NBER WORKING PAPER SERIES

YOUTH DRAIN, ENTREPRENEURSHIP AND INNOVATION

Massimo Anelli
Gaetano Basso
Giuseppe Ippedico
Giovanni Peri

Working Paper 26055
http://www.nber.org/papers/w26055

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2019

We are grateful to Emanuele Ciani, Federico Cingano, Xavier Jaravel, Salvatore Lo Bello, Francesca Lotti, Sauro Mocetti, Paolo Sestito, Eliana Viviano and Josef Zweimueller for helpful comments and to seminars participants at Bank of Italy, Copenhagen Business School, CESifo Area Conference on Employment and Social Protection, 2018 ESPE conference, ifo Institute, Univeristat de Barcelona, Free University of Bozen, Paris School of Economics, UC Davis Migration Research Cluster, 2018 Workshop on Migration in Edinburgh and 2018 Brucchi Luchino workshop. We thank Rosario Ballatore (Bank of Italy), 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 or of the National Bureau of Economic Research. Any errors or omissions are the sole responsibility of the authors.

Youth Drain, Entrepreneurship and Innovation
Massimo Anelli, Gaetano Basso, Giuseppe Ippedico, and Giovanni Peri
NBER Working Paper No. 26055
July 2019
JEL No. J61,M13,O3

ABSTRACT

Migration outflows, especially of young people, may deprive an economy of entrepreneurial energy and innovative ideas. We exploit exogenous variation in emigration from Italian local labor markets to show that between 2008 and 2015 larger emigration flows reduced firm creation. The decline affected firms owned by young people and innovative industries. We estimate that for every 1,000 emigrants, 10 fewer young-owned firms were created over the whole period. A simple accounting exercise shows that about 60 percent of the effect is generated simply by the loss of young people; the remaining 40 percent is due to a combination of selection of emigrants among highly entrepreneurial people, negative spillovers on the entrepreneurship rate of locals, and negative local firm multiplier effect.

Massimo Anelli
Department of Social and Political Sciences
Bocconi University
Via Sarfatti 25
Milan 20136
Italy
massimo.anelli@unibocconi.it

Gaetano Basso
Directorate General for Economics,
Statistics and Research
Bank of Italy
Via Nazionale 91
00184 Rome
Italy
gaetano.basso@bancaditalia.it

Giuseppe Ippedico
Department of Economics
University of California Davis
One Shields Avenue
Davis, CA 95616
gippedico@ucdavis.edu

Giovanni Peri
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
gperi@ucdavis.edu

1 Introduction

Young people are often vehicles of innovation and economic change. This may be because their recently acquired human capital is complementary with new technologies or because they are more willing to take risks and invest in uncertain and "more disruptive" projects. Therefore, large cohorts of young people in an economy may spur innovative ability and firm creation. Young managers, professionals and entrepreneurs are important drivers of growth as they contribute to the Schumpeterian "creative destruction" by introducing new practices and technologies in the productive system (Liang et al., 2018; Acemoglu et al., 2017).

Advanced economies, where fast aging of the population is reducing the size of the young cohorts, risk a decline in creative and entrepreneurial energies. This issue is particularly relevant for several Southern European countries where population is aging and young cohorts are shrinking due to demographic transition as well as to emigration towards the rest of Europe. While emigration of Southern European youth has been significant for a long time (Schivardi and Schmitz, 2018), the Great Recession has hit Southern Europe much harder than Northern Europe and has provided strong incentives for young individuals to leave their countries as employment opportunities worsened significantly. The decline in the number of young people may imply a slowdown in the creation of new firms and start-ups. It may also hinder innovation, slowing technological progress and economic growth and, hence, depress labor demand.[1] The key question we ask in this paper is: does a decline in young people reduce firm creation? And what sectors and type of firms are most affected?

[1] A similar concern expressed in developing countries, and called "brain drain" is that their young, highly-educated individuals leave the country, lured by better opportunities in richer economies. This leaves the economy and society of origin deprived of human capital and innovative capacity (Docquier, 2014).

To answer these questions we focus on Italy, a Southern European country characterized by a surge in the number of young people emigrating since the onset of the Great Recession. Panel (a) of Figure 1 shows the scale and the sudden nature of the increase in the overall emigration rate, for Italy. The emigration rate, shown from year 2005, increased very sharply starting in 2010, and tripled by 2015, going from a yearly 0.07% to 0.18% of the population in working age. Over the entire period 2008-2015 the cumulative flows recorded by the administrative data sum up to 0.9% of the entire population.[2] Emigration was common to all age groups, but its rate was especially large among young individuals. Panel (b) of Figure 1 shows the emigration rates among young people (aged 25-44) and older people (aged 45-64) revealing that for the first group the rate grew more significantly and it was more than double that of the second group as of 2015.[3]

Estimating the causal effect of emigration on local economic outcomes is challenging. A least squares regression of local outcomes on local emigration rates would produce biased estimates. The most relevant threats are reverse causality, as people are more likely to leave poorly performing regions and omitted variable bias, as several unobserved factors may push people to emigrate and cause poor performance of firms. Measurement error in emigration flows, due for instance to under-reporting of changes in administrative residence, may also contribute to the bias. To overcome these issues, we adopt an instrumental variable strategy in the spirit of Anelli and Peri (2017), using what we call "pull-driven emigration". We construct the network links of the diaspora from each local labor market in Italy to a destination country, in a baseline year, that we choose to be year 2000. Then, we interact this bilateral network with the economic performance of the *destination*

[2] Comparable statistics on nationals' emigration flows across countries are extremely hard to obtain. A report from the Portuguese Observatory of Emigration (2015) indicates that the cumulative outflows of Portuguese citizens between 2011 and 2014 reached about 485 thousand people, or about 1.2 percent of the Portuguese population per year.

[3] Emigration rates were also larger among college graduates (Appendix Figure A1); however, the increase during this period was comparable for college and non-college graduates.

countries. In this way, we capture the facts that people residing in municipalities with stronger network ties to countries with strong economic performance during the period 2008-2015, were more likely to emigrate. This instrumental variable allows us to leverage cross-sectional variation of emigration rates driven only by "pull factors" of emigration. By construction this IV is independent of any location-specific push factor, such as the economic conditions in the local labor market, which are likely to be correlated with local outcomes. We also perform a battery of tests showing that our pull-driven emigration instrument is uncorrelated with previous local economic outcomes. This is consistent with the assumed exclusion restriction that the instrument is not correlated with unobservable and persistent local economic trends.

Our results show that emigration – instrumented by its pull-driven component– caused a decline in the number of existing firms. The depressive effect on firm creation was driven by fewer firm births, rather than by more firm deaths. This is consistent with the loss of managerial and entrepreneurial capital that drives firm creation. Moreover we observe a negative effect in the creation of firms whose owners and managers were 45 years old or younger. Importantly, we show that the instrument has stronger explanatory power for local labor markets with younger population at the baseline, and the instrument is positively associated with an increase in the average age of the remaining population. The estimated (local) treatment effect reveals that in an average-sized local labor market, which lost 1,724 individuals to emigration, 178 fewer firms were created and managed by young owners over the period 2008-2015. As reference, during that period, 2,470 firms managed by young owners were created in the average local labor market. This is a significant effect. To better characterize the underlying mechanisms, we present a simple accounting exercise that shows that about 60% of this effect can be attributed to the change in the age composition of the local labor market – the pure *demographic effect*– triggered by

emigration. Given the pre-emigration entrepreneurship rate among young people, 107 out of 178 missing firms were indeed lost simply due to young individuals leaving the local labor market. Selection of emigrants among individuals with a higher entrepreneurship-rate, and negative spillovers onto other entrepreneurs, explain up to 35% of the overall effect. The remaining 5% can be attributed to diminished demand for goods and services due to a lower population level. As further evidence indicating the potential innovative role of the young people who left, we find a decline in the number of innovative start-ups (operating in technology-intensive sectors). Finally, we find that local labor markets with higher emigration rates exhibit a decline in overall employment and a negative change in the share of qualified workers (i.e., white collars).[4]

The rest of the paper is organized as follows. Section 2 frames our contribution in the existing literature. Section 3 describes the main data and trends for emigration and firm creation in Italian local labor markets. Section 4 introduces the empirical specification and the 2SLS identification strategy, and discusses its validity. Section 5 presents the main results, and Section 5.5 reports several robustness checks. Section 6 concludes the paper.

2 Literature Review

We know little about how a decline in the number of young people affects the economic performance of an economy, and yet this phenomenon has very important implications for the economies of both rich and developing countries. In developing countries, the emigration rate of the high-skilled population, often referred to as *brain drain*, has increased significantly in the last two decades (see Docquier and Rapoport, 2012; Mayr and Peri,

[4] The worsening of the composition of the workforce is an additional mechanism that could explain the drop in entrepreneurship, as more skilled workers could have contributed to the formation of new productive firms absent migration. However, such a mechanism, which is captured by both the "selection" and the "local multiplier" effect in our decomposition, is hard to quantify given that we observe only few characteristics of the migrants and of the workforce in the data.

2009; Docquier et al., 2014; Di Giovanni et al., 2015). This trend significantly reduces the number of young people – who are more likely to migrate – depriving the countries-of-origin of their economic and creative contributions. In developed economies, the dramatic fertility decline experienced in recent decades is causing the size of young cohorts to markedly shrink. Our analysis takes advantage of a sharp increase in emigration in Italy which can be attributed to economic motivations along with free mobility within the EU. We exploit the strongly selective nature of this episode, affecting young (25-44 years old) individuals much more than any other group, to estimate its impact on firm creation and innovative start-ups. Our results contribute to several branches of the economic literature.

A growing body of literature shows that radical innovations are adopted by young managers, and that their presence in a firm proxies for innovation intensity as they sort into more innovative firms (Acemouglu et al., 2017).[5] There is also evidence that entrepreneurship peaks at an early age (Kopecky, 2017). Using cross-sectional variation Liang et al. (2018) show that there is a significant positive relationship between the share of *young* people in a country (or region) and its entrepreneurship rate. The share of young also correlates positively with productivity (Ciccarelli et al., 2017), growth (Engbom, 2019) and birth rate of start-ups (Karahan et al., 2019).

A more established literature shows that brain drain from developing countries may negatively impact growth (despite positive general equilibrium effects through remittances, return migration and innovation exchanges; Docquier et al., 2014). In the context of developed countries, there is little causal evidence on the effects of emigration on origin economies, an exception being recent work on Mexico and eastern Europe that points to short-term positive wage effects for stayers (due to a labor supply shortage; Mishra, 2007; Elsner, 2013a, 2013b; Dustmann et al., 2015). A paper closely related to ours is Giesing

[5]Part of this effect can be due to the investments that young professionals make in R&D and technology-complementary human capital (Barker and Mueller, 2002; MacDonald and Weisbach, 2004).

and Laurentsyeva (2017) who find negative effects on regional TFP due to labor mobility after the Eastern EU enlargnment of the 2000s.[6] A related literature, beginning with Rauch (1999), shows that links between origin and destination countries may have a causal impact on trade flows (Rauch, 2001; Rauch and Trindade, 2002; Combes et al., 2005; Peri and Requena-Silvente, 2010; Javorcik et al., 2011; Bratti et al., 2018a; Bratti et al., 2018b). We discuss whether these established trade patterns may pose a threat to our identification in Section 5.5.

Finally, we contribute to a broader literature on the positive role of labor mobility in absorbing demand shocks (Blanchard and Katz, 1992; Bayer and Smets, 2015; Arpaia et al., 2016; Dao et al., 2017; Basso et al., 2018). In this paper we only identify a short-run negative effect of emigration, but declines in entrepreneurship and innovative ability may also affect longer term growth prospects. Hunt and Gauthier-Loiselle (2010) show that immigrants to the US are more likely to be active in patenting and innovation than comparable natives. This paper confirms that migrants may be positively selected among those with high entrepreneurial and innovative potential.

3 Data

3.1 Data on Firms and Local Labor Markets

We combine data on the emigration flows (described later in this section) with firm-level data on the universe of all Italian firms, obtained from the Chambers of Commerce and also with data from the social security administration (INPS) on local employment and wages.[7]

[6] In his master dissertation Ippedico (2017) also looked at the relationship between migration and firm creation. In their work on the start-up deficits, Karahan et al. (2019) use variation on immigration across US states and relate it to firm entry.

[7] Both the Chambers of Commerce and INPS data identify the location of the firm by its headquarters. The vast majority of Italians firms have only one establishment, so the main location will typically

The data from the Chambers of Commerce include information on entry and exit of firms and demographic characteristics of owners, shareholders and managers for each firm for the period 2005-2015. We use this latter piece of information to classify firms with a majority of owners and managers aged below 45, which we refer to as "young-owned firms". INPS data cover the period 1990-2015, and collect information on the yearly number of employees (broken down by broad occupation category, i.e., apprentices and blue collars, white collars and managers), average monthly wage, industry and geographic location of each employer.

