002	0.004
Controls	X	X	X	X	X
Province FE	X	X	X	X	X

Notes: OLS estimates, each coefficient is from a separate regression. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the share of pre-2000 emigrants to each of the 5 top destination countries described in each column, relative to the LLM population in 2000. The independent variables are the main LLMs observable characteristics, namely the change in stock, cumulative entry and exit of firms between 2005-2008, unemployment rate and value added per capita in 100,000 euros in 2005. All regressions include 110 province FEs. Standard errors are clustered at the province level (110 clusters).

Table 1.6: First stage regressions

	(1)	(2)	(3)
VARIABLES	Emig Rate	Emig Rate	Emig Rate
Pull IV	0.430	0.442	0.435
	(0.081)	(0.081)	(0.113)
Unemp Rate 2005	-3.831	0.936	3.912
	(1.720)	(2.258)	(3.336)
GDP PC 2005	1.020	1.183	1.338
	(0.271)	(0.199)	(0.368)
Observations	686	686	686
R-squared	0.138	0.245	0.400
F-excl. instrument	28.311	29.564	14.851
Mean Emig Rate	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696
Mean Pull IV	0.046	0.046	0.046
S.d. Pull IV	0.049	0.049	0.049
FE	-	Region	Province

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the cumulative emigration rate between 2008 and 2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$, and normalized to have mean zero and unit variance. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level. Column (1) includes no fixed effects while columns (2) and (3) include region (20) and province (110) FEs respectively. Standard errors are clustered at the province level (110 clusters).

Table 1.7: Effect of emigration rates on change in stock, cumulative entry and exit of firms

	(1)	(2)	(3)
	All Firms	All Firms	All Firms
	Δ Stock	$\sum$ Births	$\sum$ Deaths
VARIABLES	2008-15	2008-15	2008-15
Emig Rate	-0.414	-0.432	-0.018
	(0.155)	(0.196)	(0.189)
Observations	686	686	686
R-squared	0.175	0.527	0.241
F-excl. instr.	14.851	14.851	14.851
Avg. Baseline Outcome	0.790	9.078	8.288
Mean Emig Rate	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696
Controls	X	X	X
Province FE	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in firm stock (1), cumulative firm entry (2) and exit (3) between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Table 1.8: Effect of emigration rates on young-owned firms and innovative start-ups

VARIABLES	(1) Young Firms Δ Stock 2008-15	(2) Young Firms $\sum$ Births 2008-15	(3) Young Firms $\sum$ Deaths 2008-15	(4) Start-Ups $\sum$ Births 2008-15
Emig Rate	-0.242	-0.234	0.008	-0.004
	(0.114)	(0.133)	(0.161)	(0.001)
Observations	686	686	686	686
R-squared	0.342	0.476	0.471	0.326
F-excl. instr.	14.851	14.851	14.851	14.851
Avg. Baseline Outcome	-0.316	6.493	6.809	0.010
Mean Emig Rate	2.648	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696	1.696
Controls	X	X	X	X
Province FE	X	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in stock (1), cumulative entry (2) and exit (3) of firms owned and managed by under 45 ("Young firms") between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. In column (4), the dependent variable is the number of innovative start-ups created between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) and multiplied by 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit of firms owned and managed by under 45 in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years, while in column (4) is the average outcome in the 2008-2015 period. Standard errors are clustered at the province level (110 clusters).

Table 1.9: Effect of emigration rates on change in LLM employment

VARIABLES	(1) Δ Employees 2008-15	(2) Δ Emp/Pop 2008-15	(3) Δ Avg. Size 2008-15	(4) Δ Wage Bill 2008-15
Emig Rate	-0.046	-0.024	-0.015	-0.019
	(0.020)	(0.020)	(0.025)	(0.022)
Observations	686	686	686	686
R-squared	0.194	0.212	0.241	0.264
F-excl. instr.	14.851	14.851	14.851	14.851
Avg. Outcome 2005	16709.0	0.3	5.5	348.6
Mean Emig Rate	2.648	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696	1.696
Controls	X	X	X	X
Province FE	X	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in LLM employment (1), employment to population ratio (2), average firm size (3) and total wage bill in 100,000 euros (4) between 2008-2015, as a fraction of each outcome in 2005. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Standard errors are clustered at the province level (110 clusters). Data sources: INPS (2017) and Istat (2017).

Table 1.10: Effect of emigration rates on change in LLM skills

VARIABLES	(1) Δ Blue Coll 2008-15	(2) Δ White Coll 2008-15	(3) Δ Managers 2008-15
Emig Rate	-0.018	-0.058	-1.090
	(0.027)	(0.028)	(1.043)
Observations	686	686	584
R-squared	0.199	0.232	0.188
F-excl. instr.	14.851	14.851	6.432
Avg. Outcome 2005	8950.1	6737.4	191.7
Mean Emig Rate	2.648	2.648	2.544
S.d. Emig Rate	1.696	1.696	1.369
Controls	X	X	X
Province FE	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in LLM employment of blue collar workers (1), white collars (2) and managers (3) between 2008-2015, as a fraction of each outcome in 2005. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Standard errors are clustered at the province level (110 clusters).

Table 1.11: Robustness checks

VARIABLES	(1) Controlling for Immigration $\sum$ Births	(2) Excluding border provinces $\sum$ Births	(3) Controlling for tradable share $\sum$ Births
Emig Rate	-0.388	-0.367	-0.493
	(0.178)	(0.195)	(0.221)
Immig Rate 05-08	0.879		
	(0.106)		
Tradable sh. 2005			-4.092
			(2.732)
Observations	686	590	686
R-squared	0.614	0.508	0.517
F-excl. instr.	15.550	13.039	13.692
Avg. Baseline Outcome	9.078	9.166	9.078
Mean Emig Rate	2.648	2.457	2.648
S.d. Emig Rate	1.696	1.554	1.696
Controls	X	X	X
FE	Province	Province	Province

Notes: 2SLS estimates. In columns (1) and (3), the sample is composed of 686 local labor markets (LLMs), while in column (2) the sample is composed of 590 LLMs, excluding those in the provinces sharing a border with foreign countries. The dependent variable is the cumulative firm entry between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) and multiplied by 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. In column (1), we also include the cumulative immigration rate between 2005-2008 as a percentage of LLM population 25-64 years old (average 2005-2008). In column (3) we also control for the share of LLM firms in tradable sectors in 2005. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Standard errors are clustered at the province level (110 clusters).

Appendices

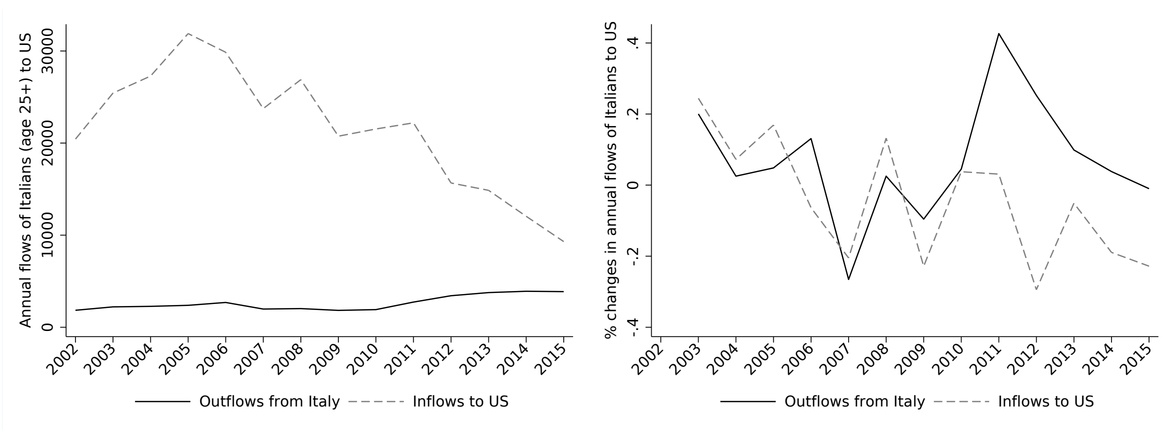
A.I Additional figures

In this section we present the additional figures discussed in the main text. Figure 1.8 shows that the outflows of Italians towards the US are underestimated if we compare Italian administrative and US Census Bureau American Community Survey (ACS) data (Ruggles et al. 2019). The year-to-year changes in the two datasets follow closely each other up to the most recent years (likely due to the inability of the ACS survey to capture recent immigration from Italy).

Figure 1.9 shows the pre-trends and the post-emigration wave change in the stock of firms in a reduced-form event-study graph. The estimates suggest that the largest decline in the number of firms occurs in the years after 2011, consistent with what we observe from the raw data in Figure 1.6.

Finally, Figure 1.10 shows how the correlation between the IV and the emigration rate, once we partial out the control variables (unemployment rate and value added per capita) and the 110 province FEs, is not driven by the presence of outliers.

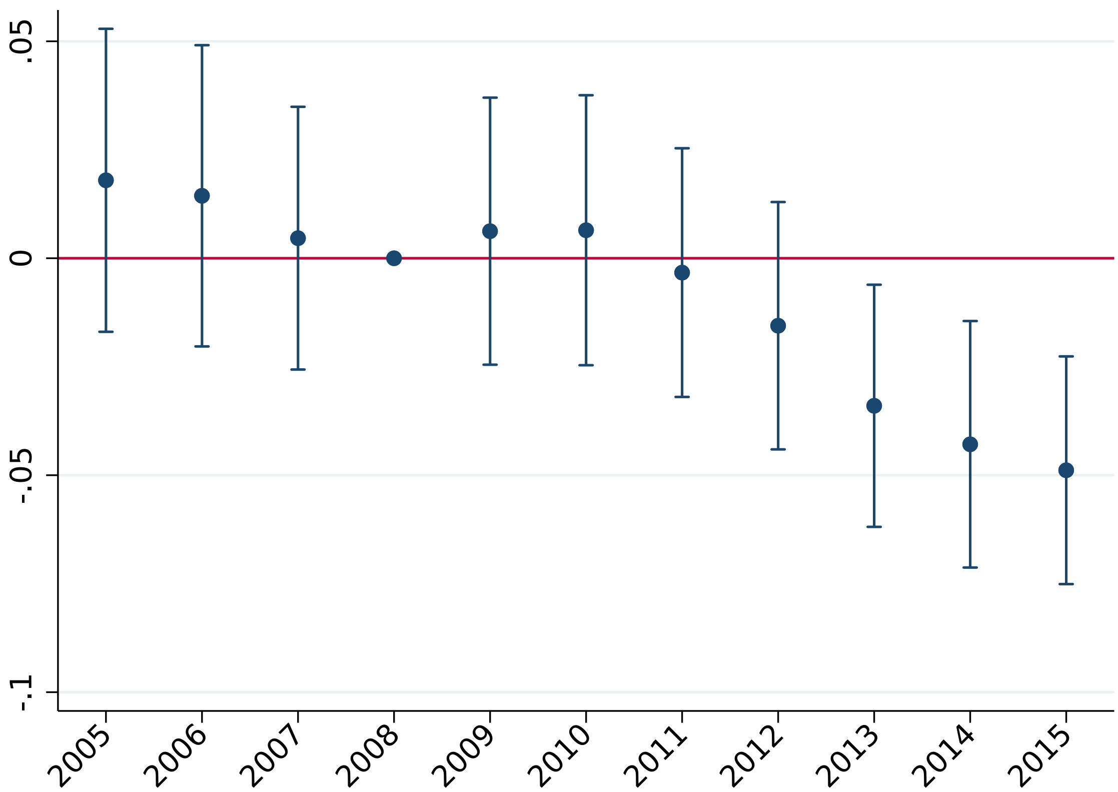
Figure 1.8: Recorded Emigration and Inflows of Italians to the US



(a) Annual emigration of Italians to US (b) % changes in annual flows of Italians to US

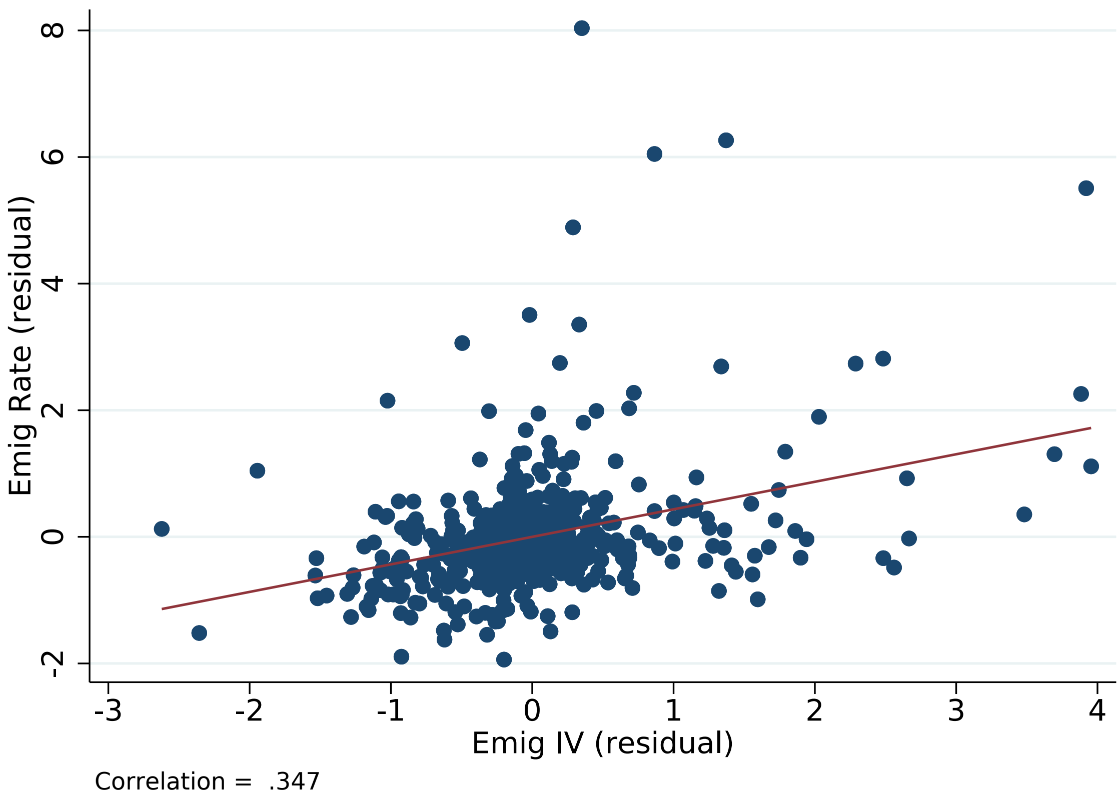
Notes: In Figure (a), the black solid line shows the annual outflows of Italians to the United States recorded in the AIRE-Istat data, while the grey dashed line shows the corresponding annual inflows of Italians to the US according to the American Community Survey (ACS) data. Figure (b) shows the percentage changes in the annual flows from the two data sources.

Figure 1.9: Event study: effect of the instrument interacted with years fixed effects on stock of firms, 2005-2015



Notes: The graph plots the coefficients γ_τ of the interaction between the instrument and year fixed effects from the regression: $y_{l,t} = \alpha + \beta Pull_l + \sum_{\tau \neq 2008} \gamma_\tau Pull_l * \mathbb{I}(\tau = t) + \xi X_{l,2005} + \phi_p + \lambda_t + \psi_{p,t} + \varepsilon_{l,t}$, where the outcome $y_{l,t}$ is the stock of firms in LLM l in year t as a fraction of LLM population 25-64 years old (average 2005-2008), $Pull_l$ is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $X_{l,2005}$ include unemployment rate and value added per capita in 100,000 euros in 2005, ϕ_p are province fixed effects, λ_t are year FEs and $\psi_{p,t}$ are province-by-year FEs. Standard errors are clustered at the province-by-year level and bars show confidence intervals at the 5-percent level.

Figure 1.10: First stage scatter plot correlation



Notes: The scatter plot shows the correlation between the Pull IV and the Emigration Rate after partialling out the control variables (unemployment rate and value added per capita) as well as the 110 province FEs. Both variables are normalized to have mean zero and unit variance.

A.II Accounting for under-registration in AIRE-Istat emigration data

In this section we validate the 2.6 adjustment factor used in the empirical analysis. To circumvent the issue that not all Italian emigrants report their change of residence by registering in the AIRE, we compare yearly outflows of Italians recorded by Istat-AIRE to the yearly inflows of Italians to three among the top-5 destination countries of Italian emigrants, namely the UK, Switzerland and the United States. For the UK, we obtained administrative data from the UK Social Security Registry based on "National Insurance number allocations to adult overseas nationals entering the UK" (NINo 2018), which include all individuals applying to work in the UK or to claim any benefit or tax credit. For Switzerland, we use Federal Statistical Office (BFS 2018) administrative data based on the migration registry (PETRA-STATPOP), which include only permanent residents ("ständige Wohnbevölkerung"). For the US, we use weighted survey data from the American Community Survey (ACS) (Ruggles et al. 2019), using information on the year of arrival and country of birth.

Table 1.12 compares the emigration flows registered in the Italian AIRE-Istat data to the immigration flows registered by each foreign source respectively. The variable *Factor* shows the ratio between the immigration and emigration flow in each year. The data shows that emigration flows are systematically under-reported in the AIRE-Istat data, in almost every year and for all the three countries considered. In the last two columns, we construct a weighted average of the correction factors (weighted by the emigration flows). If we include the US (penultimate column), the average correction factor for the period 2009-15 is about 3.48. However, as the ACS data are survey-based and thus less reliable than the administrative sources from UK and Switzerland, in the last column we only consider the two latter countries, for which the average correction factor ranges between 2.62 and 3.23 over the period 2009-15 and is 2.87 on average. Based on these results, we use the minimum value of the average correction factor, 2.6, to adjust upwards the emigration flows between 2009-15 throughout our empirical analysis.

Table 1.12: Correction factor based on destination country data

	United Kingdom			Swizterland			United States			Emig-weighted	Emig-weighted
Year	Emig	Immig	Factor	Emig	Immig	Factor	Emig	Immig	Factor	Avg Factor	Avg Factor - No US
2002	2400	7717	3.22	4587	5961	1.30	1846	20439	11.07	3.86	1.96
2003	2740	8122	2.96	6021	5820	0.97	2216	25435	11.48	3.59	1.59
2004	3097	8180	2.64	5068	5859	1.16	2272	27282	12.01	3.96	1.72
2005	4003	10361	2.59	4911	5622	1.14	2382	31892	13.39	4.24	1.79
2006	4561	11048	2.42	5271	5689	1.08	2694	29865	11.09	3.72	1.70
2007	5033	15735	3.13	3647	8540	2.34	1979	23746	12.00	4.51	2.80
2008	5474	16460	3.01	4165	10025	2.41	2029	26887	13.25	4.57	2.75
2009	**4981**	**16876**	**3.39**	**4097**	**8668**	**2.12**	**1835**	**20749**	**11.31**	**4.24**	**2.81**
2010	**5167**	**18461**	**3.57**	**4522**	**10226**	**2.26**	**1918**	**21532**	**11.23**	**4.33**	**2.96**
2011	**5317**	**24882**	**4.68**	**5669**	**10651**	**1.88**	**2736**	**22200**	**8.11**	**4.21**	**3.23**
2012	**7293**	**26599**	**3.65**	**8238**	**14098**	**1.71**	**3427**	**15668**	**4.57**	**2.97**	**2.62**
2013	**12756**	**44120**	**3.46**	**9663**	**17662**	**1.83**	**3766**	**14870**	**3.95**	**2.93**	**2.76**
2014	**13332**	**51210**	**3.84**	**10151**	**19006**	**1.87**	**3910**	**12055**	**3.08**	**3.00**	**2.99**
2015	**17248**	**58653**	**3.40**	**11227**	**18894**	**1.68**	**3871**	**9306**	**2.40**	**2.69**	**2.72**
								Average	**2009-15**	**3.48**	**2.87**

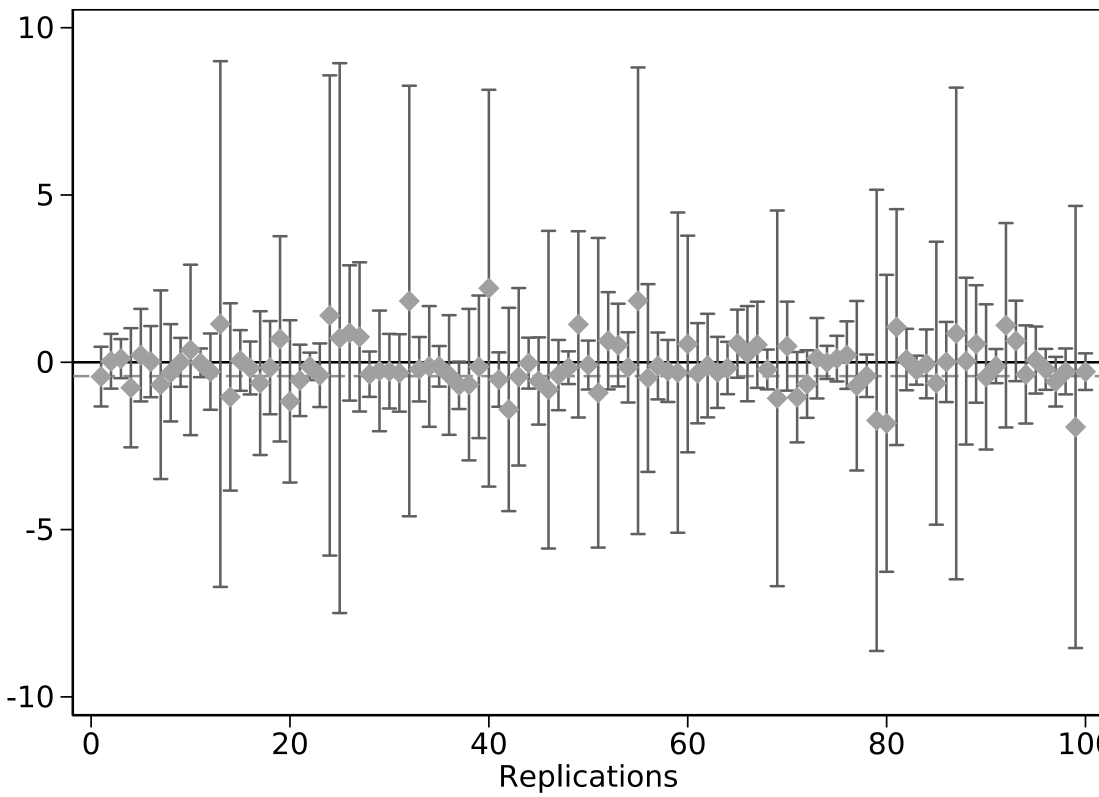
A.III Instrument validity: Additional checks

In this Section we report four figures that corroborate our identification strategy and inference. First, as reported in the main text, we cluster the standard errors at the province level: independence across labour markets is needed under an identifying assumption based on the exogeneity of the emigration networks and our identification is effectively within provinces (Adão, Kolesár, and Morales 2019; Goldsmith-Pinkham, Sorkin, and Swift 2020). To validate this choice, we follow similar exercises proposed by Adão, Kolesár, and Morales (2019) and implemented in Fouka, Mazumder, and Tabellini (2020), and perform two placebo exercises. We replace the shifters and the shares, respectively, with random numbers extracted from $N(0,5)$. The two exercises confirm that the clustered standard errors are valid and, if anything, too conservative: in the case of shifters, only 0.4 percent wrongly reject the null hypothesis of $\beta = 0$ at the 10 percent level (Figure 1.11); in the case of shares randomization, none of the 500 replications is significantly different from zero (Figure 1.12).

Then, we perform two additional exercises to validate the identification strategy. Differently from assigning a random shift as above, which effectively does not allow to identify any effect, here we first hold the emigration networks (the shares) fixed and we permute the GDP changes (the shifts); then, we do the opposite (hold the shares fixed and permute the shifters).[41] Considering that the country-of-emigration shares drive most of the identifying variation (Table 1.4), the random permutation of the shifts still allow us to identify our results (Figure 1.13). To the contrary, randomly permuting the shares while keeping the right shifts does not allow to identify any effect, consistent with our identifying assumption (Figure 1.14).

Finally, Tables 1.13, 1.14 and 1.15 show the pre-trends as discussed in Section 1.3.2 of the main text.

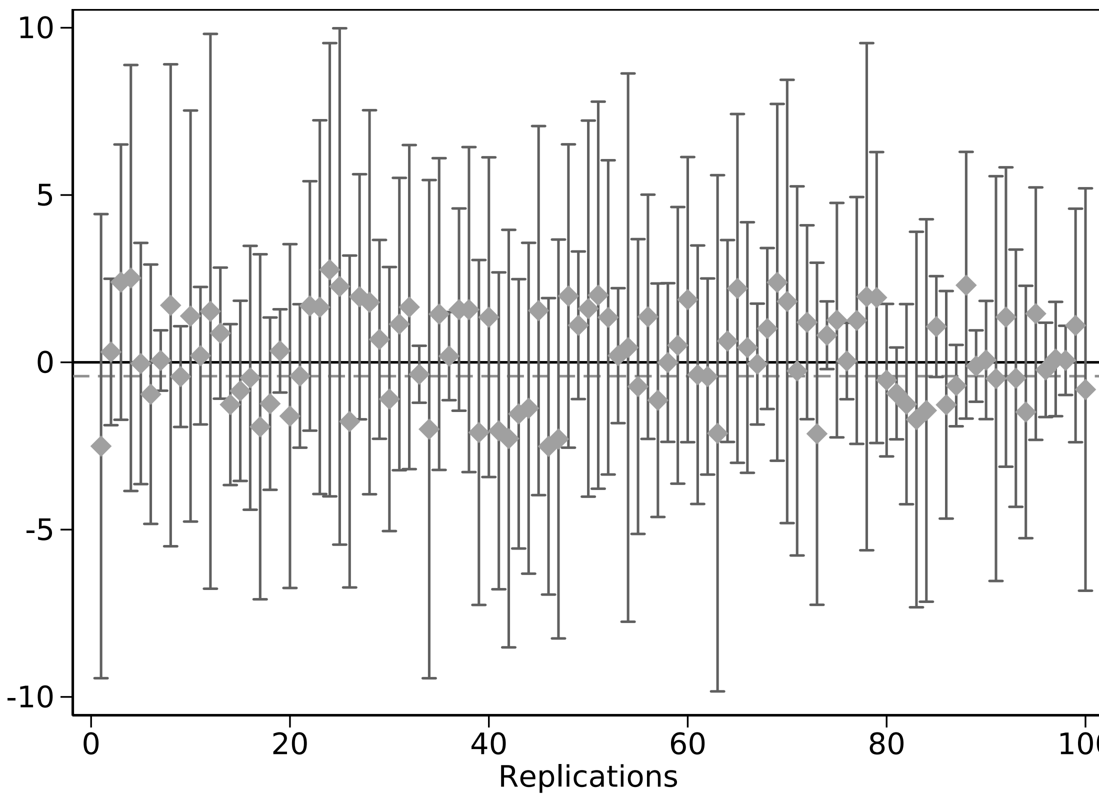
Figure 1.11: Randomization of the Shifter components



Notes: The graph reports the emigration effect estimates obtained in 100 different random draws of the GDP shifters using our baseline specification (the exercise is based on 500 replications, but for visualization clarity only 100 are reported in the graph). The estimated coefficients are significant 0.4 percent of the times at 10 percent level, and never significant at 5 percent level.

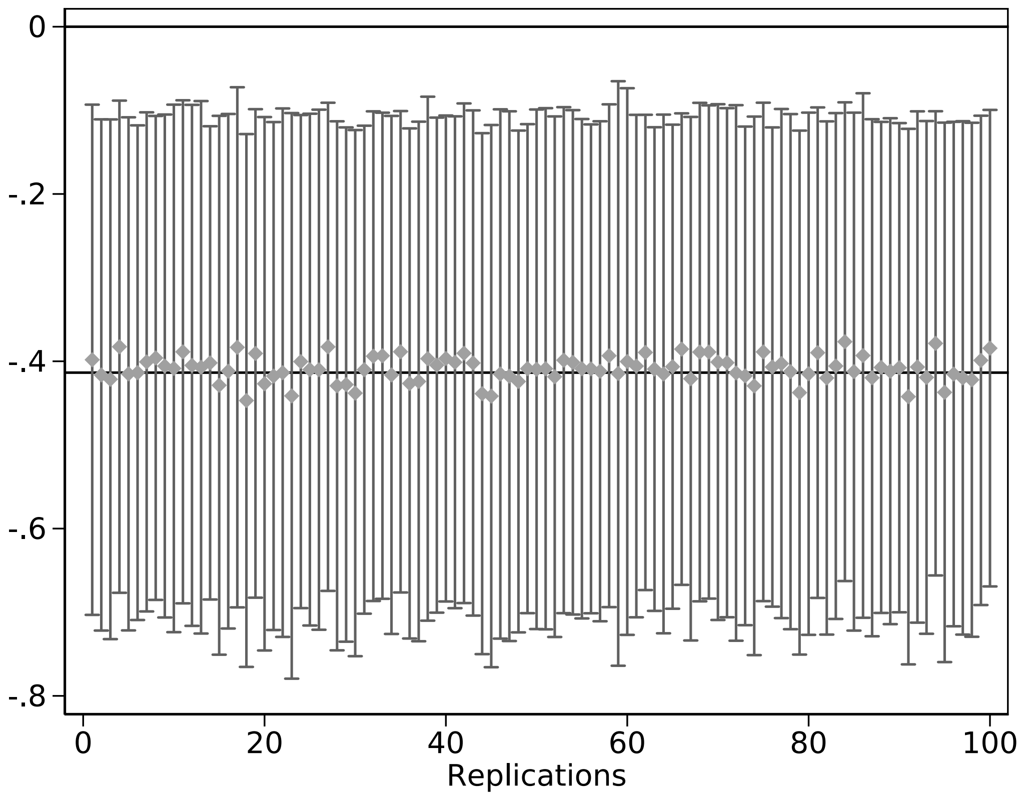
[41] In practice, we randomly reshuffle the observed shifts 500 different times, run our main specification, retrieve and plot the estimate of the emigration effect and its confidence interval (for visualization clarity, the graph reports only the first 100 replications).

