Title
Emigration and Entrepreneurial Drain

Journal
American Economic Journal Applied Economics, 15(2)

ISSN
1945-7782

Authors
Anelli, Massimo
Basso, Gaetano
Ippedico, Giuseppe
et al.

Publication Date
2023-04-01

DOI
10.1257/app.20210194

Emigration and Entrepreneurial Drain

Massimo Anelli (Bocconi, CESifo, IZA)
Gaetano Basso (Bank of Italy) Giuseppe Ippedico (UC Davis)
Giovanni Peri (UC Davis and NBER)*

June 7, 2023

Abstract

Emigration of young, highly educated individuals may deprive origin countries of entrepreneurs. We identify exogenous variation in emigration from Italy by interacting past diaspora networks and current economic pull factors in destination countries. We find that a one standard deviation increase in the emigration rate generates a 4.8% decline in firms creation in the local labor market of origin. An accounting exercise decomposes the estimated effect into four components: subtraction of individuals with average entrepreneurial propensity, selection of young and college-educated among emigrants, negative spillovers on firm creation and selection on unobservable characteristics positively associated with entrepreneurship.

Key Words: Emigration, Brain Drain, Entrepreneurship, Demography, Innovation, EU Integration.
JEL Codes: J61, O3, M13.

---

*Massimo Anelli, Bocconi University, email: massimo.anelli@unibocconi.it; Gaetano Basso, Bank of Italy, email: gaetano.basso@bancaditalia.it, Giuseppe Ippedico, UC Davis, email: gippedico@ucdavis.edu, Giovanni Peri, UC Davis, email: gperi@ucdavis.edu. This paper was previously circulated under the titles "Youth Drain, Entrepreneurship and Innovation" and "Does Emigration Drain Entrepreneurs?". We are grateful to Emanuele Ciani, Federico Cingano, Xavier Jaravel, Andrea Linarello, Salvatore Lo Bello, Francesca Lotti, Sauro Mocetti, Paolo Sestito, Eliana Viviano and Josef Zweimueller for helpful comments and to seminars participants at the Bank of Italy, Copenhagen Business School, CESifo Employment and Social Protection Conference, 2018 ESPE, ifo Institute, Univeristat de Barcelona, Free University of Bozen, Paris School of Economics, UC Davis Migration Cluster, 2018 Edinburgh Migration Workshop and 2018 Brucchi Luchino. We thank Rosario Ballatore (Bank of Italy), Francesca Licari (Istat), Desideria Toscano and Annamaria Grispo (Italian Ministry of Interior) and Gina Carbonin (SOGEI) for valuable data provision. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Bank of Italy. Any errors or omissions are the sole responsibility of the authors.

Entrepreneurship 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 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 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.[1] While emigration was

[1] 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.

occurring across all age groups, Figure 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.

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.[2] The common push in the post-2008 period, generated by the recession, interacted with the pre-existing

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

networks in economically more successful countries, generates the post-2008 variation of predicted emigration in our IV.[3] 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 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,

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

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 2009), (Waldinger 2010), (Docquier and Rapoport 2012), (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 2017), 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.[4]

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 as job-creators

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

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 I describes the main data and trends for emigration and firm creation in Italian local labor markets. Section II introduces the empirical specification, then describes the 2SLS identification strategy and discusses its validity. Section III presents the main results, and Section IV discusses several additional results. Section V reports the main robustness checks. Section VI concludes the paper.

I Data

I.A 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 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 II.C. 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).

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[5], 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 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 characterizing the formal registration process. An analysis based on data from the Switzerland Statistical Office (BFS 2018) show similar patterns (Figure 4).[6]

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

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

[6]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.

the magnitude of the effects relative to the size of the emigration rate, as we do in Section III below.

I.B 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[7] of firms and demographic characteristics of owners, shareholders and executives of each firm over the period 2005-2015.[8] 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 include 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.[9]

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.[10] There are 686 LLMs in Italy

[7]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.

[8]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.

[9] 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.

[10] 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

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.

II 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). The main explanatory variable is the cumulative outflow of Italians who are 25 to 64 years old from 2008 to 2015, indicated as $\sum_{t=2008}^{2015} m_{l,t}$.[11] 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 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:

$$\frac{\Delta y_l}{pop_{l,pre}} = \alpha + \beta \frac{\sum_{t=2008}^{2015} m_{l,t}}{pop_{l,pre}} + \phi_p + \gamma X_{l,2005} + \varepsilon_l \tag{1}$$

If the size of migration outflows were distributed randomly across LLM, the OLS

include province fixed effects in the main empirical specification.