Our unit of analysis is the local labor market (LLM), defined using the national statistical institute (ISTAT) 2001 definition. According to ISTAT, LLMs are clusters of municipalities with commuting patterns that are mainly within LLMs rather than across LLMs. Similar to US Commuting Zones, they include geographic aggregations such that people reside and work within them. They are used as a proxy for local labor markets. There are 686 LLMs in Italy covering the whole national territory. Usually they do not cross provincial boundaries. We focus our analysis on the period 2005-2015, and consider the period 2008-2015 as the "treatment" period, when emigration increased substantially. We can control for pre-2008 economic performance (e.g., LLMs unemployment rate and value added obtained from ISTAT) and test the correlation of the instrument with pre-2008 trends.

3.2 Data on Emigration Flows

We obtained data on emigration flows from each Italian municipality from ISTAT. The data, which cover the period 2002-2015, come from administrative sources and are aggre-

encompass the whole firm.

gated into municipality of origin-country of destination-age group cells.[8] We also obtained emigration flows directly from the registry of Italians residing abroad (AIRE; Anelli and Peri, 2017) that cover more years in the past (from 1990 to 2014) and contain information about country of destination. These two features allow us to construct the networks of emigrants as of 2000.

Despite the fact that emigrants are required by the law to register their residence abroad within six months from the date of emigration, there are concerns about under-registration and not all changes of residence may be recorded by the Italian authorities – thus not showing up in our data. Figure 2 compares the outflows of Italians as registered by the AIRE registries and the registration of Italians recorded by the UK social security system. The UK data indicate that the level of outflows from Italy to the UK is underestimated by about two thirds (panel (a)), while the year-to-year changes follow closely those of the UK social security registrations, but with one year of lag (panel (b)), which is consistent with the six month window to communicate the new residence abroad and the bureaucratic delays characterizing the formal registration process, which involves communications between the consulate and the municipality of origin. An analysis based on data from the Switzerland Immigration Agency show similar patterns (Figure 3).[9] According to the UK and Switzerland data, actual emigration flows are about 2.6 times that registered by official records: it will be important to account for such difference in measurement when interpreting the magnitudes of our estimates.

[8] The data also contain information on educational attainment, which we use in additional analyses not reported. We restrict the data to include only Italian citizens emigrating abroad although all the results are robust to the inclusion of non-citizens.

[9] A similar analysis has been carried out for the US using the 2005-2015 American Community Survey and the information on the year of arrival and the country of birth. For the earliest years the US data register inflows of Italian born migrants similar, if not larger, than those seen migrating to the UK and Switzerland. However, the survey nature of the data does not allow to precisely estimate the immigration of Italians in the most recent years, nor the year-to-year changes. In any case, the analysis based on the US data also confirms qualitatively the evidence based on the UK and Swiss administrative data.

Table 1 shows the stock of emigrants from Italy by country of destination as of 2000, and their cumulative flows also decomposed into age groups post-2008. The table shows two important points. First, the top destination countries have slightly changed over time: while Italians always emigrated towards Germany, Switzerland and France, in recent years they also moved more towards better-performing countries (e.g., U.K. and the United States) rather than to countries with strong historic ties (e.g., Argentina and Belgium). Second, we already saw in aggregate that the recent emigration flow was stronger for young people.[10]

4 Empirical Specification and Identification

If we could rely on a randomly distributed outflow of young people across local economies, it would be possible to estimate the causal treatment effect of such outflows on firm creation and innovation with a simple OLS regression. However, emigration is likely to be correlated to local observable and unobservable economic and social conditions which are likely to be correlated with our outcomes of interest. Moreover, due to under-reporting of the change of residence status, it is likely that we measure emigration with error. Nevertheless, it is useful to start by looking at the correlation between the change in the firm stock (and in cumulative entry and exit of firms) across local labor markets indexed by l (indicated Δy_l in equation (1) for brevity), and the cumulative migration outflows, $\sum_{t=2008}^{2015} m_{l,t}$ (normalized by the baseline population in 2000 and multiplied by 100) using an OLS regression. In this baseline specification, we also control for a set of pre-determined and observable local labor market characteristics, $X_{l,2004}$, namely GDP per capita and unemployment rate in 2004, to account for economic and demographic performance of the area before the emigration

[10] The table also reports the countries' GDP performances relative to Italy, which will be discussed in Section 4.1.

event. We also include, in different specifications, region and province fixed effects, ϕ_p, which capture unobserved shocks common within clusters of LLMs. We thus estimate the following equation:

$$\Delta y_l = \alpha + \beta \frac{\sum_{t=2008}^{2015} m_{l,t}}{pop_{l,2000}} \cdot 100 + \phi_p + \gamma X_{l,2004} + \varepsilon_l \tag{1}$$

Table 2 reports the OLS estimates of the coefficient β. The main outcomes, represented by birth, death and changes in number of firms, do not appear to be significantly correlated with the emigration of young individuals. OLS estimates, however, could be biased by several factors. First, the error term ε_l includes unobserved area-specific variations in economic, demographic and social factors that may be positively or negatively correlated with emigration. One the one hand, if more successful and resilient localities tend to have more youth emigration (as it appears to be the case from the descriptive statistics shown in Figure 4) the OLS would be biased positively. On the other hand, if young individuals disproportionately leave less economically attractive labor markets, this would produce a negative correlation between emigration and entrepreneurship. Second, the measurement error in emigration could bias the coefficient toward zero.

To reduce these potential biases of the OLS estimates, we exploit variation in migration flows driven by historical networks and due to pull, rather than push, factors. We thus develop an IV strategy.

4.1 Identification: The IV Approach

The key intuition for the instrumental variable, which is based on the work by Anelli and Peri (2017), is that municipalities have connections with specific foreign countries through their networks of past-residents who emigrated to those countries. This network is likely to generate flows of information and opportunities, through personal and family

connections between the foreign country and the municipality of origin. These connections are especially valuable during an economic downturn if those countries provide favorable economic opportunities, thus attracting residents of the municipality abroad. Hence we leverage the interaction between such networks and economic attractiveness of destination countries. First, to proxy for these pre-determined networks, we count the number of people who emigrated from the municipalities in local labor market l to each foreign country c before year 2000, as a percentage of the LLM population in 2000.[11] It is reasonable to think that these networks are stronger the larger is the size of the community of emigrants in the receiving country. We then interact these shares with the real GDP growth of foreign countries during the period 2008-2015[12]. Summing across destination countries results in a LLM-specific pull factor, which captures the intensity of the economic attractiveness exerted by foreign countries during the 2008-15 period to each specific LLM. The variable is defined as follows:

$$Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008}) \tag{2}$$

In expression (2), the first term $sh_{l,c,2000}$ is the number of people from local labor market l who live in country c as of year 2000, as share of the LLM population in year 2000. This factor captures the relative size of the historic network between local labor market l and each specific foreign country c. The second term $GDP_c^{2015}/GDP_c^{2008}$ is the real GDP growth in country c during the period 2008-2015, which includes the deep recession that hit the Mediterranean economies much more then the rest of the European Union. This factor captures the relative, country-specific "pull factor", which proxies the economic

[11] Importantly, while the AIRE registry was started in 1990, all Italians emigrated in previous decades were required to register to continue having consular services. This has allowed to construct historical network using virtually all the first-generation Italian emigrants and not just those emigrating after 1990.

[12] Source: IMF International Financial Statistics. We are able to match more than 100 destination countries that comprise almost all emigration outflows.

incentives for moving to country c during the considered period. Table 1 summarizes the variation in GDP growth between 2008 and 2015 for the main destination countries. The variable defined in (2) is used as instrument for the actual emigration rate, $\frac{\sum_{t=2008}^{2015} m_{l,t}}{pop_{l,2000}}$, our main variable of interest in the estimating equation (1).

The interaction between these two factors, the strength of expatriates network in year 2000 and the economic attractiveness of destination countries in 2008-2015, is likely to be uncorrelated with unobserved factors specific to the local labor markets that may drive both business creation and emigration. Threats to identification remain, however. For instance, past economic shocks, if they are persistent and affected past emigration as well as current firm creation, may constitute such a threat. To increase confidence in the exclusion restriction of our IV strategy, we examine the correlation between pre-2008 trends of the main outcomes of interest (namely in 2005-2008) and the IV capturing predicted emigration. Table 3 reports the pre-period changes in local labor markets with below median predicted emigration (left column) and in LLMs with above median predicted emigration in the post-period. Such distinction allows to compare LLMs with low predicted emigration (as induced by the pull IV) to those with more predicted emigration and to assess whether these two sets of locations differ systematically in their pre-2008 trends. Reassuringly, the two groups of LLMs look remarkably similar in all respects. While this simple check suggests that pre-2008 trends are similar for LLMs with high or low levels of emigration as predicted by our instrument, we conduct more systematic tests of the validity of our IV strategy below, in section 4.3.

4.2 First Stage Results

In Table 4 we report the first stage results when using the "pull factor", $Pull_l$, as our instrument and the emigration rate as the dependent variable. In the regressions we control

for GDP per capita and the unemployment rate in 2004, and we include region fixed effects in column (2) and province fixed effects in column (3). These controls capture predetermined economic conditions in the LLMs of origin. The estimates in the first row of Table 4 show that the pull factor (*Emigration IV*) has a significant predictive power for actual emigration and the size of the coefficient is stable across specifications at around 3.8. The first stage F-statistics lie between 12 and 26, well above the standard rule of thumb value of 10, below which weak instrument concerns would arise.

Among the three specifications, the one including the province fixed effects is the most restrictive as it leverages variation only within provinces; that is, the fixed effects account for all unobservable characteristics common to local labor markets within the same province. In the rest of the paper we use this more demanding specification. Figure 4 shows the geographic variation which we are leveraging. The maps plot the variation in emigration rates (panel (a)) from local labor markets within each province (in bold) as well as the variation in emigration as predicted by the IV pull factor (panel (b)).

4.3 Instrument Validity

In order to strengthen our confidence that the constructed "pull-driven" IV is not correlated to unobserved economic trends at the local level, we perform a series of checks. First, we can exploit information on the period before 2008 (2005-2008) when there was very little emigration (this would be a pre-treatment period using the difference-in-difference terminology). To have a first visual check, we start by dividing the Italian local labor markets into those with high pull-driven emigration rates (above the median) and those with low pull-driven emigration rates (below the median) as measured by our IV. Then Figures 5, 6(a) and 6(b) show, respectively, the three main outcomes of interest: changes in stock of firms and cumulative entries and exits of firms, all normalized by the existing stock

of firms in 2005. All three graphs show that the LLM with high "predicted emigration" (solid line) and those with low "predicted emigration" (dashed line) have similar pre-2008 trends. The number of firms and the average firm creation and firm destruction (as of share of firms in 2005) were moving together prior to 2008 for these two groups of LLMs. However, as the Italian economy starts under-performing with respect to the other major European economies around 2009, and more clearly since 2010, the stock of firms in LLMs with large predicted emigration rates fell behind that of LLMs with few emigrants. Importantly for the validity of our identification strategy, the economies with high or low predicted emigration behave similarly in the pre-2008 period. This is consistent with the instrument being uncorrelated with unobserved and persistent economic factors affecting the firm creation outcomes.

Along the same lines, in order to check the correlation of our instrument with the pre-2008 trends for firm creation, we report the results of two more formal tests. In Table 5 we regress the pre-2008 change in the stock, birth and death of firms on the emigration predicted by the pull-driven instrument post-2008. The test confirms formally that the IV does not predict the pre-2008 firm creation rate – the change in stock for the universe of firms (panel A) or for those owned and managed by people under 45 (panel B). We find, however, a significant negative effect on pre-2008 firm deaths.[13] One could also be concerned that our instrument captures other dimensions of demographic change. Further results, reported in Table 4 Panel B, shows that the instrument predicts emigration abroad and not internal outflows (while it marginally predicts internal past immigration). The analysis is interesting in itself, as it indicates that pull forces from abroad, as identified by

[13] The negative effect on pre-period firm deaths indicates that in areas with more predicted emigration during the Great Recession there were relatively *fewer* firm exits before the recession, which would imply the instrument is correlated with good economic performance pre-2008. Note also that, in a regression not reported but available upon request, we test the first stage effect of the IV on the pre-2008 migration outflows: although positive, the relationship is quite weak and falls below the conventional F-statistics threshold of 10.

our IV, are not substituting or complementing internal migration flows.