Figure 1.12: Randomization of the Share Components



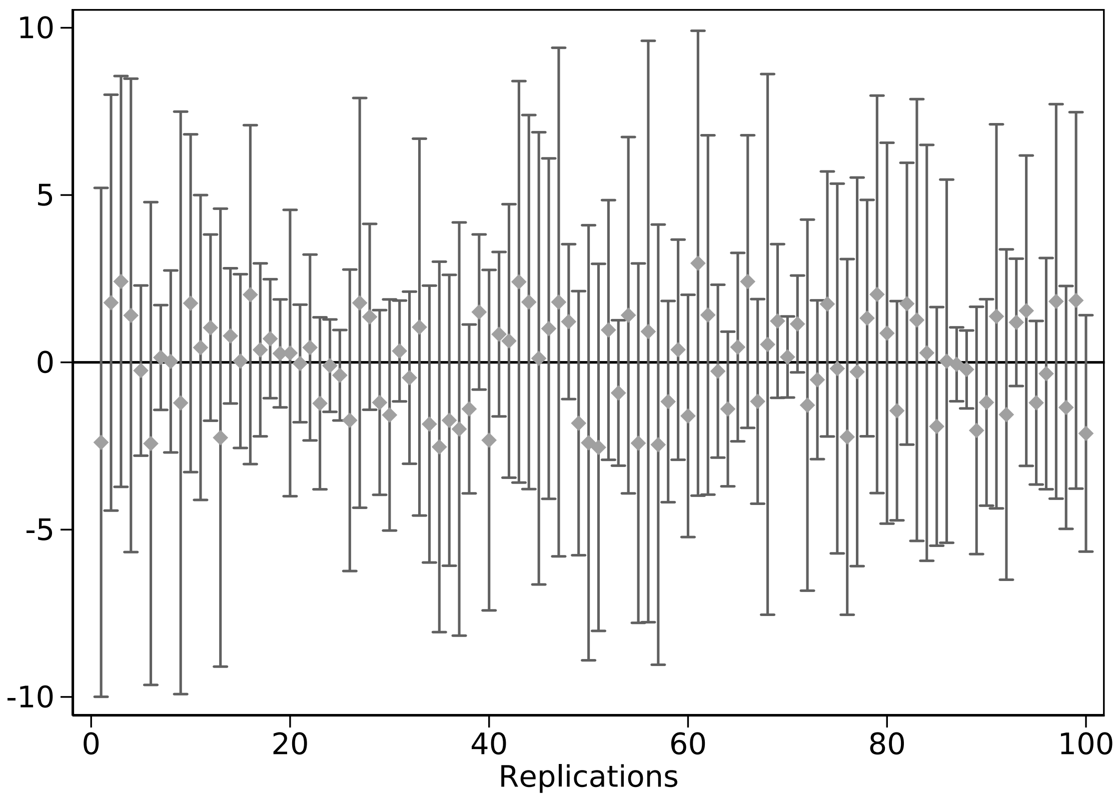
Notes: The graph reports the emigration effect estimates obtained in 100 different random draws of the emigration network shares using our baseline specification (the exercise is based on 500 replications, but for visualization clarity only 100 are reported in the graph). The estimated coefficients are never significant at 10 percent level.

Figure 1.13: Permutation of the Shifter components



Notes: The graph reports the emigration effect estimates obtained in 100 different random permutations of the GDP shifters using our baseline specification (the exercise is based on 500 replications, but for visualization clarity only 100 are reported in the graph). The average estimated coefficient is -0.43 and the estimated coefficients from all replications are significant at 5 percent level.

Figure 1.14: Permutation of the share components



Notes: The graph reports the emigration effect estimates obtained in 100 different random permutations of the emigration network shares using our baseline specification (the exercise is based on 500 replications, but for visualization clarity only 100 are reported in the graph). The average estimated coefficient is -0.989 and the estimated coefficients are significant 0.2% of the times at 5 percent level, and never at the 1 percent level.

Table 1.13: Instrument validity: effect of the instrument on pre-shock change in stock and flows of Young-owned firms (2005-08)

VARIABLES	(1) Young Firms Δ Stock 2005-08	(2) Young Firms $\sum$ Births 2005-08	(3) Young Firms $\sum$ Deaths 2005-08
Pull IV	0.015	-0.002	-0.017
	(0.049)	(0.048)	(0.028)
Observations	686	686	686
R-squared	0.222	0.530	0.340
Avg. Outcome	-0.135	2.783	2.918
Mean Pull IV	0.046	0.046	0.046
S.d. Pull IV	0.049	0.049	0.049
Controls	X	X	X
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in stock (1), cumulative entry (2) and exit (3) of firms owned and managed by under 45 ("Young firms") between 2005-2008 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$, and normalized to have mean zero and unit variance. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Standard errors are clustered at the province level (110 clusters).

Table 1.14: Instrument validity check: effect of the instrument on pre-shock change in LLM employment (2005-08)

VARIABLES	(1) Δ Employees 2005-08	(2) Δ Emp/Pop 2005-08	(3) Δ Avg. Size 2005-08	(4) Δ Wage Bill 2005-08
Pull IV	-0.007	-0.001	-0.008	-0.009
	(0.010)	(0.010)	(0.009)	(0.011)
Observations	686	686	686	686
R-squared	0.282	0.260	0.177	0.194
Avg. Outcome 2005	16709.0	0.3	5.5	348.6
Mean Pull IV	0.046	0.046	0.046	0.046
S.d. Pull IV	0.049	0.049	0.049	0.049
Controls	X	X	X	X
Province FE	X	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in LLM employment (1), employment to population ratio (2), average firm size (3) and total wage bill in 100,000 euros (4) between 2005-2008, as a fraction of each outcome in 2005. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$, and normalized to have mean zero and unit variance. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Standard errors are clustered at the province level (110 clusters).

Table 1.15: Instrument validity check: effect of the instrument on pre-shock change in LLM skills (2005-08)

VARIABLES	(1) Δ Blue Coll 2005-08	(2) Δ White Coll 2005-08	(3) Δ Managers 2005-08
Pull IV	0.002	-0.011	0.293
	(0.011)	(0.012)	(0.286)
Observations	686	686	584
R-squared	0.323	0.137	0.135
Avg. Outcome 2005	8950.1	6737.4	191.7
Mean Pull IV	0.046	0.046	0.039
S.d. Pull IV	0.049	0.049	0.040
Controls	X	X	X
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in LLM employment of blue collar workers (1), white collars (2) and managers (3) between 2005-2008, as a fraction of each outcome in 2005. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$, and normalized to have mean zero and unit variance. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Standard errors are clustered at the province level (110 clusters).

A.IV IV diagnostics for the consulate-based IV

The following Tables, 1.16 and 1.17, replicate the main tests proposed by Goldsmith-Pinkham, Sorkin, and Swift (2020) for the IV based on destination regions (Italian consulates abroad) rather than on the countries. Similarly to what shown in Table 1.4, Table 1.16 shows that the cross-sectional components of the pull emigration instrumental variable is driven by networks of Italian emigrants towards German and Swiss regions. Interestingly, the estimated coefficients of Stuttgart/Friburg and Dortmund/Koln, that alone make up about 40 percent of the IV variation, are close to each other (-0.498 and 0.355) and close to the main estimate at the consulate (-0.432) and country level (-0.433). Table 1.17 shows the correlations between the share of emigrants towards the most relevant regions and the main labor market characteristics: we fail to find statistically significant correlations with observable LLM characteristics, similarly to what shown in the country level analysis in the main text.

Table 1.16: Pull IV diagnostics (destination regions IV)

Panel A: Negative and positive weights

	Sum	Mean	Share
$\hat{\alpha}_c \leq 0$	-0.007	0	0.007
$\hat{\alpha}_c > 0$	1.007	0.008	0.993

Panel B: Correlations

	$\hat{\alpha}_c$	G_c	$\hat{\beta}_c$	$\hat{F}_c$	$Var(NTWK_c)$
$\hat{\alpha}_c$	1.0000				
G_c	-0.0735	1.0000			
$\hat{\beta}_c$	0.0072	0.0694	1.0000		
$\hat{F}_c$	0.0150	0.0296	0.0050	1.0000	
$Var(NTWK_c)$	0.7561	-0.1236	0.0043	-0.0041	1.0000

Panel C: Top 5 destination regions

	$\hat{\alpha}_c$	G_c	$\hat{\beta}_c$	$\hat{F}_c$	95% C.I.
Stuttgart/Friburg	0.250	1.075	-0.491	7.33	(-3.20, -0.10)
Zurich	0.105	1.01	-0.145	12.63	(-0.30, 0.00)
Dortmund/Koln	0.090	1.075	-0.343	2.64	(-∞,∞)
Lugano	0.076	1.010	-0.009	6.45	(-0.40, 0.10)
France	0.075	1.007	-0.364	3.56	(-∞, 1.10)

Notes: The table reports the Pull IV diagnostics as suggested by Goldsmith-Pinkham, Sorkin, and Swift (2020). Panel A reports the sum, the mean and the share of negative and positive Rotemberg weights $\hat{\alpha}_c$. Panel B reports correlations between the weights ($\hat{\alpha}_c$), the 2008-2015 destination region/country GDP growth (G_c), the just-identified coefficients ($\hat{\beta}_c$), the first stage F-statistics for the just-identified instruments ($\hat{F}_c$) and the variance in the emigrant networks across destination region/country ($Var(NTWK_c)$). Panel C reports the top five destination regions according to the Rotemberg weights. The 95% CI are the weak instrument robust confidence intervals obtained with the Chernozhukov and Hansen (2008) method with a range from -10 to 10 ($(-\infty, \infty)$ indicates that the CI is undefined). The coefficients $\hat{\beta}_c$ are based on the regression of Table 1.7, column (1), where the outcome is the change 2008-2015 in the stock of firms per capita, and control variables include LLM value added per capita and unemployment rate in 2005 as well as 110 province FEs. We computed the Rotemberg decomposition using Goldsmith-Pinkham, Sorkin, and Swift (2020)'s Stata package.

Table 1.17: Relationship between destination regions' emigration networks and pre-period LLMs characteristics

VARIABLES	(1) Share to Stuttgart/Friburg	(2) Share to Zurich	(3) Share to Dortmund/Koln	(4) Share to Lugano	(5) Share to France
Δ Stock	-0.002	-0.001	0.001	-0.001	0.002
	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
$\sum$ Births	0.139	-0.039	0.043	-0.054	-0.085
	(0.100)	(0.031)	(0.058)	(0.022)	(0.084)
$\sum$ Deaths	0.006	-0.001	0.000	-0.001	-0.010
	(0.007)	(0.002)	(0.002)	(0.002)	(0.008)
Unemp Rate 2005	0.018	-0.011	0.035	-0.012	-0.006
	(0.049)	(0.020)	(0.036)	(0.009)	(0.028)
GDP PC 2005	-0.006	-0.005	-0.004	-0.005	-0.012
	(0.004)	(0.003)	(0.003)	(0.003)	(0.009)
Observations	653	666	651	645	683
Avg. Outcome	0.004	0.003	0.003	0.002	0.006
Controls	X	X	X	X	X
Province FE	X	X	X	X	X

Notes: OLS estimates, each coefficient is from a separate regression. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the share of pre-2000 emigrants to each of the 5 top destination regions described in each column, relative to the LLM population in 2000. The independent variables are the main LLMs observable characteristics, namely the change in stock, cumulative entry and exit of firms between 2005-2008, unemployment rate and value added per capita in 100,000 euros in 2005. All regressions include 110 province FEs. Standard errors are clustered at the province level (110 clusters).

A.V Additional robustness checks

One may be concerned that our instrument is correlated with internal migration flows. While these should not be correlated with pull factors from abroad, the network of emigrants may be correlated with the internal flows and with local push factors. For instance, LLMs with high emigration rates to foreign countries could also exhibit substantial emigration to other Italian LLMs, and the latter may reduce firm creation, violating the exclusion restriction. We thus test whether our estimates are robust to this potential threat. In Table 1.18, columns (1) and (2), we report the results of a placebo first stage regression where we regress internal migration outflows and inflows on our emigration Pull IV. The effects are not statistically significant, suggesting that the instrument does not predict internal emigration to or immigration from other LLMs in Italy. In column (3), we test whether there is a direct substitution effect by regressing foreign immigration inflows on the instrument. Reassuringly, the instrument does not have a statistically significant effect on foreign immigration flows. Furthermore, as shown in Table 1.11, our main estimates are robust to the inclusion of immigration as a control variable.

Another concern is that our estimates may be capturing the effect of trade linkages, which may be correlated to migration flows. For this reason, in Table 1.19 we report the results of our main regression on firm entry by firm creation between tradable and non-tradable sectors. The largest impact of emigrants on firm creation is for non-tradable sector firms. This indicates that the emigration flows we are analyzing do not seem to be particularly linked to international trade activity.

In Tables 1.20 and 1.21, we estimate the first stage and the main specification using the 1992 (rather than 2000) emigration shares when constructing the IV. Results are very similar to those of Tables 1.6 and 1.7 albeit less precise, as mentioned in Section 1.3.4. In Tables 1.22, 1.23, 1.24 and 1.25, we include the lag of the outcome variables among the set of controls. In all cases the main results continue to hold. In Tables 1.26, 1.27 and 1.28, we show the results of regressing each outcome of Table 1.7 with different sets of controls and fixed effects: no controls in columns (1)-(3) and controlling for LLM unemployment rate and GDP per capita in 2005 in columns (4)-(6), and no fixed effects in columns (1) and (4), 20 regions fixed effects in columns (2) and (5) and 110 province fixed effects in columns (3) and (6). All results are qualitatively similar across the different specifications as long as we include at least some controls or fixed effects.

Table 1.18: Placebo first stage regression on internal migration flows and immigration

	(1)	(2)	(3)
VARIABLES	Internal Emig	Internal Immig	Immig Rate 05-08
Pull IV	0.078	-0.087	-0.022
	(0.069)	(0.052)	(0.040)
Observations	686	686	686
R-squared	0.405	0.451	0.710
F-excl. instrument	1.249	2.839	0.306
Mean Outcome	10.084	9.166	2.222
S.d. Outcome	2.377	3.233	1.255
Mean Pull IV	0.046	0.046	0.046
S.d. Pull IV	0.049	0.049	0.049
Controls	X	X	X
FE	Province	Province	Province

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). In columns (1) and (2), the dependent variable is the cumulative emigration and immigration rate of Italian citizens 25-64 years old to and from different LLMs in Italy between 2008-2015 respectively, while in column 3 the dependent variable is the cumulative immigration rate of foreign citizens 25-64 years old from abroad between 2005-2008. All the outcomes are as a fraction of the LLM population 25-64 years old (average 2005-2008), and normalized to have mean zero and unit variance. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$, and normalized to have mean zero and unit variance. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level. Column (1) includes no fixed effects while columns (2) and (3) include region (20) and province (110) FEs respectively. Standard errors are clustered at the province level (110 clusters).

Table 1.19: Effect of emigration rates on cumulative firm entry, in tradable and non tradable sectors

VARIABLES	(1) Tradable $\sum$ Births 2008-15	(2) Non Tradable $\sum$ Births 2008-15
Emig Rate	-0.069	-0.363
	(0.028)	(0.178)
Observations	686	686
R-squared	0.547	0.513
F-excl. instr.	14.851	14.851
Avg. Baseline Outcome	0.664	8.414
Mean Emig Rate	2.648	2.648
S.d. Emig Rate	1.696	1.696
Controls	X	X
Province FE	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the cumulative firm entry in tradable (1) and nontradable (2) sectors between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Table 1.20: Robustness check: first stage regressions, IV based on 1992 shares

VARIABLES	(1) Emig Rate	(2) Emig Rate	(3) Emig Rate
Pull IV (1992 shares)	0.311	0.323	0.305
	(0.069)	(0.073)	(0.091)
Unemp Rate 2005	-2.210	1.996	3.926
	(1.710)	(2.420)	(3.449)
GDP PC 2005	0.863	1.040	1.201
	(0.255)	(0.176)	(0.310)
Observations	686	686	686
R-squared	0.079	0.194	0.362
F-excl. instrument	20.460	19.444	11.248
Mean Emig Rate	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696
Mean Pull IV	0.028	0.028	0.028
S.d. Pull IV	0.031	0.031	0.031
FE	-	Region	Province

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the cumulative emigration rate between 2008 and 2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The independent variable is the predicted emigration rate based on the shares of pre-1992 emigrants to different countries relative to the LLM population in 1992 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$, and normalized to have mean zero and unit variance. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level. Column (1) includes no fixed effects while columns (2) and (3) include region (20) and province (110) FEs respectively. Standard errors are clustered at the province level (110 clusters).

Table 1.21: Robustness check: effect of emigration rates on change in stock and flows of firms, IV based on 1992 shares

	(1)	(2)	(3)
	All Firms	All Firms	All Firms
	Δ Stock	$\sum$ Births	$\sum$ Deaths
VARIABLES	2008-15	2008-15	2008-15
Emig Rate	-0.507	-0.631	-0.124
	(0.208)	(0.302)	(0.258)
Observations	686	686	686
R-squared	0.170	0.478	0.241
F-excl. instr.	11.248	11.248	11.248
Avg. Baseline Outcome	0.790	9.078	8.288
Mean Emig Rate	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696
Controls	X	X	X
Province FE	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in firm stock (1), cumulative firm entry (2) and exit (3) between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-1992 emigrants to different countries relative to the LLM population in 1992 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Table 1.22: Robustness check: effect of emigration rates on change in stock and flows of firms (2008-2015) controlling for lagged outcomes (2005-08)

VARIABLES	(1) All Firms Δ Stock 2008-15	(2) All Firms $\sum$ Births 2008-15	(3) All Firms $\sum$ Deaths 2008-15
Emig Rate	-0.294	-0.403	-0.121
	(0.181)	(0.152)	(0.135)
Δ Stock	1.129		
	(0.007)		
$\sum$ Births		1.659	
		(0.078)	
$\sum$ Deaths			1.157
			(0.018)
Observations	686	686	686
R-squared	0.963	0.839	0.965
F-excl. instr.	14.853	14.865	14.843
Avg. Baseline Outcome	0.790	9.078	8.288
Mean Emig Rate	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696
Controls	X	X	X
Province FE	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in firm stock (1), cumulative firm entry (2) and exit (3) between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Additionally, we control for the change in stock, cumulative entry and exit of firms between 2005-2008 as a fraction of LLM population 25-64 years old (average 2005-2008). The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Table 1.23: Robustness check: effect of emigration rates on change in stock and flows of Young-owned firms (2008-15) controlling for lagged outcomes (2005-08)

	(1)	(2)	(3)
	Young Firms	Young Firms	Young Firms
	Δ Stock	$\sum$ Births	$\sum$ Deaths
VARIABLES	2008-15	2008-15	2008-15
Emig Rate	-0.267	-0.226	0.050
	(0.147)	(0.120)	(0.132)
Δ Stock	0.736		
	(0.065)		
$\sum$ Births		1.496	
		(0.083)	
$\sum$ Deaths			1.068
			(0.194)
Observations	686	686	686
R-squared	0.658	0.810	0.792
F-excl. instr.	15.018	14.928	14.856
Avg. Baseline Outcome	-0.316	6.493	6.809
Mean Emig Rate	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696
Controls	X	X	X
Province FE	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in stock (1), cumulative entry (2) and exit (3) of firms owned and managed by under 45 ("Young firms") between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Additionally, we control for the change in stock, cumulative entry and exit of firms owned and managed by under 45 between 2005-2008 as a fraction of LLM population 25-64 years old (average 2005-2008). The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Table 1.24: Robustness check: effect of emigration rates on LLM employment (2008-15) controlling for lagged outcome (2005-08)

VARIABLES	(1) Δ Employees 2008-15	(2) Δ Emp/Pop 2008-15	(3) Δ Avg. Size 2008-15	(4) Δ Wage Bill 2008-15
Emig Rate	-0.045	-0.024	-0.014	-0.018
	(0.020)	(0.021)	(0.025)	(0.022)
Δ Employees	0.061			
	(0.087)			
Δ Emp/Pop		0.120		
		(0.095)		
Δ Avg. Size			0.048	
			(0.115)	
Δ Wage Bill				0.018
				(0.063)
Observations	686	686	686	686
R-squared	0.197	0.216	0.242	0.264
F-excl. instr.	15.013	15.251	13.962	14.583
Avg. Outcome 2005	16709.0	0.3	5.5	348.6
Mean Emig Rate	2.648	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696	1.696
Controls	X	X	X	X
Province FE	X	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in LLM employment (1), employment to population ratio (2), average firm size (3) and total wage bill in 100,000 euros (4) between 2008-2015, as a fraction of each outcome in 2005. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Additionally, we control for the percentage change in each outcome between 2005-2008. Standard errors are clustered at the province level (110 clusters).

Table 1.25: Robustness check: effect of emigration rates on LLM skills (2008-15) controlling for lagged outcome (2005-08)

VARIABLES	(1) Δ Blue Coll 2008-15	(2) Δ White Coll 2008-15	(3) Δ Managers 2008-15
Emig Rate	-0.018	-0.055	-1.168
	(0.027)	(0.028)	(1.074)
Δ Blue Coll	0.002		
	(0.079)		
Δ White Coll		0.155	
		(0.103)	
Δ Managers			0.050
			(0.250)
Observations	686	686	584
R-squared	0.199	0.240	0.184
F-excl. instr.	15.293	14.746	6.361
Avg. Outcome 2005	8950.1	6737.4	191.7
Mean Emig Rate	2.648	2.648	2.544
S.d. Emig Rate	1.696	1.696	1.369
Controls	X	X	X
Province FE	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in LLM employment of blue collar workers (1), white collars (2) and managers (3) between 2008-2015, as a fraction of each outcome in 2005. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. Additionally, we control for the percentage change in each outcome between 2005-2008. Standard errors are clustered at the province level (110 clusters).

Table 1.26: Robustness check: effect of emigration rates on change in stock of firms, different controls and fixed effects

VARIABLES	(1) All Firms Δ Stock 2008-15	(2) All Firms Δ Stock 2008-15	(3) All Firms Δ Stock 2008-15	(4) All Firms Δ Stock 2008-15	(5) All Firms Δ Stock 2008-15	(6) All Firms Δ Stock 2008-15
Emig Rate	-0.384	-0.715	-0.573	-0.550	-0.647	-0.414
	(0.369)	(0.206)	(0.191)	(0.213)	(0.196)	(0.155)
Observations	686	686	686	686	686	686
R-squared		0.019	0.156		0.030	0.175
F-excl. instr.	26.223	31.313	13.469	28.311	29.564	14.851
Avg. Baseline Outcome	0.790	0.790	0.790	0.790	0.790	0.790
Mean Emig Rate	2.648	2.648	2.648	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696	1.696	1.696	1.696
Controls	-	-	-	X	X	X
FE	-	Region	Province	-	Region	Province

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). In columns (1)-(3) no controls are included, while in columns (4)-(6) we control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level. In columns (1) and (4) no fixed effects or controls are included, in columns (2) and (5) we include 20 region fixed effects, and in columns (3) and (6) we include 110 province fixed effects. The dependent variable is the change in firm stock between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Table 1.27: Robustness check: Effect of emigration rates on cumulative firm entry, different controls and fixed effects

VARIABLES	(1) All Firms $\sum$ Births 2008-15	(2) All Firms $\sum$ Births 2008-15	(3) All Firms $\sum$ Births 2008-15	(4) All Firms $\sum$ Births 2008-15	(5) All Firms $\sum$ Births 2008-15	(6) All Firms $\sum$ Births 2008-15
Emig Rate	-1.301	-0.631	-0.514	-0.769	-0.577	-0.432
	(0.368)	(0.242)	(0.222)	(0.263)	(0.240)	(0.196)
Observations	686	686	686	686	686	686
R-squared		0.207	0.493		0.240	0.527
F-excl. instr.	26.223	31.313	13.469	28.311	29.564	14.851
Avg. Baseline Outcome	9.078	9.078	9.078	9.078	9.078	9.078
Mean Emig Rate	2.648	2.648	2.648	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696	1.696	1.696	1.696
Controls	-	-	-	X	X	X
FE	-	Region	Province	-	Region	Province

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). In columns (1)-(3) no controls are included, while in columns (4)-(6) we control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level. In columns (1) and (4) no fixed effects or controls are included, in columns (2) and (5) we include 20 region fixed effects, and in columns (3) and (6) we include 110 province fixed effects. The dependent variable is the cumulative firm entry between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Table 1.28: Robustness check: effect of emigration rates on cumulative firm exit, different controls and fixed effects

VARIABLES	(1) All Firms $\sum$ Deaths 2008-15	(2) All Firms $\sum$ Deaths 2008-15	(3) All Firms $\sum$ Deaths 2008-15	(4) All Firms $\sum$ Deaths 2008-15	(5) All Firms $\sum$ Deaths 2008-15	(6) All Firms $\sum$ Deaths 2008-15
Emig Rate	-0.917	0.083	0.059	-0.219	0.070	-0.018
	(0.462)	(0.258)	(0.187)	(0.293)	(0.256)	(0.189)
Observations	686	686	686	686	686	686
R-squared		0.095	0.238	0.016	0.095	0.241
F-excl. instr.	26.223	31.313	13.469	28.311	29.564	14.851
Avg. Baseline Outcome	8.288	8.288	8.288	8.288	8.288	8.288
Mean Emig Rate	2.648	2.648	2.648	2.648	2.648	2.648
S.d. Emig Rate	1.696	1.696	1.696	1.696	1.696	1.696
Controls	-	-	-	X	X	X
FE	-	Region	Province	-	Region	Province

Notes: 2SLS estimates. The sample is composed of 686 local labor markets (LLMs). In columns (1)-(3) no controls are included, while in columns (4)-(6) we control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level. In columns (1) and (4) no fixed effects or controls are included, in columns (2) and (5) we include 20 region fixed effects, and in columns (3) and (6) we include 110 province fixed effects. The dependent variable is the cumulative firm exit between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the cumulative emigration rate between 2008-2015, i.e. the number of Italian citizens aged 25-64 migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), and normalized to have mean zero and unit variance. The instrument is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005-2008) as a fraction of population 25-64 years old in the LLM (average 2005-2008) times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

A.VI Additional tables

Table 1.29 presents reduced form estimates by regressing the 2008-2015 change in the main outcomes of interest directly on our IV. These reduced form results are a useful benchmark to compare the magnitude of the main effects with the pre-trends shown in Table 1.3.

Table 1.29: Reduced-form: Effect of the instrument on change in stock and flows of firms

VARIABLES	(1) All Firms Δ Stock 2008-15	(2) All Firms ∑ Births 2008-15	(3) All Firms ∑ Deaths 2008-15
Pull IV	-0.180	-0.188	-0.008
	(0.053)	(0.084)	(0.090)
Observations	686	686	686
R-squared	0.186	0.572	0.241
Avg. Outcome	-0.123	7.867	7.990
Mean Pull IV	0.046	0.046	0.046
S.d. Pull IV	0.049	0.049	0.049
Controls	X	X	X
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the change in firm stock (1), cumulative firm entry (2) and exit (3) between 2008-2015 as a fraction of LLM population 25-64 years old (average 2005-2008) times 100. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries relative to the LLM population in 2000 interacted with GDP growth of each country between 2008-2015, $Pull_l = \sum_c NTWK_{l,c} * G_c$. Controls include unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level as well as 110 province FEs. The average baseline outcomes are the change in firm stock, cumulative firm entry and exit in the pre-period (2005 to 2008) as a fraction of average pre-period population 25-64 years old in the LLM times 100, annualized (i.e., divided by 3 years) and multiplied by 7 years. Standard errors are clustered at the province level (110 clusters).

Chapter 2

The Effects of Returnees' Tax Schemes on High-Skilled Migration in Italy (with Jacopo Bassetto)

2.1 Introduction

Emigration of young and high-skilled individuals is an increasingly relevant concern for many countries (Docquier and Rapoport 2012b; Docquier, Ozden, and Peri 2014). "Brain drain" is especially detrimental if outflows are not compensated by an equivalent inflow of high-skilled workers from abroad, determining a net loss in countries' human capital (Boeri et al. 2012). Once a phenomenon mostly characterizing developing countries, in the last decade there has been an increase in brain drain from developed countries such as Southern European countries. These experienced substantial emigration flows especially after the Schengen treaty introduced free mobility of labor within the European Union (Dorn and Zweimuller 2021).

The surge in emigration flows has motivated many European countries to introduce preferential tax schemes to attract expatriates and foreign nationals, by granting fiscal incentives to high-skilled individuals who move their residence to the country[1]. The

[1]Preferential tax schemes have been introduced in the Netherlands (1985), Denmark (1991), Finland (1999), Sweden (2001), France (2004), Spain (2005), Portugal (2009), and more recently in Italy (2011). Source: Kleven et al. (2020).

goal of these schemes is to reduce wage differentials between countries, and to compensate for the fact that high skilled migrants do not internalize the positive (negative) fiscal and human capital externalities that their relocation exerts in the destination (origin) location (e.g. Moretti 2004). Yet, there is limited empirical evidence on the effectiveness of tax incentives in attracting high-skilled migrants, particularly in a context of brain drain.

