CESifo WORKING PAPERS

8388
2020

June 2020

Does Emigration Drain Entrepreneurs?

Massimo Anelli, Gætano Basso, Giuseppe Ippedico, Giovanni Peri

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8388*

Does Emigration Drain Entrepreneurs?

Abstract

Emigration of young, motivated individuals may deprive countries-of-origin of entrepreneurs. We isolate exogenous variation in a large emigration wave from Italy between 2008 and 2015 by interacting diaspora networks with economic pull factors in destination countries, and find that larger emigration rates reduced firm creation and innovative start-ups. We estimate that for every 100 emigrants, 26 fewer firms were created. An accounting exercise shows that 37 percent of the effect was due to the disproportionate loss of young people. The remaining effect was due to selection into emigration of highly entrepreneurial individuals, as well as negative spillovers on firm creation.

JEL-Codes: J610, H700, O300, M130.

Keywords: emigration, demography, brain drain, entrepreneurship, innovation, EU integration.

Massimo Anelli
Bocconi University / Milan / Italy
massimo.anelli@unibocconi.it

Gætano Basso
Bank of Italy / Rome / Italy
gaetano.basso@bancaditalia.it

Giuseppe Ippedico
UC Davis / CA / USA
gippedico@ucdavis.edu

Giovanni Peri
UC Davis / CA / USA
gperi@ucdavis.edu

June 17, 2020
This paper was previously circulated under the title “Youth Drain, Entrepreneurship and Innovation”. We are grateful to Emanuele Ciani, Federico Cingano, Xavier Jaravel, Andrea Linarello, Salvatore Lo Bello, Francesca Lotti, Sauro Mocetti, Paolo Sestito, Eliana Viviano and Josef Zweimueller for helpful comments and to seminars participants at 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), Francesca Licari (Istat), Desideria Toscano and Annamaria Grispo (Italian Ministry of Interior) and Gina Carbonin (SOGEI) for valuable data provision. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Bank of Italy. Any errors or omissions are the sole responsibility of the authors.

1 Introduction

By taking risks and investing in uncertain and "disruptive" projects, entrepreneurs bring innovation and change in the economy. Entrepreneurship requires a high degree of initiative, risk-taking and adaptability to new situations. Often these abilities are found in young people (Liang *et al.* 2018; Acemoglu *et al.*, 2017), as they are more willing to introduce new practices and new technologies and bear higher risk in order to generate "creative destruction", often leading to growth.

Interestingly, research has shown that those abilities and traits conducive to entrepreneurship also increase the probability to emigrate. Jaeger *et al.* (2010) show that migrants have lower risk aversion than non-migrants, and Bütikofer and Peri (2020) show that individuals with higher level of adaptability and cognitive ability are more likely to emigrate. Hence countries and regions experiencing significant emigration rates may be at risk of losing a substantial amount of entrepreneurial potential, with negative consequences on firm- and job-creation.

This issue has long been a concern in developing countries, and has recently become so in Southern Europe, too, where younger cohorts are shrinking due to demographic transition and because emigration surged after 2010. In fact, since 2009 many individuals—especially young ones (Schivardi and Schmitz, 2018)—have left their countries as a consequence of free mobility in Europe (since 1992) due to the deep recession caused by the sovereign debt crisis that hit Southern Europe much harder than Northern Europe.

In this paper we investigate the causal effect of emigration on firm creation, innovative start-ups and labor demand in the area-of-origin. We also try to understand the role played by emigrants' age composition and their selection. Our empirical analysis focuses on Italy, a country that experienced a dramatic surge in the number of

individuals who emigrated since the onset of the Great Recession. Panel (a) of Figure 1 shows the sharp increase in the emigration rate starting in 2010, almost tripling by 2015. Over the period 2008-2015 the cumulative flows of emigrants recorded by administrative data amount to a loss of almost 1% of the Italian population.[1] While emigration was common across all age groups, its rate was especially high among young individuals (aged 25-44) as revealed by Panel (b).[2]