Finally, the pull IV resembles the traditional Bartik/shift-share IV in that it combines variation in the cross-sectional distribution of emigrants-stock by destination country with the countries' aggregate economic performance. Recent work by Borusyak et al. (2018) and Goldsmith-Pinkham et al. (2018) show the sufficient conditions needed for identification. On the one hand, our instrument does not seem to violate the conditions for identification regarding the time-varying shock because these are determined outside the Italian economy (one of the strategies suggested by Borusyak et al., 2018). On the other, it could be that the emigration networks that drive the pull factor are strong for few countries only and these are correlated with some characteristics of the origin area. Following the suggestions of Goldsmith-Pinkham et al. (2018), we analyze the cross-sectional components of the IV, including the weight associated with each country in the instrument, and how the stock of emigrants in the main destination countries correlates with the main observable characteristics of the local labor market of origin. Appendix Tables A1 and A2 report the results of such tests: although emigration to Germany drives a large part of the variation of the pull factor-IV, we conclude that there is no systematic reason to believe that our identification strategy would be violated.[14]

5 Main Results

5.1 LATE and Characterization of the Shock

Before getting to the main results, it is important to better understand the characteristics of the local labor markets most affected by the instrument, i.e. the "compliers" LLMs. Those have migration flows due to the external pull factors, and they would not have if the pull factor was absent (or, alternatively, if the emigrant network was empty). Column (1)

[14] See the Appendix for more comments and a description of the tests.

of Table 6 reports again the main first stage result (column 3, Table 4), while columns (2) and (3) show the specification on the subsample of LLMs with a low median age and on the subsample of LLM with a high median age (measured in the period prior to the emigration wave)[15]. Confirming the descriptive facts about emigration presented in Table 1 and Figure 1, the LLMs most affected by the pull factor are those with a relatively low median age. The estimated first stage coefficient is about twice as large for relatively younger LLMs than for relatively older ones. Following a LATE interpretation of the instrumental variable identification, the effects we are going to estimate are identified locally for LLMs with higher shares of young workers – who are potentially also those individuals with higher entrepreneurial skills and more likely to start a new firm.

In Table 7 we then report the results of three regressions (reduced form IV, OLS and 2SLS) in which the outcome is the ratio of 25-44 years old to 45-64 population. Controlling for province fixed effects, economic conditions at the baseline and pre-trends in the ageing structure, we observe that the instrumental variable has a direct negative effect on the relative size of the younger cohort in LLMs. This effect carries through if we run the 2SLS specification (column 3), but it is not observed in the OLS estimates (column 2). This table, although merely suggestive, confirms the idea that the recent emigration wave was more prominent among young people and that the variation we leverage produces a decline in the number of young versus older people, which is consistent with the idea that young people respond more to the economic pull captured by the IV. Another exercise to characterize complier LLMs is based on geography rather than on demographic characteristics (Appendix Table A3). It shows that LLMs affected by the pull factor are more likely to be in the North of the country. This exercise is relevant because actual and predicted emigration plotted in the maps of Figure 4 differ somewhat because of their

[15] The median age of an Italian LLM in the pre-period is 43.6 years and the distribution is close to a Normal with average 43.8 and standard deviation 2.8 years (minimum 35.3 and maximum 57.9).

geographical location. This is consistent with the idea that stronger economic pull had a larger impact on emigration from the North and is not only correlated with long-term trends in emigration which are more prominent in the South.

5.2 Effects on Firm Creation

Panel A of Table 8 shows the results of estimating equation (1) using the change in the stock of active firms as dependent variable, as well as the creation of new and the destruction of existing firms. The coefficients reported are from 2SLS regressions where the endogenous migration flow is instrumented with the pull factor. The dependent variable is the change in total number of firms over the post period (2008-2015) in column (1), cumulative firm births in column (2), and cumulative firm deaths in column (3). All the outcomes are standardized by the stock of active firms in each local labor market in 2005. Standard errors are clustered at the province level.

The estimates indicate that in areas with larger emigration flows between the period 2008-2015 the number of firms declined. This effect is driven by fewer firm births (that is, less firm creation) rather than more firm deaths: on average, for a one percentage point increase in emigration rate there has been a 8.8 percent decline in firms created as share of 2005 LLM firms. A simple back of the envelope calculation, which also accounts for the under-reporting of emigrants in the official data discussed above, indicates that over the whole period, on average, 1,724 individuals left the average LLM and 617 fewer firms have been created.[16]

The small and non-significant coefficient of emigration on the number of firm deaths is reassuring. First, emigration of young people is more logically associated with a decline

[16] The back of the envelope simply relates the emigration flow, aggregated over the period and adjusted by the misreporting factor of 2.6, with the estimated effect on each outcome in an average local labor market.

in potential firm-creation. Second, a prominent channel through firm-failure could suggest a reverse channel of causation, namely people left local labor markets where firms were closing.

The fact that firm creation was particularly affected is consistent with the idea that emigration drained potential entrepreneurs, likely reducing the creation of new firms in the area, an hypothesis we further test later. Further evidence on this is provided in panel B of Table 8. To zoom specifically into the role of young people in starting new firms, we look at the creation and destruction of firms whose owners and managers are younger than 45 (i.e., the same age threshold we use throughout the paper to define "young individuals"). The age of owners and managers is reported in the data from the Chambers of Commerce and we use this information to construct a synthetic measure that identifies a firm as "owned and managed" by young people if the majority of owner-managers are under 45. We then look at the number, creation and destruction of these firms "owned and managed" by young people. The results in panel B, which mirror those of panel A, indicate that emigration reduced the creation of firms whose owners and managers are 45 or younger. According to the estimates, absent emigration (as induced by our pull instrument) there would have been 178 more firms created with respect to an average of 2,470 firms created in the average LLM (or of 2,750 "young-managed" firms for every 10,000 existing in 2005). The effect on "young-managed" firms is, on average, about 30 percent of the total change in new firms created observed.

It is important to consider that this estimated effect of emigration (178 fewer new "young-managed firms") is not simply the result of potential entrepreneurs leaving the region, but captures the aggregate, reduced form effect resulting from multiple channels. With a simple "accounting" exercise, we can decompose the total effect into three components.

The first part is simply due to the fact that some of the young individuals who emigrated would have started a business. This is a pure *demographic effect* and we evaluate it assuming the average probability of starting a firm, r (entrepreneurship rate), for people in the age group of the emigrants, as of 2005. The second part is the effect of departures on the entrepreneurship rate of those who are left. This may be due to the selection of emigrants among those with higher entrepreneurship relative to the average, hence a *selection effect*. Additionally, it may be due to a *peer spillover effect*, the potential externality that emigrants might have on those who did not migrate. For instance, a computer programmer might become an entrepreneur if her friend with a great entrepreneurial idea stays, but may not start a business if her friend leaves the region. Finally, the third part of the effect is a local multiplier. As fewer firms are created and fewer people remain in the area, this decreases the local demand for goods and services reducing the opportunity for additional firms to be created. We call this latter the *local multiplier effect*. Hence, in equation notation we can decompose the aggregate effect of emigration on the creation of young-managed firms as follows:

$$\Delta YoungFirmBirths = \overbrace{\underbrace{\Delta Young}_{-849} \times \underbrace{r_{pre}}_{0.018*7}}^{\text{Demographic}} + \overbrace{\underbrace{Young_{pre}}_{23480} \times \underbrace{\Delta r}_{-0.0004*7}}^{\text{Selection+Spillover}} + \overbrace{\underbrace{\rho}_{\text{Multiplier}} \times [\underbrace{\Delta Young \times r_{pre}}_{\text{Demographic}} + \underbrace{Young_{pre} \times \Delta r}_{\text{Selection+Spillover}}]}^{\text{Local Multiplier Effect}} \tag{3}$$

$$-178 = \overbrace{-107}^{\text{Demographic}} + \overbrace{-63}^{\text{Selection+Spillover}} + \overbrace{0.05 \times [-107 - 63]}^{\text{Local Multiplier Effect}} \tag{4}$$

$$100\% = \overbrace{60\%}^{\text{Demographic}} + \overbrace{35\%}^{\text{Selection+Spillover}} + \overbrace{5\%}^{\text{Local Multiplier Effect}} \tag{5}$$

The decomposition above shows simple imputations that allows us to give a quantitative assessment of each channel and to have an estimate of the local multiplier coefficient.

We estimate the demographic effect by calculating the annualized entrepreneurship rate r for individuals aged 25-44 before the Great Recession (between 2005 and 2008) which is equal to 1.8 percent per year. We then multiply the average number of individuals aged 25-44 who left the representative average-sized local labor market (849) by the annualized pre-recession entrepreneurship rate r cumulative over the 7 year period for which we estimate our aggregate effect. This calculation generates a decrease of -107 "young-managed" firms. Therefore, roughly 60% of the aggregate estimated effect (-178) is imputable to the simple demographic channel. Then, we evaluate the effect on the entrepreneurship rate of people left in the LLM, over the 2008-2015 period, $\Delta r = r_{2015} - r_{2008}$. We do this by estimating our main equation, using 2SLS, with this change as outcome. Δr captures both the change in the entrepreneurship rate due to selection and that due to peer spillovers. The estimated effect of emigration on Δr corresponds to a -0.04 percentage point change in entrepreneurship rate per year. This coefficient multiplied by the average number of young people in the LLM, 23,480, produces a decrease by 63 firms. This accounts for about 35% of the aggregate effect. Then, the residual difference between the total effect and the sum of the two above effects is the "local multiplier effect" (i.e., -8 firms corresponding to 5% of the aggregate effect). Exploiting the simple accounting equation presented above, it is possible to calculate the local firm multiplier ρ implied by our downstream effect: for every 100 firms not created due to the main effects (demographic *plus* selection *plus* spillover), around 5 firms are not created because of the diminished local multiplier.

We acknowledge that part of both the demographic and the local multiplier effects may be due to a loss of workforce rather than of entrepreneurs. The magnitude of this effect is hard to quantify given the limited information we have on those who left and on the characteristics of the employees. In this paper we are focused on the effect on firm creation, and we interpret the reduced demand for qualified labor that we show below as an impact

of fewer firms and hence fewer job opportunities for those people (see section 5.4 and Table 10).

5.3 Effects on Innovative Firms

As entrepreneurship and firm creation are engines to introduce new technologies and to create new jobs, the loss due to emigration may hinder growth through these channels. Such a loss can be particularly damaging if it is also associated with less innovation and slower technological and productivity growth. We analyze the potential impact on innovation, or at least on economic activity in innovative sectors, by focusing on the creation of start-ups, namely newly created firms that operate in technology-intensive sectors and are not spin-offs of larger established firms. We call this group of new firms "innovative start-ups" as they are those more likely to embody genuinely new technologies and ideas.[17] Table 9 shows the results of equation (1) estimated using the number of start-ups in each LLM in the post period as dependent variables. One limitation of our data is that we observe only the net cumulative number of such firms: given that innovative start-ups tend to be short-lived, our variable captures those which were able to survive during the entire period.

The estimated coefficient is statistically significant and indicates that the larger is the migrant outflows from Italian LLMs, the less likely is the creation of innovative start-ups. While on average there were five additional innovative start-ups for every 10,000 existing firms in a local labor market, areas with emigration rates one percentage point higher than the average had essentially zero innovative start-ups. Emigration seems associated with a

[17] Data on start-ups come from the *Registry of Innovative Start-ups*, a special section of the Italian firms registry. 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 at least to 15% of revenues (or costs, if higher); iii) they hold at least a patent of innovative nature. If accepted, these firms benet from favourable fiscal treatments and simplified labor regulations. Firms can maintain this status up to 5 years after registration provided their revenues do not exceed 5 million euros.

worrying decline in the creation of innovative firms which are responsible for job creation and growth. Given 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 et al., 2015) or the UK, and their significant contribution to innovation in the destination countries (see Kerr and Lincoln, 2010), it is very likely that the corresponding effect in countries of origin could be a slowdown of innovation. Moreover, this result has important consequences for job creation, as the youngest firms are those responsible for most job creation (Haltiwanger et al., 2013).

A second, related, exercise splits the main analysis reported in Table 8 for two subgroups: firms acting in high- and in low-value-added (VA) sectors (over the total number of firms present in the LLM in 2005). The results are very informative (Table 9): emigration is impacting the creation of both low- and high-VA firms. This shows that emigration does not only impact low-VA firm dynamism, but has a significant impact on the creation of high-VA firms. When comparing the magnitude of the effects of emigration on high- and low-VA firms, back of the envelope calculations point to a larger decline among low-VA sector firms in absolute value (482 fewer firms created versus 8 in the high-VA in response to an emigration outflows of 1,700 people). However, the initial stock of high-VA firms is very low (around 8.5 percent in 2005). To better appreciate the magnitude of the effect on high-VA firms, the point coefficients indicate that the decline in high-VA firms is as large as the average change over the period. That is, in a local labor market affected by a one percentage point higher emigration outflow, the creation of high-VA firms declines by 1.3 firms for every 100 existing firms at the beginning of the period. This decline is about the same size as the average change of high-VA firms over the period, but has opposite sign. In other words, LLMs with high emigration fully reversed the creation of high-VA firms that LLMs had on average over the same period. For comparison, the decline of low-VA firms

following a one percentage point emigration rate is only one third in size with respect to the average change increase of low-VA firms over the period. These considerations point to a significant role of emigration for creation of firms operating in high-VA sectors.

5.4 Effects on Employment and Skill Composition

The evidence presented suggests that the outflow of young people deprived local economies of entrepreneurs. This also potentially reduced local employment opportunities by lowering job creation. Ex-ante, we can think that for each person who leaves the municipality less than one job may be lost (if the person was not employed or if she is replaced by previously unemployed stayers). To test this we inquire into the effect of emigration on local labor market employment, and decompose this effect by workers' qualification. We show the estimates of this regression in Table 10. The coefficient in column (1) indicates that the associated decline in employees, as a percentage of initial employment, is about 9 percentage points for a one percent increase in emigration – that is, about 1,200 workers for the average emigration rate. Column (2) shows the corresponding effect on the average firm size that is equivalent in magnitude, although not statistically significant.