What are the effects of granting tax discounts to high-skilled immigrants and return migrants? On the one hand, a growing empirical literature finds large mobility responses of high earners in response to tax differentials (see Kleven et al. (2020) for a survey), suggesting that preferential tax schemes can be effective to attract high-skilled individuals. On the other hand, as young college graduates are not necessarily high earners at the beginning of their careers, their responsiveness to tax incentives may be limited. Further, if tax incentives are mainly taken up by infra-marginal individuals who would have returned anyway, or if a large fraction of these benefits are passed through to employers by reducing returnees' gross wages, then the cost of providing tax incentives may exceed the benefits from a public finance perspective. Last, as other countries may react by introducing tax incentives, these schemes can result in sub-optimal tax competition between jurisdictions.

In this paper, we aim to shed light on the effectiveness of tax incentives to mitigate brain drain. Our empirical analysis focuses on Italy, which has experienced a steady increase in the outflows of young and high-skilled individuals during the last decade. As Figure 2.1 shows, emigration has increased dramatically since the onset of the Great Recession and especially after the Sovereign Debt crisis hit in 2011, while return migration flows, which in the early 2000s were nearly balancing outflows, were systematically lower. Furthermore, the surge in outflows was concentrated among young (under 40 years old) and highly educated (college graduates) individuals, as shown in Figure 2.2.

Besides being characterized by a context of brain drain, Italy provides a suitable empirical setting to study the effects of tax incentives on migrations for at least two reasons. The first is the availability of administrative data on international migration:

while most countries accurately record immigration flows, data on emigration and return migration is usually more sparse and less reliable. Italy is an exception in that it collects detailed administrative records on emigrants and returns, which we describe in Section 3.2. The second is the existence of plausibly exogenous tax variation to identify a causal effect. In fact, as migration choices of individuals are likely correlated with unobserved economic conditions of destination and origin locations, selection and omitted variable bias might hinder identification of the effects of tax incentives.

In late 2010, Italy introduced a preferential tax scheme for young high-skilled expatriates who moved their residence back to Italy. Specifically, this scheme granted a generous 70-80% reduction of taxable labor income[2] for expatriates who returned to Italy starting from 2011, but only if they had a college degree and they were born after January 1, 1969 (i.e. if they were under 41 years old in 2010).[3] The joint presence of these eligibility requirements create suitable quasi-experimental variation to identify the elasticity of return migration to tax incentives.

In our empirical analysis we exploit these quasi-experimental conditions in a Diff-in-Diff and Triple DiD strategy. Using administrative data on international migration of Italian citizens, we find that eligibility for the 2010 tax scheme significantly increased return migration rates of young and high-skilled Italian expatriates. In our favorite specification - a Triple DiD regression where we allow for origin-country-specific differential trends between treated and control groups - eligible individuals are between 22-41% more likely to return after the introduction of the scheme, implying that between 18 and 29% of eligible returnees would not have returned absent the incentives. The estimated effect is driven by Italians returning from other European countries - such as Belgium, France, Germany, Switzerland and the UK. We then use social security data on the universe of Italians workers in Germany (which is the main origin

[2]The reduction was 70% for men and 80% for women.

[3]Eligibility also required at least 2 years of stay abroad and 2 years of *pre-residence* in Italy before expatriating. While all EU citizens were eligible for the scheme, this latter requirement limits dramatically the fraction of eligible foreign nationals. For this reason, in this paper we focus only on Italian citizens.

country of Italian returnees). We estimate a similar effect on the probability that Italians expatriates in Germany leave the registry, and we uncover relatively homogeneous effects across the wage distribution of Italian workers in Germany, suggesting that mobility responses to tax incentives may be a broader phenomenon not limited to top earners.

Our paper contributes to three main strands of literature. First, we contribute to the growing public finance literature on mobility responses to taxation. Previous research has focused on the mobility of top earners either within national boundaries (Agrawal and Foremny 2019; Moretti and Wilson 2017; Schmidheiny and Slotwinski 2018; Liebig, Puhani, and Sousa-Poza 2007; Young and Varner 2011) or across countries (Akcigit, Baslandze, and Stantcheva 2016; Kleven et al. 2014; Kleven, Landais, and Saez 2013b; Muñoz 2019) in response to tax changes. For instance, Kleven et al. (2014) estimate a large elasticity of high-earner foreigners to a 1991 preferential tax scheme in Denmark. We show that a tax incentives which does not exclusively apply to top earners can nonetheless trigger substantial migration responses. Second, our paper is related to the literature on the determinants of return migration (Dustmann and Görlach 2016; Adda, Dustmann, and Görlach 2016). A prominent example is Del Carpio et al. (2016), who study the effects of a program to attract Malaysian nationals living abroad that offers tax deductions upon return. We complement this literature by studying how Italians expatriates eligible for the tax schemes respond to a large shock in net wage differentials between their home country and the destination countries. Finally, we speak to the literature studying the economic effects of brain drain (Anelli et al. 2020; Giesing and Laurentsyeva 2017b), brain return (Mayr and Peri 2009b) and brain gain (Fackler, Giesing, and Laurentsyeva 2018), and to papers investigating the role of migration policies on the mobility of high-skilled individuals (Kato and Sparber 2013; Kerr et al. 2017a; Czaika and Parsons 2017; Boeri et al. 2012). We provide the first evaluation of a large policy predominantly targeting high-skilled nationals residing abroad.

The remainder of this paper unfolds as follows. In Section 2.2 we describe our empirical setting by documenting the context of brain drain and illustrating the main features of the Italian tax schemes for return migrants. In Section 3.2 we describe our data sources and we show descriptive statistics. In Section 3.3 we explain our identification strategy and in Section 3.4 we present our results. Section 2.6 concludes with a discussion and future extensions to this work.

2.2 Setting

2.2.1 The Italian Brain Drain

Already for the late 1990s, Becker, Ichino, and Peri (2004) documented the peculiarity of Italy as a "net exporter of brains" among EU countries: "the tendency of Italian college graduates to move abroad does not seem to be balanced by a corresponding tendency of foreign college graduates to move into the country". Using data from the OECD DIOC on the stock of immigrants and natives by education level in 2010, in Figure 2.3 we confirm that this picture is still accurate one decade later: while the share of high-educated Italians living abroad is comparable to many other European countries such as the UK, Belgium, France and Germany, Italy stands out by ranking at the bottom of OECD countries in terms of the share of high-educated among its foreign-born population. Furthermore, survey evidence from Saint-Blancat (2019) reveals that such a brain imbalance is especially pronounced among scientists and researchers, thus confirming the severity of the Italian brain drain at the very top of the skill distribution.

If the stocks in 2010 already reveal the severity of the Italian brain drain, the picture becomes much worse once we take into account the surge in migration flows during the following decade: as shown in Figure 2.2, the share of new emigrants among young and tertiary educated Italians has more than tripled since 2010, while the increase was much smaller among older and lower educated individuals. The preferred destination is Western Europe – in line with the easiness to travel and relocate that EU citizens enjoy within the Schengen area – and Germany is consistently among the top-3 destination countries of recent emigrants (Figure 2.5) and the main origin country of return migrants (Figure 2.6).

2.2.2 Tax incentives for returnees

In this context of brain drain, Italy has tried to reverse the negative trend by granting fiscal incentives to high-skilled return migrants. In late 2010, the Parliament approved Law 238/2010 *Controesodo* ("counter-exodus"; henceforth the "2010 Reform") which introduced, starting from 2011, the first tax incentives applying to all return migrants regardless of their occupation[4], as long as they held a university degree and they were born on or after January 1st 1969. In addition, eligibility required a EU citizenship, a minimum of 2 years of *pre-residence* in Italy before emigration, as well as at least 2 years of stay abroad before returning. Under this scheme, only a fraction between 20% (for women) and 30% (for men) of returnees' labor income[5] is subject to income taxation, starting from the year of return to Italy until a set date of expiration of the scheme, which was postponed several times.[6]

The regime was substantially modified by a decree in 2015, the D.Lgs. 147/2015 *Impatriati* ("back to homeland"; henceforth the "2015 Reform"), which applies to those who returned starting from 2016. Specifically, it increased the taxable share, removed the birth-cohort requirement and set a fixed maximum duration of 5 fiscal years for receiving the benefits. Further, it increased the required stay abroad before moving to Italy from 2 to 5 years and it also removed the required 2-year pre-residency in Italy, effectively allowing non-Italian EU citizens to be eligible (as well as non-EU citizens, as long as they were not moving their residence from a tax-free country). The 2015 regime was further modified in 2019, when a new reform removed the college degree requirement, lowered the taxable share and increased the maximum duration beyond

[4]Back in 2003, the very first tax scheme has been implemented for researchers and university professors relocating to Italy. This scheme grants a 90% income tax exemption for 3 years (later extended to 4 and then to 6 years), and it is still in place as of today. The main eligibility criterion, besides holding a research position, is the length of stay abroad, two years at least.

[5]The exemption applies to employee and self-employed labor income as well as to business income.

[6]The law initially stated that the incentives were to be in place until December 31, 2013, which is about 5 years from when the Law was first discussed in the Parliament (January 20, 2009). As the legislative process took almost two years, in late 2011 the incentives were extended until December 31, 2015. Then, in late 2014 they were further extended to December 31, 2017, but this applied only to those who returned to Italy no later than December 31, 2015, as those who returned afterwards were subject to the new regime of D.Lgs. 147/2015.

5 years.

Figures 2.7 and 2.8 summarize all the relevant characteristics of the tax schemes, by showing the timing, size and main eligibility criteria (Figure 2.7) as well as the duration (Figure 2.8) of the preferential tax schemes. Figure 2.7 shows that, while until 2010 young high-skilled returnees were subject to full taxation of their income (just like the rest of the population), for those who returned starting from 2011, only a fraction between 20-30% of their income is taxable, resulting in dramatically lower average and marginal tax rates (which we simulate and show in Figures 2.9 and 2.10 below). This advantageous scheme applied to college graduates born on or after 1969 as long as they returned to Italy until 2015. From 2016 onward, the 2015 scheme was in place, which was less generous in terms of the percentage of tax exemption (initially 70% but then lowered again to 50% in 2017) but also more straightforward and generous in terms of the duration of the incentives (5 fiscal years regardless of the year of migration). In fact, while the duration of the 2010 scheme was eventually extended until 2017 (solid triangles in Figure 2.8), at the time when most individuals made their migration decisions, the duration was expected to be lower than 5 years (as shown by the light triangles).

How do these taxable shares translate in terms of average and marginal income tax rates? This is important in order to estimate a migration elasticity as well as to compare the Italian schemes with those of other countries. To fix ideas, let w denote annual before-tax labor earnings. Absent the incentives, after-tax earnings c are given by:

$$c = w - T(w)$$

where $T(w)$ is a non-decreasing step-wise function that determines the income tax due as a function of gross earnings. Let now s denote the share of income subject to income tax for individuals eligible for a given tax scheme. With the incentives, net earnings are now given by:

$$c = w - T(sw)$$

where s is equal to 0.2 and 0.3 for eligible women and men respectively under the 2010 tax scheme, and to 0.5 under the 2015 scheme.

In Figures 2.9 and 2.10 we simulate the actual reduction in the average and marginal income tax rates due to the 2010 scheme for different levels of before-tax earnings.[7] The solid lines plot the average and marginal tax rates[8] without the incentives ($T(w)/w$ and $T'(w)$ respectively), while the dotted lines plot the average and marginal tax rates with the incentives ($T(sw)/w$ and $T'(sw)$ respectively). For simplicity we use $s = 0.25$, a simple average between the taxable share for women and men under the 2010 scheme. The solid lines show the progressivity of the Italian tax system, whose effective marginal tax rate above the no-tax area (8,000 euros) is roughly 30% up to 28,000 of gross annual earnings, rising to 40% until 100,000 and to almost 45% above 100,000 euros. More interestingly, the dotted lines show that the effective marginal tax rate under the incentives is virtually zero until about 35,000, below 10% up to 120,000 and just 11% above that amount. Therefore, the scheme reduces the average tax rate by as much as 28-33 percentage points and the marginal tax rate by 34-37 p.p. for all taxpayers above the extended non-tax area created by the tax incentives, which is slightly above 50,000 euros. Appendix Figures 2.18 and 2.19 show that the tax rates reduction is remarkably similar if we include the compulsory social security contributions (payroll taxes) paid by the employee, whose rates are unaffected by the tax schemes.[9]

The focus of this paper is on the 2010 Reform for several reasons. First, it was the first and most substantial set of tax incentives for high-skilled returnees, which generated a dramatic drop in the fraction of the returnee's income subject to taxation, and applied to all employed and self-employed workers regardless of their occupation. Second, the joint presence of a birth-cohort threshold as well as the educational attain-

[7]In Appendix Figures 2.20 and 2.21 we show the corresponding graphs for the 2015 scheme.

[8]The graphs are simulated for a representative single taxpayer (the tax unit in Italy is the individual) taking into account all the standard deductions of the Italian tax system. Also, we assume for simplicity that individual earnings are entirely composed by employment, self-employment and business income (henceforth "labor earnings").

[9]Likewise, Appendix Figures 2.22 and 2.23 show the corresponding figures for the 2015 scheme.

ment requirement creates a suitable quasi-experimental setting.[10] Third, as we observe the education level and the birth cohort of return migrants, we are able to identify their eligibility in the data. Last, the reform provides a generous tax discount which is sizable throughout the entire income distribution, and therefore it is not limited to top earners. This is important since these incentives specifically targeted younger individuals, who are often in the early stages of their careers and thus not necessarily high earners.[11] For these reasons, in our Difference-in-Difference strategy, the pre-period will be until year 2010 while the post period will be from 2011 onward.[12]

2.3 Data

2.3.1 Administrative data on migration flows (Istat)

The Italian National Statistical Institute (Istat) collects information from civil registries on all changes of residence, both within Italy and to/from abroad, which are among the best available measures of migration flows. These administrative, individual-level records include information on year of migration, origin and destination (Italian municipality or foreign country) as well as several demographic variables such as date of birth, birthplace, gender, education level, citizenship and marital status at the time of migration.[13] In this paper we use an aggregate version of these data. Specifically, we obtained data from Istat on the number of Italian citizens returning to Italy from abroad, by year of migration (2002-2018), birth cohort (5-year groups), education (less

[10]While the 2015 Reform eliminated the birth-cohort requirement, we still include individuals born before 1969 who returned after 2015 in the control group, in order to have consistent eligibility requirement throughout the pre- and post-period. Results are nonetheless very similar if we include them among the treated, which suggests that the birth-cohort requirement was not as binding anymore in 2015 as it was in 2010.

[11]As shown in Figure 2.10, the largest reduction in the marginal tax rate occurs for an individual earning between 31,000 and 42,000 euros - which is close to the starting gross wage of high-skilled individuals in Italy- as their marginal tax rate would drop from 39% to zero.

[12]We use all the data until the last available year, which is either 2017 or 2018. Results are robust to excluding years 2016 and onward, when the 2015 Reform with different eligibility requirement and exempt income percentage was in place, as well as to excluding year 2011, as some minor eligibility requirements were not clear until mid-2011, when *Agenzia dell'Entrate* (the Italian fiscal authority) issued some clarifications.

[13]Access to the full individual-level microdata is restricted. Researchers can apply for access, which must happen in the Istat cold rooms in Italy with several restrictions.

than high school, high school and college), sex, foreign country of origin and a foreign-born indicator.

The Istat data on international migration of Italian citizens are based on the enrollment and dis-enrollment from the *Anagrafe degli Italiani Residenti all'Estero* (AIRE; Registry of Italians Residing Abroad).[14] Italian citizens are required by law to change their residence whenever they migrate abroad for more than 6 months, which involves a dis-enrollment from the civil registry of their municipality of origin and the enrollment in the AIRE registry. The main benefit of enrolling is that foreign income is not subject to income taxation in Italy, in addition to access to voting from abroad and consular services. Once they return to Italy, they are dis-enrolled from the AIRE registry and enrolled in the civil registry of their destination municipality, which is our measure of return migration. Istat collects all these individual records and aggregates them into emigration (from Italy to abroad) and return migration (from abroad to Italy) flows.

Two main limitations characterize the Istat data. The first is under-reporting. In fact, despite the substantial benefits to enroll in the AIRE registry, there is evidence that a large fraction of Italians do not enroll when they move abroad (Anelli et al. 2020), and, consequently, they do not appear in the return migration data. In Section 2.3.3 below we assess the extent of under-reporting by comparing the return migration flows from Germany recorded in the Istat data with the corresponding statistic from Destatis, the German statistical institute. While this is an important limitation of the Istat data, it does not constitute a problem for our identification strategy as long as it is not differential between eligible and non-eligible individuals, as we discuss in Section 3.3.

A second limitation of the Istat migration data is that we do not have direct information on the eligibility for tax incentives. Nonetheless, we can infer their eligibility quite precisely from their birth cohort, sex and education level.[15] Eligibility for the

[14]We refer to **?**) and Anelli et al. (2020) for a discussion about these data.

[15]While education is self-reported at the time of migration, there is no incentive to misreport the educational attainment to qualify for the tax scheme, as the residence change form which the Istat data is based on has purely statistical and registry purposes, and it is not used by the tax administration

2010 Law requires, in addition to being born after 1969 as well as having a college degree, to have resided in Italy for at least 2 years and then to have spent at least 2 years abroad. While we do observe birth cohorts and educational level precisely enough to identify eligible and non-eligible individuals, unfortunately we do not observe whether they had resided in Italy before emigrating, nor how much time they spent abroad, nor whether individuals are in the labor force and thus earn a positive income or not. While these should not create threats to identification (as long as they are not differential between the treated and control groups), it is likely that measurement error could attenuate our estimates. For these reasons, we limit our analysis to Italian citizens born in Italy, as by definition they must have spent some time in Italy before going abroad, and to individuals 25-64 years old, as they are more likely to be in the labor force. Furthermore, we drop individuals with educational level below high school to have a better control group for college graduates, which are the affected group by the reform.

Figure 2.6 shows the number of returnees over time for each of the top-5 countries of origin. Germany is the top origin country during the entire period, accounting for about 4,000 individuals returning each year, followed closely by the United Kingdom and Switzerland especially in later years. The predominance of Germany in the migration flows of Italians motivates our focus on the Germany-to-Italy migration flows.

2.3.2 German social security data (IEB)

To complement our analysis, we use German social security data (Integrated Employment Biographies; henceforth IEB) from IAB to investigate the effects of eligibility for tax incentives for Italian citizens who are abroad. For this purpose, Germany is a particularly suitable destination country for at least three reasons. First, it is the second largest Italian community abroad (second only to Argentina) with over 800,000 registered Italians (AIRE 2018). Second, together with the UK it is the most frequent destination for recent Italian emigrants (Figure 2.5) and the top country of origin of re-

to determine actual eligibility, which instead relies on information from employers.

turnees (Figure 2.6). Last, migration pattern of Italian to and from Germany strongly resemble the overall outflows.

For this paper, we obtained access to the universe of Italian citizens[16] in the German social security data from 2000 to 2018. The data include detailed information on their employment and unemployment spells, including wage, occupation, sector, firm characteristics and precise location of work. The main limitation of these data is that we do not observe individuals migration behavior. For this reason, we assume that Italian citizens between 25 and 55 years old who leave the register (excluding working students and retired individuals) are return migrants to Italy. In Section 2.3.3 below we corroborate the soundness of this assumption by comparing the annual flows of Italians leaving the social security registry with official migration statistics from Germany and Italy.

In Table 2.1 we show the main characteristics of two subgroups of interest, namely Italians 25-64 years old who entered to and who exited from the social security registry between 2000-2017. The first group is relevant because it includes migrants who arrive to Germany before, during and after the recession. The second group is the group of interest, mostly target of the reform. Italian migrants are more likely to be male, around 20% has a college degree or higher certificate. The average duration of leavers is around 2.5 years, which suggests that migration is rather temporary and possibly circular.

2.3.3 Comparison between Istat and IEB data

In Appendix Figure 2.24 we compare migration flows from Germany to Italy as proxied by three different data sources: the Italian migration data (Istat), the IEB data on the Italians that leave the German social security registry, and the German migration data (Destatis). Two main takeaways emerge from this exercise. First, a comparison between the migration flows recorded by Destatis (red-circles) and Istat (green-triangles) suggests that the under-reporting in the Istat data is substantial, as found by Anelli

[16]We consider only individuals with Italian nationality as the most frequent nationality value.

et al. (2020). Nevertheless, since the Destatis flows include all non-German nationals while the Istat data includes only Italian citizens, the difference between the two series is an upper bound of the true difference. Second, a comparison between the Destatis data and the IEB data (dotted blue-diamonds) reveals that, despite the Italians who leave the IEB data are much more than the Italians who return to Germany as per the Istat data, the leavers flow is still smaller than the migration flows recorded by Destatis. This suggests that, even if some of the Italians who disappear from the German social security data might have migrated elsewhere than Italy, or perhaps might have exited the labor market without leaving Germany, this is unlikely to be a major issue in our setting. If anything, the IEB might instead be missing some Italians (e.g. students) who might have been in Germany and then returned to Italy without ever appearing in the social security data.

Overall, while both data sources have limitations and thus both imperfect proxies for return migration, these limitations are nonetheless very different in terms of their underlying causes (under-reporting for the Istat data and imperfect proxy of return migration for the IEB data). Therefore, it is reassuring that we find very similar result in our empirical analysis, as we show in Section 3.4.

2.3.4 Additional migration data

While Istat data provide a measure of migration flows, there is no information on the stock of Italians abroad. Therefore, we complement the Istat data with migration statistics from destination countries using the OECD *Database on Immigrants in OECD and non-OECD Countries* (DIOC) as well as the IAB Brain Drain dataset (Brücker, Capuano, and Marfouk 2013), which allow us to measure the stock of Italian emigrants in each destination country by educational group. The OECD and IAB Brain Drain datasets collect information on migrant stocks from national censuses and are disaggregated by gender, education and age.

2.4 Empirical strategy

2.4.1 Effects of eligibility for tax incentives on return migration

To study the effect of tax incentives for high-skilled immigration, we rely on the quasi-experimental conditions created by the 2010 Reform which introduced a preferential tax scheme for young high-skilled returnees. Specifically, to identify the effect of tax incentives, we exploit the joint presence of two eligibility requirements - being born in 1969 or after (birth cohort requirement) as well as having a college degree (education requirement). The pre-period will be until 2010 while the post-period is after 2011.

We begin by estimating a simple DiD model, similarly to Kleven et al. (2014). Using repeated cross-sectional data on the annual return migration flows of Italian citizens (Istat administrative data), we first collapse data at the treated by year level and estimate the following specification:

$$logN_{cet} = \gamma Treated_{ce} + \eta Treated_{ce} * Post_t + \lambda_t + \epsilon_{cet} \tag{2.1}$$

where $logN_{cet}$ is the log count of returnees in birth cohort c and education level e relocating to Italy in year t, $Treated_{ce} = 1(c \geq 1969) * 1(e = college)$ is an indicator for the group eligible for the tax incentives, $Post_t = 1(t \geq 2011)$ is indicator for the years when tax incentives were in place and λ_t denotes year fixed effects. Under the parallel trend assumption, namely, that absent tax incentives the eligible and non-eligible groups would have had similar trends in the likelihood of returning, η identifies the reduced-form, intention-to-treat (ITT) effect of eligibility for tax incentives on return migration.

In Section 3.4 we present the results of estimating Equation (2.1) and we show some visual checks on the parallel trend assumption. However, several threats could pose a challenge to a causal interpretation of the estimated effect. For instance, if labor demand for college graduates in Italy was less impacted than demand for high school graduates by the Sovereign Debt crisis in 2011, we would overestimate the effect of tax incentives. Further, group-specific labor demand shocks may have also differentially affected emigration flows, mechanically increasing returns among these broadly defined groups.

To deal with these threats, we enrich our simple DiD model in three ways. First, we exploit the fact that eligibility combines an age as well as an education requirement to estimate a Triple DiD model:

$$logN_{cet} = \eta Treated_{ce} * Post_t + \lambda_t + \mu_c + \mu_e + \mu_{ce} + \mu_{ct} + \mu_{et} + \epsilon_{cet} \tag{2.2}$$

where we allow the various birth cohorts and education groups to be on different trends with the inclusion of birth cohort-by-year ($\mu_{c,t}$) and education-by-year ($\mu_{e,t}$) fixed effects. Under this Triple DiD specification, we need a weaker version of the parallel trend assumption, which requires that the relative outcomes between college and high school graduates among the eligible cohorts (those born in or after 1969) would have been on the same trend as the relative outcomes among the non-eligible cohorts (born before 1969).

Second, we use all the information available in the Italian administrative data to estimate a richer specification across cells defined by birth cohort c, education level e, gender g, origin country o and year of migration t:

$$logN_{cegot} = \eta Treated_{ce} * Post_t + \lambda_t + \mu_c + \mu_e + \mu_g + \mu_o + \mu_{ct} + \mu_{et} + \mu_{gt} + \mu_{ot} + \epsilon_{cegot} \tag{2.3}$$

This specification is much richer, as it allows returnees from each origin country - who experience origin-specific labor market conditions and hence different opportunity costs of moving back to Italy - to be on different levels (μ_o) and trends (μ_{ot}). In other words, the identifying variation now stems from comparing eligible and non-eligible individuals *within* origin countries, thus partialling out any time invariant characteristic of foreign countries as well as country-specific trends. We weight these regressions by the stock of Italian expatriates in each cell (cohort, education, gender) residing in each foreign country as of 2010 (constructed using the OECD DIOC data)[17] - a measure of the size of the group "at risk" of returning to Italy - and we cluster standard errors at the cohort-education-gender-origin level to account for within-cell serial correlation.

Last, as the number of returnees mechanically increases with the stock of individuals abroad, we use the return migration rate, i.e. the number of returnees in each year t as a percentage of the stock of Italian expatriates in each cell in 2010, as an alternative outcome in all speficiations.

2.4.2 Effects on leavers from German social security data

As a second approach to evaluate the effects of tax incentives on return migration, we use the universe of Italian migrants working in Germany. This second approach is complementary to the first along two dimensions. First, we are able to precisely identify the eligible group, both for stayers and for migrants who leave the German labor market in each year. We can therefore replicate the analysis with Italian migration data using individual-level data on Italian citizens who leave the German social security register. Second, exploiting the panel structure of the data and the detailed labor market information we can characterize the last spells before leaving and investigate whether tax incentives changed the selection of return migrants and the length of stay in Germany.

To investigate the effect on returns, we estimate the following equation:

$$Pr[L_{it}] = \alpha + \gamma Treated_i + \beta(Treated_i * Post_t) + \psi' X_{it} + \lambda_t + \epsilon_{it} \tag{2.4}$$

where $Pr[L_{it}]$ is a dummy for individual i leaving the German labor market in quarter t, $Treated_i$ is a dummy for being eligible to the tax incentives (born after 1969 and with tertiary education degree), $Post_t$ is a dummy for the timing of the Controesodo Law (equal to 1 for 2011 and after), X_i is a vector of individual-level controls, λ_t are year fixed effects and ϵ_{it} is the error term. The parameter of interest

[17] Unweighted regressions deliver very similar results.

is β which identities the post-reform difference in the probability of leaving between treated and control migrants relative to the pre-reform difference. Similarly to the first approach, we also estimate a Triple DiD combining the two eligibility criteria (education and birth cohort). One threat to identification in the German data is that changes in return migration might be related to shocks specific to Germany that affect all migrants. For this reason, we also use an alternative control group, Spanish nationals born on after 1969 and with college education to estimate Equation 2.3. Finally, to investigate changes in the characteristics of leavers, we consider only the last employment spell of migrants who leave the register and estimate Equation 2.3 using the log length of stay in Germany as outcome. In all regressions, standard errors are clustered at the individual level to take into account the panel structure of the data.

2.5 Results

2.5.1 Visual evidence on return migration by eligibility status

Figure 2.11 shows the evolution of return migration flows by eligibility for tax incentives. Two main trends emerge from the graph. First, the treated and control series were on parallel trends before 2011, providing confidence on the validity of our parallel trend assumption. Second, the slope of the return migration flows of eligible individuals increase sharply after 2011 relative to the non-eligible group. In Figure 2.12, we further breakdown the non-eligible group by plotting separate series for college graduates born before 1969 (blue-diamonds), as well as high school graduates born on or after 1969 (yellow-circles) and before 1969 (green-circles). It is reassuring that all groups are on fairly parallel trends before 2011, and none of the control groups displays the slope change experienced by the eligible group after 2011.

To give a sense of which cohorts are driving the divergence shown in the previous graphs, in Figure 2.13 we plot the age distribution of returnees, separately for four groups: college graduates before (average 2006-2010) and after (average 2012-2016) the reform, as well as high school graduates before and after the reform. On the x-axis we have each 5-year birth cohort intervals (from older to younger cohorts), while the y-axis shows the yearly average number of returnees from each cohort. The figure reveals several patterns. First, return migration tends to peak at about 35 years old (consistently with the German data shown in Table 2.1), suggesting that the age requirement of the reform (being born after 1969) was binding for a non-negligible share of returnees. Second, while there is a mechanical leftward shift for both college (eligible) and high school (non-eligible) graduates post-reform due to the fact that we observe return migration 6 years after relative to the pre-reform period, the former (college graduates) shows a considerable excess mass for the cohorts born after 1969, i.e. to the right of the vertical line.