[11]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.

estimate of equation (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 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 I, 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 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.

II.A 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.[12] 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:

$$Pull_l = \sum_c NTWK_{l,c} * G_c \tag{2}$$

In expression (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).[13] It captures the size of the 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

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

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

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 summarizes the variation in GDP growth from 2008 to 2015 for the main countries-of-destination. The variable defined in equation (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). 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.

II.B 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 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.[14]

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[15]. In Table 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.[16] Thus, the estimates of Table 3 are consistent with our identifying assumption, that the IV is not correlated with pre-2008 firm creation and

[14] 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.

[15] 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.

[16] 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.

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 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[17] 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.

II.C 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 uncorre-

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

lated with the error term.[18] 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 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 4 and 5 report the main results of diagnostic tests as suggested in (Goldsmith-Pinkham, Sorkin, and Swift 2020). Table 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) with per-capita stock of firms as the dependent variable ($\hat{\beta}_c$).[19] 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

[18] (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).

[19] 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).

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 4 reports the five destination country shares receiving the highest weight and, hence, driving most of the identifying variation. The share of emigrants to 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 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), 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.[20] Panel D of Table 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

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

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 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 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 III) 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.

II.D First stage and compliers characterization

In Table 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 6 show that the $Pull_l$ has a significant predictive power for actual emigration. The size of the coefficient is stable across specifications.[21] 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

[21] 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 6, the instrument power is slightly lower, consistent with an older expatriate network being less relevant for emigrants in 2008-2015.

concerns would arise.[22]

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 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 analogously to Local Average Treatment Effects (LATE-like).[23] 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 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,

[22] 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.

[23] 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 II.C, the Rotemberg weights in our setting are non-negative, thus limiting the extent of non-convex weights in our Bartik-style estimate.

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.

III Main results

III.A Effects on firm creation

Table 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.[24] 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 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

[24] 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.

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 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.[25] 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 III.B 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 the checks in Section II.B 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

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

those in Table 7.

III.B 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 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.[26]

[26] This is because one standard deviation of the emigration rate which generates a 0.432 decline–the

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).[27] 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.[28] 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

coefficient in Table 7– of firms per 100 people, which multiplied by (44,805/100) equals 194.

[27]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.

[28]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 III.C.

the firm-creation rate of college-educated individuals was 2.7% (versus 1.2% for non-tertiary educated), selection on education explains the loss of 19 additional firms.[29,30]

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.[31] 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

[29]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%.

[30]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.

[31]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 2011) find positive selection on schooling, (Parey et al. 2017) shows selection on pre-migration earnings, and (Patt et al. 2021) on occupational skills.

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 due to *Residual Selection and Spillovers*:

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

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

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

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).[32] 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.[33]

[32] 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.

[33] 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. 2017)), confirming

III.C 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 firms "innovative start-ups," as they are those more likely to embody new technologies and ideas.[34]

In Table 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 8, which mirror those of Table 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.

---

that their departure exerts negative spillover effects on the local economies-of-origin.

[34] 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.

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.[35] 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. 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.

IV 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

[35] 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.

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 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.[36] 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 together, these results do not suggest that the departure of emigrants was associated with a tightening of the labor market.[37] Instead, the overall picture is more consistent with the idea that emigration reduced labor demand as much as labor supply.

Furthermore, in Table 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,

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

[37] 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 10 below) together with the relative demand and supply effects.

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.

V 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 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 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.[38]

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 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 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.[39]

VI 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,

[38] 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.

[39] 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.

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.

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.

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

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.

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).

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

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.

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.

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

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.

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.

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. 2017. “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. 2011. “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.” *The Review of Economics and Statistics* 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.

Kerr, Sari Pekkala, William R. Kerr, Caglar Ozden, and Christopher Parsons. 2017. “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.

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.

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

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

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.

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

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).

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): 12491298.

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.