Estimating the causal effect of emigration on local economic outcomes is challenging. The main threats are reverse causality, as people are more likely to leave poorly performing regions; and omitted variable bias, as several unobserved factors push people to emigrate and may also affect firm creation. Moreover, measurement error in recording emigration flows—resulting from delays and under-reporting in changes of residence—could potentially attenuate the relationship between emigration and firm creation, especially in the short-run. To overcome these issues, we adopt an instrumental variable strategy in the spirit of Anelli and Peri (2017), and Fouka *et al.* (2018), and construct a proxy for "pull-driven" emigration. To do so, we exploit the intensity of the diaspora from each Italian local labor market to each destination country (or region, in a more detailed version of the instrument), in a pre-emigration wave year—specifically, 2000. We then interact this measure of bilateral network intensity with the economic performance of *destination countries*. In this way, we capture the economic pull exerted by each destination economy during the period 2008-2015 and we assume it to be stronger in local labor markets with stronger pre-existing network ties to that economy. This instrumental variable allows us to leverage cross-sectional variation of emigration rates driven by "pull factors", and independent of location-specific

[1]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 485,000 people, or about 1.2% of the Portuguese population.

[2]Throughout this period emigration rates were also larger among college graduates (Appendix Figure A1); however, the increase was similar for college and non-college graduates.

push factor, such as the economic conditions in the labor market-of-origin. Moreover, following Goldsmith-Pinkham *et al.* (2018), we test that the share of emigrants to different countries in 2000, and the IV itself, are uncorrelated with local economic and demographic trends predating the 2008-2015 period. Our IV passes the validity tests suggested by Goldsmith-Pinkham *et al.* (2018), which implies that it is not correlated with unobservable and persistent local economic trends.

Our results show that emigration—when instrumented with the pull-driven IV described above—produced a decline in the number of existing firms, driven by lower birth rate and unchanged death rate of firms. This is consistent with a loss of the entrepreneurial capital that drives firm creation. To quantify this effect, in an average-sized local labor market 313 fewer firms were created over the period 2008-2015, while 1,187 individuals emigrated (or about 2.65 percent of the working age population, after correcting for the under-reporting by the official statistics, as shown in Section 4). As a point of reference, during that period a cumulative 3,735 firms were created in the average local labor market. The emigration effect is therefore large, reducing the number of new firms by 8 percent. We then present an accounting exercise to understand whether the estimated effect is consistent with a simple subtraction of individuals with average entrepreneurial ability, or if it implies a selection of highly entrepreneurial individuals and/or negative entrepreneurial spillover on the remaining people. Such an accounting exercise reveals that the simple "subtraction" of young individuals, given their average firm-creating ability, explains more than one third of the overall effect; the selection of emigrants among individuals with a high entrepreneurship rate and the presence of negative entrepreneurship spillovers on non-migrants make up the remaining two thirds. Given the substantial role of the emigration of young people in explaining the drain in firm creation, we explore the demographic channel further and find a strong negative effect of emigration on the creation of firms whose owners and executives

were younger than 45 years. As further evidence indicating the potential innovative role played by the people who left, we find a very significant decline in the number of innovative start-ups operating in technology-intensive sectors. Finally, we study the potential effects of this emigration wave on overall employment and its composition. We find that local labor markets with higher emigration rates exhibit a decline in overall employment and a drop in the share of highly-skilled workers (i.e., non-production workers, as defined by the Italian National Statistical Institute, ISTAT). This indicates that, in spite of the decline in labor supply due to the emigration of working-age individuals, the employment-population ratio did not increase: if anything, it declined. Lacking wage adjustment, which we do not detect in our analysis, the departure of several people should create job opportunities for those left behind, and likely increase the employment-population ratio in the labor markets of origin. The fact we do not observe such a tightening of the labor markets is consistent with a simultaneous loss in labor demand, which would be implied by the decline in firm-creation.