2.5.2 Effects of eligibility status on return migration

DiD results - We then confirm the visual evidence presented above by estimating a Difference-in-Differences model. In Table 2.2, Columns 1-4, we show the results of a simple DiD regression with multiple periods (Equation 2.1) by collapsing data into eligibility status by year cells, thus mimicking the trends shown in Figures 2.11 and 2.12. In Column 1 we pool all control groups into the non-eligible group, while in Columns 2 and 3 we use only college graduates born before 1969 and high school graduates born on or after 1969 as control groups respectively. The estimates confirm the graphical evidence shown in Figure 2.11: the coefficient of the interaction term $Treated_{ce} * Post_t$ is positive and statistically significant, suggesting (after the exponential of η) that eligible individuals are 56% more likely to return in the post period relative to non-eligible individuals, regardless on the control group considered. This translates to a fraction of marginal individuals of about 36%, i.e. the percentage of eligible who returned after 2011 that would not have returned absent the tax incentives.

In terms of elasticity of return migration with respect to the average net-of-tax rate, the estimated η translates to a flow elasticity of about 1.22.[18] Compared to the

[18]We consider a reduction in the average income tax rate of 30 p.p. corresponding to a gross income of 75,000 euros (see Figure 2.9). As the average net-of-tax rate before the incentives is about 0.32, we get $\log(1.56)/\log(0.98/0.68) \approx 1.22$.

literature on migration responses to taxation, our estimates are comparable to the short-term flow elasticities for foreigners estimated by Kleven et al. (2014), but they are larger than their estimates for expatriates returning to Denmark. This is consistent with Italian tax incentives mostly targeting Italian expatriates, as Italy is experiencing a brain drain which was not the case in the Danish context.

In Column 4, we estimate the same specification of Column 1, but the outcome variable is now return migration from Germany only. The results are similar to the estimates pooling all countries of origin, thus confirming that the trends in return migration from Germany by eligibility status resemble the overall trends.

Triple DiD results - One may worry that treated individuals could have been on different trends than non-eligible individuals post-reform, e.g. due to a differential impact of the Great Recession on demand for college graduates relative to high school graduates in Italy. As eligibility in our setting is based on both age and education, a Triple DiD allows us to include different intercepts by for each birth cohort and education level as well as differential trends by education and by cohort by including their interaction with year fixed effects (as shown in Equation 3.1). The results are shown in Column 5 (all origin countries) and Column 6 (Germany only), which shows a similar estimate of the key interaction term when pooling all countries - suggesting that cohort- and education-specific trends do not affect the DiD estimates - although a somewhat smaller estimate using only Germany as origin country.

Return migration rates - In Table 2.3 we run the same regressions as Table 2.2 but we use as outcome variable the return migration rate, i.e., the flow of returnees relative to the stock of Italians abroad as of 2010, rather than the log count of returnees. While we lose most non-OECD countries because of data availability on the Italian expatriates stocks, this specification has several advantages. First, it allows us to compare more directly the results estimated with the German social security data, where the outcome variable is the probability of leaving the registry (which proxies the probability of return), which we discuss in Section 2.5.3. Second, as the return migration flows are a function of the stock of Italian expatriates in foreign countries, it is important to account for the distribution of eligible vs non-eligible individuals abroad before the introduction of tax incentives, which we measure using the OECD DIOC data on the stock of Italian immigrants from destination countries Censuses as of 2010. Reassuringly, we estimated effects are similar to the estimates on the log counts: eligibility increases the return migration rates from all countries of origin by 41-43% relative to the baseline, and return migration rates from Germany by 29-38%, depending on the control group used.

Origin country variation - As European countries were unevenly affected by the double-dip Recession, which in turn could have influenced the opportunity cost of moving back to Italy differentially between eligible and non-eligible individuals, origin country-specific labor demand shocks may be partially driving the effects estimated in Tables 2.2 and 2.3. For these reasons, in Tables 2.4 and 2.5, we show the results of estimating Equation 2.3, which leverages within-origin-country variation to estimate the effect of eligibility on log return migration and return migration rates (respectively). To deal with serial correlation within groups - which now are defined by birth cohort, education level, gender and origin country - we cluster standard errors at the cohort-education-gender-origin country level. We also weight regressions by the stock of Italian expatriates in each cell as of 2010.

Looking at Columns 1 and 2 of Table 2.4 - where we only include year fixed effects and year, cohort, education, gender and origin FEs respectively - we see that coefficients are very similar to those in Table 2.2 (likewise for Table 2.5). In Columns 3-4, however, the inclusion of subgroup-specific trends, and specifically of origin country by year fixed effects, reduces the coefficients on log return migration by almost half, translating in a 22% increase relative to the baseline. On the other hand, the coefficients on return migration rates are much less affected by country-specific trends: the treated-post interaction increases return migration rates by 0.304 percentage points, or about 41% of the average baseline return migration rate (0.764).

Heterogeneity and Robustness checks - In Table 2.6, we then perform some sensitivity checks by estimating Equation 2.3 using different subsamples. Column 1 is the baseline - the same specification of Column 3 of Table 2.5. In Columns 2-3, we keep European countries and Switzerland and EU countries only. Coefficients are slightly larger, suggesting that countries within the free-mobility Schengen area are driving the estimated effects. Last, in Columns 4-5, we show that results are robust to excluding years 2016-2018 - when returnees are subject to the new regime of the 2015 Reform (thus without the birth-cohort requirement) - as well as to the exclusion of year 2011, as some eligibility requirements of the 2010 Reform were not clear until mid-2011.

Last, to further understand which groups of individuals and countries of origin are driving the estimated effects, in Figure 2.14 we show some heterogeneity by estimating separate DiD regressions for each demographic group that we observe in the Italian data. While for women we estimate a slightly larger effect (consistent with the larger incentive), the coefficients for men and women are not statistically significantly different from each other. Finally, we plot the coefficients for the top countries of origin of Italian returnees. We estimate the largest and most precise effects on return migrants from other EU countries, such as Germany, Switzerland, UK, France and Belgium, consistent with the fact that expatriates to countries not distant from Italy are more responsive the shock in net wage differentials created by the tax incentives.

2.5.3 Effects on the probability of leaving the German registry

In this section we present results of estimating Equation 2.3 with the IEB data on Italian workers in Germany. The outcome variable is the individual probability of leaving the register in year $t+1$, conditional on being in the register in year t. Treated and control groups are constructed in the same as for analysis with the Italian migration data.

In Figures 2.15 and 2.16, we first show event-study graphs by plotting the coefficients of the interaction between the Treated dummy and year FEs. In the top graph we include all Italian workers while in the bottom graph we only include workers whose last spell in the data is an employment spell. While we do not see any pre-trend prior to 2011, the probability that the eligible group leaves the registry is significantly higher after 2011 relative to the controls, regardless of whether we use all control groups, high school graduates born on or after 1969 or college graduates born before 1969. The estimated effects are still significant but about half in size when we focus on employed workers only, which is not surprising, as it is likely that Italian workers who leave the country might transition through an unemployment spell first.

Table 2.7 shows the corresponding DiD results, pooling all control groups in Column 1, and separately for the two usual control groups, namely college graduates born before 1969 (Column 2) and high school graduates born on or after 1969 (Column 3) as control groups respectively. Similar to the event study graphs, the table is organized in two panels: in the top Panel (A), we include all workers while in the bottom Panel (B) we only include workers whose last spell in the data is an employment spell. We find a positive and statistically significant effect of the tax scheme on the probability of leaving the register. For Panel A, pooling all control groups in Column 1, we estimate a 1.1 percentage point increase in the probability of leaving the register after the 2010 Reform for the eligible group relative to the controls, which corresponds to a 32% increase relative to the baseline. The coefficients are similar by using the young high school graduates (Column 2) and the college graduates born before 1969 (Column 3) as control groups, with the effect ranging between 29-35% of the baseline probability of leaving. Reassuringly, these effects are remarkably similar to the ones we showed in Tables 2.2 and 2.3 using the Italian data on return migration from Germany, which suggests that the probability of leaving the register is likely a good proxy for the probability of returning to Italy.

Finally, to shed some light on the characteristics of workers at the margin of migrating in response to the tax incentives, we estimate the DiD model separately for different subgroups of employed workers. Figure 2.17 show graphically the results of this exercise. Interestingly, we find relatively homogeneous effects across the wage distribution of Italian workers in Germany, with a slightly larger effect for workers with below-median wages, consistently with the fact marginal returnees are relatively young, thus likely at the beginning of their careers. Furthermore, we find a stronger responsiveness of workers in small-medium-enterprises (10-19 employees) and in medium-sized firms (20-99). Last, the largest effects are estimated for individuals working in sectors such as IT and Communication and Finance, which is in line with the literature on migration responses to taxation (e.g. Kleven et al. (2014)).

2.6 Discussion and further extensions

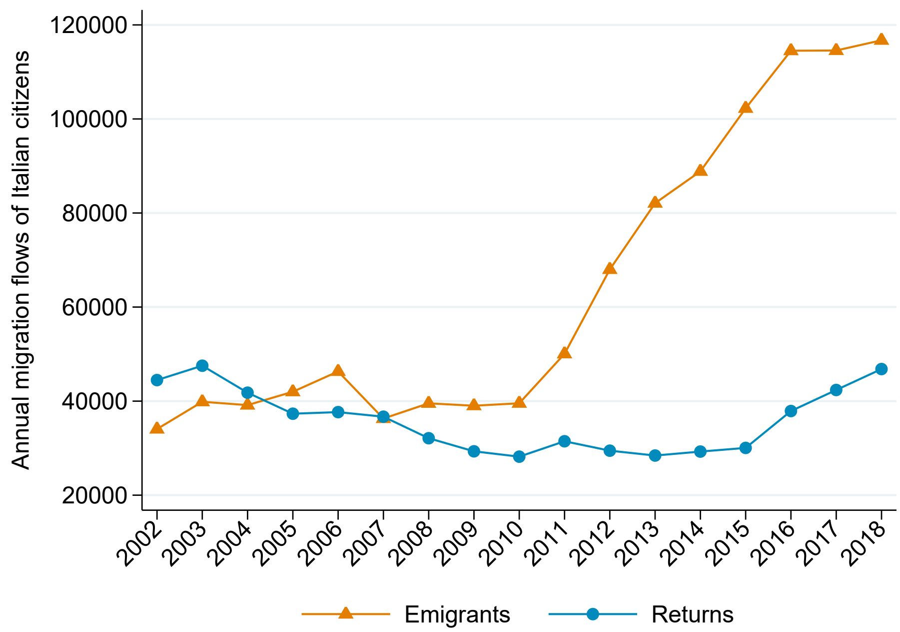
Large emigration flows of young and highly educated individuals have characterized the recent history of many countries. While governments worry about reversing brain drain, few effective policies have been adopted. In this paper we investigate the effects of a unique policy to counteract emigration introduced by the Italian government in 2010. The reform granted large tax discounts to Italian migrants moving back to Italy, as long as they spent at least two years abroad, they had a college degree and they were born on or after 1969. Exploiting the discontinuities in eligibility criteria in a Difference-in-Differences strategy on Italian administrative data on return migration flows, we find that return rates for the eligible group increase by between 22-56%, depending on the specification. We then focus on Germany, the main destination country for Italian expatriates, and estimate similar effects on the probability that Italian workers return to Italy, as proxied by leaving the German social security data. Interestingly, we find that marginal returnees are mostly in the lowest half of the wage distribution in Germany, which is consistent with the fact that marginal returnees are often young college graduates at the onset of their careers. For both the Italian and the German analyses, results are robust to alternative definitions of the control group and sample restrictions. Overall, our findings suggest that tax-schemes-induced mobility can be a broader phenomenon than relocation of top earners in specific occupations (e.g. inventors in Akcigit, Baslandze, and Stantcheva 2016 or football players in Kleven, Landais, and Saez 2013b) and may result in a substantial reallocation of human capital across sending and receiving regions.

A few limitations of our study are worth highlighting and suggest some caution in interpreting our estimates. First, as we do not observe actual take-up, our estimates are intention-to-treat effects of eligibility for tax incentives on return migration. Second, we do not observe in the German data whether Italian citizens who leave the social security register actually return to Italy, although the similarity with the estimates using the Italian migration data suggest that the vast majority do. Last, and most important for policy implications, we do not observe in our data for how long eligible individuals who return stay in Italy nor their earnings, which are a crucial pieces of information to thoughtfully evaluate the effectiveness of tax incentives in reducing brain drain and on their effects on public finance of the destination country.

To conclude, many countries have enacted or are enacting preferential tax schemes to attract high-skilled expatriates and foreigners, particularly in a context of brain drain. Our findings suggest that tax incentives may be an effective policy tool to influence migration choices of workers at the margin, although more research (and more data) is needed to estimate the net social welfare benefit of tax schemes as well as the welfare implications for countries that lose workers as a result, which could inform the design of preferential tax schemes in the future.

Figures

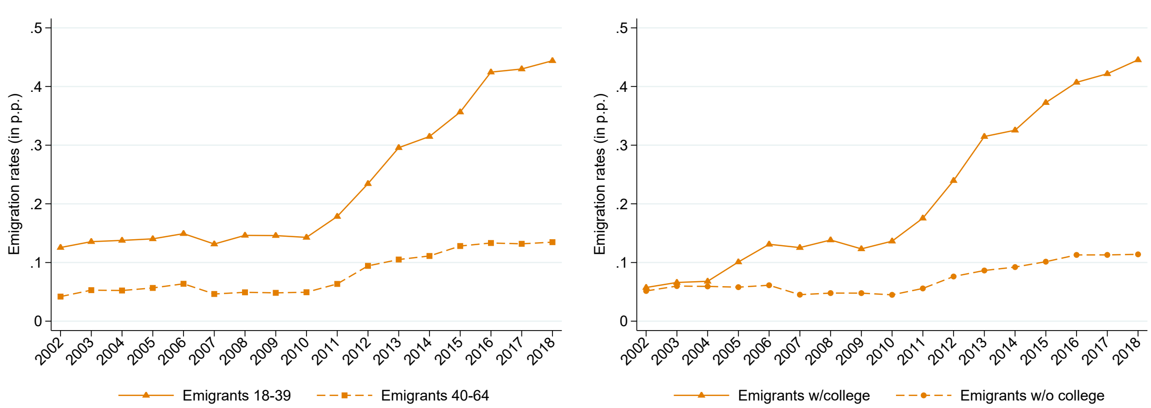
Figure 2.1: Annual emigration and immigration flows of Italians



Notes: annual outmigration (orange triangles) and immigration (blue circles) flows of Italian citizens. Source: authors' elaboration on Istat data.

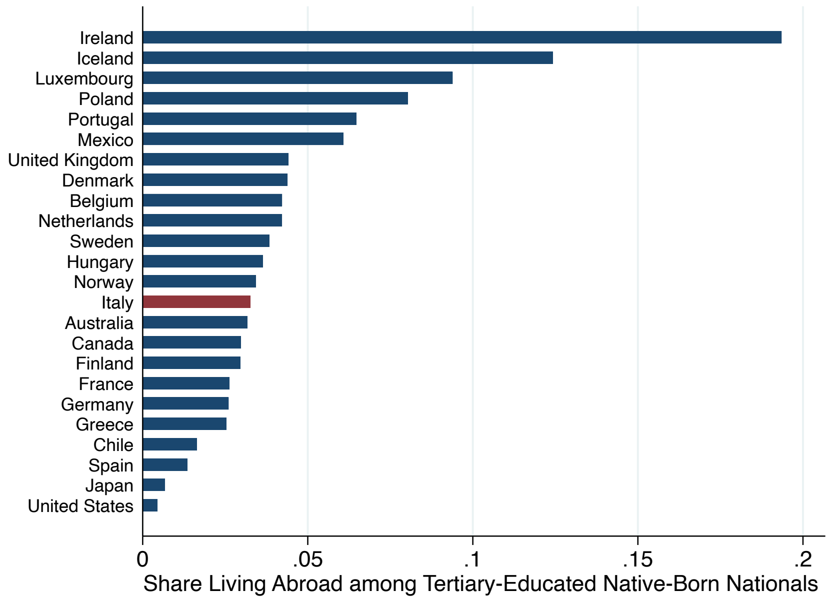
Figure 2.2: Annual emigration flows of Italians, by age and education group

(a) By age (b) By education



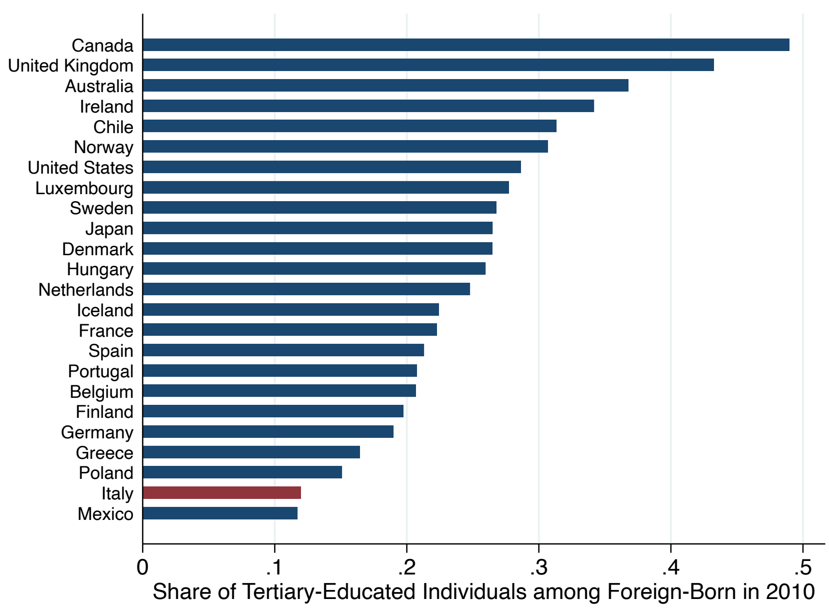
Notes: annual outmigration flows of Italians. Flows are in % of residents in each age/education group as of 2011 Census, multiplied by 100, therefore in percentage points. Source: authors' elaboration on Istat data.

Figure 2.3: Share living abroad among high-skilled nationals in 2010



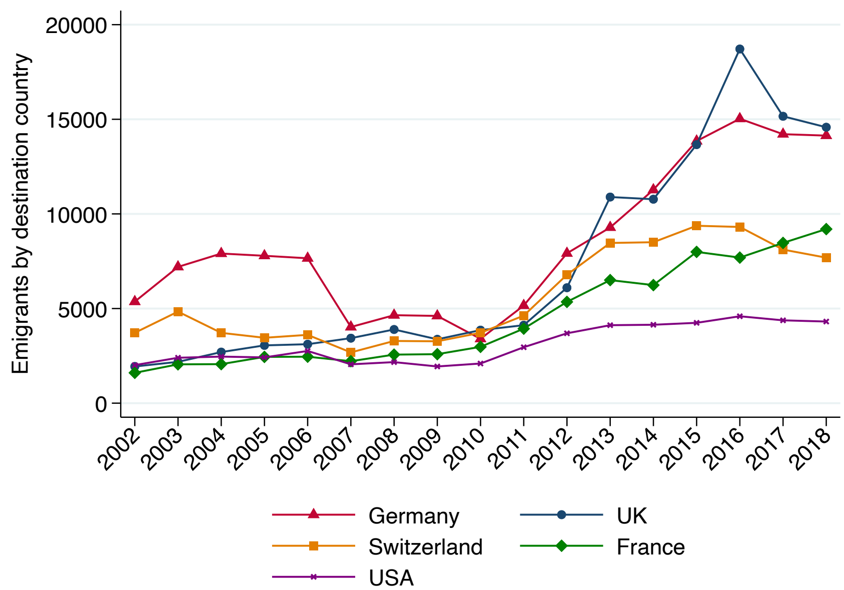
Notes: The graph plots the share of tertiary-educated native-born nationals of each country living abroad relative to the tertiary-educated native-born nationals both living abroad and in the country as of 2010. Source: authors' elaboration on OECD DIOC data.

Figure 2.4: Share of high-skilled among immigrants in 2010



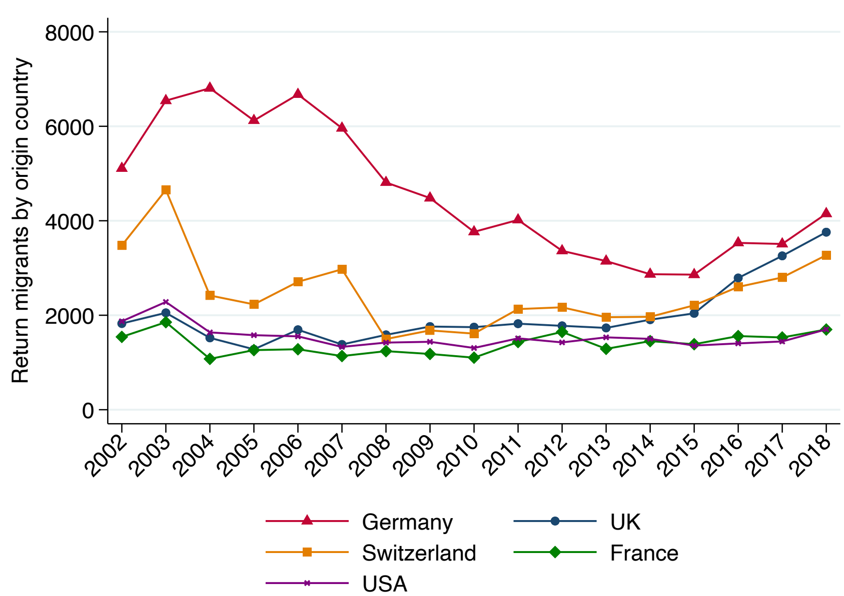
Notes: The graph plots the share of tertiary-educated among foreign born individuals in each country as of 2010. Source: authors' elaboration on OECD DIOC data.

Figure 2.5: Emigration flows of Italians to the top-5 destination countries



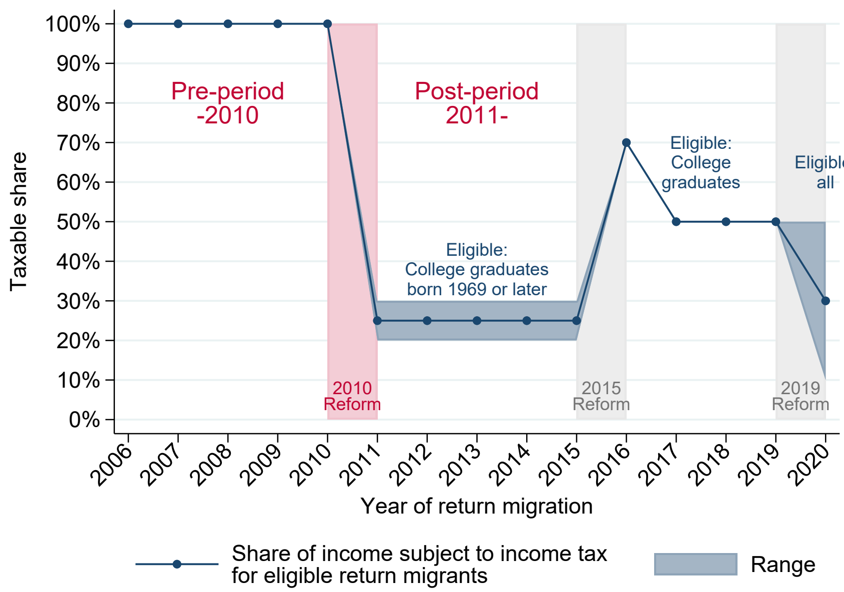
Notes: The graph plots the number of Italian citizens born in Italy migrating to each of the top-5 foreign countries of destination in each year. Source: authors' elaboration on Istat data.

Figure 2.6: Return migration flows of Italians from the top-5 origin countries



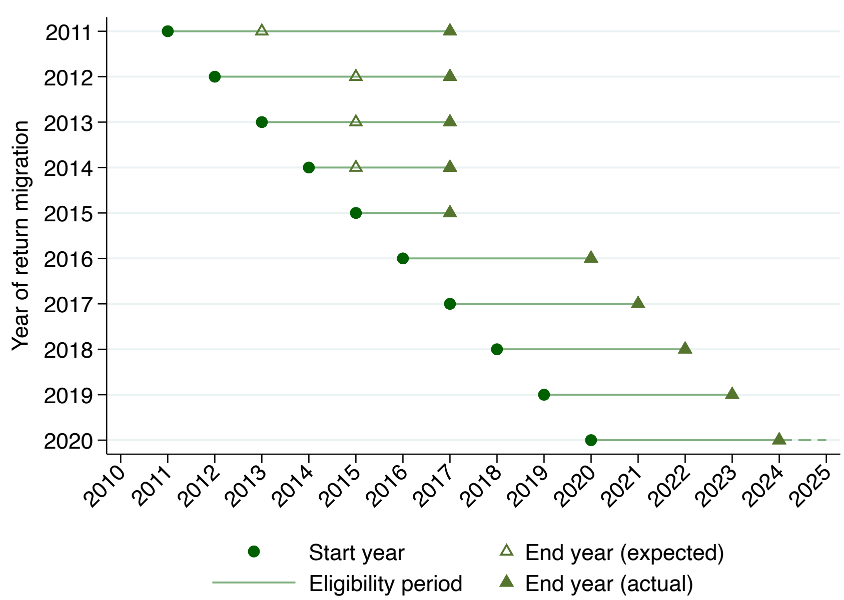
Notes: The graph plots the number of Italian citizens born in Italy moving to Italy from each of the top-5 foreign countries of origin. Source: authors' elaboration on Istat data.

Figure 2.7: Timing, size and eligibility of tax incentives for high-skilled returnees



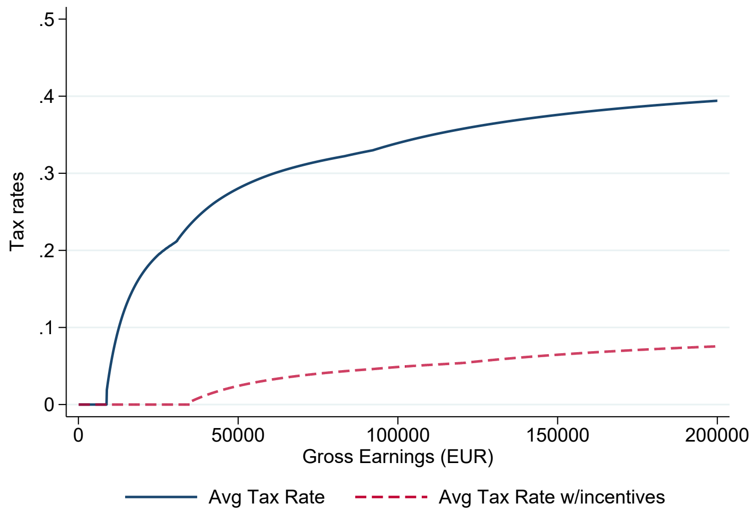
Notes: The graph shows the share of labor income subject to income tax for return migrants eligible for the different tax schemes, depending on the year of return to Italy (on the horizontal axis).

Figure 2.8: Duration of tax incentives for high-skilled returnees



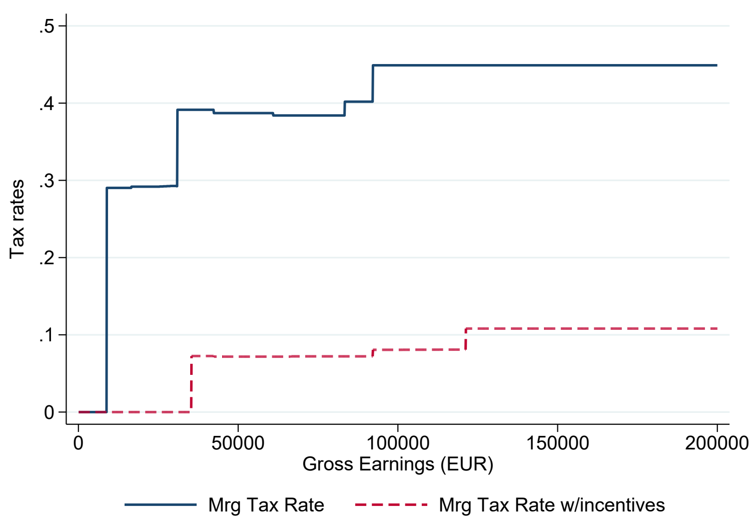
Notes: The graph shows the start and the end year (as well as the expected end year, if different), depending on the year of return to Italy (on the vertical axis).