To better interpret the average effect on employment we perform a similar back of the envelope calculation as the ones proposed in Table 8. Backing out the number of employees lost in an average local labor market, we find a sizeable effect, yet smaller than a 1-to-1 decline: 1,230 fewer employees are observed for 1,724 people (of all working age) who left. We find this effect large, but reasonable, as most of the people who emigrate are working age and it is generally costly to emigrate abroad, especially for low-income, unemployed or inactive people.

Columns (3) to (5) of Table 10 explore in some detail the effects of emigration rates on the labor market skill structure by estimating the effects on the change in the number

of workers, separately by broad occupation group. We distinguish between blue collars, white collars and managerial jobs. We find that, while there is a small non-significant negative coefficient on the number of blue collar workers in the labor market, the negative effect is substantially larger for white collar workers. The association of emigration with a (negative) change in managers is very large but imprecisely estimated and it is not statistically significant. The coefficients indicate a possibly vicious cycle in the decline of highly skilled workers and managers in areas where the youth drain takes place.[18]

Still, there could be a complementary channel that goes through local demand: as more people leave an area, fewer opportunities are created for non-tradable activities. Appendix Table A5 reports the effects of emigration on firms in the tradable and the non-tradable sectors. The coefficients are larger for non-tradable than for tradables firms: in response to a one percentage point higher emigration rate, we find one percentage point fewer tradable firms created versus eight percentage points fewer non-tradable firms created. However, as the number of firms in the non tradable sector is much larger than those in the tradable sector, relative to the average, the effect on firm creation is larger for tradable firms. These are usually firms employing more skilled workers and using more advanced technology, so this result is in line with the high skill content of emigrants.

5.5 Robustness checks

To ensure that our identification strategy is effectively capturing variation in emigrant flows due only to pull forces, we want to rule out some possible alternative channels and

[18] On the one hand fewer workers in the local labor market, if there are issues of crowding or decreasing returns, may increase wages at least in the short-run (Dustmann et al., 2015). On the other hand, the larger loss of skilled young people may reduce local productivity, and the composition of workers will change towards lower wage workers. The combination of selection and productivity effects can therefore produce negative effects on local average wages. This is what we find in Appendix Table A4 for both the total wage bill and average individual wages. It seems therefore that the drain of young productive workers drive wages down, consistent also with the change in skill composition presented in Table 10.

confounding factors.

First, as we focused on net emigration flows, we are neglecting substitution effects between emigrants, Italian internal migrants and foreign-born immigrants. These groups can relocate into areas experiencing migration outflows and re-equilibrate their demographic structure. So, we augment the main analysis by adding as a control the immigration inflows in both the first and second stage. Table 11 shows that the main results hold strongly when adding this control.

The map in Figure 4(a) shows that migration outflows were stronger from border regions, which are possibly different from the rest of the LLMs because their connection with foreign countries in terms of commuting patterns and local trade. Related, a large literature shows that emigrants establish trade relations with their origin countries and possibly with firms in their birth region (Rauch, 1999, 2001). This latter concern would hardly be a threat to our identification strategy as more emigration to a certain country would possibly imply more economic activity and a stronger firm creation in the municipality of origin, especially in the tradable sector, something that we do not find in the data. Yet, this channel might lead to underestimation of the negative effect of emigration on entrepreneurship. We tackle this concern in several ways. First, in our preferred specifications trade patterns are absorbed by the province fixed effects as long as they are common to a province. Second, we test the robustness of our estimates to the exclusion of border provinces, which also have strong trade relations with border countries. The results are presented in Table 11 Column 2: the point estimates barely change, reassuring that our main conclusions are not biased by the presence of specific channels in border regions. Finally, in Table A6 we control for the initial share of tradable firms in the local labor markets and the change in the value of exports from each LLM in the period of interest:

the main conclusions continue to hold.[19]

Tables A7-A11 report additional checks. For each outcome, we run the main 2SLS specification, but also including as controls, in turn, the pre-2008 outcome (whether it is the change of firm stock, the cumulative entries, the cumulative exits, local employment, etc.). The main results are confirmed for all the outcomes.

The set of tests and checks performed support the validity of our IV strategy and, therefore, the identification of causal effects of emigration on firm creation.

6 Conclusions

In this paper, we provide empirical evidence on an important question about which we know very little: what happens to firm creation if the cohort of young individuals becomes smaller? We do this by taking advantage of an emigration wave from Italy, mainly concentrated among young people, and by using an instrumental variable to isolate pull, rather than push, factors. We then combine data on emigrants at the local labor market level with data on firm creation and on new start-ups operating in technology-intensive sectors. The IV-induced variation in the emigration rates across local economies, which exploits past emigration networks and GDP growth in destination countries, is concentrated in areas with a younger demographic structure. Moreover, the validity of our identification strategy is supported by several tests, including the independence of the IV with pre-2005 local trends in firm creation and economic outcomes.

Our results indicate that Italian LLMs that lost more young to emigration experienced less firm creation. Moreover, we observe a smaller number of innovative start-ups in those areas and, in turn, a decrease in local job creation and a drop in the share of highly qualified

[19] Additional regressions, which are not currently available for publication, show that emigration flows do not predict a change in export values.

workers. We then provide a quantitative assessment of how strong are different channels linking emigration to firm creation. We find that about 60 percent of the loss in firm creation can be attributed to a pure demographic effect, as emigration deprives an area of young people. The remaining 40 percent is mainly due to selection of immigrants among people who are more entrepreneurial than average, and to spillovers on the entrepreneurship of people who remain. The results are consistent with ideas put forth by Liang et al. (2018), Acemoglu et al., (2017) and Engbom (2019), namely that demography, and specifically the lack of young managers and young entrepreneurs, may negatively affect firm creation, innovation and ultimately labor demand. Such negative effects of emigration on the economic outcomes of markets of origin may be in part overturned if "brain gain" benefits materialize (e.g., destination-origin trade and innovation networks or higher local investments in human capital; see Beine et al, 2001, and Docquier and Rapoport, 2012). Moreover, the long-term benefits of economic mobility and the role of out-migration in re-equilibrating asymmetric shocks are well acknowledged, but beyond the scope of this paper.

References

[1] Acemoglu, Daron, Ufuk Akcigit and Murat Alp Celik, 2017. "Young, Restless and Creative: Openness to Disruption and Creative Innovations," NBER Working Papers 19894 (updated), National Bureau of Economic Research, Inc.

[2] Altonji, Joseph G. and David Card, 1991. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," NBER Chapters, in: Immigration, Trade, and the Labor Market, pages 201-234 National Bureau of Economic Research, Inc.

[3] Anelli, Massimo and Giovanni Peri, 2017. "Does Emigration Delay Political Change? Evidence from Italy during the Great Recession." *Economic Policy*, vol. 32(91), pages 551–596.

[4] Arpaia, Alfonso, Aron Kiss, Balazs Palvolgyi and Alessandro Turrini, 2016. "Labour mobility and labour market adjustment in the EU," *IZA Journal of Migration*, vol. 5, pages 1-21.

[5] Barker, Vincent L. and George C. Mueller, 2002. "CEO Characteristics and Firm R&D Spending," *Management Science*, 48(6), pages 782-801.

[6] Basso, Gaetano, Francesco D'Amuri and Giovanni Peri, 2018. "Immigrants, Labor Market Dynamics and Adjustment to Shocks in the Euro Area," NBER Working Papers 25091, National Bureau of Economic Research, Inc.

[7] Beine, Micheal, Frédéric Docquier and Hillel Rapoport, 2001. "Brain drain and economic growth: theory and evidence," *Journal of Development Economics*, vol. 64, pages 275–289.

[8] Beyer, Robert C. M. and Frank Smets, 2015. "Labour market adjustments and migration in Europe and the United States: how different?" *Economic Policy*, Volume 32, Issue 84, pages 643682.

[9] Blanchard, Olivier J. and Larry F. Katz, 1992. "Regional Evolutions." *Brookings Papers on Economic Activity*, vol. 1, pages 1-75.

[10] Borusyak, Kirill, Peter Hull, and Xavier Jaravel, 2018. "Quasi-Experimental Shift-Share Research Designs," NBER Working Papers 24997, National Bureau of Economic Research, Inc.

[11] Bratti, Massimiliano and Chiara Conti, 2018a. "Immigration, Population Diversity and Innovation of Italian regions," *Regional Studies*, vol. 52(7), pages 934-947.

[12] Bratti, Massimiliano, Luca De Benedictis and Gianluca Santoni, 2018b. "Immigrant Entrepreneurs, Diasporas and Exports," IZA DP 11280.

[13] Ciccarelli, Carlo, Matteo Gomellini and Paolo Sestito, 2017. "Demographics and Productivity Growth: The Case of Italy," mimeo.

[14] Combes, Pierre-Philippe, Lafourcade Miren and Mayer Thierry, 2005. "The trade-creating effects of business and social networks: Evidence from France," *Journal of International Economics*, vol. 66(1), pages 129.

[15] Dao, Mai, Davide Furceri and Prakash Loungani, 2017. "Regional labor market adjustment in the United States: trend and cycle," *The Review of Economics and Statistics*, vol. 99(2), pages 243-257.

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

[17] Docquier, Frédéric, 2014. "The brain drain from developing countries," IZA World of Labor, no. 31.

[18] Docquier, Frédéric and Hillel Rapoport, 2012. "Globalization, Brain Drain, and Development", *Journal of Economic Literature*, 50(3), pages 681-730.

[19] Docquier Frédéric, Çağlar Özden and Giovanni Peri, 2014. "The Labour Market Effects of Immigration and Emigration in OECD Countries," *Economic Journal*, Royal Economic Society, vol. 124(579), pages 1106-1145.

[20] Dustmann, Christian, Tommaso Frattini and Anna Rosso, 2015. "The Effect of Emigration from Poland on Polish Wages," *Scandinavian Journal of Economics*, Wiley Blackwell, vol. 117(2), pages 522-564.

[21] Elsner, Benjamin, 2013a. "Does Emigration Benefit the Stayers? Evidence from EU Enlargement," *Journal of Population Economics*, vol. 26(2).

[22] Elsner, Benjamin, 2013b. "Emigration and wages: The EU enlargement experiment," *Journal of International Economics*, Elsevier, vol. 91(1), pages 154-163.

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

[24] Giesing, Yvonne and Nadzeya Laurentsyeva, 2017. "Firms Left Behind: Emigration and Firm Productivity," CesIfo Working Papers, no. 6815.

[25] Goldsmith-Pinkham, Paul, Isaac Sorkin, Henry Swift, 2018. "Bartik Instruments: What, When, Why, and How," NBER Working Papers, no. 24408 (updated), National Bureau of Economic Research, Inc.

[26] Grogger, Jeffrey and Gordon H. Hanson, 2011. "Income Maximization and the Selection and Sorting of International Migrants", *Journal of Development Economics*, vol. 95 (1), pages 42-57.

[27] Hunt, Jennifer and Marjolaine Gauthier-Loiselle, 2010. "How Much Does Immigration Boost Innovation?," *American Economic Journal: Macroeconomics* vol. 2, pages 3156.

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

[29] Javorcik, Beata, Caglar Özden, Mariana Spatareanu and Cristina Neagu, 2011. "Migrant networks and foreign direct investment," *Journal of Development Economics*, vol.94, pages 231241.

[30] Karahan, Fatih, Benjamin Pugsley and Aysegül Sahin, 2018. "Demographic Origins of the start-up Deficit," NBER Working Papers 25874, National Bureau of Economic Research, Inc.

[31] 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*, University of Chicago Press, vol. 28(3), pages 473-508.

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

[33] Liang, James, Hui Wang and Edward P. Lazear, 2018. "Demographics and Entrepreneurship," *Journal of Political Economy*, vol. 126(S1), pages S140-S196.

[34] MacDonald, Glenn and Michael S. Weisbach, 2004. "The Economics of Has-beens," *Journal of Political Economy*, 2004, vol. 112, no. 1, pages S289-S310.

[35] 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*, De Gruyter, vol. 9(1), pages 1-52.

[36] Mishra, Prachi, 2007. "Emigration and wages in source countries: Evidence from Mexico," *Journal of Development Economics*, Elsevier, vol. 82(1), pages 180-199.

[37] Peri, Giovanni and Francisco Requena-Silvente, 2010. "The trade creation effect of immigrants: evidence from the remarkable case of Spain," *Canadian Journal of Economics*, 2010, vol. 43(4), 1433-1459.

[38] Peri, Giovanni, Kevin Shih and Chad Sparber, 2015. "STEM Workers, H-1B Visas, and Productivity in US Cities," *Journal of Labor Economics*, University of Chicago Press, vol. 33(S1), pages S225 - S255.

[39] Portuguese Observatory of Emigration, 2015. "Emigração Portuguesa Relatório Estatístico 2015," Technical report (available at `https://bit.ly/2U1Bb6E`, last checked on April 7, 2019).

[40] Rauch, James E., 1999. "Networks versus markets in international trade," *Journal of International Economics*, vol. 48(1), pages 735.