Our paper offers three main contributions related to different branches of the literature. First, extending the literature on the effects of migration on countries-of-origin, this paper is the first to analyze closely the effects on entrepreneurship, and to provide a credible identification thanks to higher quality firm and emigration data, and a credible IV. While our shift-share IV is not completely new, we innovate by exploiting a sudden emigration episode and by constructing a pull-driven shift component. Moreover, following the recent contributions of Borusyak et al. (2017) and Goldsmith-Pinkham et al. (2018), we contribute by making explicit the identification assumptions in this context. Related papers analyzing the impact of emigration of the high-skilled population (often referred to it as *brain drain*) on developing country economies are Waldinger (2010), Mayr and Peri (2009), Docquier and Rapoport (2012), Docquier *et al.* (2014), Di Giovanni *et al.* (2015). Less is known about the effects of emigration

on developed economies. A paper related to ours is Giesing and Laurentsyeva (2017): they find that high-skilled emigration from Eastern Europe after the EU enlargement of the 2000s had negative effects on firms' TFP in the countries-of-origin. Anelli and Peri (2017) and Ippedico (2017) looked at the relationship between emigration and political outcomes and local firms at a more aggregate level without deepening either the mechanisms or the identification.[3] Most of the previous brain drain and migration literature considered emigration as a decline in the country-of-origin supply of labor for a constant labor demand schedule. Such an effect could increase the wages and employment opportunities of those who remain, at least in the short-run (Mishra 2007; Elsner 2013a, 2013b; Dustmann *et al.* 2015). By focusing on firm-creation, our paper shows instead that emigration can reduce labor demand, thus identifying for the first time an unexplored economic effect on countries-of-origin.

Second, our paper connects with the research analyzing the role of young people on innovation and starting up new firms (Acemouglu *et al.* 2017, Barker and Mueller 2002 and MacDonald and Weisbach 2004). Our paper is the first to analyze the causal impact of emigration on reducing the number of young people and their innovative entrepreneurial role. Related literature shows a positive relationship between the share of *young* people in a country (or region) and entrepreneurship rates (Liang *et al.* 2018), productivity (Ciccarelli *et al.* 2019), growth (Engbom 2019) and birth rate of start-ups (Karahan *et al.* 2019). If innovative entrepreneurship is higher at a young age (Kopecky 2017), the loss of young people may be associated with a loss of growth and innovative ideas.

Finally, our paper brings new evidence that emigrants are positively selected among those with high entrepreneurial and innovative potential, complementing the findings in the literature on immigration-innovation and on immigration-entrepreneurship. Hunt

[3] Another related paper is Karahan *et al.* (2019) who use variation in immigration across US states and relate it to firm entry.

and Gauthier-Loiselle (2010), Kerr and Lincoln (2010) and Moser *et al.* (2014) show that immigrants to the US are more likely to be active in patenting and innovation than comparable natives. Similarly, as reviewed in Fairlie and Lofstrom (2015), a significant number of studies finds that immigrants in the US have a higher probability of being self-employed and starting firms, relative to natives. This evidence points to a positive selection of immigrants among innovators and entrepreneurs.

The rest of the paper is organized as follows. Section 2 describes the main data and trends for emigration and firm creation in Italian local labor markets. Section 3 introduces the empirical specification and describes the 2SLS identification strategy, and discusses its validity. Section 4 presents the main results, and Section 5 discusses several additional results. Section 6 reports the main robustness checks. Section 7 concludes the paper.

2 Data

2.1 Data on Emigration

We obtained data on emigration flows from each municipality from the Italian National Statistical Institute (Istat). The data, which cover the period 2002-2015, are from administrative sources and are aggregated into year-of-emigration by municipality-of-origin by country-of-destination by age-group cells.[4] We also obtained data on the stock of emigrants directly from the Registry of Italians Residing Abroad (AIRE; Anelli and Peri, 2017), which records all individuals permanently emigrated between 1990 and 2014, who were still abroad as of 2015, and includes precise information about the destination country (and region), the municipality-of-origin, and the year of emigration.

[4]The data also contain incomplete information on educational attainment, which we use in additional analyses not reported. See also the online Appendix.

These features allow us to construct the historic networks of emigrants as of the year 2000.