Figure 2.9: Average income tax rates with and without incentives of the 2010 tax scheme



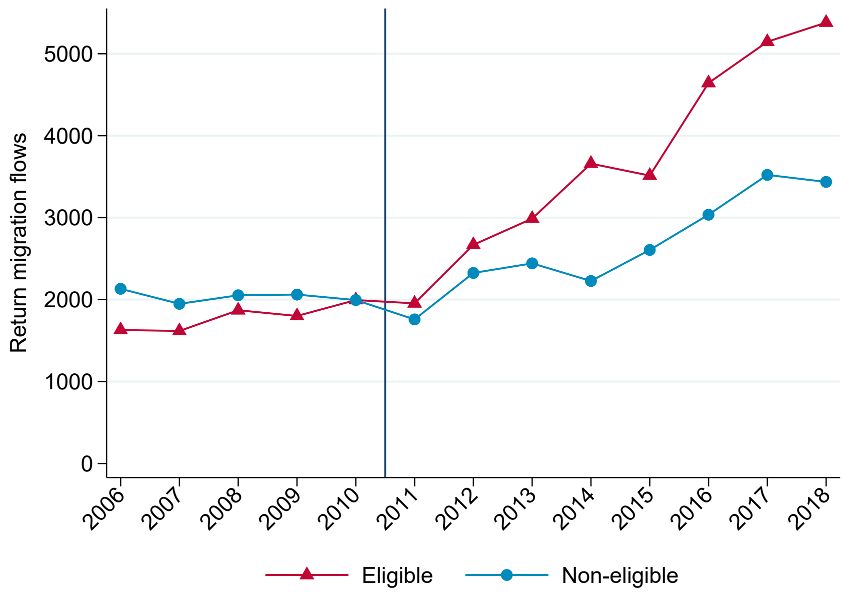
Notes: The figure plots the average income tax rates based on the 2010 Italian tax schedule for an individual with no dependents (source: OECD Taxing Wages 2010). The fiscal incentive used is a 25% share of taxable income (Law 238/2010), i.e. an average between 20% (women) and 30% (men). For the tax rates with the tax incentives, gross earnings are assumed to be entirely deriving from employee labor income, self-employed labor income and/or business income.

Figure 2.10: Marginal income tax rates with and without incentives of the 2010 tax scheme



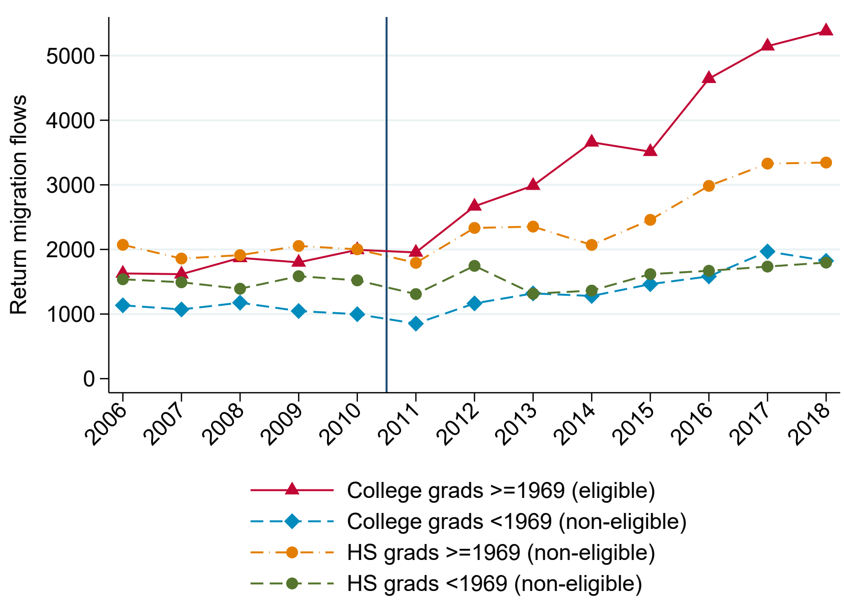
Notes: The figure plots the marginal income tax rates based on the 2010 Italian tax schedule for an individual with no dependents (source: OECD Taxing Wages 2010). The fiscal incentive used is a 25% share of taxable income (Law 238/2010), i.e. an average between 20% (women) and 30% (men). For the tax rates with the tax incentives, gross earnings are assumed to be entirely deriving from employee labor income, self-employed labor income and/or business income.

Figure 2.11: Returns migration flows, by eligibility for tax incentives



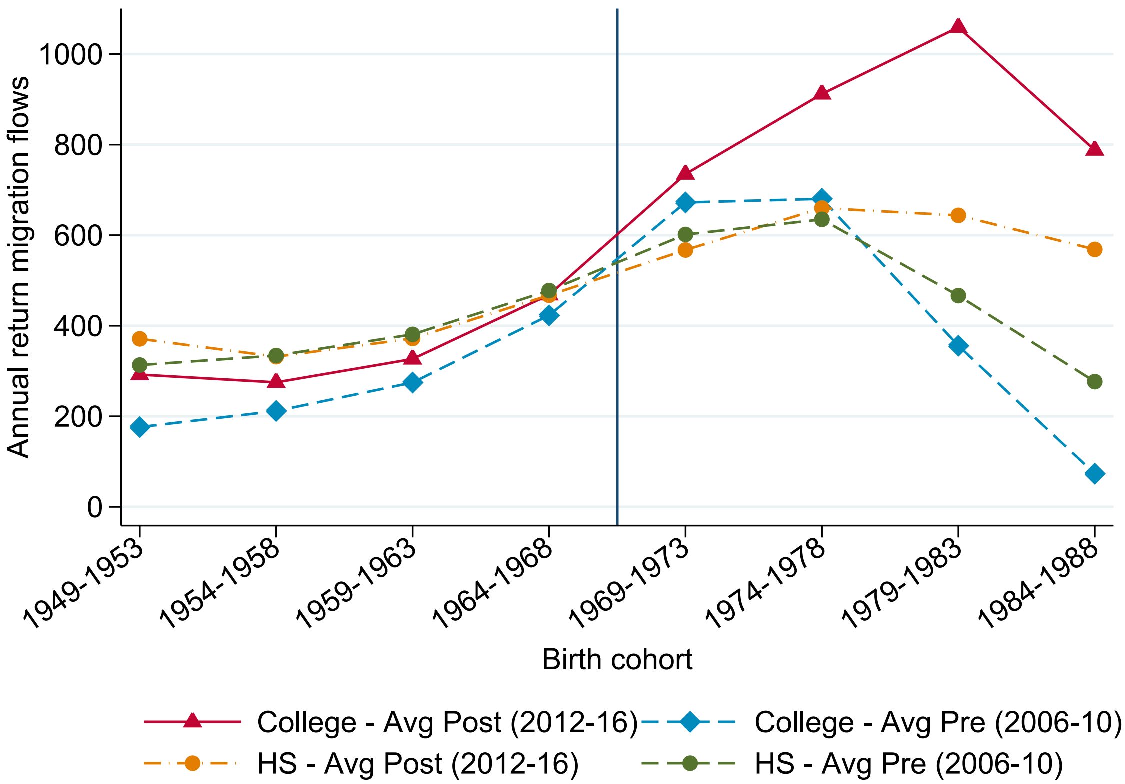
Notes: Non-eligible series is standardized to match the eligible in year 2010. Eligible: college graduates born in 1969 or after. Non-eligible: college graduates born before 1969 and high school graduates born after 1969. Source: authors' elaboration on Istat data on the universe of Italian citizens who move their residence from abroad (AIRE) to Italy in each year. We exclude individuals born abroad as well as individuals born before 1944 (thus older than 66 years old in 2010) and after 1989 (thus younger than 21 years old in 2010).

Figure 2.12: Returns migration flows, by education level and birth cohort group



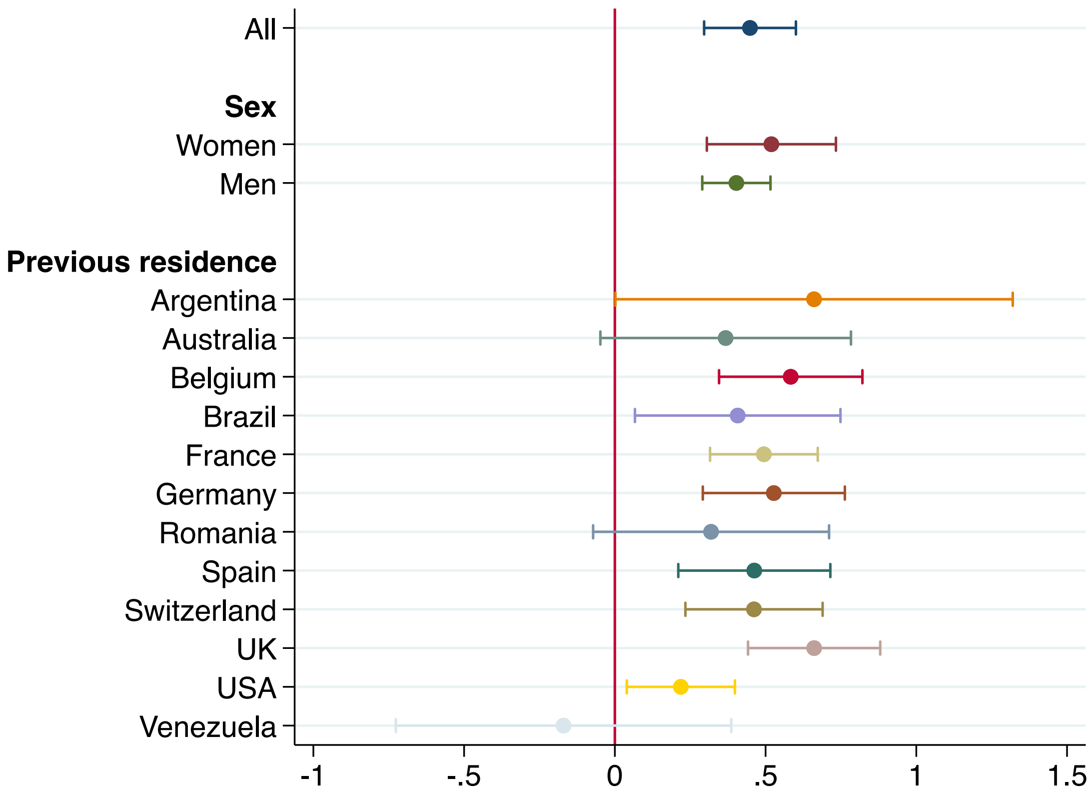
Notes: Source: authors' elaboration on Istat data on the universe of Italian citizens who move their residence from abroad (AIRE) to Italy in each year. We exclude individuals born abroad as well as individuals born before 1944 (thus older than 66 years old in 2010) and after 1989 (thus younger than 21 years old in 2010).

Figure 2.13: Age distribution of returnees, separately by eligibility status and time period before (2006-2010) and after (2012-2016) the reform



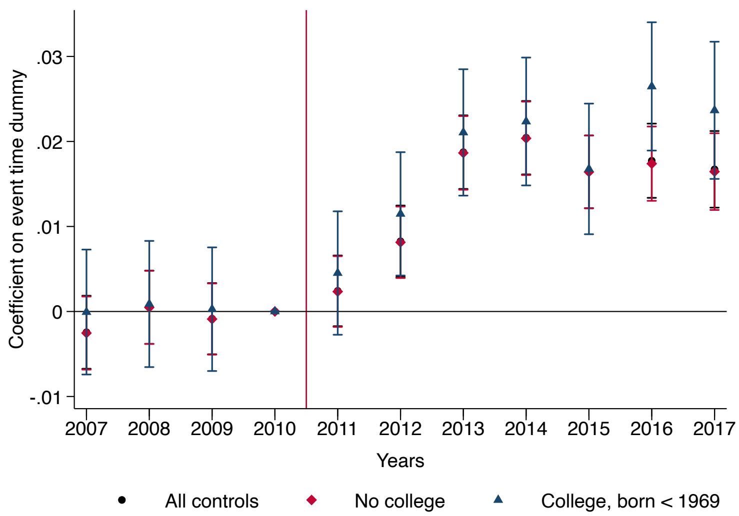
Notes: The figure plots the average annual return migration flows, by birth cohort, for four combination of education and time period: college graduates returning after and before 2011, as well as high school graduates returning after and before 2011. Source: Istat data on the universe of native-born Italian citizens who move their residence from abroad (AIRE) to Italy in each year.

Figure 2.14: Heterogeneous effects of tax incentives on returns



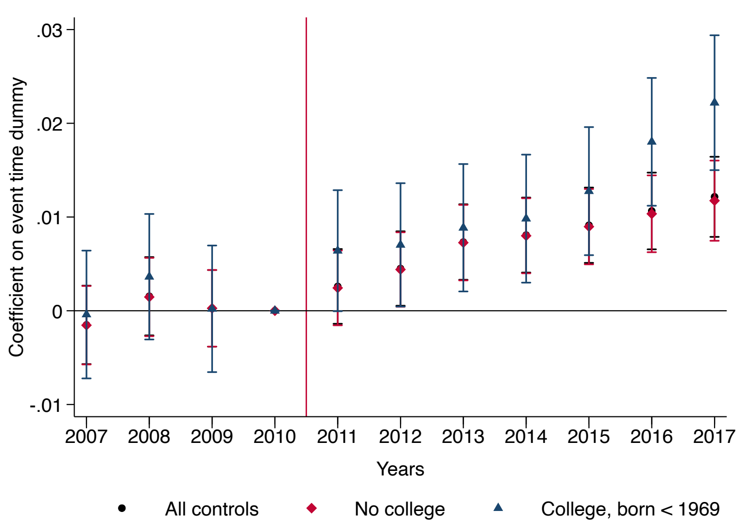
Notes: DiD estimates. Dependent variable is log returns. Bars denote 95% C.I.
Sample is all Italian citizens born in Italy, with at least a high school diploma, born between 1944 and 1989 and returning to Italy between 2006 and 2018. Source: authors' elaboration on Istat data.

Figure 2.15: Event study plot for the difference in probability of leaving the German social security registry, by eligibility for tax incentives in Italy



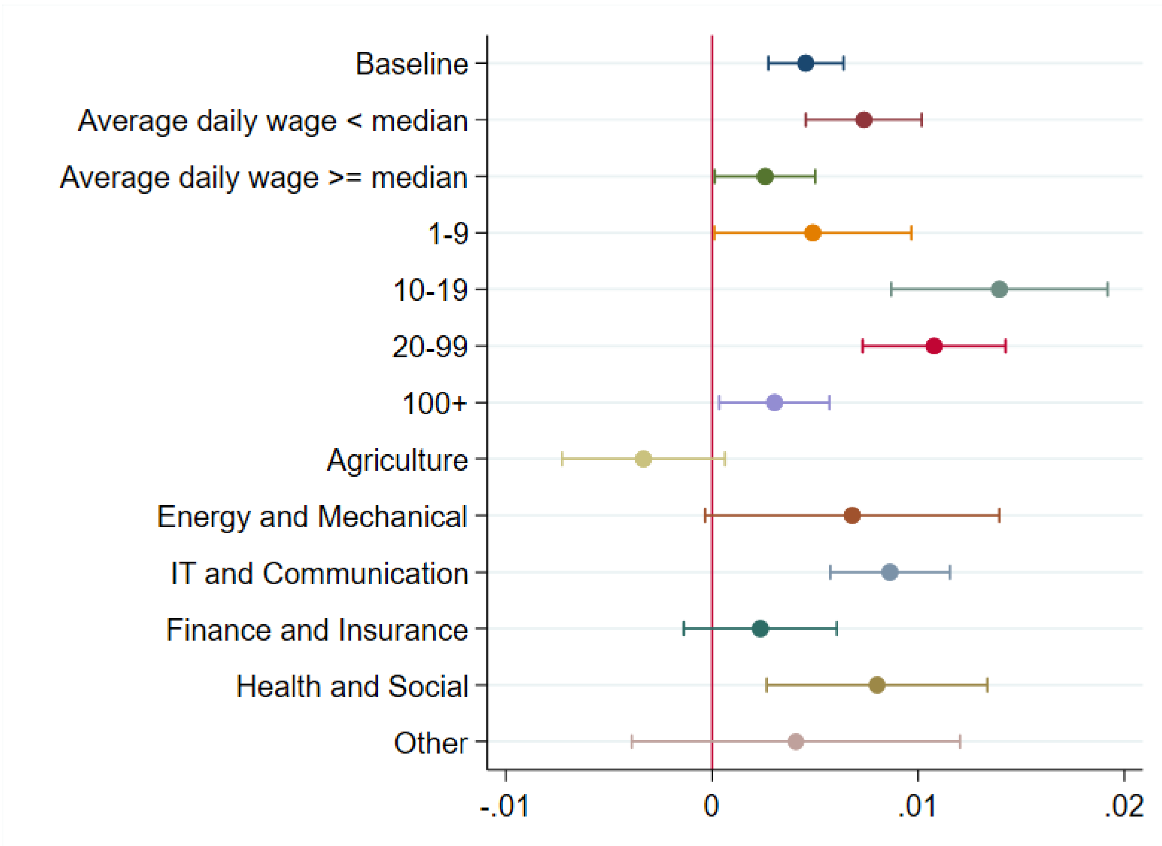
Notes: The figure plots the difference in the probability to return, as proxied by exit from the registry of Italian citizens. The treated group is college graduates born on 1969 or after. Control groups are Italians with high school diploma born after 1969 ("No college"), Italians with university degree born before 1969 ("College, born<1969") and both ("All controls"). Source: IEB data.

Figure 2.16: Event study plot for the difference in probability of leaving the German social security registry, by eligibility for tax incentives in Italy - Employed only



Notes: The figures plot the difference in the probability to return, as proxied by exit from the registry of Italian citizens whose last spell in the registry is an employment spell. The treated group is college graduates born on 1969 or after. Control groups are Italians with high school diploma born after 1969 ("No college"), Italians with university degree born before 1969 ("College, born<1969") and both ("All controls"). Source: IEB data.

Figure 2.17: Heterogeneous effects of tax incentives on returns from Germany



Source: IEB data. Notes: the figure displays the point estimates and confidence intervals of the DiD coefficient for the baseline estimate and separately for subgroups of Italians, breakdown by whether their wage is below or above median, by firm size (1-9, 10-19, 20-99 and 100+ employees) and by sector. All regressions control for age at arrival, year and years in the register. Age of arrival is the age at first entry in the register. Sectors are aggregated from the German WZ08 Classification. Only Italians aged 23 or above and with a higher education or high school (or VET) degree are included. Bars denote 95% C.I.

Tables

Table 2.1: Characteristics of Italians in the German Social Security Data

Italians entered between 2000 and 2017, age at entry 25-64		
	Entered	Left
Female	39.41	39.21
Mean age	35.30	37.22
Degree	19.23	14.87
Mean duration in the register (years)	3.86	2.32
Total individuals	208 156	104 652
Total establishments	44 155	40 294

Notes: the table displays basic characteristics of Italian workers in the private sector in Germany. Age of entry between 25 and 64. Nationality identified based on the mode value of the nationality variable. The restriction on age of entry aims at reducing the risk of considering also Italians born in Germany. Only individuals who entered from 2000 onwards are included. Mean age is the mean age at entry and the mean age at exit respectively. The total number of firms is based on the number of unique firm identifiers in which the Italian workers considered have worked for at least one spell. Source: authors' elaboration on the universe of Italians in the German social security data (IEB).

Table 2.2: DiD and Triple DiD effect of eligibility for tax scheme on Log Return Migration

VARIABLES	(1) DD Control: All	(2) DD Control: Coll<1969	(3) DD Control: HS≥1969	(4) DD Origin: Germany	(5) Triple DD Control: All	(6) Triple DD Origin: Germany
Treated $*$ Post	0.448***	0.446***	0.450***	0.527***	0.447***	0.342***
	(0.069)	(0.070)	(0.075)	(0.107)	(0.118)	(0.122)
Observations	26	26	26	26	208	208
R-squared	0.971	0.988	0.960	0.930	0.955	0.930
Avg Outcome	7.482	7.482	7.482	5.544	5.776	3.948
exp^{η}	1.566	1.562	1.568	1.694	1.564	1.407
Year FE	X	X	X	X	X	X
Treated dummy	X	X	X	X		
Cohort FE					X	X
Educ FE					X	X
Cohort-by-Educ FE					X	X
Cohort-by-Year FE					X	X
Educ-by-Year FE					X	X

Notes: Observations (Columns 1-4 - DD): treatment status by year (2006-2018) cells. Observations (Columns 5-6 - Triple DD): education (high school and college) by birth cohort (8 five-year groups from 1949 to 1988) by by year (2006-2018) cells. The dependent variable is the log count of Italian citizens, born in Italy between 1949 and 1988 and with at least a high school diploma, moving to Italy from abroad in year t (Istat data). "Treated" is a dummy equal to 1 if birth year is equal or greater than 1969 and education level is college and "Post" is a dummy equal to 1 for the post period years (2011 and after). In Columns 1 and 4,5,6, we include both high school graduates and college graduates born before 1969 in the control group, while in Columns 2 and 3 we only include college graduates born before 1969 and high school graduates born on or after 1969 respectively. In Columns 4 and 6, we only include return migrants from Germany. Robust standard errors in parenthesis. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 2.3: DiD and Triple DiD effect of eligibility for tax schemes on Return Migration Rates

VARIABLES	(1) DD Control: All	(2) DD Control: Coll<1969	(3) DD Control: HS≥1969	(4) DD Origin: Germany	(5) Triple DD Control: All	(6) Triple DD Origin: Germany
Treated $*$ Post	0.470***	0.455***	0.488***	0.612***	0.349***	0.631***
	(0.073)	(0.071)	(0.090)	(0.160)	(0.074)	(0.210)
Observations	26	26	26	26	182	169
R-squared	0.955	0.980	0.898	0.907	0.988	0.932
Avg Outcome	1.091	0.913	1.369	1.458	0.857	1.118
Year FE	X	X	X	X	X	X
Treated dummy	X	X	X	X		
Cohort FE					X	X
Educ FE					X	X
Cohort-by-Educ FE					X	X
Cohort-by-Year FE					X	X
Educ-by-Year FE					X	X

Notes: Observations (Columns 1-4 - DD): treatment status by year (2006-2018) cells. Observations (Columns 5-6 - Triple DD): education (high school and college) by birth cohort (8 five-year groups from 1949 to 1988) by by year (2006-2018) cells. The dependent variable is the return migration rate of Italians abroad, and is equal to the count of Italian citizens, born in Italy between 1949 and 1983 and with at least a high school diploma, moving to Italy from abroad in year t (Istat data), divided by the stock of Italian expatriates as of 2010 (OECD DIOC data). "Treated" is a dummy equal to 1 if birth year is equal or greater than 1969 and education level is college and "Post" is a dummy equal to 1 for the post period years (2011 and after). In Columns 1 and 4,5,6, we include both high school graduates and college graduates born before 1969 in the control group, while in Columns 2 and 3 we only include college graduates born before 1969 and high school graduates born on or after 1969 respectively. In Columns 4 and 6, we only include return migrants from Germany. Robust standard errors in parenthesis. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 2.4: DiD and Triple DiD effect of eligibility for tax schemes on Log Return Migration - Origin Country variation

VARIABLES	(1) Log(Returns)	(2) Log(Returns)	(3) Log(Returns)	(4) Log(Returns)
Treated $*$ Post	0.407***	0.439***	0.202***	0.204***
	(0.097)	(0.097)	(0.061)	(0.065)
Observations	7,368	7,368	7,368	7,368
R-squared	0.097	0.813	0.861	0.915
Avg Outcome	2.544	2.544	2.544	2.544
exp^{η}	1.503	1.551	1.224	1.226
Year FE	X	X	X	X
Cohort		X	X	X
Educ		X	X	X
Gender		X	X	X
Origin		X	X	X
Year by C E G O			X	X
C by E by G by O				X

Notes: Observations: birth cohort c by education e by gender g by country of origin o by year of migration y cells. The dependent variable is the return migration rate of Italians abroad, and is equal to the count of Italian citizens, born in Italy between 1949 and 1983 and with at least a high school diploma, moving to Italy from abroad in year t (Istat data). All columns include Year fixed effects and a Treated dummy. Columns 2 also includes cohort, education, gender and origin countries fixed effects. Columns 3 and 4 are Triple DiD specifications as we include year by cohort, year by education, year by gender and year by origin countries fixed effects (in both columns) as well as all indicators for all the two-way combinations of cohort, education, gender and origin countries (Column 4) Observations are weighted by the stock of Italian expatriates in each cohort-education-gender-origin country cell as of 2010, based on the OECD DIOC data. Standard errors are clustered at the cohort-education-gender-origin country level. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 2.5: DiD and Triple DiD effect of eligibility for tax scheme on Return Migration Rates - Origin Country variation

VARIABLES	(1) Return Rate	(2) Return Rate	(3) Return Rate	(4) Return Rate
Treated $*$ Post	0.511***	0.490***	0.304***	0.306***
	(0.099)	(0.099)	(0.077)	(0.082)
Observations	7,368	7,368	7,368	7,368
R-squared	0.074	0.501	0.563	0.843
Avg Outcome	0.746	0.746	0.746	0.746
Year FE	X	X	X	X
Cohort		X	X	X
Educ		X	X	X
Gender		X	X	X
Origin		X	X	X
Year by C E G O			X	X
C by E by G by O				X

Notes: Observations: birth cohort c by education e by gender g by country of origin o by year of migration y cells. The dependent variable is the number of Italian citizens, born in Italy between 1949 and 1988 and with at least a high school diploma, moving to Italy from abroad in year t (Istat data), divided by the stock of Italian expatriates as of 2010 (OECD DIOC data). All columns include Year fixed effects and a Treated dummy. Columns 2 also includes cohort, education, gender and origin countries fixed effects. Columns 3 and 4 are Triple DiD specifications as we include year by cohort, year by education, year by gender and year by origin countries fixed effects (in both columns) as well as all indicators for all the two-way combinations of cohort, education, gender and origin countries (Column 4) Observations are weighted by the stock of Italian expatriates in each cohort-education-gender-origin country cell as of 2010, based on the OECD DIOC data. Standard errors are clustered at the cohort-education-gender-origin country level. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 2.6: Robustness: DiD and Triple DiD effect of eligibility for tax scheme on Return Migration Rates - Origin Country variation

VARIABLES	(1) Return Rate Baseline	(2) Return Rate EU + CH only	(3) Return Rate EU only	(4) Return Rate No 2016-18	(5) Return Rate No 2011
Treated ∗ Post	0.312***	0.372***	0.450***	0.273***	0.321***
	(0.101)	(0.137)	(0.127)	(0.093)	(0.106)
Observations	7,368	5,117	4,753	5,529	6,822
R-squared	0.215	0.215	0.205	0.201	0.215
Avg Outcome	0.746	0.999	1.034	0.701	0.756
Year FE	X	X	X	X	X
Cohort	X	X	X	X	X
Educ	X	X	X	X	X
Gender	X	X	X	X	X
Year by C E G	X	X	X	X	X

Notes: Observations: birth cohort c by education e by gender g by country of origin o by year of migration y cells. The dependent variable is the number of Italian citizens, born in Italy between 1949 and 1988 and with at least a high school diploma, moving to Italy from abroad in year t (Istat data), divided by the stock of Italian expatriates as of 2010 (OECD DIOC data). All columns include year, cohort, education, gender and origin countries fixed effects, as well as year by cohort, year by education, year by gender and year by origin countries fixed effects. In Column 2, we only keep migration from European Union countries and Switzerland, and only EU countries in Column 3. In Columns 4-5, we respectively exclude years 2016-2018 and 2011 from the regressions. Observations are weighted by the stock of Italian expatriates in each cohort-education-gender-origin country cell as of 2010, based on the OECD DIOC data. Standard errors are clustered at the cohort-education-gender-origin country level. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

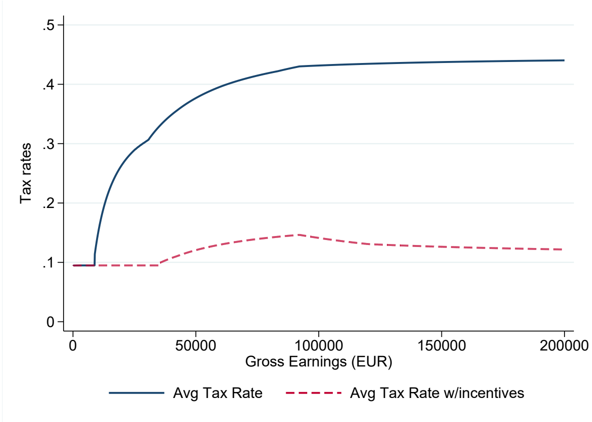
Table 2.7: DiD effect of eligibility on the probability of leaving the German social security registry

	(1) All controls	(2) No college	(3) CollegeBorn<1969
Panel A: All workers			
Treated*Post	0.011 ***	0.010 ***	0.012 ***
	[0.0009]	[0.0010]	[0.0016]
Mean outcome (treated at t = 0)	0.034	0.034	0.034
Observations	1,851,074	1,788,001	272,301
Individuals	221,278	213,442	42,578
R2	0.0136	0.0138	0.0201
Panel B: Only employed			
Treated*Post	0.005 ***	0.004 ***	0.006***
	[0.0009]	[0.0009]	[0.0015]
Mean outcome (treated at t = 0)	0.028	0.028	0.028
Observations	1,587,216	1,532,365	239,818
Individuals	209,221	201,982	39,522
R2	0.0114	0.0115	0.0151

Notes: Source: IEB. Notes: the outcome variable is the probability of leaving the register compared to being still in the register at *t+1*. All migrants are included as long as their highest reported educational level is either high school diploma (and VET studies) or university degree and if the mode of the nationality variable is "Italian". Controls include gender, age at entry in the register, birth cohort and year fixed effects. Standard errors are clustered at the individual level. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

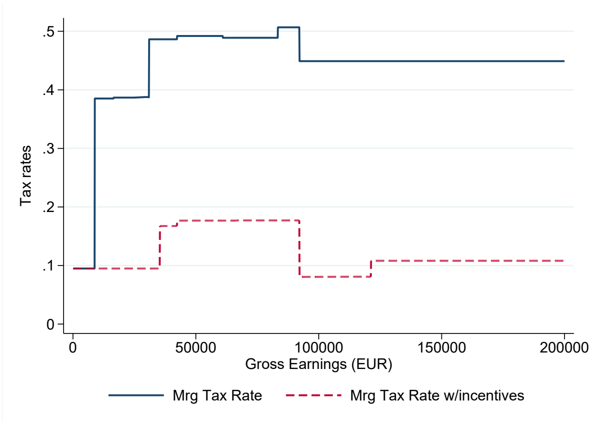
Appendix

Figure 2.18: Average income tax rates including employee compulsory social security contributions, with and without incentives of the 2010 tax scheme



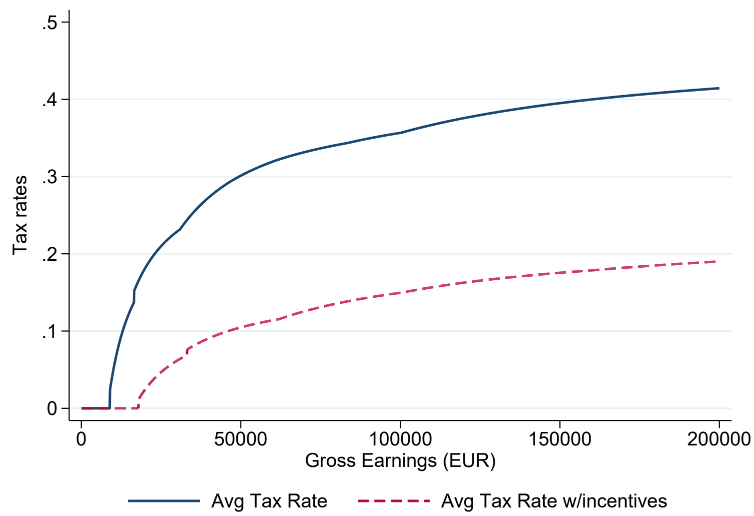
Notes: The figure plots the average income tax rates, including payroll taxes paid by employees, based on the 2010 Italian tax schedule for an individual with no dependents (source: OECD Taxing Wages 2010). The fiscal incentive used is a 25% share of taxable income (Law 238/2010), i.e. an average between 20% (women) and 30% (men). For the tax rates with the tax incentives, gross earnings are assumed to be entirely from employee labor income, self-employed labor income and/or business income.