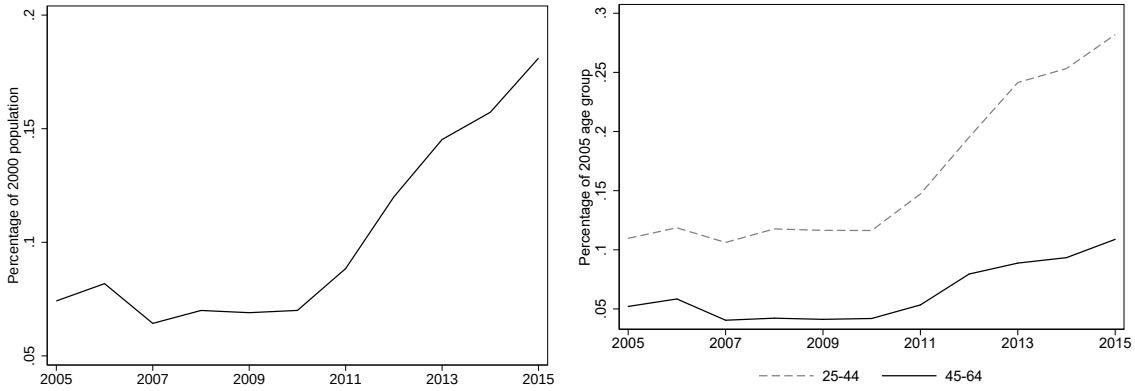
[41] Rauch, James E., 2001. "Business and social networks in international trade," *Journal of Economic Literature*, vol. 39(1), pages 11771203.

[42] Rauch, James E. and Vitor Trindade, 2002. "Ethnic Chinese networks in international trade," *Review of Economics and Statistics*, vol. 84(1), pages 116130.

[43] Schivardi, Fabiano and Tom Schmitz, 2018. "The IT Revolution and Southern Europe's Two Lost Decades," CPER Discussion Papers, no. 12843.

Figures

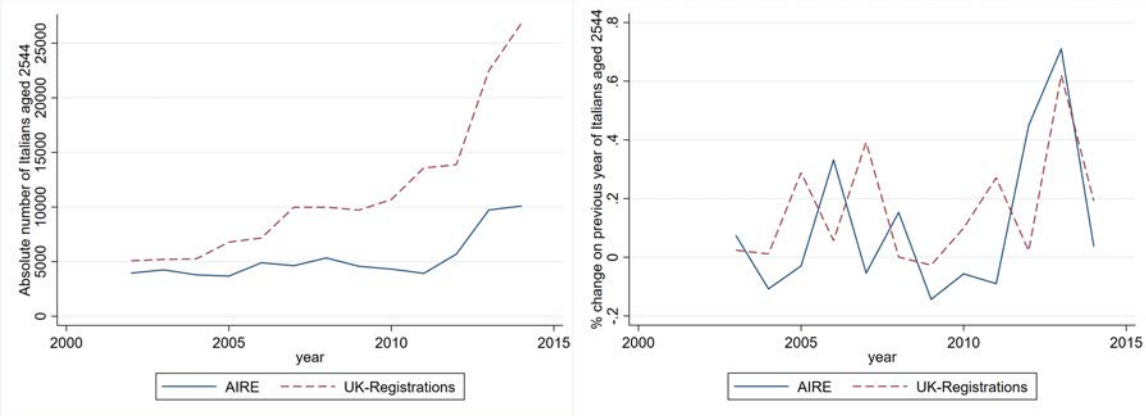
Figure 1: Emigration flows, percentage of 2005 population, 2005-2015



(a) Annual overall outflows

(b) Annual outflows, by age group

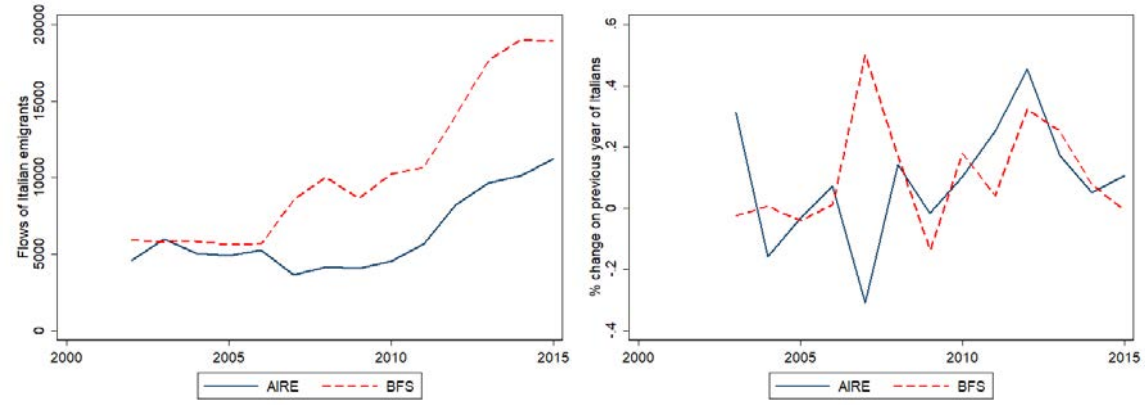
Figure 2: Recorded Emigration and UK Registered Italian Inflows



(a) Annual inflows from AIRE and UK Social Security Registry, 2005-2015

(b) Annual changes from AIRE and UK Social Security Registry, 2005-2015

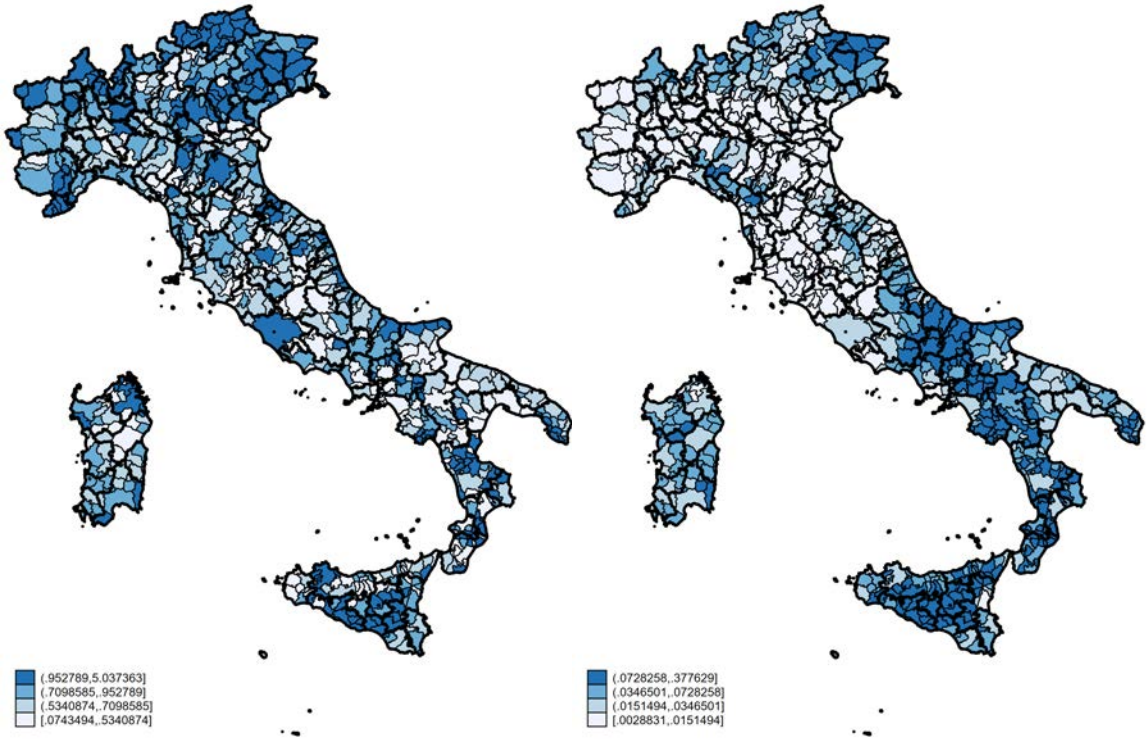
Figure 3: Recorded Emigration and Switzerland Registered Italian Inflows



(a) Annual inflows from AIRE and Switzerland BFS Registry, 2002-2015

(b) Annual changes from AIRE and Switzerland BFS Registry, 2002-2015

Figure 4: Actual and Predicted Emigration from Italian LLMs



(a) Emigration (b) Predicted Emigration

Figure 5: Firm stock in predicted high and low emigration LLMs, 2005-2015

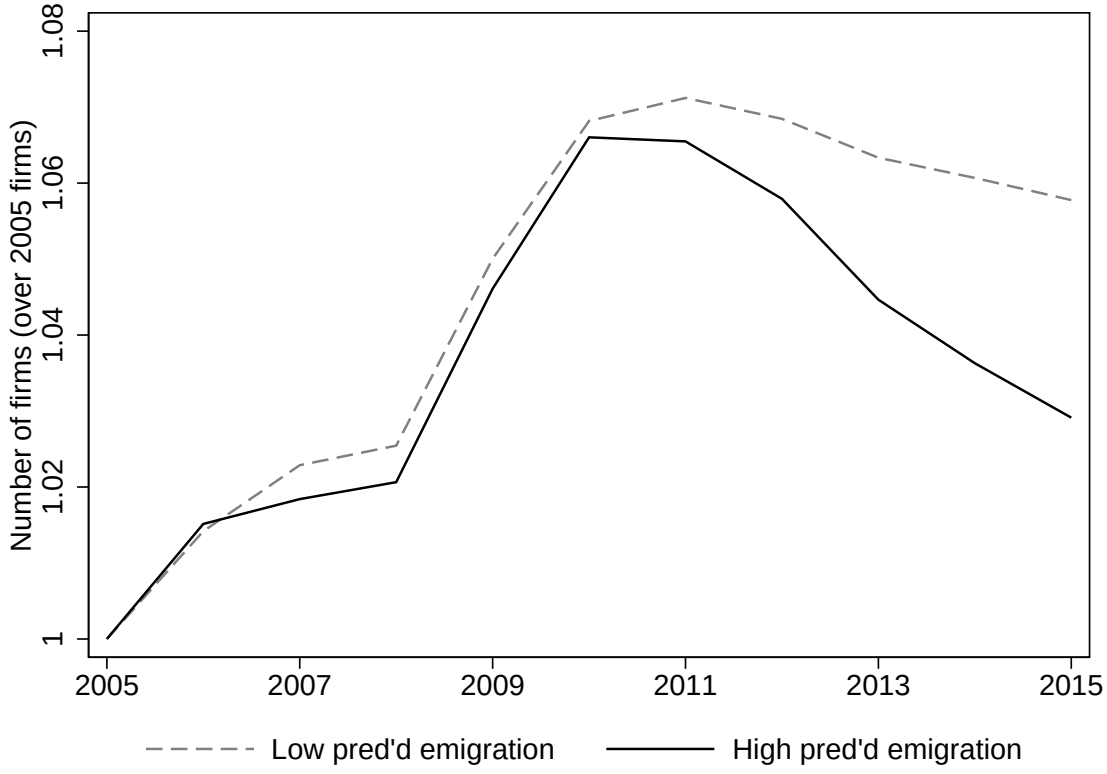
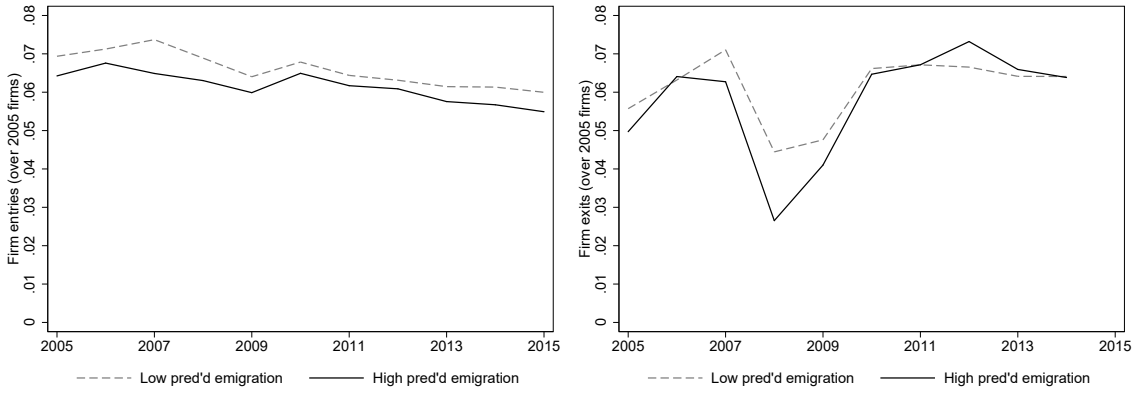


Figure 6: Firm flows in predicted high and low emigration LLMs, 2005-2015



(a) Firm births

(b) Firm deaths

Tables

Table 1: Emigration by country of destination, top 5 countries: 2000 stock, GDP performance and 2005-2015 flows

	(1)	(2)
Panel A		
Top countries in 2000	Stock of Emigrants	GDP 2015/2008
Germany	286,570	1.07
Switzerland	228,725	1.09
France	165,244	1.04
Belgium	117,935	1.06
Argentina	99,506	1.11
Panel B		
Top countries in $2008-15$	Flows	% 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 main countries in terms of size of the emigration network as of 2000 as measured in the AIRE data, and the GDP growth between 2008 and 2015 calculated on IMF data (out of a total of 92 countries considered). For reference, both UK and US growth was 1.19 in the period of interest and GDP growth in Italy was 0.93. Panel B reports the flows of emigrants to the top destination countries in the periods 2008-2015 and the share of 25-44-year old measured in the ISTAT-AIRE data.