Table 1 shows the stock of emigrants from Italy by country-of-destination as of 2000 in Panel A, and the cumulative emigration flows between 2008-2015 in Panel B, by age group. The table reveals two important trends. First, the top destination countries have slightly changed over time. While large numbers of Italians have always migrated to Germany, Switzerland, and France, in recent years Italians have often moved to high-performing countries (e.g., the United Kingdom and the United States) rather than to countries with strong historic and cultural ties to Italy (e.g., Argentina and Belgium). Given their significant role, the distribution of historical emigrants to Germany and Switzerland will turn out to be crucial to identify the pull-driven migration in 2008-2015. Second, as we already saw in aggregate, the table shows that young people are substantially over-represented in the recent migration flows.

Despite the fact that Italian emigrants are required by law to register abroad within six months from the date of emigration (and have significant financial incentives to do so), there is anecdotal evidence of under-registration, at least in the early years after emigration, as not all changes of residence may be timely recorded by the Italian authorities. Figure 2 compares the outflows of Italians to the UK registered by the AIRE-Istat data and the registration of Italian immigrants recorded in the UK social security registry. The UK data indicate that outflows from Italy to the UK are underestimated by about two thirds (panel (a)), while the year-to-year changes follow closely those of the UK social security registrations with one year of lag (panel (b)). This lag is consistent with the six-month window to communicate the new residence abroad and with 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).[5] For our analysis, this delay is less of an issue since we rely on variation over multiple years by constructing long-differences, still measurement error can be non negligible. In Appendix A.II, using these destination-country sources we estimate that actual emigration flows of Italian are plusibly about 2.6 times larger than those registered in the AIRE-Istat records. It is important to account for such under-counts when interpreting the magnitude of the effects relative to the size of the emigration rate. Moreover measurement error due mainly to delays and imperfect registration of temporary migrants is a further reason to use IV estimation. Measurement error is likely to be much smaller on the measure of long-term existing networks of Italians abroad (those emigrated before year 2000) as those numbers are not affected by delays nor by the presence of temporary migrants. Hence the cross sectional distribution of long-term (pre-2000) Italian emigrants across municipalities, which we use to construct the instrument, is likely a very precise measure of the Italian diaspora.

2.2 Data on Firms, Employment and Local Labor Markets

We combine data on the emigration flows with firm-level data on the universe of all Italian firms, obtained from the Chambers of Commerce, and with data from the social security administration (INPS) on employment and wages. Data from the Chambers of Commerce include information on birth and death of firms and demographic characteristics of owners, shareholders and executives of each firm over the period 2005-2015.[6] We use this latter piece of information to classify firms with a majority of owners and executives under age 45, which we refer to as "young-owned firms". Our data include

[5] We performed a similar analysis for the US using data from the American Community Survey, which we show in Figure A2 in the Appendix. Despite the fact that the survey nature of the data does not allow to precisely estimate the immigration of Italians, the analysis based on the US confirms, qualitatively, the evidence based on the UK and Swiss administrative data.

[6] We consider as a birth the appearance of a newly constituted firm in any given year, provided it survive at least through the end of the year.

all firms, independently of their legal status, some of which may be multi-plant (but the vast majority has only one establishment). The INPS data cover the period 1990-2015, and include information on the yearly number of employees (divided by broad occupation category, i.e., *apprentices*, production workers, often referred to as "blue collar" workers, *non-production workers*, often referred to as "white collar" workers and *managers*), their average monthly wage, industry and geographic location of the employer.[7]

Our unit of analysis is the local labor market (LLM), defined using the Istat 2001 definition. According to Istat, LLMs are geographic clusters of municipalities with commuting patterns mainly internal to the cluster, an analogue definition to that of Commuting Zones (CZ) for the United States. They are used as a proxy for local labor markets and they are a partition of Italian provinces.[8] There are 686 LLMs in Italy covering the whole national territory. We focus our analysis on the period 2005-2015, considering the period 2008-2015 as the "treatment" period, as emigration increased suddenly and substantially in those years.