Figure 2.19: Marginal income tax rates including employee compulsory social security contributions, with and without incentives of the 2010 tax scheme



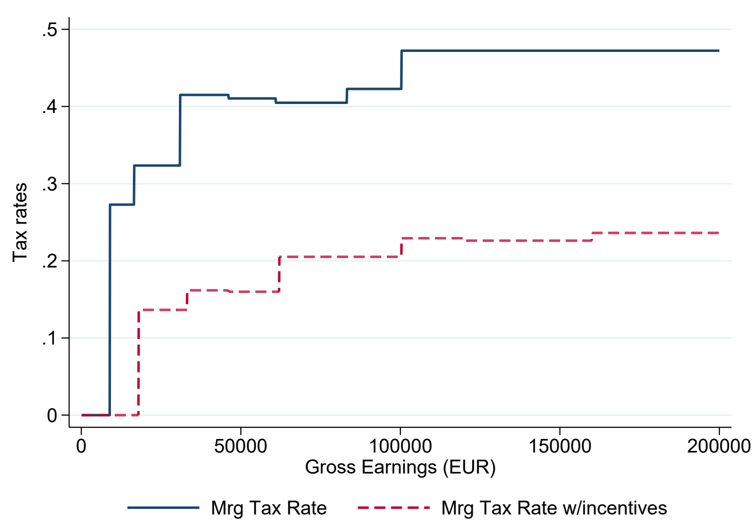
Notes: The figure plots the marginal income tax rates, including payroll taxes paid by employees, based on the 2010 Italian tax schedule for an individual with no dependents (source: OECD Taxing Wages 2010). The fiscal incentive used is a 25% share of taxable income (Law 238/2010), i.e. an average between 20% (women) and 30% (men). For the tax rates with the tax incentives, gross earnings are assumed to be entirely from employee labor income, self-employed labor income and/or business income.

Figure 2.20: Average income tax rates with and without incentives of the 2015 tax scheme



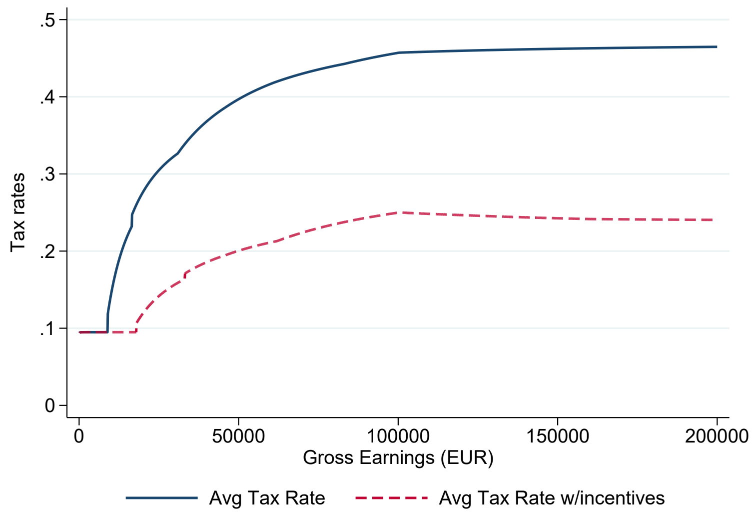
Notes: The figure plots the average income tax rates based on the 2017 Italian tax schedule for an individual with no dependents (source: OECD Taxing Wages 2017). The fiscal incentive used is a 50% share of taxable income (D.Lgs. 147/2015). For the tax rates with the tax incentives, gross earnings are assumed to be entirely deriving from employee labor income, self-employed labor income and/or business income.

Figure 2.21: Marginal income tax rates with and without incentives of the 2015 tax scheme



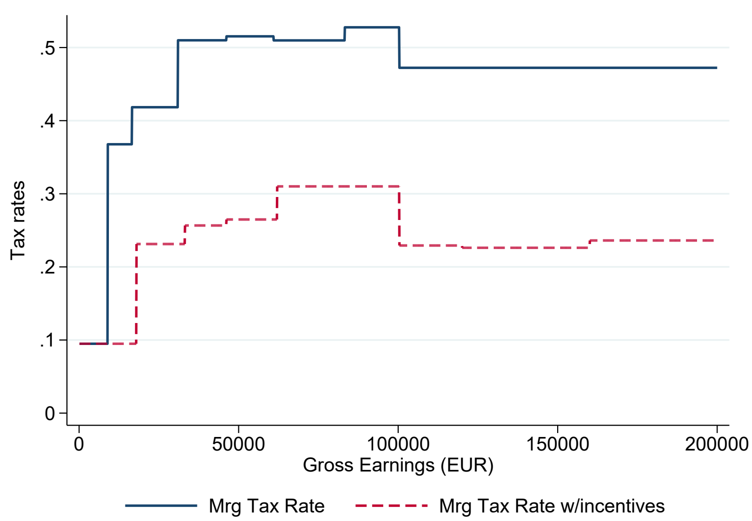
Notes: The figure plots the marginal income tax rates based on the 2017 Italian tax schedule for an individual with no dependents (source: OECD Taxing Wages 2017). The fiscal incentive used is a 50% share of taxable income (D.Lgs. 147/2015). For the tax rates with the tax incentives, gross earnings are assumed to be entirely deriving from employee labor income, self-employed labor income and/or business income.

Figure 2.22: Average income tax rates including employee compulsory social security contributions, with and without incentives of the 2015 tax scheme



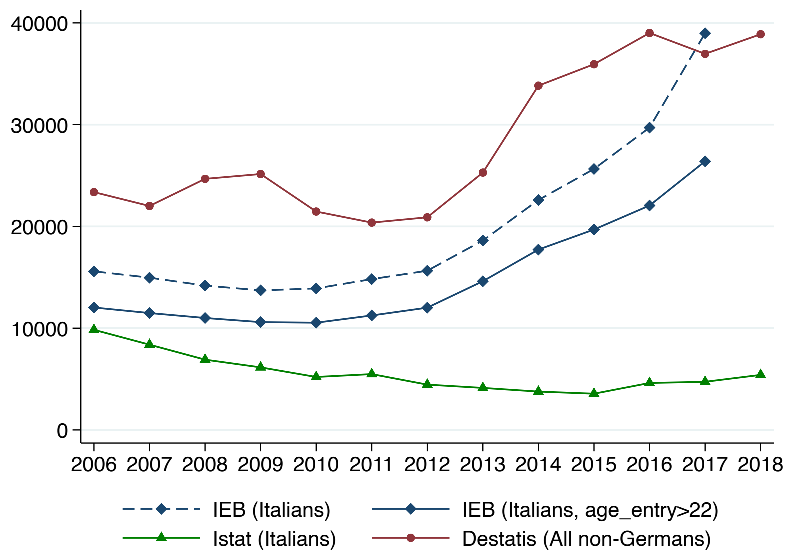
Notes: The figures plots the average income tax rates, including payroll taxes paid by employees, based on the 2017 Italian tax schedule for an individual with no dependents (source: OECD Taxing Wages 2017). The fiscal incentive used is a 50% share of taxable income (D.Lgs. 147/2015). For the tax rates with the tax incentives, gross earnings are assumed to be entirely from employee labor income, self-employed labor income and/or business income.

Figure 2.23: Marginal income tax rates including employee compulsory social security contributions, with and without incentives of the 2015 tax scheme



Notes: The figure plots the marginal income tax rates, including payroll taxes paid by employees, based on the 2017 Italian tax schedule for an individual with no dependents (source: OECD Taxing Wages 2017). The fiscal incentive used is a 50% share of taxable income (D.Lgs. 147/2015). For the tax rates with the tax incentives, gross earnings are assumed to be entirely from employee labor income, self-employed labor income and/or business income.

Figure 2.24: Annual migration flows from Germany to Italy, by data source



Notes: The green-triangles lines "Istat" plots the number of Italian citizens moving from Germany to Italy in each year, as recorded by Istat (the Italian statistical institute). The red-circles line "Destatis" plots the number of non-German citizens moving from Germany to Italy in each year, as recorded by Destatis (the German statistical institute). The blue-diamonds line "IEB" plots the number of Italian citizens (solid line) and the number of Italian citizens who appeared in the German social security registry after the age of 22 (dotted line) leaving the registry in each year, as recorded in the IEB data.

Chapter 3

Do local tax differentials affect internal migration?

3.1 Introduction

Do individuals move within countries in response to tax incentives? While a growing empirical literature finds that international migration of top earners is highly responsive to tax differentials (e.g. Kleven, Landais, and Saez 2013a; Muñoz 2019; Kleven et al. 2020), fiscal incentives may also play a role to explain mobility within countries and migration of non top earners. There are several reasons why this may be the case. First, internal migration is less costly relative to international migration and usually unrestricted. Second, while career opportunities as well as personal relationships are likely to be the main determinants of individuals' mobility choices, fiscal incentives may play a role especially for workers who have access to job opportunities in different locations (e.g. within a large firm) and for individuals owning multiple properties located in different jurisdictions. Importantly, for these incentives to play a role, there must be some variation in tax rates across locations. Furthermore, these differences must be somehow not entirely reflected in the quality of public goods, otherwise observed sorting in response to taxation may just reflect differences in preferences for public good provisions (à la Tiebout 1956).

A number of recent works has studied internal mobility responses to local tax differentials in the context of different countries. An important distinction drawn in the literature is between countries like the United States, where in most states local income taxation is based on the state where the individuals work (employment-based), and European countries where local income taxation is based on the location where individuals elect their residence (residence-based). Mobility responses have been documented in both contexts. In the US, Moretti and Wilson (2017) and Akcigit, Baslandze, and Stantcheva (2016) use patent data to show that inter-state migration of star scientists is highly sensitive to marginal tax rates differentials, while a few studies on millionaire taxes (Young and Varner 2011, Young et al. 2016 and Varner, Young, and Prohofsky 2018) find relatively smaller mobility responses. Agrawal and Hoyt (2017) show that in those metropolitan statistical areas that cross state boundaries and where there are reciprocity agreements (which make local income taxation residence-based), the high-tax side has fewer high-income individuals commuting interstate to work and shorter commuting times. In Europe, Agrawal and Foremny (2019) show that local taxes have a significant effect on the location choices of taxpayers between Spanish regions, conditional on moving and especially among richer taxpayers. Liebig, Puhani, and Sousa-Poza (2007), Martinez (2017) and Schmidheiny and Slotwinski (2018) find migration responses to local tax rates within Switzerland.

In this paper, we contribute to this strand of literature by studying the migration response to tax differentials between local jurisdictions in Italy. The Italian context offers a number of advantages for studying this question. The first is the availability of administrative data on bilateral transfers of residence at a very fine geographical level (approximately 8,000 by 8,000 municipalities). As local taxation is based on individuals' municipality of residence, these bilateral migration flows are the most accurate measure to study the effects of local taxation on the mobility of tax bases. The second is the context of fiscal decentralization: local jurisdictions are allowed to levy income taxes (in addition to the central government) and property taxes. Importantly, local income and property tax differentials between municipalities have increased as a result of a several national-level reforms, which created significant tax variation both over time and over space in our period of analysis, as we document in Section 3.2. The closest paper to ours is Rubolino (2021), who also studies mobility responses to local income tax differentials in Italy. We complement his findings by studying mobility responses at a fine geographical level (the municipality) and by also studying the responsiveness to local property tax differentials.

Exploiting the bilateral structure of our migration data to estimate regressions across municipality-pairs over time, in which we are able to control for all time-invariant determinants of migration flows between each specific pair as well as for time-varying origin-specific shocks, we uncover a large migration response to local tax differentials. First, consistent with the literature and with findings in Rubolino (2021), we estimate a large and precise migration elasticity to local *income* tax differentials during the period 2009-2015, which includes a reform in 2011 that allowed local jurisdictions to increase the level and the progressivity of their income taxation (Figure 3.1).

Our second and most novel finding is the large responsiveness to local *property* tax differentials. While properties (real estate) are an immobile asset, a series of reforms starting in 2009 introduced a preferential property taxation on the primary residence, i.e. the property where the individual homeowner has establishing their residence, relative to a non-primary residence (Figure 3.3). Interestingly, we find that, after the reforms, individuals systematically move their residence towards municipalities with *higher* property tax rates relative to the municipality of origin, unlike in the pre-reform period. This finding is consistent with at least two explanations. First, as individuals owning multiple properties[1] benefit from a preferential tax treatment on their primary residence, electing their primary residence in the high property tax location reduces their overall tax liability.[2] Second, to the extent that property taxes are incorporated into property values (Oates 1973), after the primary residence becomes exempted, prospective buyers who plan to establish their residence in the property benefit from

[1]Homeowners in Italy own on average 1.29 properties.

[2]This is the case if the two properties have a similar value or if the property of primary residence has a higher value.

a lower market price, ceteris paribus.[3]

Further, we document that post-reforms tax-induced migration is mostly a long-distance phenomenon, as opposed to the pre-reform period when it was mostly a local phenomenon. Taken together, our results point towards a large migration responsiveness to changes in local taxation. Importantly, the large, sudden and long-distance responsiveness in the period 2009-15 relative to the 2002-08 period, seems to suggest that the estimated migration response to local tax differentials entails a relocation of local tax bases (which is what our internal migration measure is capturing) which may not necessarily correspond to an actual relocation of individuals, which would happen if for instance individuals change their job location in response to local taxation. In other words, while our estimates do suggest that tax bases are highly responsive to local tax differentials, we are not able to differentiate the spatial mobility of human capital in response to tax differentials. Still, these elasticities are relevant for local policymakers, as they constrain their ability to raise local tax rates and the extent to which they can implement a progressive tax schedule. In a future extension of this work, we plan to use a different measure of internal migration which is more likely to capture human capital mobility, which would allow us to study the migration responsiveness of human capital to local tax differentials.

The remainder of this work is structured as follows. In Section 3.2 we desribe the institutional background and data sources. In Section 3.3 we discuss our empirical strategy and in Section 3.4 our results. Finally, Section 3.5 offers some concluding remarks and future avenues for this work.

3.2 Institutional background and data

In the following paragraphs we provide an overview of the Italian institutional setting and we describe the data sources used in this paper, including some descriptive figures and summary statistics. We begin with a brief overview of local taxation in Italy and

[3] In line with this hypothesis, Oliviero and Scognamiglio (2019) find that property tax values dropped by 2.7% after 2012 in municipalities that increased property taxes by one standard deviation.

we then turn to discuss our measure of internal migration.

3.2.1 Local taxation

Local taxation in Italy is articulated in two levels, regional and municipal. The 20 regions have their own local government and are responsible for public good provision in specific domains, mainly education and healthcare. These services are mainly funded by income taxes levied at the regional level (*Addizionale Regionale*), which tops the national-level income tax (*Imposta sul Reddito delle Persone Fisiche*, IRPEF), and by transfers from the central government, which guarantee a minimum level of public good provision whenever resources raised locally are not sufficient. The nearly 8,000 municipalities are also responsible for public good provisions (e.g. local transportation, garbage disposal, social services, nursing schools) which are financed by an additional income tax (*Addizionale Comunale*), levied on top of regional and national income taxes, property taxes (*Imposta Municipale Unica*, IMU) and also transfers from the central government. The following paragraphs describe the two main domains of local taxation, income taxation (both regional and municipal) and property taxation (municipal only).

Local income taxation

Local income taxes (henceforth, LIT) vary both over time and across space. As of 2016, regions are allowed to set income tax rates between 1.23% and 3.33% while municipalities up to 0.8% (0.9% for the capital Rome). Thus, the regional and municipality combined tax rates nominally vary between 1.23% and 4.23%. However, because local taxation mimics the progressivity of national tax system and often includes a wider no-tax area, the resulting average tax rates range from a minimum of 0.9% to a maximum of 3.08%, with a mean of about 2.1% (combined municipality and region).[4]

Figure 3.1 shows that local income tax rates have increased over time. Specifically, average tax rates increased substantially after two reforms in 2007 and in 2011, which

[4]The national income tax rates are 23% for incomes below 15,000 euros, 27% between 15,001-28,000 euros, 38% between 28,001-55,000 euros, 41% between 55,001-75,000 euros and 43% above 75,000 euros throughout the 2007-2018 period, with a no tax area for incomes under 8,000 euros.

allowed local governments to implement more progressive taxation schemes (Rubolino 2021). These changes also altered substantially the distribution of municipalities in terms of relatively low and relatively high tax rates. Figure 3.2 plots the geographical distribution of local income tax rates across Italian municipalities for two years, 2006 and 2016. Comparing the two maps reveals two patterns. First, while the regions in the top quartile of income tax rates in 2006 and 2016 are substantially the same (e.g. Latium and Campania), the distribution of regions has nonetheless changed substantially: regions like Emilia-Romagna and Tuscany moved from the bottom to the second quartile, while other regions such as Veneto have seen a decrease in their income tax rates relative to other regions. Second, within region variation has increased between 2006 and 2016 as well, as it is clearly visible for the southernmost island of Sicily.

Income taxation at the local level is based on the municipality of residence of the individual taxpayer (residence-based taxation), regardless of the municipality where income is produced (source-based taxation), such as the job location. The definition of residence for fiscal purposes coincides with the individual's residence in the civil registry[5], which determines the individual's voting polling station for national elections, assigns the right to vote in local elections (municipal and regional) and grants access to local public services such as healthcare.

Local income tax rates and tax bases are obtained from the Ministry of Economics and Finance (MEF). This publicly available data source includes several municipality-level variables such as the stock of taxpayers, total incomes by source (labor, capital, pension etc.), as well as region and municipality income tax rates, for each year between 2002-2018. We use these data to construct average local income tax rates and tax bases (local incomes), which are summarized in Table 3.1.

[5]For individuals owning multiple properties, their fiscal and civil residence also coincides with the location where they establish their "primary residence" for property tax purposes, which we discuss in the next section.

Local property taxation

Local property taxes (henceforth, LPT) are levied only at the municipality level - in contrast with local income taxes - and nearly all real properties such as real estate, commercial buildings and land are subject to local property taxes.[6] The value of the tax base is the *cadastral value* of the property, a measure of the expected income that the property should generate. Cadastral values are a sluggish measure of property values, often far below the market values, especially in coastal and mountain areas with a high shares of second homes.[7] In fact, as cadastral values are determined at the time of construction or after a major renovation and they are rarely updated, they are *de facto* a decreasing function of the age of the building (Bianchi, Giorcelli, and Martino 2021): for instance, older properties located in the historical centers of towns have often lower cadastral values than newer buildings in peripheral areas.

Besides cadastral value, the key determinant of individuals' property tax liability is whether the property is their primary residence (henceforth P1) or a secondary home (henceforth P2), such as a rental unit or a vacation house. Property tax rates are generally lower on primary residences, which also benefit from a nonrefundable deduction which reduces the tax liability, often to zero for properties with lower cadastral values.

Specifically, for an individual property i located in municipality m, the property tax liability in year t is then determined by the formula:

$$LPT_{imt} = \begin{cases} \lambda \cdot CV_{im} \cdot \tau_{mt}^{P1} - Ded_{mt} & \text{if } i = \text{primary} \\ \lambda \cdot CV_{im} \cdot \tau_{mt}^{P2} & \text{if } i \neq \text{primary} \end{cases}$$

where λ is a fixed multiplier[8], CV_{im} is the cadastral value, τ_{mt}^{P1} and τ_{mt}^{P2} are the property tax rate of municipality m in year t for primary (P1) and non-primary (P2) residences

[6] The main exception are buildings and land owned by the Catholic Church due to the Lateran Treaty between Italy and the Vatican in 1929.

[7] Source: https://www.lavoce.info/archives/91212/catasto-a-chi-conviene-che-resti-com-e/

[8] The multiplier is equal 168 throughout our analysis period, which is the product between the general revaluation rate of cadastral values, 1.05, and the specific multiplier applying to residential properties, 160.

respectively and Ded_{mt} is the deduction for primary residence. Thus, municipalities take as given the cadastral values of the properties located in their jurisdiction, and they are allowed to set τ_{mt}^{P1}, τ_{mt}^{P2} and Ded_{mt} within certain boundaries set by the central government.

Property taxes are especially salient in Italy, which is why they have been subject to a number of modifications in the past decades (Messina and Savegnago 2014). Two main reforms occurred during our period of analysis. The first major change was in 2008, when the Berlusconi government eliminated property taxes on primary residences starting from 2009.[9] The second happened in 2012 as a part of the fiscal consolidation undertaken by the Monti government in response to the Sovereign Debt crisis. Specifically, this reform reinstated the tax on primary residences while introducing two different ranges for τ_{mt}^{P1} and τ_{mt}^{P2} (the latter being higher) and also set a more generous and fixed deduction for primary residences. While the primary residence tax was formally abolished again in 2013-2014, in practice it was replaced by a new tax on local services (*Tributo per i servizi indivisibili*, TASI) which is basically a property tax under a different name (Oliviero and Scognamiglio 2019).

In Figure 3.3, we plot the evolution over time of property tax rates τ_{mt}^{P1} and τ_{mt}^{P2} and deductions Ded_{mt} for the median municipality and for municipalities at the 5th and 95th percentiles of the yearly distribution. From the graph, we can identify two specific periods based on the overall trends in local property taxation. The first period (until 2008) is characterized by a relatively homogeneous tax treatment for all properties, regardless of their primary residence status. Before 2009, in fact, all properties are subject to property taxes, with LPT rates ranging from 0.004 to 0.007 and with slightly lower rates for primary residences, which also benefit from a deduction of at least 100 euros. The second period (from 2009 onwards) is instead characterized by a preferential property tax treatment on primary residences. In 2009-2011, primary residences are completely exempted, resulting in a substantial difference in the tax

[9] The only exception was luxury buildings, which were still taxed at the pre-2009 primary residence rate and still benefited from the deduction.

treatment of properties depending on their primary residence status. After 2012, while all properties are again subject to local taxes, primary residences enjoy much lower tax rates (between 0.002 and 0.006) than secondary dwellings (between 0.0066 and 0.0106) as well as a 200 euros deduction which is twice the amount of the pre-2012 period.

In Figure 3.4, we then show the geographic variation of property tax rates, separately by primary residence status and for two years, 2008 (before the first reform) and in 2013 (after the second reform). The maps offer a number of insights about local property taxation. First, they reveal a considerable spatial variation in property tax rates, both between and within regions, particularly after the second reform (Figures (c) and (d)). Second, they show that, for nearly all municipalities, primary residences benefit from a preferential tax treatment relative to secondary houses after the reforms.

Property taxation is based on the municipality where the property is physically located. Therefore, properties are an immobile asset, unlike incomes that may react to local income taxation by relocating to another municipality. Nevertheless, the preferential tax treatment of primary residences may create incentives to relocate individual residence in response to property taxes, especially after 2009, when the wedge between property taxes on primary and non-primary residences increased substantially. For instance, an individual owning two properties with similar cadastral values in different municipalities could reduce their tax liability by locating their primary residence in the municipality where property tax rates are *higher*. Further, since the tax unit in Italy is the individual, couples owning multiple properties may have a even stronger incentive by relocating the fiscal residence of one or both spouses, again especially after 2009. The existence of these incentives motivates our focus the migration response to local property taxes, in addition to local income taxation.

Local property tax rates were obtained from the Institute for Finance and Local Economy (IFEL). This municipality-level dataset includes property tax rates both on primary and non-primary residences, as well as the deduction amount for primary residences, for each year between 2001 and 2017. Property tax bases (cadastral values) were instead obtained from the Italian Revenue Agency (*Agenzia delle Entrate*), and include municipality-level information on the total cadastral values and the number of real estate units, for each year starting from 2013 and for each type of unit (individual residences, business properties, etc.). We use these data to construct average cadastral values for individual residences.[10] Table 3.1 summarizes the key variables on local property taxation which are used in our analysis.

3.2.2 Internal migration and transfers of residence

Our main outcome of interest is the migration of tax bases across local jurisdictions in Italy. As explained in the previous paragraphs, the location of the individual's residence determines the municipality where their income is taxed, as well as the location where an individual homeowner is subject to a preferential property tax. For this reason, in this paper we use transfers of residence between municipalities as a measure of internal migration. While transfers of residence are likely an imperfect proxy for the actual relocation of human capital across municipalities, they are nonetheless a precise measure of relocation of tax bases across municipalities. Specifically, we use administrative data on individual residence changes from the Italian National Statistical Institute (Istat), which collects information from civil registries on all transfers of residence, which in turn determine population counts of the resident population at the municipality level. These data cover the period 2002-2015 and are further breakdown by age and sex.

Figure 3.5 offers a few stylized facts about internal migration, as proxied by transfers of residences. Figure (a) shows that about the majority of moves are within regions (about 70%) rather than between regions: about 100 individuals move from/to the average municipality to/from another municipality within the same region, while about 40 individuals move to/from a municipality in a different region. Figure (b) reveal that the overwhelming majority of within-region moves are between municipalities located within the same province[11], thus very close to each other. In Figure (c), instead, we

[10] Specifically, we use the two most common cadastral categories for individual residences: civil housing (*Abitazioni di tipo civile* - A02) and economic housing (*Abitazioni di tipo economico* - A03).

[11] There are about 100 provinces in Italy, which are clusters of nearby municipalities within regions.

breakdown between-region moves by their direction. South to North/Center regions account for the majority of inter-regional transfers, while the opposite flow (North/Center to South) is about half. Last, in Figure (d) we plot the average distance for intra-regional and inter-regional moves. The former are relatively short distance, covering 22-23 km on average, as opposed to the long distance for inter-regional moves, 465 km on average, which has been somewhat declining over time.

Last, Figure 3.6 plots immigration and emigration rates (in percentage of the population) of Italian municipalities for two periods, 2002-08 and 2009-15. The maps show that Northern and Central regions (especially Piedmont, Emilia-Romagna, Tuscany and Latium) experience much higher immigration and emigration rates than Southern regions. Last, the map shows large migration rates for municipalities located around large urban areas (e.g. Rome, Florence, Milan, Bologna, Genoa, Catania), where most individuals resident in the small surrounding municipalities are likely to have their job location.

3.3 Empirical strategy

To estimate the effects of local tax rates on mobility across municipality, we exploit the bilateral structure of the migration data to estimate regressions across municipality pairs over time. Let o denote the municipality of origin and d the municipality of destination of migrants, and let $\tau^{Y}_{m,t}$ and $\tau^{P2}_{m,t}$ denote respectively the local income tax rate and the local property tax rate (on non-primary residences, P2) of municipality m in year t. Following a methodology widely used in the literature of mobility responses to local taxation (e.g. Moretti and Wilson 2017), we estimate the following regression:

$$log\frac{mig_{odt}}{pop_{oot}} = \alpha log\frac{(1-\tau^{Y}_{dt})}{(1-\tau^{Y}_{ot})} + \beta log\frac{(1-\tau^{P2}_{dt})}{(1-\tau^{P2}_{ot})} + \gamma' log(\frac{Z_{dt}}{Z_{ot}}) + \lambda_t + \mu_{od} + \epsilon_{odt} \tag{3.1}$$

where mig_{odt}/pop_{oot} is the share of population of municipality o that moves to municipality d relative to the population share of o that does not move, $(1-\tau^{Y}_{dt})/(1-\tau^{Y}_{ot})$ and $(1-\tau^{P2}_{dt})/(1-\tau^{P2}_{ot})$ measure how lower the local income and property tax rates are in d relative to o, Z_{dt}/Z_{ot} is the ratio of control variables in the two municipalities (e.g. average local income), λ_t are time fixed effects and μ_{od} are municipality pair fixed effects, which absorb all the time-invariant characteristics of each municipality pairs.[12] The identifying variation in these model comes from changes in local tax rates between municipality pairs over time. To deal with potential time-varying location-specific unobserved shocks, which could be a threat to identification, in some specifications we also include region-pair by year fixed effects, to allow for heterogeneous trends across each pair of origin-location regions, as well as origin municipality by year and destination municipality by year dummies, which capture time-varying origin-specific and destination-specific unobserved shocks. Following Moretti and Wilson (2017), we cluster standard errors three-way at the municipality pair, origin municipality by year and destination municipality by year (Cameron, Gelbach, and Miller 2011).