Table 2: OLS regression of LLMs firm dynamics on observed emigration rates

VARIABLES	(1) All Firms Δ Stock 2008-15	(2) All Firms $\sum$ Births 2008-15	(3) All Firms $\sum$ Deaths 2008-15
Panel A			
Emig Rate	-0.008	0.005	0.013**
	(0.006)	(0.006)	(0.005)
R-squared	0.004	0.001	0.014
Panel B			
Emig Rate	0.005	0.021***	0.016***
	(0.007)	(0.007)	(0.006)
Emig Rate*Young	-0.027**	-0.036***	-0.009
	(0.011)	(0.011)	(0.010)
R-squared	0.072	0.104	0.039
Panel C			
Emig Rate	-0.000	0.015*	0.015**
	(0.004)	(0.008)	(0.006)
Unemp.Rate 2004	0.214	0.764***	0.550***
	(0.144)	(0.230)	(0.149)
GDP 2004	0.205***	0.278*	0.073
	(0.078)	(0.146)	(0.070)
R-squared	0.567	0.460	0.599
Avg. Outcome	0.005	0.344	0.339
Avg. Treatment	8518.980	8518.980	8518.980

Notes: OLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables are the changes in firm stock (column 1), cumulative firm entry (column 2) and exit (column 3) between 2008 and 2015 as a fraction of the stock of firms in 2005. The independent variable is the cumulative emigration rate between 2008 and 2015 as a fraction of the LLM population in 2000, times 100. In Panel A there are no control variables. In Panel B we include the indicator Young defined as average age in 2002 below median, and we interact it with emigration. In Panel C we control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level and we include province (110) FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 3: Pre-period local labor market characteristics, 2005-2008

	Below median predicted emigration	Above median predicted emigration
Firm stock over 2005 firm stock	0.023	0.025
	(0.032)	(0.038)
Under 45 owned firms over 2005 firm stock	-0.006	-0.002
	(0.022)	(0.024)
Entry of firms over 2005 firm stock	0.196	0.199
	(0.039)	(0.041)
Exit firms over 2005 firm stock	0.167	0.169
	(0.023)	(0.025)
Entry of under 45 owned firms over 2005 firm stock	0.185	0.193
	(0.040)	(0.039)
Exit of under 45 owned firms over 2005 firm stock	0.077	0.080
	(0.020)	(0.018)
Change in LLM employees over 2005 employees	0.116	0.121
	(0.118)	(0.115)
Change in avg firm size	0.021	0.119
	(0.810)	(0.840)
Change in number of managers to blue collars	-0.004	-0.014
	(0.245)	(0.091)

Notes: LLM averages and standard deviations in parenthesis, 2005-2008. The 686 LLMs are split in two subsamples, one above and one below the median emigration rate as predicted by the pull-instrument. Further details in the text.

Table 4: First stage regression

Panel A

	(1) Emig Rate	(2) Emig Rate	(3) Emig Rate
Emig. IV	3.722***	3.883***	3.803***
	(0.771)	(0.758)	(1.077)
Unemp.Rate 2004	-1.720**	0.784	1.791
	(0.736)	(1.114)	(1.621)
GDP 2004	0.564***	0.566***	0.665***
	(0.141)	(0.111)	(0.195)
Observations	686	686	686
R-squared	0.131	0.243	0.397
F-excluded instrument	23.312	26.243	12.458
Avg. Outcome	0.825	0.825	0.825
FE	-	Region	Province

Panel B

	(1) Internal Emig	(2) Immigration 05-08
Emig. IV	-0.661	-1.259*
	(2.624)	(0.690)
Observations	686	686
R-squared	0.400	0.677
F-excluded instrument	0.064	3.329
Avg. Outcome	8.496	1.316
Province FE	X	X

Notes: OLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). In Panel A, the dependent variable is the cumulative emigration rate between 2009 and 2015 as a fraction of the LLM population in 2000, times 100. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level. Column 1 include no fixed effects while Columns 2 and 3 include region (20) and province (110) FEs respectively. In Panel B, Column 1, the dependent variable is the cumulative emigration rate between 2009 and 2015 to different LLMs in Italy, while in Column 2 the dependent variable is the change in the stock of immigrants from abroad between 2015 and 2009, as a fraction of LLM population in 2000. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 5: Instrument validity check: Effect of emigration rates on pre-shock change in stock and flows of firms (2005-08)

	(1) All Firms Δ Stock 2005-08	(2) All Firms $\sum$ Births 2005-08	(3) All Firms $\sum$ Deaths 2005-08
Panel A: All firms			
Emig. IV	-0.062	-0.081	-0.019
	(0.051)	(0.064)	(0.023)
R-squared	0.483	0.495	0.608
Avg. Outcome	0.025	0.198	0.173
Avg. Treatment	0.825	0.825	0.825
Panel B: Firms with ownership under 45			
Emig. IV	0.016	-0.087	-0.103***
	(0.047)	(0.063)	(0.032)
R-squared	0.449	0.474	0.483
Avg. Outcome	-0.004	0.190	0.193
Avg. Treatment	0.825	0.825	0.825
Observations	686	686	686
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables are the changes in firm stock, cumulative firm entry and exit between 2005 and 2008 as a fraction of the stock of firms in 2005, in each column respectively. In Panel A we include all firms while in Panel B we include only firms owned and managed by individuals under 45 years old, which we discuss in Section 5. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported), as well as for 110 province FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 6: First stage regression: Complier LLMs, by median age in 2005

	(1) All	(2) Rel. Younger	(3) Rel. Older
Emig. IV	3.803***	5.084***	2.449**
	(1.077)	(1.695)	(0.995)
Observations	686	343	343
R-squared	0.397	0.472	0.477
F-excl. instr.	12.460	9.000	6.060
Avg. Outcome	0.825	0.899	0.750
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables is the emigration rate in LLMs with a relatively young and a relatively old demographic structure based on their median age in the pre-period (2005) in Columns 2 and 3 respectively. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported), as well as 110 province FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 7: Emigration and local labor market age composition

	(1) Δ 25-44/45-64	(2) Δ 25-44/45-64	(3) Δ 25-44/45-64
Emigration IV	-0.100***		
	(0.028)		
Emigration rate		0.002	-0.024***
		(0.003)	(0.007)
$\Delta_{pre-period}$ 25-44/45-64	0.739***	0.746***	0.761***
	(0.014)	(0.014)	(0.015)
Observations	686	686	686
R-squared	0.898	0.896	0.880
Avg Ratio 09-15	0.918	0.918	0.918
Avg Ratio 05-08	1.099	1.099	1.099
Avg Emigr IV	0.053		0.053
Avg Emig		0.828	0.828
Province FE	Yes	Yes	Yes
Method	OLS	OLS	2SLS

Notes: OLS, reduced form IV and 2SLS estimates, in each column respectively. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables is the age composition odds ratio, i.e. the ratio between the average number of residents aged 25-44 between 2009 and 2015 over the average number of residents aged 25-64 in the same period. The independent variables are the cumulative emigration rate between 2008 and 2015 as a fraction of the LLM population in 2000, the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$, and the emigration rate instrumented with the predicted emigration rate, in each column respectively. We control for the age composition odds ratio for the period 2005-2008, unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported), as well as 110 province FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 8: Effect of emigration rates on change in stock and flows of firms

	(1) All Firms Δ Stock 2008-15	(2) All Firms $\sum$ Births 2008-15	(3) All Firms $\sum$ Deaths 2008-15
Panel A: All firms			
Emig Rate	-0.053***	-0.088**	-0.035
	(0.018)	(0.038)	(0.026)
R-squared	0.427	0.125	0.447
F-excl. instr.	12.458	12.458	12.458
Avg. Outcome	0.005	0.344	0.339
Avg. Treatment	0.825	0.825	0.825
Back of Env Avg. Emig.	1724	1724	1724
Back of Env Firms	-373	-617	-244
Panel B: Firms with ownership under 45			
Emig Rate	-0.018**	-0.061**	-0.043
	(0.009)	(0.030)	(0.028)
R-squared	0.415	0.300	0.289
F-excl. instr.	12.458	12.458	12.458
Avg. Outcome	-0.089	0.275	0.364
Avg. Treatment	0.825	0.825	0.825
Back of Env Avg. Emig.	1724	1724	1724
Back of Env Firms	-53	-178	-125
Observations	686	686	686
Province FE	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables are the changes in firm stock, cumulative firm entry and exit between 2008 and 2015 as a fraction of the stock of firms in 2005, in each column respectively. In Panel A we include all firms while in Panel B we include only firms owned and managed by individuals under 45 years old. The independent variable is the cumulative emigration rate between 2008 and 2015 as a fraction of the LLM population in 2000 times 100, instrumented by the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported) as well as 110 province FEs. The back of the envelope report the number of emigrants from the average local labor market over the period adjusted by the misreporting factor of 2.6 and the estimated effect on each outcome in an average LLM. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 9: Effect of emigration rates on innovation

	(1) Start-Ups $\sum$ Births 2008-15	(2) High VA $\sum$ Births 2008-15	(3) Low VA $\sum$ Births 2008-15
Emig Rate	-0.049***	-0.013***	-0.074**
	(0.017)	(0.004)	(0.035)
Observations	686	686	686
R-squared	0.241	0.168	0.194
F-excl. instr.	12.458	12.458	12.458
Avg. Outcome	0.052	0.013	0.330
Avg. Outcome 2005	0.000	0.085	0.915
Avg. Treatment	0.825	0.825	0.825
Back of Env Avg. Emig.	1724	1724	1724
Back of Env Firms		-8	-479

Notes: 2SLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). In Column 1, the dependent variable is the number of innovative start-ups created between 2010 and 2016 as a fraction of the stock of firms in 2005 times 10,000. In Columns 2 and 3, the dependent variables are the changes are the changes in cumulative firm entry between 2008 and 2015 as a fraction of the stock of firms in 2005, in high and low value added sectors respectively. The independent variable is the cumulative emigration rate between 2008 and 2015 as a fraction of the LLM population in 2000 times 100, instrumented by the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported) as well as 110 province FEs. The back of the envelope report the number of emigrants from the average local labor market over the period adjusted by the misreporting factor of 2.6 and the estimated effect on each outcome in an average LLM. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 10: Effect of emigration rates on change in local labor market employment

	(1) Δ Employees% 2008-15	(2) Δ Avg. Size 2008-15	(3) Δ Blue Coll 2008-15	(4) Δ White Coll 2008-15	(5) Δ Managers 2008-15
Emig Rate	-0.089**	-0.089	-0.032	-0.110*	-3.187
	(0.041)	(0.217)	(0.054)	(0.058)	(3.296)
Observations	686	686	686	686	584
R-squared	0.198	0.278	0.197	0.233	0.162
F-excl. instr.	12.458	12.458	12.458	12.458	4.379
Avg. Outcome	-0.110	-0.042	-0.117	-0.013	0.248
Avg. Outcome 2005	16709.0	5.5	8950.138	6737.377	163.226
Avg. Treatment	0.825	0.825	0.825	0.825	0.799
Estimated Emig.	1724	1724	1724	1724	1672
Back of Env. Effect	-1229.9	-0.4	-233.5	-612.4	-415.8
Province FE	X	X	X	X	X

Notes: 2SLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables are the change in employment (as a share of initial number of employees in each local labor market in 2005), average size and workers by qualifications between 2008 and 2015, in each column respectively. The independent variable is the cumulative emigration rate between 2008 and 2015 as a fraction of the LLM population in 2000 times 100, instrumented by the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported) as well as 110 province FEs. The back of the envelope report the number of emigrants from the average local labor market over the period adjusted by the misreporting factor of 2.6 and the estimated effect on each outcome in an average LLM. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 11: Robustness checks: controlling for immigration and excluding border provinces

Panel A: Second stage		
	(1)	(2)
	All Firms	All Firms
	$\sum$ Births	$\sum$ Births
	2008-15	2008-15
Emig Rate	-0.072**	-0.093**
	(0.031)	(0.042)
Immig Rate	0.015***	
	(0.003)	
Observations	686	590
R-squared	0.262	0.086
F-excl. instr.	15.083	11.165
Avg. Outcome	0.344	0.350
Avg. Treatment	0.825	0.767
Back of Env Avg. Emig.	1724	1605
Back of Env Firms	-502	-591
Province FE	X	X