3 Empirical Specification and Identification

In our empirical specification, the main outcome is the change in the stock of firms (equal to the difference between entries and exits in the same period) in local labor markets, which are indexed by l. This variable is indicated as Δy_l in equation (1). The main explanatory variable is the cumulative outflow of Italians aged 25-64 between 2008

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

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

and 2015, indicated as $\sum_{t=2008}^{2015} m_{l,t}$.[9] Both variables are normalized by the average 25-64 years old resident population in years 2005-2008, right before the beginning of the emigration surge. We call this variable $pop_{l,pre}$. This normalization produces the emigration rate in the area-of-origin, l, in terms of initial population. In the baseline specification, we control for a set of observable LLM characteristics pre-dating 2008 accounting for the economic performance of the area before the emigration surge ($X_{l,2005}$; these are the 2005 GDP per capita and unemployment rate). We also include either twenty regions or 110 provincial fixed effects (ϕ_p) that capture time-invariant unobserved economic, demographic and institutional factors common to a region or a province.[10] We thus estimate the following equation:

$$\frac{\Delta y_l}{pop_{l,pre}} = \alpha + \beta \frac{\sum_{t=2008}^{2015} m_{l,t}}{pop_{l,pre}} \cdot 100 + \phi_p + \gamma X_{l,2005} + \varepsilon_l \tag{1}$$

If the size of migration outflows were distributed randomly across local economies, the OLS estimate of equation (1) would deliver the causal effect of emigration on the number of firms. This is, however, unlikely. Such outflows are correlated with local economic and social conditions, which in turn might affect our outcomes of interest. On the one hand, if localities with more intense entrepreneurial and economic activity tend to have a stronger connection with foreign economies and possibly more migrants as a consequence of this (notice in Figure 4 many LLMs in Northern Italy, the more economically entrepreneurial part of the country, have large emigration rates), the OLS estimates would be biased upward, towards finding a positive correlation between emigration and entrepreneurial intensity. On the other hand, if individuals are more likely to leave labor markets when labor demand declines and economic activity slows,

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

[10] As noted above, in case a local labor market crosses province boundaries the fixed effect is assigned to the province or region where most of the population resides.

then there would be a negative correlation between emigration and entrepreneurship and thus a downward bias towards a negative effect. Moreover, because of delays and missing reports for short-term migration, the measures of emigration rates over the years 2008-2015 could have large measurement error, biasing the estimated coefficient towards zero. All these reasons suggest the existence of potential bias in the OLS coefficient, although its direction is unclear. Hence we should not focus on the OLS estimates of the β coefficients in Table 2, which indicate no significant correlation between the LLM emigration rate and changes in firm stock, entry, or exit. To correct the omitted variable and measurement error biases of OLS estimates, we exploit variation in migration flows driven by historical networks (which are measured more precisely) and due to recent economic pull factors, which are less correlated with local economic conditions in the place-of-origin.

3.1 Identification: The IV Approach

The key intuition for the instrumental variable is that LLMs have connections with specific foreign countries through their networks of past-residents who emigrated to those countries. These networks are likely to generate flows of information, job offers and job referrals through personal and family connections between the diaspora in the foreign country and individuals living in the LLM-of-origin. Such networks exert a strong attraction effect on potential migrants from countries exhibiting stronger economic performance. Building on this intuition, we interact the intensity of pre-existing networks with the economic success of destination countries in 2008-2015. In practice, we count the number of people who emigrated from each LLM l to each foreign country c before year 2000, as a percentage of the LLM population in 2000. Then, we interact these percentages with the growth of real GDP per capita in foreign countries during

the period 2008-2015.[11] Summing across destination countries results in an economic pull factor, exerted by foreign countries during the 2008-15 period, on each specific LLM. The variable is defined as follows:

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

In expression (2), the first term, $NTWK_{l,c}$, is the number of Italians from LLM l living in country c since before year 2000, as a share of the LLM population in year 2000. It captures the size of the historic diaspora from LLM l in country c, which affects the potential for subsequent emigration outflows from l to c. The second term, $G_c = GDP_c^{2015}/GDP_c^{2008}$, is the growth of real GDP per capita in country c during the period 2008-2015 (which includes the Great Recession and the sovereign debt crisis disproportionately hitting the Southern European countries). This term captures the destination-specific "pull factor", which proxies the economic 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 countries-of-destination. The variable defined in equation (2) is used as instrument for the actual emigration rate, $\frac{\sum_{t=2008}^{2015} m_{l,t}}{pop_{l,pre}}$, which is the main explanatory variable in the estimating equation (1).