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

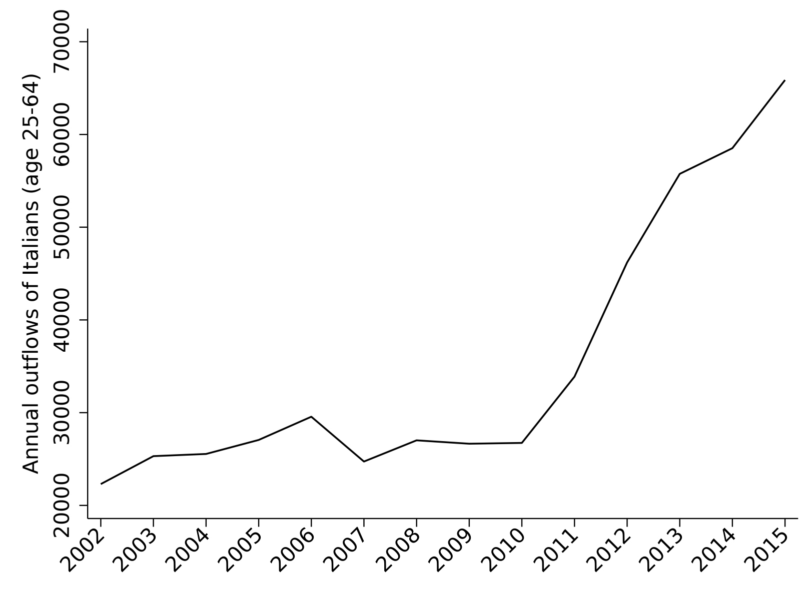
Waldinger, Fabian. 2010. “Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany.” *Journal of Political Economy* 118 (4): 787–831.

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).

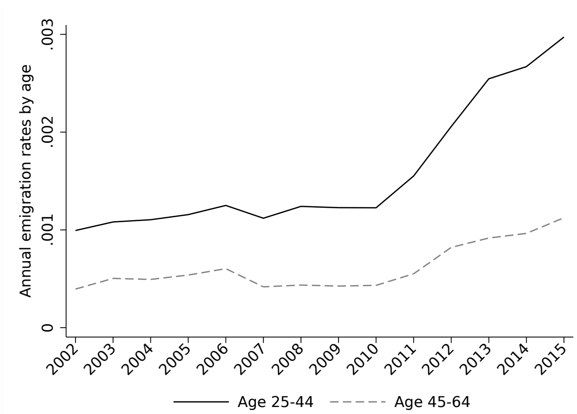
Figures

Figure 1: Emigration flows of Italians 25-64 years old

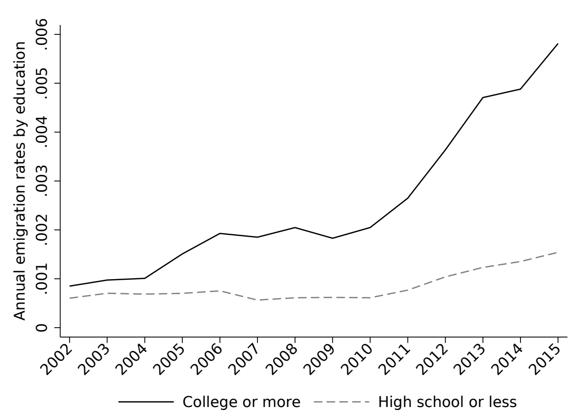


Notes: Annual outflows of Italian citizens 25-64 years old. Source: AIRE-Istat.