Panel A: First stage		
	(1)	(2)
	Emig Rate	Emig Rate
Emig. IV	4.166***	3.728***
	(1.073)	(1.116)
Immig Rate	0.085***	
	(0.017)	
Observations	686	590
R-squared	0.426	0.313
F-excluded instrument	15.083	11.165
Avg. Outcome	0.825	0.767
Province FE	X	X

Notes: 2SLS (Panel A) and first-stage OLS (Panel B) estimates. In Column 1, the sample is composed of 686 Local Labor Markets, while in Column 2 the sample is composed of 590 Local Labor Markets (LLMs), excluding those in the provinces at the boundary of Italy. In Panel A, the dependent variable is the change in cumulative firm entry between 2008 and 2015 as a fraction of the stock of firms in 2005. The independent variable is the cumulative emigration rate between 2008 and 2015 as a fraction of the LLM population in 2000 times 100, instrumented by the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. In Column 1, we also include the cumulative immigration rate between 2008 and 2015 as a percentage of LLM population in 2000. In both columns, we further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported) as well as 110 province FEs. In Panel B, the dependent variable is the Emigration rate defined above and the independent variables are $Pull_l$, unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported) as well as province FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

A Appendix for Online Publication

A.I Additional figures

Figure A1: Absolute share of 2005 population emigrating by education level, 2005-2015

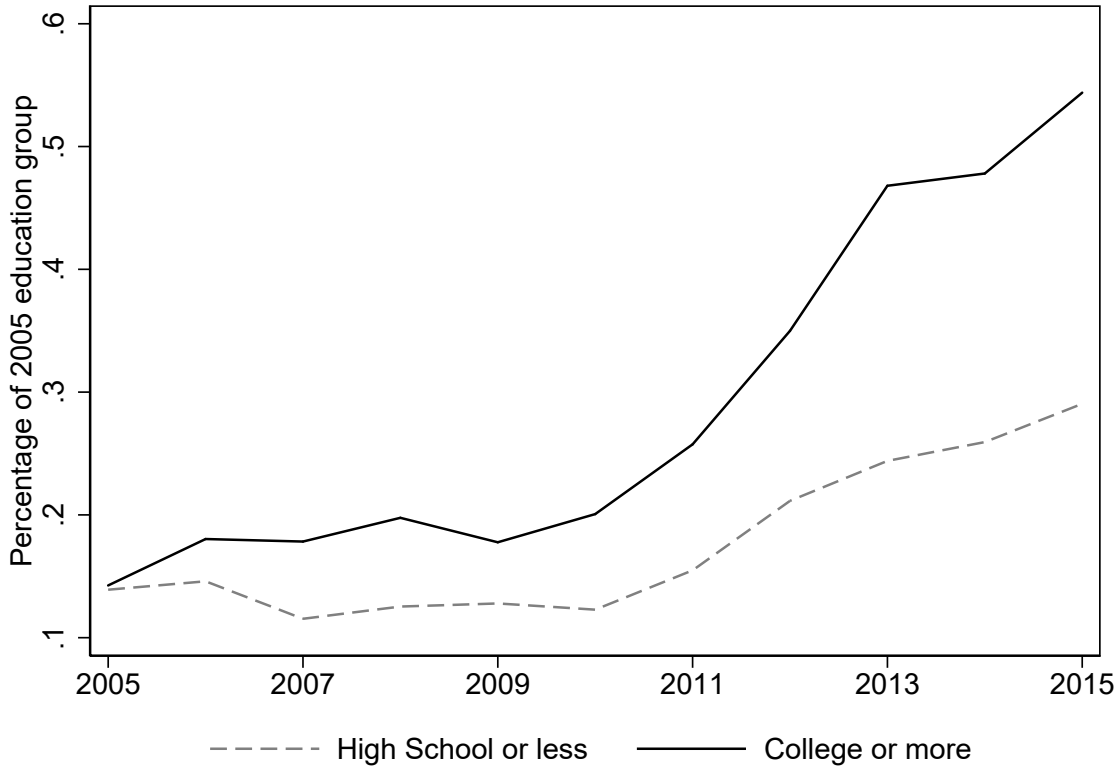
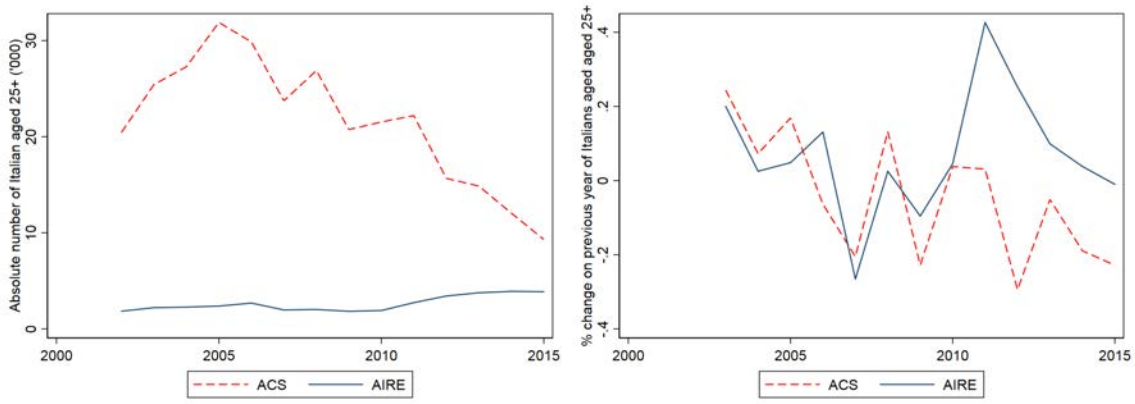


Figure A2: Recorded Emigration and US Registered Italian Inflows



(a) Annual inflows from AIRE and US Census ACS, 2005-2015

(b) Annual changes in inflows from AIRE and US Census ACS, 2005-2015

A.II IV diagnostics

The following tables, A1 and A2, report the main tests proposed by Goldsmith-Pinkham et al. (2018). Table A1 reports the diagnostics of the cross-sectional components of the pull emigration forces instrumental variable. The βs are estimated on the change in the stock of firms in each local labor market. As it can be noted in panel A, the share of the weights below zero is very small indicating that our estimates are likely to be little sensitive to misspecification. More important is panel C that reports the five top destination countries in terms of weights, i.e., the countries that matters the most for our identifying variation. To a certain extent it is worrisome to observe that Germany weighs up to 44 percent of the total instrument variation. On the other hand, the estimated coefficients for the top five destinations are all negative and close to each other and to the main estimates. Note that we do not report diagnostics by time period because our variation spans only the period 2008-2015.

Table A2 instead relates the stock of emigrants in the five main destination countries

with the main observable characteristics of the origin local labor markets (unemployment rate and GDP per capita in the baseline year, and emigration flows by demographic group in the pre-period 2005-2008). As Germany has a prominent weight in our IV variation, we are concerned that emigration to Germany is concentrated in labor markets with peculiar characteristics. From this Table, however, we see no systematic correlation between the destination country networks and the origin area characteristics in the baseline period.

One might be concerned that our instrument is correlated with internal migration flows. While these should not be correlated with pull factors from abroad, the network of emigrants may be correlated with the internal flows and with local push factors. For instance, local labor markets with high emigration rates to foreign countries could also exhibit substantial emigration to other Italian LLMs, and the latter may reduce firm creation, violating the exclusion restriction. We thus test whether our estimates are robust to this potential threat. In Table 4 Panel B we report the results of our first-stage equation where internal migration outflows is the outcome variable. The effect is not statistically significant: this indicates that the instrument predicts emigration abroad and not internal flows. Moreover, the analysis is interesting in itself, as it indicates that pull forces from abroad, as identified by our IV, are not substituting or complementing internal migration flows. In column (2) we test whether there is a direct substitution effect by regressing immigration inflows on (instrumented) emigration. We estimate a negative and marginally statistically significant effect. While this may imply that areas with lower business dynamism – triggered by higher emigration flows – are less attractive for immigrants, our main estimates are robust to the inclusion of immigration as a control variable (Table 11).

Finally, in Table A3 we perform an exercise to characterize complier LLMs based on geography rather than on demographic characteristics (in Table 6). Dropping the province fixed effects, Table Table A3 shows whether LLMs affected by the pull factor are more

likely to be in the South or the North of the country. This exercise is relevant because actual and predicted emigration plotted in the maps of Figure 4 differ somewhat because of their geographical location. While predicted emigration was slightly higher in the South, where historically there has been more emigration and hence stronger networks, actual emigration in the considered period was stronger in the North despite better economic conditions. The table shows that the first stage coefficient is higher for Northern than for Southern LLMs. This is also consistent with the idea that stronger economic pull had a larger impact on emigration from the North and is not only correlated with long-term trends in emigration which are more prominent in the South. We also emphasize that in the IV specification with province fixed effects we use variation within a province, which is much more specific to local networks and does not rely on generic differences between North and South, which are also correlated with economic performance.

Table A1: Emigration pull factor IV diagnostics

Panel A

	Sum	Mean	Share
αs≤ 0	-0.034	-0.001	0.446
αs> 0	1.034	0.020	0.554

Panel B

	αs	G	βs
αs	1.000		
G	-0.069	1.000	
βs	-0.033	-0.037	1.000

Panel C

	α	G	β
Germany	0.442	1.141	-0.043
Switzerland	0.271	1.165	-0.035
France	0.054	1.115	-0.137
Australia	0.040	1.283	-0.073
Canada	0.037	1.196	-0.096

Panel D

	Sum	Mean	p25	Median	p75	Sh< 0
$\hat{\beta}_k$	4.640	0.050	-0.129	-0.024	0.098	0.620

Notes: The table reports the emigration pull diagnostics as suggested by Goldsmith-Pinkham et al. (2018). Panel A reports the share of negative weights; panel B reports correlations between the weights, the 2008-2015 destination country GDP growth (G) and the just-identified coefficient estimates; panel C reports the top five destination countries according to the Rotemberg decomposition calculated with the Goldsmith-Pinkham et al.'s Stata package. Finally, panel D reports statistics about the dispersion in the just identified coefficients $\hat{\beta}_k$

Table A2: Relationship between country of destination emigration networks and LLMs' characteristics

	(1) Germany	(2) Switzerland	(3) France	(4) Australia	(5) Canada
Unempl rate 2004	0.076	-0.001	0.014	-0.054	-0.007
	(0.068)	(0.047)	(0.025)	(0.038)	(0.022)
GDP per cap 2004	-0.001	-0.001	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Δ_{pre} Emig 25-44	-0.021	0.029	-0.001	-0.001	0.007
	(0.022)	(0.028)	(0.017)	(0.009)	(0.009)
Δ_{pre} Emig 45-64	0.036	0.018	0.038	0.021	0.008
	(0.038)	(0.041)	(0.026)	(0.015)	(0.013)
Δ_{pre} Emig 65+	-0.066	0.048	0.004	-0.025	0.017
	(0.051)	(0.036)	(0.023)	(0.023)	(0.013)
Δ_{pre} Emig Female	0.031*	0.017	0.021	-0.002	-0.000
	(0.017)	(0.023)	(0.014)	(0.007)	(0.006)
Δ_{pre} Emig Male	0.014	-0.041	-0.019	-0.001	-0.011
	(0.024)	(0.036)	(0.018)	(0.009)	(0.008)
Observations	686	686	686	686	686
R-squared	0.492	0.386	0.427	0.498	0.343
Province FE	X	X	X	X	X

Notes: OLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables are the shares of emigrants in the five destination countries described in each column. The independent variables are the LLMs main observable characteristics. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

A.III Additional tables

Table A3: First stage regression: complier LLMs, by macro-region

	(1) All	(2) North	(3) Center	(4) South
Emig. IV	3.803***	4.443**	2.621**	3.903***
	(1.077)	(1.933)	(1.218)	(0.953)
Unemp.Rate 2004	1.791	0.268	-2.611**	0.546
	(1.621)	(7.059)	(1.176)	(1.271)
GDP 2004	0.665***	0.488**	0.472***	-0.106
	(0.195)	(0.185)	(0.088)	(0.630)
Constant	0.682***	0.779**	0.728***	0.431***
	(0.085)	(0.312)	(0.065)	(0.156)
Observations	686	235	154	297
R-squared	0.397	0.074	0.111	0.193
F-excl. instr.	12.460	5.290	4.630	16.770
Avg. Outcome	0.825	0.927	0.661	0.828
Province FE	X	NO	NO	NO

Notes: OLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables is the emigration rate in LLMs by macro-region, in each column respectively. The independent variable is the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table A4: Effect of emigration rates on change in local labor market wages

	(1) Δ Wage Bill 2008-15	(2) Δ CZ Wage 2008-15
Emig Rate	-0.034	-0.023***
	(0.045)	(0.008)
Observations	686	686
R-squared	0.264	0.098
F-excl. instr.	12.458	12.458
Avg. Outcome	-0.113	-0.007
Avg. Outcome 2005	3486.004	1644.434
Avg. Treatment	0.825	0.825
Estimated Emig.	1724	1724
Back of Env. Effect	-96.5	-30.6
Province FE	X	X

2SLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables are the change 2008-2015 in wage bill (as a share of initial wage bill in each LLM in 2005) and the change 2008-2015 in average LLM wage (as a share of initial average wage in each LLM in 2005) in each column respectively. The independent variable is the cumulative emigration rate between 2008 and 2015 as a fraction of the LLM population in 2000 times 100, instrumented by the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported) as well as 110 province FEs. The back of the envelope report the number of emigrants from the average local labor market over the period adjusted by the misreporting factor of 2.6 and the estimated effect on each outcome in an average LLM.