Equation (3.1) can be formally derived from the pairwise equilibrium condition of a Rosen-Roback or a Tiebout model, in which individuals can choose where to reside among a number of locations which differ in terms of their local tax rates. Assuming that individuals choose their municipality of residence to maximize their net income and that unobserved idiosyncratic taste shocks for each location follow a i.i.d. extreme value distribution (McFadden 1974; Grogger and Hanson 2011b), then we can derive an expression for the log-migration-odds (population of o that moves to d in year t relative to the population of o that does not move) as a function of the log net-of-tax rates differentials as well as other relative characteristics of each pair (such as local amenities, local productivity etc.), either time-invariant (absorbed by the municipality-pair fixed effects) or time-varying (absorbed by time-varying controls as well as location-specific trends).

3.4 Results

In this section we present our results from estimating Equation (3.1). Because of the importance of the 2009 reform, which introduced a preferential property tax treatment

[12]These are ordered pairs: μ_{od} is distinct from μ_{do}.

on primary residences, we breakdown the estimation sample in two distinct periods, 2002-2008 and 2009-2015.

3.4.1 Baseline: municipality pairs over time

Table 3.2 show the estimated elasticity of migration to local income and property net-of-tax rates differentials, for the period 2009-2015. Our baseline specification in Column 1 includes municipality-pairs and year fixed effects but no control variables. The effect of income taxes is statistically significant and has the expected sign: a higher net-of-tax rate in the destination relative to the origin municipalities induces higher migration from origin to destination. The estimated elasticity is large and its magnitude is comparable to the estimates in the literature: a 1% decrease in the destination tax rate relative to the origin, which corresponds to moving from the median to the maximum of the nationwide distribution of local tax rates, increases migration by 1.75%.

The coefficient of the property net-of-tax rate ratio has, on the contrary, a negative sign, which indicates that individuals relocate towards areas with higher property tax rates relative to their origin. This is consistent with the preferential tax treatment of primary residences in the post-2009 period. The coefficient is highly significant and, taken at face value, suggests an elasticity over three. While we should not over-interpret the magnitude of this coefficient, since cadastral values vary widely between neighborhoods within municipalities, it is important to notice that the tax base of local income taxation in the average municipality (about 20,000 euros in 2008, Table 3.1) is three times lower than the tax base of local property taxation (about 60,000 in 2008).

In Column 2 and in the following columns, we include the relative local average incomes between destination and origin municipalities. While including this control does not affect the estimated coefficients of LIT and LPT differentials, it is interesting to notice that its own coefficient is negative and statistically significant, suggesting that transfers of residence in the 2009-15 period are systematically from richer to poorer areas, in terms of average incomes. This is instead not the case for the period 2002-08

(Table 3.3), as we discuss below.

In Column 3, we further saturate our regressions by including region-pair by year fixed effects, which allows for differential trends for each region pair. The estimated coefficients of the tax rates ratios are still statistically significant albeit slightly lower in magnitude. In Columns 4 and 5, we include municipality of origin (destination) by year fixed effects, which absorb all origin (destination) specific time-varying determinants of internal migration. Both coefficients of LIT and LPT are robust to the inclusion of origin-year fixed effects, although the coefficient of LPT is smaller and insignificant when including destination-year fixed effects.

To corroborate the hypothesis that the estimated response to local property tax rates differential is a function of the preferential tax treatment of primary residences post-2009, in Table 3.3 we show the results of estimating the same regressions on the pre-reform period, 2002-2008. While the coefficients for local income taxes are close to zero and insignificant in all specifications, the effect for property taxes is marginally significant in most specifications and is positive, consistently with individuals moving from areas with higher property tax rates to areas with lower tax rates. Individuals in this period move from poorer to richer places, as shown by the coefficient on the average income ratio, and not much in response to tax incentives, which in the pre-2009 are very low, both in terms of local income and property taxes (Figures 3.1 and 3.3).

3.4.2 Short vs. long distance mobility

Transfers of residence between municipalities can involve a very short distance, in case of transfers within provinces (23-24 km on average), but also a long distance across the Italian peninsula, as in the case of inter-regional transfers (470 km on average), as shown in Section 3.2. For this reason, one may wonder to what extent our estimated migration responses to tax differential are driven by short vs. long distance moves. In Tables 3.4 and 3.5 we tackle this geographical heterogeneity by breaking down the estimation samples in short distance vs. long distance migrations. As for the previous tables, we run separate regressions for the periods 2009-2015 and 2002-2008. Column

1 of both tables report our baseline specification (Column 1 of Tables 3.2 and 3.3) for comparisons. In columns 2 and 3 we only include inter-regional and intra-regional moves respectively, while in columns 4 and 5 we breakdown within-region moves by whether they are between provinces (in the same region) or within the same province.

For the period 2009-15 (Table 3.4), while the coefficients are significant for all types of moves, we see that the effects of LIT and LPT differentials are stronger for longer distance moves, i.e. transfers between regions and between provinces located within the same region. The picture is different for 2002-2008 period (Table 3.5), when the only significant response occurs in the case of short-distance moves, within regional boundaries and within provinces. Overall, these results are suggestive that the increased local income tax differentials and the preferential property tax treatment of primary residences after 2009 may have contributed to inducing long-distance relocation of individuals' residences.

3.4.3 Timing of the responses and heterogeneity

In Tables 3.6 and 3.7 we investigate the timing of mobility responses to tax changes for both periods. Specifically, we show the estimates for our baseline specification on lags and leads of the migration rates, limiting the sample to a balanced panel of municipalities. The coefficients of LIT differentials are larger and more precise for the same year and for one year after, suggesting that mobility responses occur in the short-run. The same is true of LPT differentials, although we also see highly significant LPT coefficients also for the lags, which could be attributable to anticipation effects in the post-2009 periods (e.g. due to expectations about) but also to the serial correlation of property tax rates over time within each period.

Last, in Figure 3.7 we show the estimated coefficients from our baseline specification (for the period 2009-15) separately for demographic subgroups defined by sex and age. While the elasticity to local income tax differentials appear rather homogeneous across subgroups, we observe a much higher migration responsiveness to local property tax differentials for older individuals, consistent with higher home-ownership rates for these groups.[13]

3.5 Conclusions

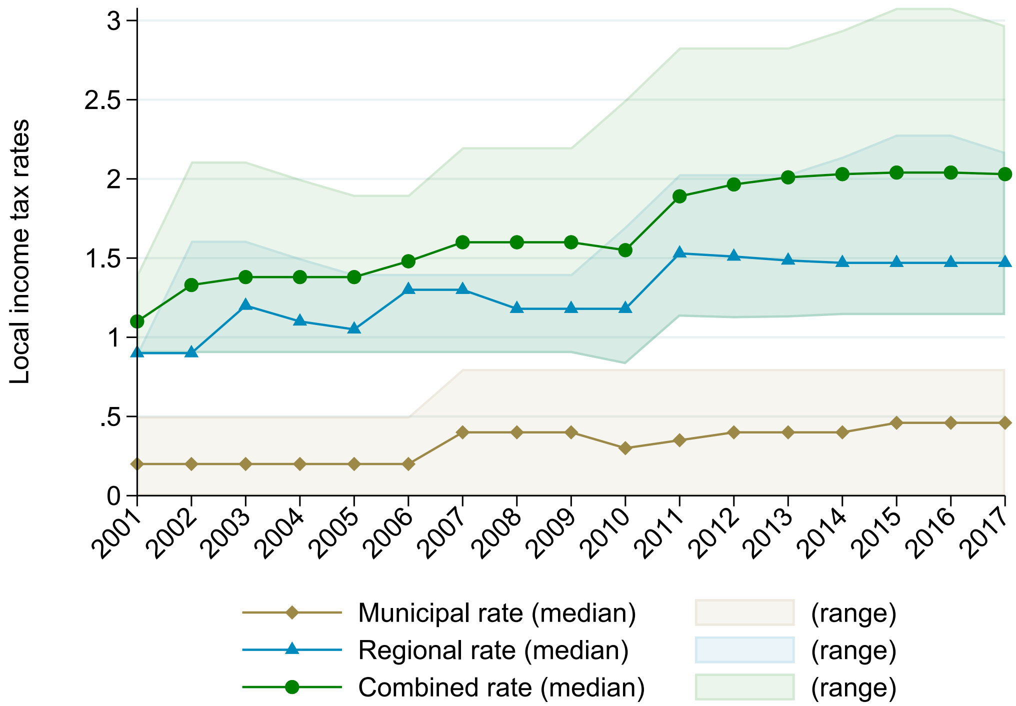
Tax differentials between jurisdictions create incentives for individual to relocate to minimize tax liability, with potentially important consequences for local economies. In this paper, we study the Italian context, which offers the availability of granular, bilateral data on transfers of residence between local jurisdictions, as well as a context of fiscal decentralization with increasing local tax differentials across locations over time. By estimating regressions across municipality-pairs over time, we show that local income and property tax differentials influence the location of tax bases across jurisdictions within the national boundaries. Specifically, we find evidence of large migration responses both to local income and property tax differentials concentrated in the post-2009 period, which was characterized by larger local income tax differentials and by a preferential property tax treatment of primary residences.

Despite our findings are suggestive of the fact that a substantial fraction of the moves in response to tax differential may be a reporting phenomenon – consistent with the findings in Rubolino (2021) –, an open question is to what extent jurisdictions that lose their tax bases due to tax-induced outmigration experience also a loss in their *human capital* stock, which could be detrimental for economic growth of local areas. To investigate this question, we plan to replicate our analysis by using a new measure of migration flows, which capture movements of human capital across municipalities. Specifically, we will use a novel administrative data source linking individuals' job location to their residence, for each municipality pairs and for all years in the 2013-2018 period. Finding out whether these human-capital related migration flows are also affected by local tax differentials could have important policy implications, as it would determine whether local policymakers should worry about tax-induced migration only in terms of its impact on local tax revenues or also on their human capital stock.

[13] Home-ownership rates in Italy exhibit a steep gradient by age: 0.7% for under 20 years old, 11% for 21-30 years old, 47% for 31-50 years old, 61% for 51-70 years old and 69% for over 70 years old.

Figures

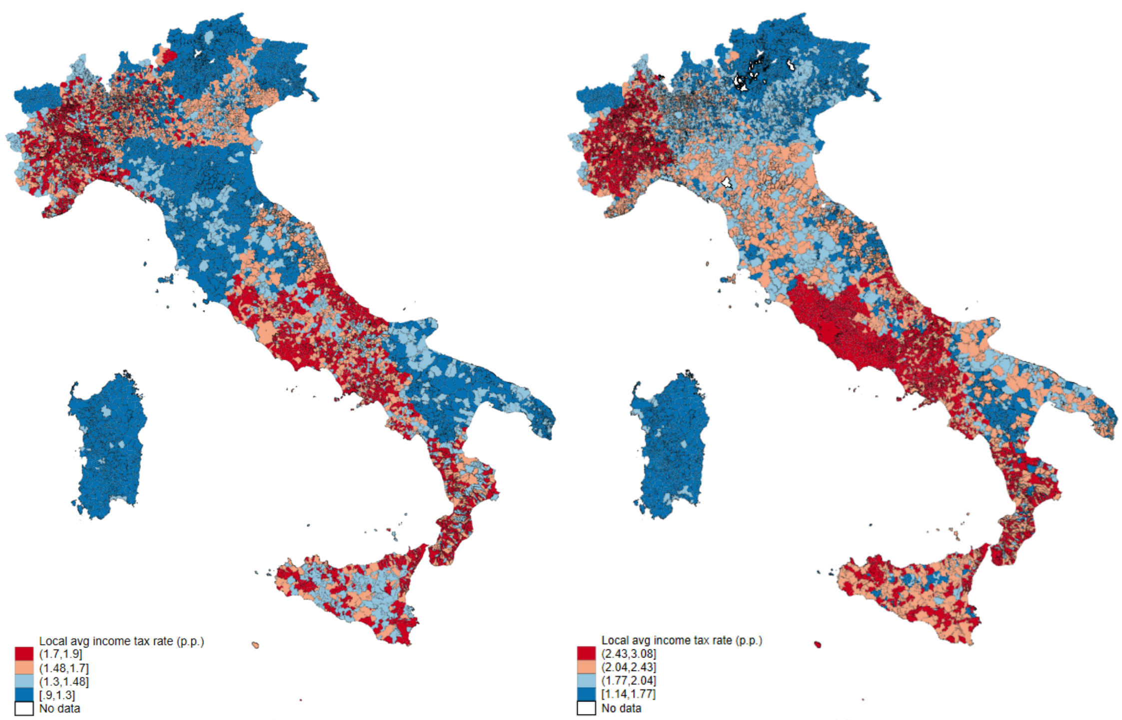
Figure 3.1: Local income tax rates over time



Notes: Average local income tax rates in percentage points. Source: Italian Ministry for Economics and Finance.

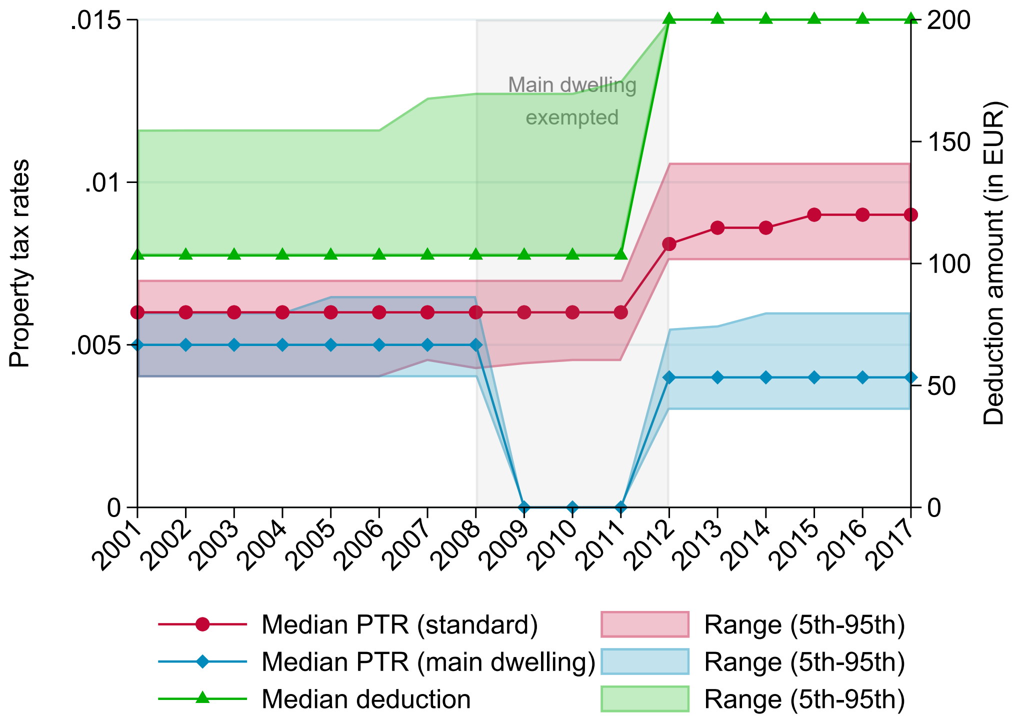
Figure 3.2: Local income tax rates, by year

(a) 2006 (b) 2016



Notes: Average local income tax rates in percentage points, combining the regional and the municipality income tax rates. Source: Italian Ministry for Economics and Finance.

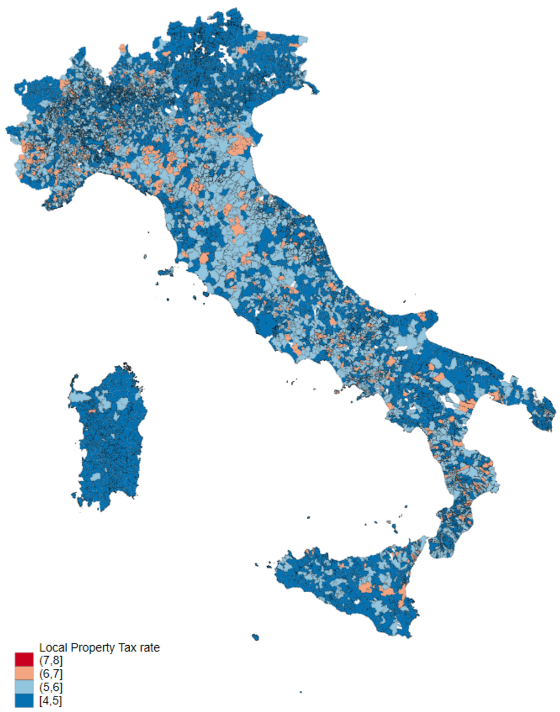
Figure 3.3: Local property tax rates over time



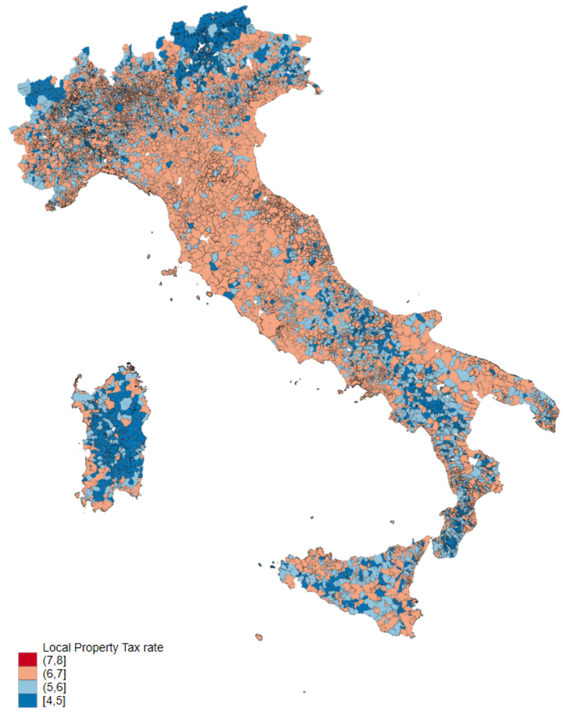
Notes: Local property tax rates are in percentage points. Source: Institute for Finance and Local Economy.

Figure 3.4: Local property tax rates, by primary residence status and year (2008 and 2013)

(a) Primary residence (2008)

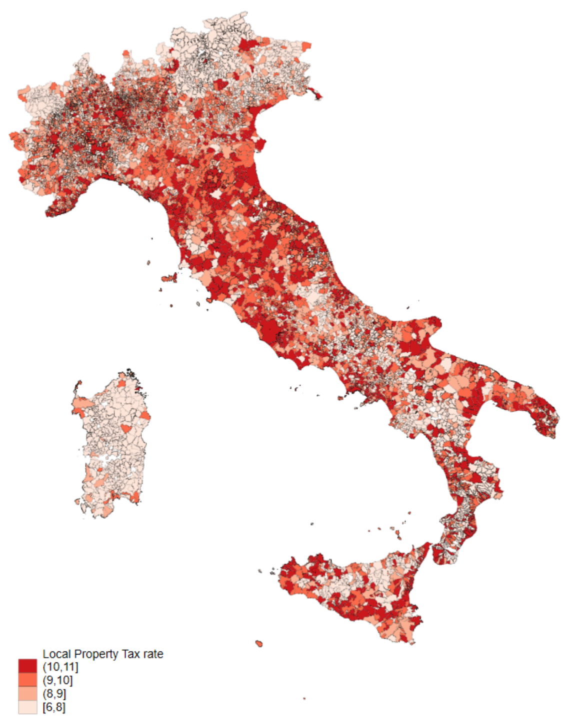


(b) Secondary dwelling (2008)



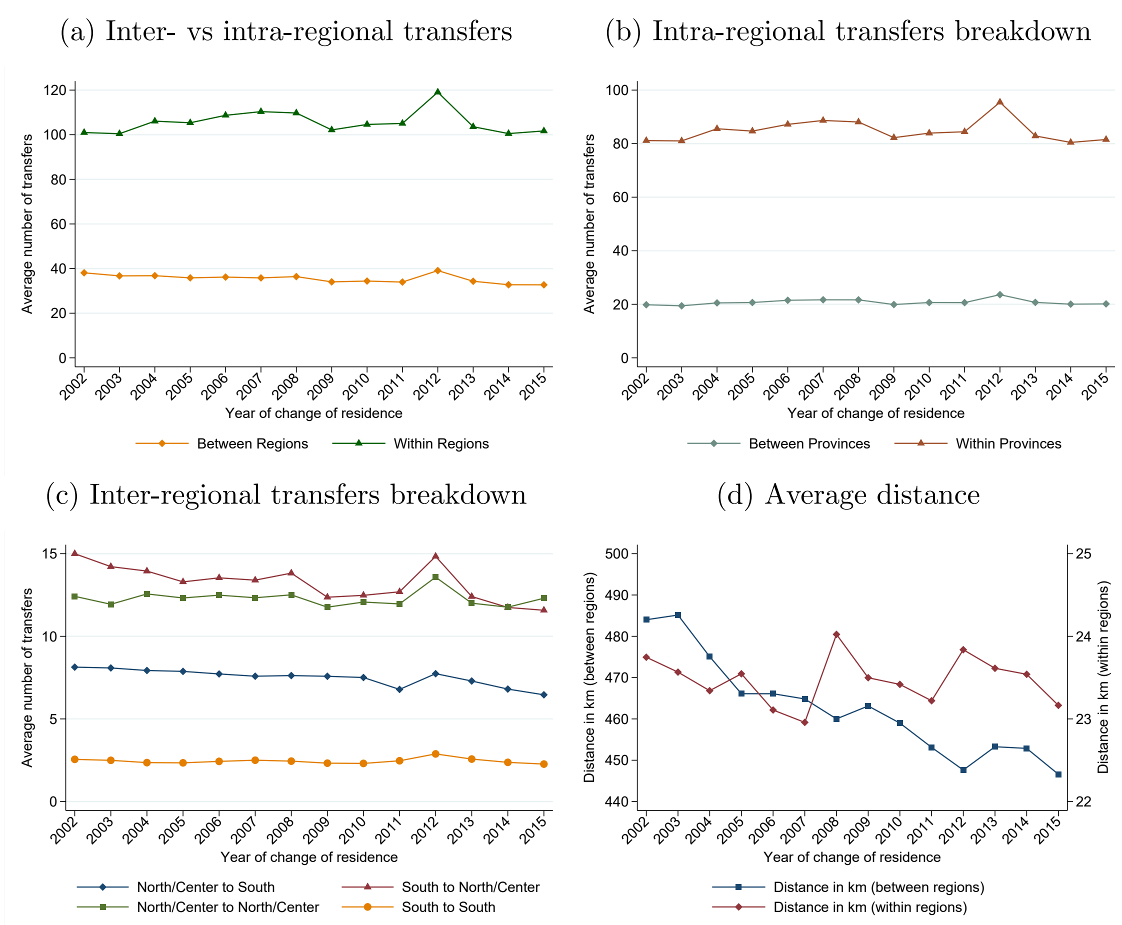
(c) primary residence (2013)

(d) Secondary dwelling (2013)



Notes: Local property tax rates are expressed per 1,000 euros. Source: Institute for Finance and Local Economy.

Figure 3.5: Trends in internal migration



(a) Inter- vs intra-regional transfers

(b) Intra-regional transfers breakdown

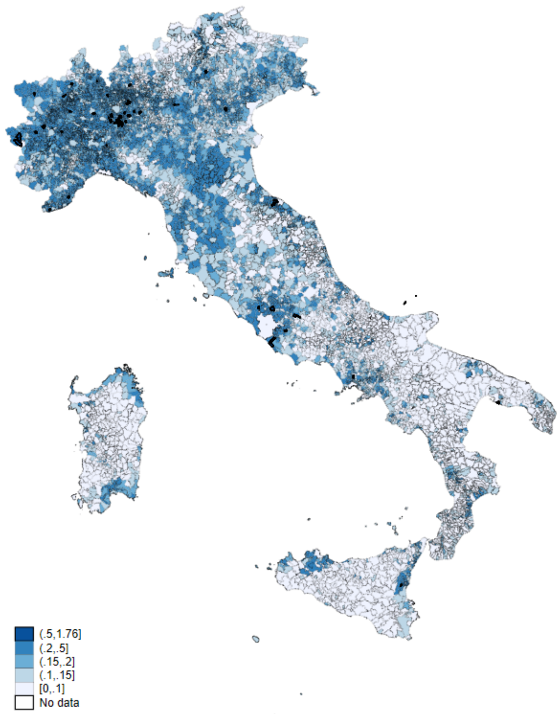
(c) Inter-regional transfers breakdown

(d) Average distance

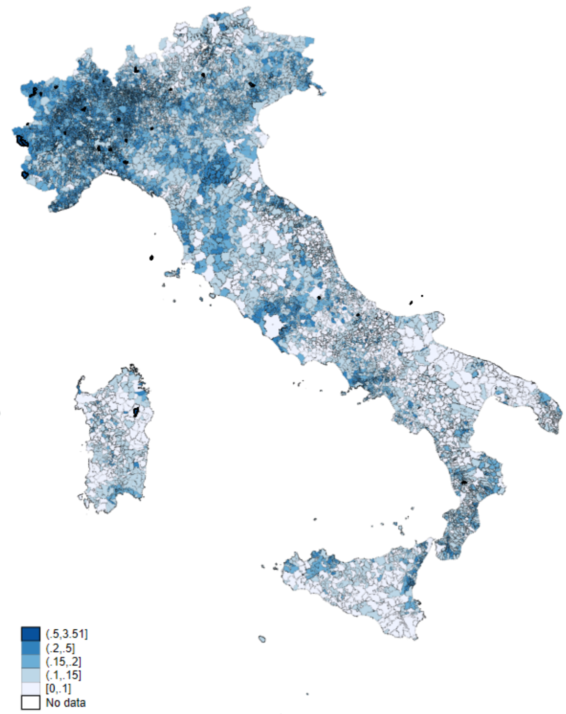
Notes: Figures (a)-(b)-(c) plot the number of transfers of residence from/to the average municipality in each year. Source: authors' elaboration on Istat data.

Figure 3.6: Internal immigration and emigration rates (transfers of residence)

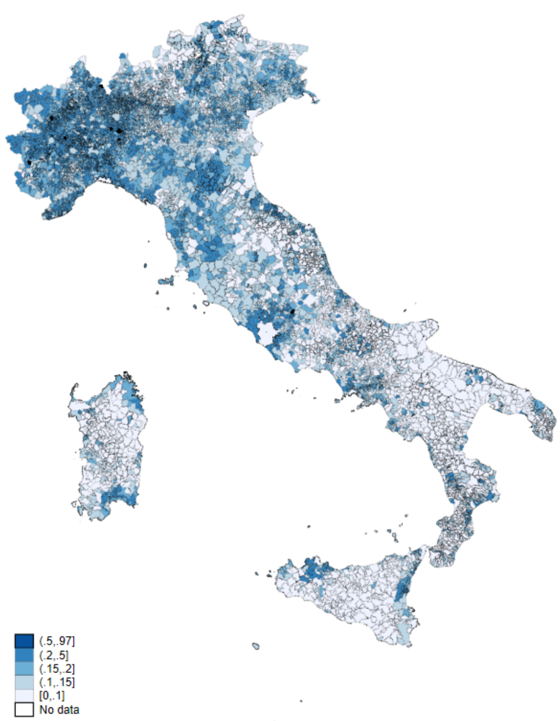
(a) Immigration 2002-2008



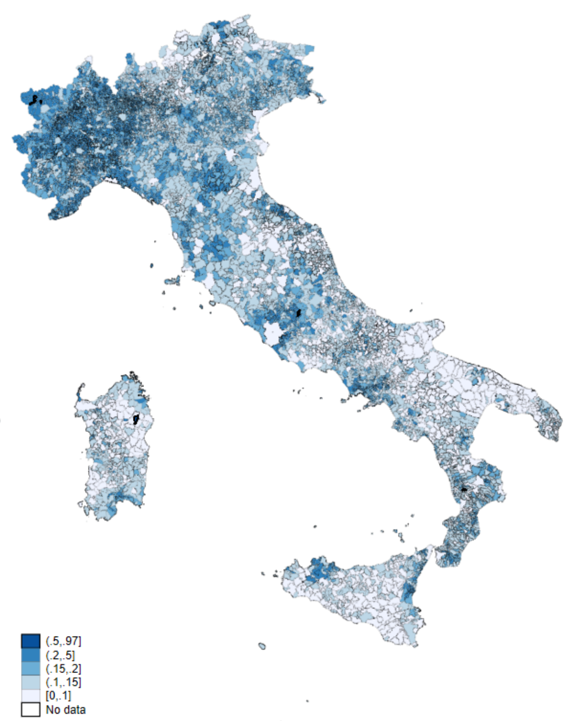
(b) Emigration 2002-2008



(c) Immigration 2009-2015



(d) Emigration 2009-2015

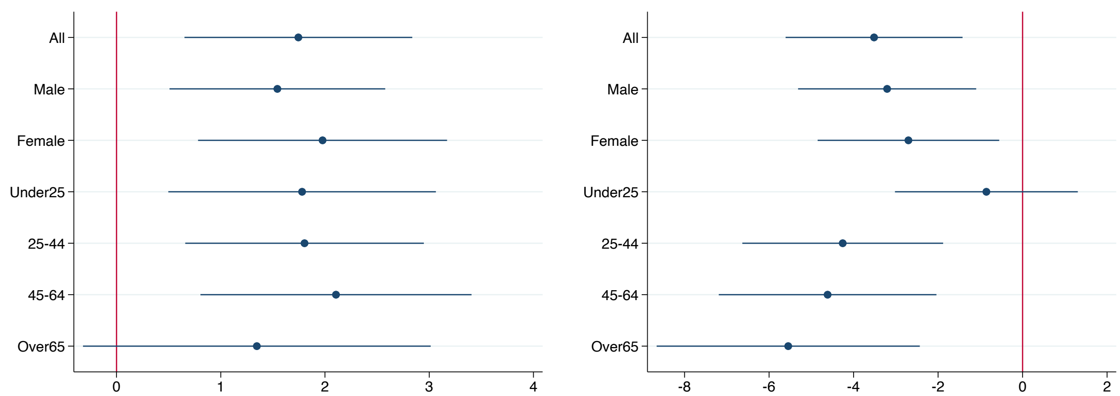


Notes: Source: authors' elaboration on Istat data.