Figure 2: Emigration rates by age and education



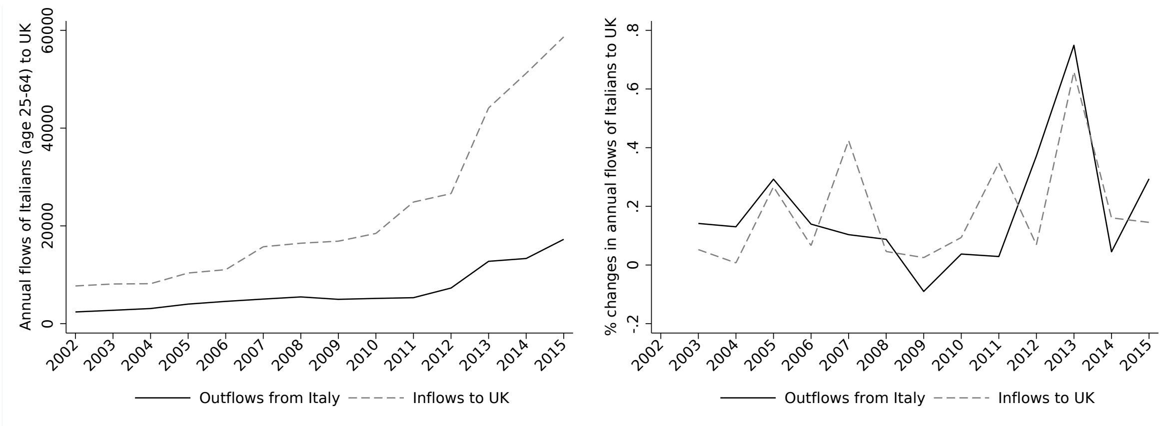
(a) Annual emigration rates, by age



(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.

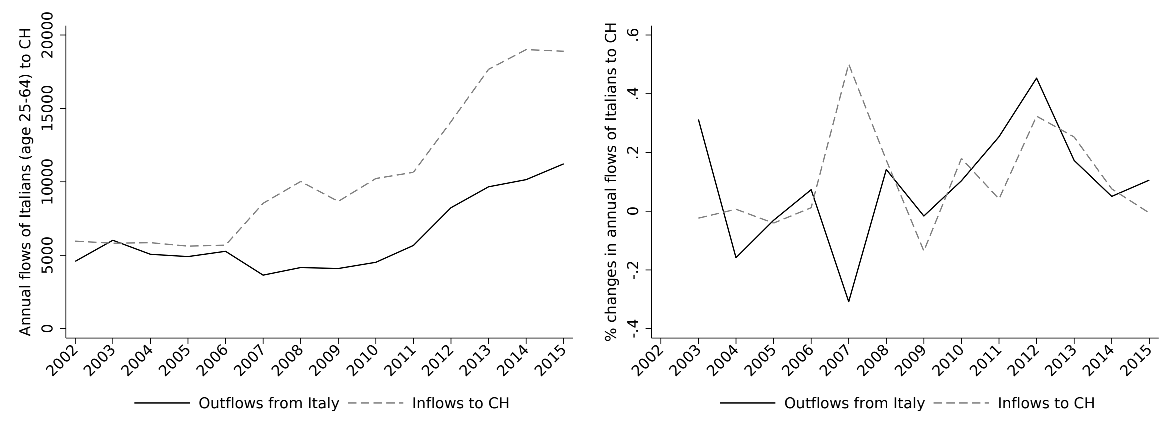
Figure 3: Recorded Emigration and Registered Inflows of Italians to the UK



(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.

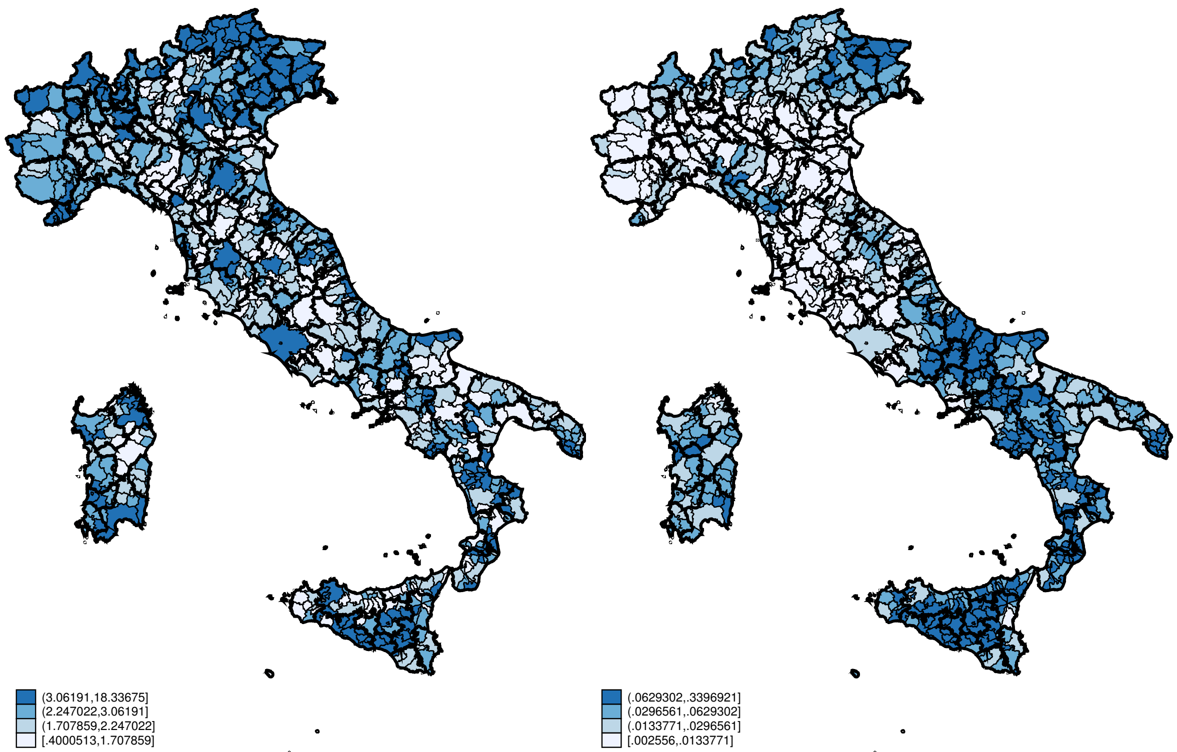
Figure 4: Recorded Emigration and Registered Inflows of Italians to Switzerland