Table A5: Effect of emigration rates on firms growth, in tradable and non tradable sectors

	(1) Tradable $\sum$ Births 2008-15	(2) Non Tradable $\sum$ Births 2008-15
Emig Rate	-0.011***	-0.076**
	(0.004)	(0.035)
Observations	686	686
R-squared	0.248	0.178
F-excl. instr.	12.458	12.458
Avg. Outcome	0.030	0.314
Avg. Outcome 2005	0.119	0.881
Avg. Treatment	0.825	0.825
Back of Env Avg. Emig.	1724	1724
Back of Env Firms	-9	-470

Notes: 2SLS estimates. The sample is composed of 686 Local Labor Markets (LLMs). The dependent variables are the changes in cumulative firm entry between 2008 and 2015 as a fraction of the stock of firms in 2005, in tradeble sectors (Column 1) and non tradable sectors (Column 2) respectively. The independent variable is the cumulative emigration rate between 2008 and 2015 as a fraction of the LLM population in 2000 times 100, instrumented by the predicted emigration rate based on the shares of pre-2000 emigrants to different countries to LLM population in 2000 interacted with real GDP growth of each country between 2008 and 2015, $Pull_l = \sum_c sh_{l,c,2000} * (GDP_c^{2015}/GDP_c^{2008})$. We further control for unemployment rate and value added per capita in 100,000 euros in 2004 at the LLM level (not reported) as well as 110 province FEs. The back of the envelope report the number of emigrants from the average local labor market over the period adjusted by the misreporting factor of 2.6 and the estimated effect on each outcome in an average LLM. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table A6: Robustness check - Main effect controlling for the share of tradeable sector in the LLM

	(1) All Firms Δ Stock 2008-15	(2) All Firms Δ Stock 2008-15	(3) All Firms Δ Stock 2008-15	(4) All Firms Δ Stock 2008-15	(5) All Firms ∑ Births 2008-15	(6) All Firms ∑ Births 2008-15	(7) All Firms ∑ Births 2008-15	(8) All Firms ∑ Births 2008-15	(9) All Firms ∑ Deaths 2008-15	(10) All Firms ∑ Deaths 2008-15	(11) All Firms ∑ Deaths 2008-15	(12) All Firms ∑ Deaths 2008-15
Emig Rate	-0.041***	-0.073***	-0.033**	-0.053***	-0.064*	-0.109***	-0.028	-0.088**	-0.023	-0.036	0.005	-0.035
	(0.015)	(0.022)	(0.016)	(0.018)	(0.033)	(0.040)	(0.026)	(0.038)	(0.024)	(0.025)	(0.021)	(0.026)
Tradable Sh. 2005	0.388***		0.405***		0.748***		0.827***		0.360***		0.422***	
	(0.111)		(0.107)		(0.167)		(0.160)		(0.073)		(0.069)	
Tradable Sh. x Emigr. 2000		1.301	-0.487			1.396	-2.253**			0.094	-1.766***	
		(1.258)	(0.832)			(1.828)	(0.966)			(0.852)	(0.660)	
Δ Trade per capita				0.001				0.001				0.001
				(0.001)				(0.001)				(0.001)
Observations	686	686	686	686	686	686	686	686	686	686	686	686
R-squared	0.509	0.306	0.540	0.427	0.320		0.475	0.125	0.535	0.438	0.632	0.447
F-excl. instr.	11.638	15.483	13.992	12.402	11.638	15.483	13.992	12.402	11.638	15.483	13.992	12.402
Avg. Outcome	0.005	0.005	0.005	0.005	0.344	0.344	0.344	0.344	0.339	0.339	0.339	0.339
Avg. Treatment	0.825	0.825	0.825	0.825	0.825	0.825	0.825	0.825	0.825	0.825	0.825	0.825
Back of Env Avg. Emig.	1724	1724	1724	1724	1724	1724	1724	1724	1724	1724	1724	1724
Back of Env Firms	-287	-512	-232	-373	-452	-766	-194	-617	-165	-254	37	-244
Province FE	X	X	X	X	X	X	X	X	X	X	X	X

Sample: Italian Local Labor Markets.
Specifications: 2SLS
Dependent variable: Changes in firm stock (col 1-4), cumulative firm entry(col 5-8) and exit (col 9-12).
Independent Variable: Emigration flows computed using AIRE registry enrollment as share of population in 2000 instrumented by our instrumental variable of the pull emigration factor based on the pre-period LLM emigration network: $\sum_{t=2008}^{2015} \widehat{\frac{m_{l,t}}{pop_{l,2000}}} = \sum_c sh_{i(l),c,2000} * (GDP_c^{2015}/GDP_c^{2008})$
Control variables: Per capita added value in 100,000 euros and unemployment rate in 2004 at local labor market level, share of firms in tradeable sectors, change in local labor market export per capita

Table A7: Robustness check - Effect of emigration rates on change in stock and flows of firms (2008-15) controlling for lagged outcome difference (2005-08)

	(1) All Firms Δ Stock 2008-15	(2) All Firms Δ Births 2008-15	(3) All Firms Δ Deaths 2008-15
Emig Rate	-0.068***	-0.073***	-0.045**
	(0.022)	(0.025)	(0.021)
ΔStock$_{pre}$	0.995***		
	(0.085)		
$\sum$Births$_{pre}$		1.792***	
		(0.085)	
$\sum$Deaths$_{pre}$			1.219***
			(0.195)
Observations	686	686	686
R-squared	0.851	0.856	0.740
F-excl. instr.	19.476	22.055	19.657
Avg. Outcome	0.005	0.344	0.339
Avg. Treatment	0.825	0.825	0.825
Back of Env Avg. Emig.	25707	25707	25707
Back of Env Firms	-475	-514	-315
Province FE	X	X	X

Sample: Italian Local Labor Markets.
Specifications: 2SLS
Dependent variable: Change in firms between 2008-2015 as a share of initial number of firms in each LLM in 2005.
Independent Variable: Emigration flows computed using AIRE registry enrollment as share of population in 2000 instrumented by our instrumental variable of the pull emigration factor based on the pre-period LLM emigration network: $\sum_{t=2008}^{2015} \widehat{\frac{m_{l,t}}{pop_{l,2000}}} = \sum_c sh_{i(l),c,2000} * (GDP_c^{2015}/GDP_c^{2008})$
Control variables: Per capita added value in 100,000 euros and unemployment rate in 2004 at local labor market level, Change in Firms between 2005-2008 as a share of initial number of firms in each commuting zone in 2005.

Table A8: Robustness check - Effect of emigration rates on change in stock and flows of firms owned by under 45 (2008-15) controlling for lagged outcome difference (2005-08)

	(1) Young Firms Δ Stock 2008-15	(2) Young Firms $\sum$ Births 2008-15	(3) Young Firms $\sum$ Deaths 2008-15
Emig Rate	-0.019**	-0.029**	0.000
	(0.009)	(0.013)	(0.013)
Δ Stock$_{pre}$	0.139		
	(0.088)		
$\sum$ Births$_{pre}$		1.415***	
		(0.053)	
$\sum$ Deaths$_{pre}$			1.588***
			(0.074)
Observations	686	686	686
R-squared	0.418	0.856	0.799
F-excl. instr.	13.302	15.622	15.183
Avg. Outcome	-0.089	0.275	0.364
Avg. Treatment	0.825	0.825	0.825
Back of Env Avg. Emig.	1724	1724	1724
Back of Env Firms	-55	-83	1
Province FE	X	X	X

Sample: Italian Local Labor Markets.
Specifications: 2SLS
Dependent variable: Change in firms between 2008-2015 as a share of initial number of firms in each LLM in 2005.
Independent Variable: Emigration flows computed using AIRE registry enrollment as share of population in 2000 instrumented by our instrumental variable of the pull emigration factor based on the pre-period LLM emigration network: $\sum_{t=2008}^{2015} \widehat{\frac{m_{l,t}}{pop_{l,2000}}} = \sum_c sh_{i(l),c,2000} * (GDP_c^{2015}/GDP_c^{2008})$
Control variables: Per capita added value in 100,000 euros and unemployment rate in 2004 at local labor market level, Change in Firms between 2005-2008 as a share of initial number of firms in each commuting zone in 2005.

Table A9: Robustness check - Effect of emigration rates on LLM employment (2008-15) controlling for lagged outcome difference (2005-08)

	(1) Δ Employees% 2008-15	(2) Δ Avg. Size 2008-15	(3) Δ Qualified Empl. 2008-15
Emig Rate	-0.087**	-0.182	-0.041
	(0.041)	(0.256)	(0.101)
Employees$_{pre}$	0.060		
	(0.087)		
Avg. Size$_{pre}$		-0.154	
		(0.195)	
Qualified Empl.$_{pre}$			-0.210
			(0.176)
Observations	686	686	686
R-squared	0.201	0.283	0.143
F-excl. instr.	12.630	12.440	12.197
Avg. Outcome	-0.110	-0.042	0.027
Avg. Outcome 2005	1.671	5.466	0.367
Avg. Treatment	0.825	0.825	0.825
Province FE	X	X	X

Sample: Italian Local Labor Markets.
Specifications: 2SLS
Dependent variable: Change in LLMs employment between 2008-2015 as a share of initial employment in each LLM in 2005.
Independent Variable: Emigration flows computed using AIRE registry enrollment as share of population in 2000 instrumented by our instrumental variable of the pull emigration factor based on the pre-period LLM emigration network: $\widehat{\sum_{t=2008}^{2015} \frac{m_{l,t}}{pop_{l,2000}}} = \sum_c sh_{i(l),c,2000} * (GDP_c^{2015}/GDP_c^{2008})$
Control variables: Per capita added value in 100,000 euros and unemployment rate in 2004 at local labor market level, Change in Firms between 2005-2008 as a share of initial number of firms in each commuting zone in 2005.

Table A10: Robustness check - Effect of emigration rates on LLM employment composition (2008-15) controlling for lagged outcome difference (2005-08)

	(1) Δ Blue Coll 2008-15	(2) Δ White Coll 2008-15	(3) Δ Managers 2008-15
Emig Rate	-0.032	-0.102*	-3.366
	(0.054)	(0.057)	(3.288)
Blue Coll_{pre}	-0.006		
	(0.079)		
White Coll_{pre}		0.162	
		(0.103)	
Managers_{pre}			0.052
			(0.267)
Observations	686	686	584
R-squared	0.197	0.242	0.153
F-excl. instr.	12.752	12.449	4.418
Avg. Outcome	-0.117	-0.013	0.248
Avg. Outcome 2005	8950.138	6737.377	163.226
Avg. Treatment	0.825	0.825	0.799
Province FE	X	X	X

Sample: Italian Local Labor Markets.
Specifications: 2SLS
Dependent variable: Change in LLMs employment composition 2008-2015.
Independent Variable: Emigration flows computed using AIRE registry enrollment as share of population in 2000 instrumented by our instrumental variable of the pull emigration factor based on the pre-period LLM emigration network: $\widehat{\sum_{t=2008}^{2015} \frac{m_{l,t}}{pop_{l,2000}}} = \sum_c sh_{i(l),c,2000} * (GDP_c^{2015}/GDP_c^{2008})$
Control variables: Per capita added value in 100,000 euros and unemployment rate in 2004 at local labor market level, Change in Firms between 2005-2008 as a share of initial number of firms in each commuting zone in 2005.

Table A11: Robustness check - Effect of emigration rates on LLM wages (2008-15) controlling for lagged outcome difference (2005-08)

	(1) Δ Wagebill 2008-15	(2) Δ CZ Wage 2008-15
Emig Rate	-0.033	-0.017***
	(0.045)	(0.006)
Wagebill_{pre}	0.016	
	(0.065)	
Δ CZ Wage_{pre}		-0.546***
		(0.032)
Observations	686	686
R-squared	0.265	0.512
F-excl. instr.	12.273	12.634
Avg. Outcome	-0.113	-0.007
Avg. Outcome 2005	3486.004	1644.434
Avg. Treatment	0.825	0.825
Province FE	X	X

Sample: Italian Local Labor Markets.
Specifications: 2SLS
Dependent variable: Change in LLMs wages.
Independent Variable: Emigration flows computed using AIRE registry enrollment as share of population in 2000 instrumented by our instrumental variable of the pull emigration factor based on the pre-period LLM emigration network: $\sum_{t=2008}^{2015} \widehat{\frac{m_{l,t}}{pop_{l,2000}}} = \sum_c sh_{i(l),c,2000} * (GDP_c^{2015}/GDP_c^{2008})$
Control variables: Per capita added value in 100,000 euros and unemployment rate in 2004 at local labor market level, Change in Firms between 2005-2008 as a share of initial number of firms in each commuting zone in 2005.