Figure 3.7: Tax-induced migration across municipality pairs, 2009-2015, by demographic subgroup

(a) Local Income NTR ratio

(b) Local Property NTR ratio



Notes: Estimated regression coefficients and 95% confidence intervals. NTR is an abbreviation for Nex-of-Tax Rate. Observations: municipality pairs by year (2009-2015). The dependent variable is the log share of movers from origin to destination municipality relative to the population in the origin that does not move. The independent variables are the log income and property net-of-tax rate ratios between destination and origin municipalities. Controls variables include municipality pairs FE and year FE. Three way clustered standard errors (municipality pair, destination municipality by year and origin municipality by year).

Tables

Table 3.1: Summary statistics on local taxation

year	Variable	N	Mean	p50	SD	p1	p99
2002	Avg LIT rate	7912	1.34	1.33	0.32	0.90	2.11
	Avg LPT rate (primary)	7920	5.00	5.00	1.11	0.00	7.00
	Avg LPT rate (non-primary)	7920	5.76	6.00	0.83	4.00	7.00
	Avg Deduction	7920	112	103	41	0	258
	Avg Local Income	7912	12,963	12,924	3,053	7,431	20,277
	Avg Cadastral Value	7901	377	368	122	163	738
	Avg Property Tax Base	7901	63,414	61,783	20,476	27,327	123,977
	Avg LIT payment	7911	176.63	171.59	66.31	70.32	331.87
	Avg LPT payment (primary)	7901	365.71	348.74	134.83	148.80	769.20
	Avg LPT payment (non-primary)	7925	205.90	199.04	111.31	0.00	511.62
2008	Avg LIT rate	7911	1.57	1.60	0.34	0.90	2.20
	Avg LPT rate (primary)	7921	5.13	5.00	0.94	0.00	7.00
	Avg LPT rate (non-primary)	7921	6.08	6.00	0.98	4.00	7.00
	Avg Deduction	7921	115	103	42	0	258
	Avg Local Income	7912	19,728	19,467	3,001	14,251	28,155
	Avg Cadastral Value	7901	377	368	122	163	738
	Avg Property Tax Base	7901	63,414	61,783	20,476	27,327	123,977
	Avg LIT payment	7911	308.48	312.10	79.71	141.49	497.45
	Avg LPT payment (primary)	7901	386.67	371.27	146.98	130.39	833.20
	Avg LPT payment (non-primary)	7924	207.36	198.70	109.43	0.00	525.18
2015	Avg LIT rate	7903	2.07	2.04	0.46	1.14	3.08
	Avg LPT rate (primary)	7620	4.23	4.00	0.88	2.00	6.00
	Avg LPT rate (non-primary)	7620	8.96	9.00	1.17	6.60	10.60
	Avg Deduction	7620	199	200	35	200	200
	Avg Local Income	7913	21,417	21,298	3,278	15,023	30,784
	Avg Cadastral Value	7901	377	368	122	163	738
	Avg Property Tax Base	7901	63,414	61,783	20,476	27,327	123,977
	Avg LIT payment	7903	441.42	437.08	115.63	209.28	732.78
	Avg LPT payment (primary)	7607	567.37	541.58	211.69	226.83	1243.75
	Avg LPT payment (non-primary)	7924	77.13	49.61	95.30	0.00	402.67

Notes: authors' elaboration on data from the Ministry of Economics and Finance and from the Institute for Finance and Local Economy. LIT and LPT denote Local Income Tax (both regional and municipal) and Local Property Tax respectively. LIT rates are in percentage points while LPT rates are in tenths of a percentage point. The nonrefundable Deduction reduces property tax liability for primary residences. Cadastral values refer to the year 2013. Property tax base is the product between the average cadastral value and the multiplier (160 times 1.05, i.e. 168). Average LIT payment is the product between the average LIT rate and the average municipal income. Average LPT payments are estimated by multiplying the LPT rates and the Property Tax Bases and subtracting the deduction for primary residences.

Table 3.2: Tax-induced migration across municipality pairs, 2009-2015

	Outcome: Log Migration Odds Ratio, Period: 2009-2015				
	(1)	(2)	(3)	(4)	(5)
Log Income NTR ratio	1.745***	1.733***	1.108**	2.285***	1.201**
	(0.557)	(0.558)	(0.465)	(0.814)	(0.523)
Log Property NTR ratio	-3.514***	-3.430***	-2.636***	-4.515***	-1.243
	(1.068)	(1.070)	(0.861)	(1.627)	(1.106)
Log Avg Income ratio		-0.094***	-0.127***	-0.094**	-0.094**
		(0.031)	(0.030)	(0.047)	(0.042)
Observations	1,346,775	1,346,775	1,346,631	1,344,633	1,343,774
R-squared	0.943	0.943	0.943	0.945	0.946
Municipality-pair FE	X	X	X	X	X
Year FE	X	X	X	X	X
Region-pair by Year FE			X		
Origin by Year FE				X	
Destination by Year FE					X

Notes: Observations: municipality pairs by year (2009-2015). The dependent variable is the log share of movers from origin to destination municipality relative to the population in the origin that does not move. The independent variables are the log income and property net-of-tax rate ratios between destination and origin municipalities. Three way clustered standard errors (municipality pair, destination municipality by year and origin municipality by year) in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 3.3: Tax-induced migration across municipality pairs, 2002-2008

	Outcome: Log Migration Odds Ratio, Period: 2002-2008				
	(1)	(2)	(3)	(4)	(5)
Log Income NTR ratio	-0.379	-0.331	1.044	0.197	-0.939
	(0.578)	(0.580)	(0.671)	(0.602)	(0.740)
Log Property NTR ratio	3.828*	3.721*	3.379*	4.566	2.701
	(2.021)	(2.014)	(1.881)	(3.011)	(2.475)
Log Avg Income ratio		0.093***	0.071***	0.084***	0.078***
		(0.018)	(0.020)	(0.025)	(0.021)
Observations	1,314,100	1,313,246	1,313,103	1,311,103	1,309,919
R-squared	0.941	0.941	0.941	0.944	0.945
Municipality-pair FE	X	X	X	X	X
Year FE	X	X	X	X	X
Region-pair by Year FE			X		
Origin by Year FE				X	
Destination by Year FE					X

Notes: Observations: municipality pairs by year (2009-2015). The dependent variable is the log share of movers from origin to destination municipality relative to the population in the origin that does not move. The independent variables are the log income and property net-of-tax rate ratios between destination and origin municipalities. Three way clustered standard errors (municipality pair, destination municipality by year and origin municipality by year) in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 3.4: Tax-induced migration across municipality pairs: between vs. within Regions and Provinces, 2009-2015

	Outcome: Log Migration Odds Ratio, Period: 2009-2015				
	(1)	(2)	(3)	(4)	(5)
	All	Between Reg	Within Reg	Between Prov	Within Prov
Log Income NTR ratio	1.745***	2.493***	1.281**	2.077***	1.354**
	(0.557)	(0.942)	(0.542)	(0.770)	(0.611)
Log Property NTR ratio	-3.514***	-5.094***	-2.758***	-4.951***	-1.698*
	(1.068)	(1.906)	(0.949)	(1.558)	(1.009)
Observations	1,346,775	468,221	878,554	740,849	605,926
R-squared	0.943	0.950	0.915	0.946	0.886
Municipality-pair FE	X	X	X	X	X
Year FE	X	X	X	X	X

Notes: Observations: municipality pairs by year (2009-2015). The dependent variable is the log share of movers from origin to destination municipality relative to the population in the origin that does not move. The independent variables are the log income and property net-of-tax rate ratios between destination and origin municipalities. Three way clustered standard errors (municipality pair, destination municipality by year and origin municipality by year) in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 3.5: Tax-induced migration across municipality pairs: between vs. within Regions and Provinces, 2002-2008

	Outcome: Log Migration Odds Ratio, Period: 2002-2008				
	(1)	(2)	(3)	(4)	(5)
	All	Between Reg	Within Reg	Between Prov	Within Prov
Log Income NTR ratio	-0.379	-0.844	2.025**	-0.684	1.910**
	(0.578)	(0.669)	(0.801)	(0.648)	(0.938)
Log Property NTR ratio	3.828*	3.827	3.486	1.532	6.065**
	(2.021)	(3.211)	(2.300)	(2.633)	(2.644)
Observations	1,314,100	478,206	835,894	732,169	581,931
R-squared	0.941	0.948	0.911	0.944	0.878
Municipality-pair FE	X	X	X	X	X
Year FE	X	X	X	X	X

Notes: Observations: municipality pairs by year (2002-2008). The dependent variable is the log share of movers from origin to destination municipality relative to the population in the origin that does not move. The independent variables are the log income and property net-of-tax rate ratios between destination and origin municipalities. Three way clustered standard errors (municipality pair, destination municipality by year and origin municipality by year) in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 3.6: Lags and leads of migration rates, 2009-2015

	Outcome: Log Migration Odds Ratio, Period: 2009-2015				
	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t-1	t-2
Log Income NTR ratio	3.295***	3.023***	0.527	0.812	2.019*
	(1.046)	(0.912)	(1.114)	(1.222)	(1.104)
Log Property NTR ratio	-6.858***	-7.007***	-2.521	-8.276***	-5.856***
	(1.963)	(1.902)	(1.810)	(2.174)	(2.008)
Observations	480,384	402,261	480,384	401,922	480,384
R-squared	0.934	0.933	0.936	0.931	0.932
Municipality-pair FE	X	X	X	X	X
Year FE	X	X	X	X	X

Notes: Observations: municipality pairs by year (2009-2015). The dependent variable is the log share of movers from origin to destination municipality relative to the population in the origin that does not move in year t+d, $d \in \{0, 1, 2, -1, -2\}$. The independent variables are the log income and property net-of-tax rate ratios between destination and origin municipalities. Three way clustered standard errors (municipality pair, destination municipality by year and origin municipality by year) in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 3.7: Lags and leads of migration rates, 2002-2008

	Outcome: Log Migration Odds Ratio, Period: 2002-2008				
	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t-1	t-2
Log Income NTR ratio	2.042	3.161*	4.025***	0.233	-2.039
	(1.495)	(1.740)	(1.516)	(1.435)	(1.286)
Log Property NTR ratio	3.525	7.845	-3.284	6.015	0.110
	(5.244)	(5.075)	(4.642)	(5.182)	(5.247)
Observations	320,892	266,929	320,892	265,354	320,892
R-squared	0.938	0.937	0.939	0.936	0.937
Municipality-pair FE	X	X	X	X	X
Year FE	X	X	X	X	X

Notes: Observations: municipality pairs by year (2002-2008). The dependent variable is the log share of movers from origin to destination municipality relative to the population in the origin that does not move in year t+d, $d \in \{0, 1, 2, -1, -2\}$. The independent variables are the log income and property net-of-tax rate ratios between destination and origin municipalities. Three way clustered standard errors (municipality pair, destination municipality by year and origin municipality by year) in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

References

Acemoglu, Daron, Ufuk Akcigit, and Murat Alp Celik. 2017. "Young, Restless and Creative: Openness to Disruption and Creative Innovations." NBER Working Paper 19894 (updated).

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019. "Shift-share designs: Theory and inference." *The Quarterly Journal of Economics* 134 (4): 1949–2010.

Adda, Jérôme, Christian Dustmann, and Joseph-Simon Görlach. 2016. "The dynamics of return migration, human capital accumulation, and wage assimilation." *Unpublished manuscript.*

Agrawal, David R., and Dick Foremny. 2019. "Relocation of the Rich: Migration in Response to Top Tax Rate Changes from Spanish Reforms." *Review of Economics and Statistics* 101(2):214–232.

Agrawal, David R., and William H. Hoyt. 2017. "Commuting and taxes: theory, empirics and welfare implications." *The Economic Journal* 128:2969–3007.

AIRE. 2015. "Registry of Italians resident abroad (AIRE)." Italian Ministry of Interior [publisher].

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva. 2016. "Taxation and the International Mobility of Inventors." *American Economic Review* 106 (10): 2930–81.

Anelli, Massimo, Gaetano Basso, Giuseppe Ippedico, and Giovanni Peri. 2020. "Does Emigration Drain Entrepreneurs?" IZA Discussion Paper No. 13390.

Anelli, Massimo, and Giovanni Peri. 2017. "Does emigration delay political change? Evidence from Italy during the Great Recession." *Economic Policy* 32 (91): 551–596.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda. 2020. "Age and High-Growth Entrepreneurship." *American Economic Review: Insights* 2 (1): 65–82.

———. 2021. "Immigration and Entrepreneurship in the United States." Forthcoming, *American Economic Review: Insights*.

Barker, Vincent L., and George C. Mueller. 2002. "CEO Characteristics and Firm R&D Spending." *Management Science* 48 (6): 782–801.

Beaudry, Paul, David A. Green, and Ben M. Sand. 2018. "In Search of Labor Demand." *American Economic Review* 108 (9): 2714–2757.

Becker, Sascha O, Andrea Ichino, and Giovanni Peri. 2004. "How Large is the "Brain Drain" from Italy?" *Giornale degli economisti e annali di economia*, pp. 1–32.

BFS. 2018. "International Migrations of the Permanent Resident Population by Citizenship, Sex, and Age, 1991-2016." Switzerland Federal Statistical Office (BFS) [publisher] (accessed July 2, 2018).

Bianchi, Nicola, Michela Giorcelli, and Enrica Maria Martino. 2021. "The effects of fiscal decentralization on publicly provided services and labor markets." NBER Working Paper 29538.

Boeri, Tito, Herbert Brücker, Frédéric Docquier, and Hillel Rapoport. 2012. *Brain drain and brain gain the global competition to attract high-skilled migrants*. Oxford University Press.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2021. "Quasi-Experimental Shift-Share Research Designs." Forthcoming, *The Review of Economic Studies*.

Brücker, H., S. Capuano, and A. Marfouk. 2013. "Education, gender and international migration: insights from a panel-dataset 1980-2010." mimeo.

Burchardi, Konrad B, Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J Terry. 2020. "Immigration, Innovation, and Growth." NBER Working Paper 27075.

Bütikofer, Aline, and Giovanni Peri. 2021. "How Cognitive Ability and Personality Traits Affect Geographic Mobility." *Journal of Labor Economics* 39 (2): 559–595.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2011. "Robust inference with multiway clustering." *Journal of Business & Economic Statistics* 29 (2): 238–249.

Chernozhukov, Victor, and Christian Hansen. 2008. "The reduced form: A simple approach to inference with weak instruments." *Economics Letters* 100 (1): 68–71.

Ciccarelli, Carlo, Matteo Gomellini, and Paolo Sestito. 2019. "Demography and Productivity in the Italian Manufacturing Industry: Yesterday and Today." CEIS Research Paper 457, Tor Vergata University.

Czaika, Mathias, and Christopher R Parsons. 2017. "The gravity of high-skilled migration policies." *Demography* 54 (2): 603–630.

Del Carpio, Ximena V, Caglar Ozden, Mauro Testaverde, and Mathis Wagner. 2016. "Global Migration of Talent and Tax Incentives." World Bank Policy Research Working Paper 7875.

Di Giovanni, Julian, Andrei A. Levchenko, and Francesc Ortega. 2015. "A Global View Of Cross-Border Migration." *Journal of the European Economic Association* 13 (1): 168–202.

Docquier, Frédéric, Çağlar Ozden, and Giovanni Peri. 2014. "The labour market effects of immigration and emigration in OECD countries." *The Economic Journal* 124 (579): 1106–1145.

Docquier, Frédéric, Çağlar Özden, and Giovanni Peri. 2014. "The Labour Market Effects of Immigration and Emigration in OECD Countries." *Economic Journal* 124 (579): 1106–1145.

Docquier, Frédéric, and Hillel Rapoport. 2012a. "Globalization, Brain Drain, and Development." *Journal of Economic Literature* 50 (3): 681–730.

———. 2012b. “Globalization, brain drain, and development.” *Journal of Economic Literature* 50 (3): 681–730.

Dorn, David, and Josef Zweimuller. 2021. “Migration and labor market integration in Europe.” *Journal of Economic Perspectives* 35 (2): 49–76.

Dustmann, Christian, Tommaso Frattini, and Anna Rosso. 2015. “The Effect of Emigration from Poland on Polish Wages.” *Scandinavian Journal of Economics* 117 (2): 522–564.

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

Elsner, Benjamin. 2013a. “Does Emigration Benefit the Stayers? Evidence from EU Enlargement.” *Journal of Population Economics* 26(2):531–553.

———. 2013b. “Emigration and wages: The EU enlargement experiment.” *Journal of International Economics* 91 (1): 154–163.

Engbom, Nicklas. 2019. “Firm and Worker Dynamics in an Aging Labor Market.” Federal Reserve Bank of Minneapolis Working Paper series, no. 756.

Fackler, Thomas A, Yvonne Giesing, and Nadzeya Laurentsyeva. 2018. “Knowledge remittances: Does emigration foster innovation?” CESifo Working Paper No. 7420.

Fairlie, Robert W., and Magnus Lofstrom. 2015. “Immigration and Entrepreneurship.” CESifo Working Paper Series No. 5298.

Fouka, Vasiliki, Soumyajit Mazumder, and Marco Tabellini. 2020, October. “From Immigrants to Americans: Race and Assimilation During the Great Migration.” Harvard Business School Working Paper, No. 19-018.

Giesing, Yvonne, and Nadzeya Laurentsyeva. 2017a. “Firms Left Behind: Emigration and Firm Productivity.” CESifo Working Papers, no. 6815.

———. 2017b. “Firms left behind: Emigration and firm productivity.” CesIfo Working Papers, no. 6815.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110 (8): 2586–2624.

Grogger, Jeffrey, and Gordon H. Hanson. 2011a. "Income Maximization and the Selection and Sorting of International Migrants." *Journal of Development Economics* 95 (1): 42–57.

Grogger, Jeffrey, and Gordon H Hanson. 2011b. "Income maximization and the selection and sorting of international migrants." *Journal of Development Economics* 95 (1): 42–57.

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics* 2 (2): 31–56.

Infocamere. 2016. "Special registry of innovative start-ups." InfoCamere SCpA - Chambers of commerce digital innovation [publisher]. URL: `https://startup.registroimprese.it/` (accessed July 11, 2016).

———. 2017. "Chambers of commerce firms registry." InfoCamere SCpA - Chambers of commerce digital innovation [publisher].

INPS. 2017. "INPS firm and employment registry." Social security national institute (INPS) [publisher].

Ippedico, Giuseppe. 2017. "Are we losing our entrepreneurs? Evidence from Italian emigration during the 2008-2013 crisis." Master's thesis, Università Bocconi.

Iranzo, Susana, and Giovanni Peri. 2009. "Schooling Externalities, Technology, and Productivity: Theory and Evidence from U.S. States." *REStat* 91 (2): 420–431.

Istat. 2005. "14th Census of Population and Housing." Istituto Nazionale di Statistica (Istat) [publisher] `http://dawinci.istat.it/` (accessed September 15, 2019).

———. 2013. "Inter-census Population of Italian municipalities by age and sex, 2002-2011." Istituto Nazionale di Statistica (Istat) [publisher] `https://demo.istat.it` (accessed March 9, 2017).

———. 2014a. “15th Census of Population and Housing.” Istituto Nazionale di Statistica (Istat) [publisher] http://dati-censimentopopolazione.istat.it/Index.aspx (accessed September 15, 2019).

———. 2014b. “I profili dei nuovi imprenditori.” Technical report, Istituto Nazionale di Statistica, Roma. URL: https://www.istat.it/it/files/2016/12/I-PROFILI-DEI-NUOVI-IMPRENDITORI_focus.pdf [last accessed on February 25, 2021].

———. 2014c. “Statistical Atlas of Italian municipalities.” Istituto Nazionale di Statistica (Istat) [publisher] https://www.istat.it/it/archivio/113712 (accessed September 15, 2019).

———. 2016a. “Enrollment and dis-enrollment from the registry of resident population for transfer of residence, 2002-2015.” Istituto Nazionale di Statistica (Istat) [publisher].

———. 2016b. “Local Labor Markets boundaries.” Istituto Nazionale di Statistica (Istat) [publisher] https://www.istat.it/it/informazioni-territoriali-e-cartografiche/sistemi-locali-del-lavoro (accessed March 9, 2017).

———. 2017. “Population of Italian municipalities by age and sex, 2012-2015.” Istituto Nazionale di Statistica (Istat) [publisher] https://demo.istat.it (accessed March 9, 2017).

———. 2018. “Administrative and territorial variation of Italian municipalities.” Istituto Nazionale di Statistica (Istat) [publisher] https://www.istat.it/it/archivio/6789 (accessed September 15, 2019).

Jaeger, David A., Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin. 2010. “Direct Evidence on Risk Attitudes and Migration.” *The Review of Economics and Statistics* 92 (3): 684–689.

Karahan, Fatih, Benjamin Pugsley, and Aysegül Sahin. 2019. "Demographic Origins of the start-up Deficit." NBER Working Paper 25874.

Kato, Takao, and Chad Sparber. 2013. "Quotas and quality: The effect of H-1B visa restrictions on the pool of prospective undergraduate students from abroad." *Review of Economics and Statistics* 95 (1): 109–126.

Kerr, Sari Pekkala, William Kerr, Çağlar Özden, and Christopher Parsons. 2017a. "High-skilled migration and agglomeration." *Annual Review of Economics* 9:201–234.

Kerr, Sari Pekkala, William R. Kerr, Caglar Ozden, and Christopher Parsons. 2017b. "High-Skilled Migration and Agglomeration." *Annual Review of Economics* 9:201–234.

Kerr, William R., and William F. Lincoln. 2010. "The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention." *Journal of Labor Economics* 28 (3): 473–508.

Kleven, Henrik J., Camille Landais, Mathilde Muñoz, and Stefanie Stantcheva. 2020. "Taxation and Migration: Evidence and Policy Implications." *Journal of Economic Perspectives* 34 (2): 119–142.

Kleven, Henrik J., Camille Landais, and Emmanuel Saez. 2013a. "Superstars Taxation and International Migration of Superstars: Evidence from the European Football Market." *American Economic Review* 103(5):1892–1924.

———. 2013b. "Taxation and international migration of superstars: Evidence from the European football market." *American Economic Review* 103 (5): 1892–1924.

Kleven, Henrik J., Camille Landais, Emmanuel Saez, and Esben Schultz. 2014. "Migration and wage effects of taxing top earners: Evidence from the foreigners' tax scheme in Denmark." *The Quarterly Journal of Economics* 129 (1): 333–378.

Kopecky, Joseph. 2017. "An aging dynamo: demographic change and the decline of entrepreneurial activity in the United States." mimeo.

Liang, James, Hui Wang, and Edward P. Lazear. 2018. "Demographics and Entrepreneurship." *Journal of Political Economy* 126 (S1): S140–S196.

Liebig, Thomas, Patrick A. Puhani, and Alfonso Sousa-Poza. 2007. "Taxation and internal migration - evidence from the Swiss Census using community-level variation in income tax rates." *Journal of Regional Science* 447(4):807–836.

MacDonald, Glenn, and Michael S. Weisbach. 2004. "The Economics of Has-beens." *Journal of Political Economy* 112 (S1): S289–S310.

Martinez, Isabel. 2017. "Beggar-thy-neighbour tax cuts: Mobility after a local income and wealth tax reform in Switzerland." *Luxembourg Institute of Socio-Economic Research (LISER) Working Paper Series*, vol. 8.

Mayr, Karin, and Giovanni Peri. 2009a. "Brain Drain and Brain Return: Theory and Application to Eastern-Western Europe." *The B.E. Journal of Economic Analysis & Policy* 9 (1): 1–52.

———. 2009b. "Brain drain and brain return: theory and application to Eastern-Western Europe." *The BE Journal of Economic Analysis & Policy* 9, no. 1.

McFadden, Daniel. 1974. "The measurement of urban travel demand." *Journal of public economics* 3 (4): 303–328.

Messina, Giovanna, and Marco Savegnago. 2014. "A Prova Di Acronimo: I Tributi Locali Sulla Casa in Italia (Beyond the Acronyms: Local Property Taxation in Italy)." Bank of Italy Occasional Paper, 250.

Mishra, Prachi. 2007. "Emigration and wages in source countries: Evidence from Mexico." *Journal of Development Economics* 82 (1): 180–199.

Moretti, Enrico. 2004. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review* 94 (3): 656–690.

Moretti, Enrico, and Daniel J. Wilson. 2017. "The effect of state taxes on the geographical location of top earners: evidence from star scientists." *American Economic Review* 107(7):1858–1903.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German Jewish Émigrés and US Invention." *American Economic Review* 104 (10): 3222–55.

Muñoz, Mathilde. 2019. "Do European Top Earners React to Labour Taxation Through Migration?" Working Paper.

NINo. 2018. "National Insurance Number (NINo) Registrations To Adult Overseas Nationals Entering The UK (XI)." UK Department for Work and Pensions [publisher] (accessed April 10, 2018).

Oates, Wallace E. 1973. "The effects of property taxes and local public spending on property values: A reply and yet further results." *Journal of Political Economy* 81 (4): 1004–1008.

Oliviero, Tommaso, and Annalisa Scognamiglio. 2019. "Property tax and property values: Evidence from the 2012 Italian tax reform." *European Economic Review* 118:227–251.

Parey, Matthias, Jens Ruhose, Fabian Waldinger, and Nicolai Netz. 2017. "The selection of high-skilled emigrants." *The Review of Economics and Statistics* 99 (5): 776–792.

Patt, Alexander, Jens Ruhose, Simon Wiederhold, and Miguel Flores. 2021. "International Emigrant Selection on Occupational Skills." *Journal of European Economic Association* 19 (2): 1249–1298.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics* 33 (S1): S225–S255.

Portuguese Observatory of Emigration, 2015. "Emigraçao Portuguesa Relatório Estatístico 2015." Technical report. URL: `https://bit.ly/2U1Bb6E` (accessed on April 7, 2019).

Rauch, James E.. 1999. "Networks versus markets in international trade." *Journal of International Economics* 48 (1): 7–35.

———. 2001. "Business and social networks in international trade." *Journal of Economic Literature* 39 (1): 1177–1203.

Rubolino, Enrico. 2021. "Tax-Induced Transfer of Residence: Evidence from Tax Decentralization in Italy." Working Paper.

Ruggles, Steven, Sarah Floodand, Ronald Goeken, Erin Meyer, Jose Pacas, and Matthew Sobek. 2019. "Integrated Public Use Microdata Series (IPUMS USA): Version 9.0 [dataset]." Minneapolis, MN: IPUMS, 10, D010 (accessed September 4, 2018) URL: `http://doi.org/10.18128/D010.V9.0`.

Saint-Blancat, Chantal. 2019. "Italy: brain drain or brain circulation?" *International Higher Education*, no. 96:10–11.

Schivardi, Fabiano, and Tom Schmitz. 2020. "The IT revolution and southern Europe's two lost decades." *Journal of the European Economic Association* 18 (5): 2441–2486.

Schmidheiny, Kurt, and Michaela Slotwinski. 2018. "Tax-induced mobility: Evidence from a foreigners' tax scheme in Switzerland." *Journal of Public Economics* 167:293–324.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (5): 416–424.

Varner, Charles, Cristobal Young, and Allen Prohofsky. 2018. "Millionaire Migration in California: Administrative Data for Three Waves of Tax Reform." Working Paper.

Waldinger, Fabian. 2010. "Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany." *Journal of Political Economy* 118 (4): 787–831.

Winters, John V. 2014. "STEM graduates, human capital externalities, and wages in the U.S." *Regional Science and Urban Economics* 48 (C): 190–198.

World Bank. 2019. "GDP per capita (constant 2010 US$)." World Bank national accounts data, and OECD National Accounts data files. URL: `https://data.worldbank.org/indicator/NY.GDP.PCAP.KD` (accessed September 15, 2019).

Young, Cristobal, and Charles Varner. 2011. "Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment." *National Tax Journal* 64 (2): 255–284.

Young, Cristobal, Charles Varner, Ithai Z. Lurie, and Richard Prisinzano. 2016. "Millionaire Migration and Taxation of the Elite: Evidence from Administrative Data." *American Sociological Review* 81 (3): 421–446.