(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 5: Actual and Predicted Emigration from Italian LLMs

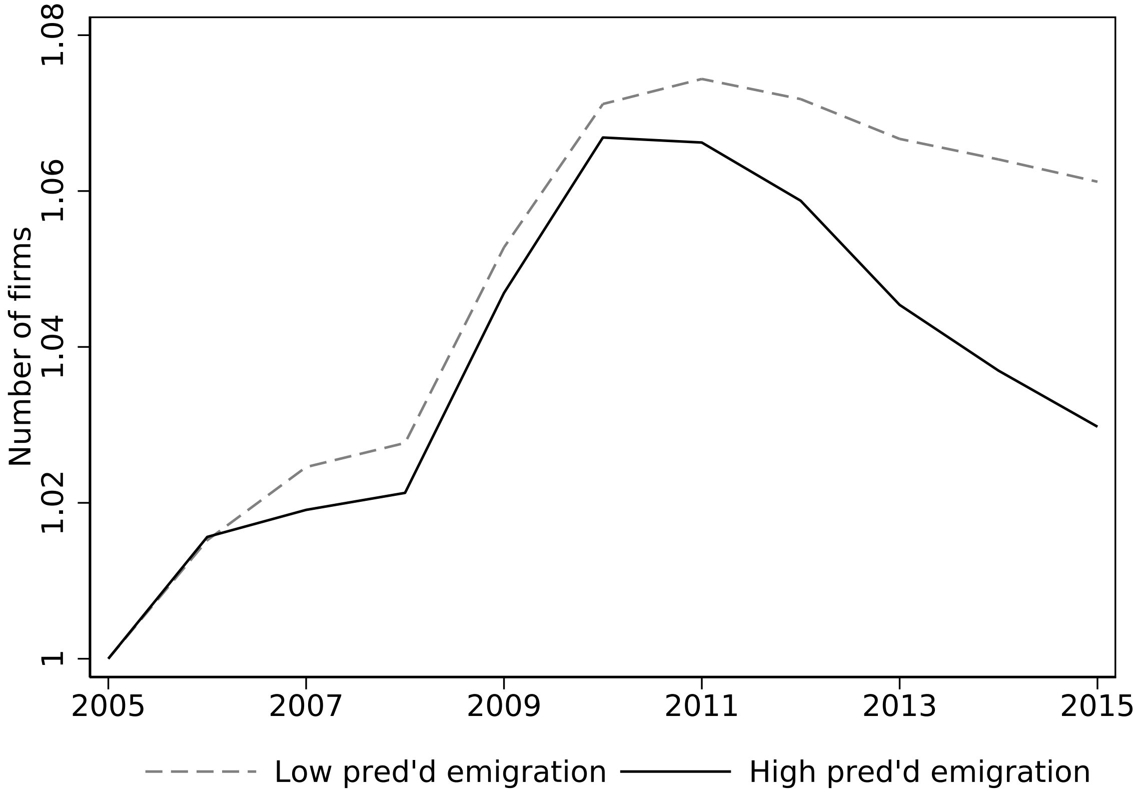


(a) Emigration

(b) Predicted Emigration

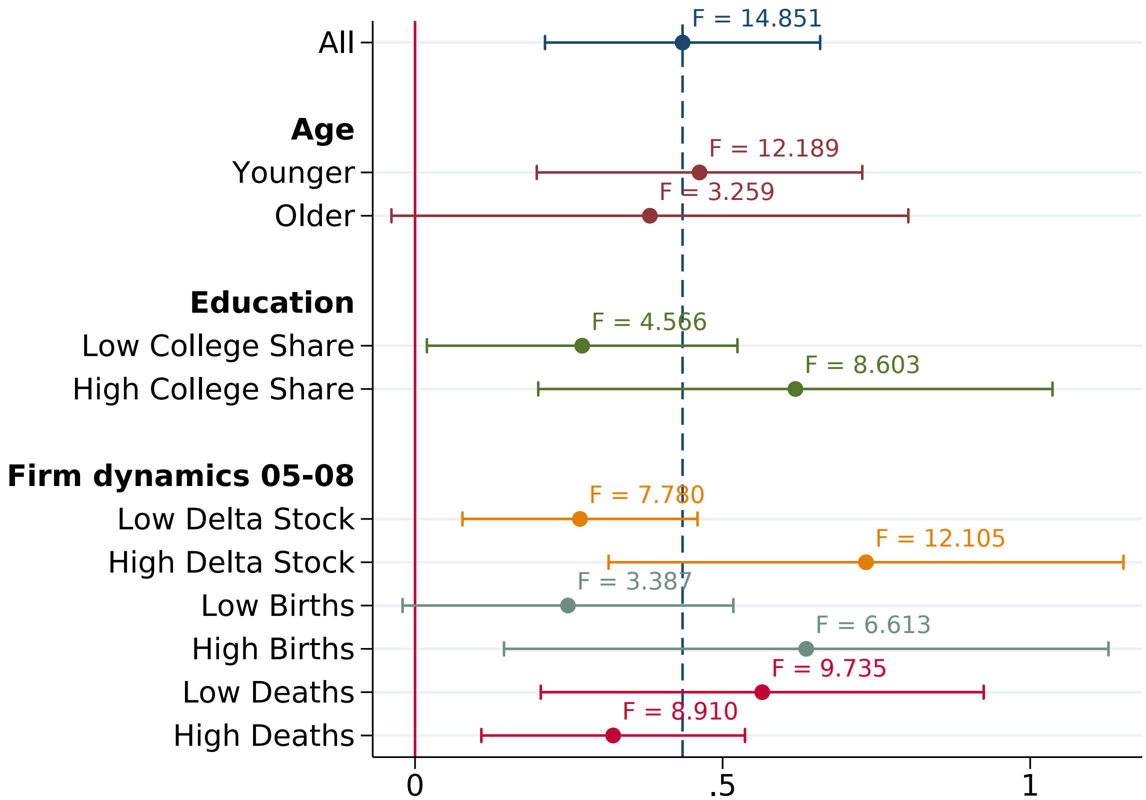
Notes: Figure (a) plots the cumulative emigration rate between 2008-2015 in percentage points, i.e. the number of Italian citizens 25-64 years old migrating abroad between 2008-2015 as a fraction of the 25-64 years old population in the origin LLM (average 2005-2008), multiplied by 100 and normalized 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 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 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 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: 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 2: OLS regressions of LLM firms dynamics on observed emigration rates

VARIABLES	(1) All Firms Δ Stock 2008-15	(2) All Firms ∑ Births 2008-15	(3) All Firms ∑ 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 3: Instrument validity: effect of the instrument on pre-shock (2005-08) change in stock, cumulative entry and exit of firms

VARIABLES	(1) All Firms Δ Stock 2005-08	(2) All Firms $\sum$ Births 2005-08	(3) All Firms $\sum$ Deaths 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 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	(-∞, 1.10)
Australia	0.039	1.064	-0.197	0.60	(-∞, ∞)
Belgium	0.029	1.005	-0.081	0.84	(-∞, ∞)

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 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 2 and 7, column (1), for comparison.

Table 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 6: First stage regressions

VARIABLES	(1) Emig Rate	(2) Emig Rate	(3) 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 7: Effect of emigration rates on change in stock, cumulative entry and exit of firms

VARIABLES	(1) All Firms Δ Stock 2008-15	(2) All Firms $\sum$ Births 2008-15	(3) All Firms $\sum$ Deaths 2008-15
Emig Rate	-0.414	-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 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 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 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 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).