3.2 Instrument Validity: Pre-trends

The key identifying assumption is that the interaction between the strength of diaspora network in year 2000 and the economic pull of destination countries in 2008-2015 is uncorrelated with unobserved factors specific to the LLMs that may affect firm creation in the same period. Threats to identification remain, however—for instance, if past economic shocks persist over time in a local labor market and they affected

[11] GDP data are obtained from the IMF International Financial Statistics database. We are able to match more than 100 destination countries that comprise almost all emigration outflows.

emigration before 2000, as well as firm creation after 2008. To increase confidence in the assumption underlying our IV, we perform several checks.

Let us first remind the reader that we include province fixed effects in our preferred specification to control for relevant economic, institutional and policy variables, as they vary substantially across locations in Italy. These capture the potential impact of policies common to areas of about 500,000 people on average (provinces tend to be homogeneous units economically and politically). Most importantly, the fixed effects force the identifying variation of the IV to be across labor markets close to each other and with very similar economic conditions. Our identification is helped by the fact that diaspora networks are LLM-specific, often driven by historical contingencies—and their geographical localization is very fine—while the economic and policy conditions within a province are rather homogeneous.

We formally check the within-provinces correlation (as we control for province effects) of our instrument with the 2005-2008 trends of the key outcomes as well as other economic and demographic variables. In Table 3 we regress the 2005-2008 change in the stock, cumulative births, and cumulative deaths of firms on the IV-predicted emigration, post-2008. The estimated coefficients are small and not statistically significant. This is consistent with our identifying assumption, that the IV does not correlate with pre-2008 firm creation and destruction rates. We also estimate similar regressions on the other outcomes that we analyze in the paper, namely the stock of firms owned by young entrepreneurs, total employment, employment-population ratio, total wage bill, and the number of blue and white collar workers and of managers. When we consider changes in those outcomes between 2005 and 2008 we never find a significant correlation with the IV (results are reported in Appendix Tables A2, A3 and A4).

An additional concern is that the constructed instrument is correlated with other dimensions of local mobility. If the IV predicts internal migration or inflows of foreign-

born into the local labor markets, then the causal interpretation of IV estimates would be problematic. In Appendix Table A7 we show that pull-driven emigration is not systematically associated with internal migration flows or immigration from abroad. There is no significant correlation of the IV with 2008-2015 internal migration flows, namely mobility of individuals to and from other local labor markets in the country, and with the immigration rate of foreign-born individuals. This is not surprising, as the countries-of-origin of immigrants to Italy (mainly from Eastern Europe and North Africa) are different from those where Italian emigrants reside.

Figure 5 offers a visual representation of the similarity of pre-2008 trends in the main outcome variable (the number of firms per-capita) and of the significance of the post-2008 "event". We first separate LLMs into two groups, based on the value of the "predicted emigration" IV, and we average the outcome for those with above-median (high) values and those with below-median (low) values. We then represent those averages between 2005 and 2015 as a solid and a dashed line, respectively, standardizing both values to one in 2005. Figure 5 shows two clear patterns. First, the two groups have similar trends up to 2009-2010, which marks the onset of the Great Recession and of the emigration episode we analyze: the average number of firms per capita were moving together for these two groups of LLMs in the pre-event period. Second, after 2009-2010 the lines start progressively diverging and they show a significant difference by the end of the considered period, 2015. Partly recalling the standard pre-trend checks of diff-in-diff estimation, this chart conveys the idea that it is reasonable to consider high and low predicted emigration areas as similar before the Great Recession and diverging post 2009-2010 when the emigration flows became substantial.

3.3 Shift-share diagnostics

The IV we construct has the structure of a traditional Bartik/shift-share. Specifically, it combines the variation in the cross-sectional distribution of emigrants' population shares by destination country (the share part) with the destination countries' aggregate economic growth (the shift part). Goldsmith-Pinkham *et al.* (2018) show that a sufficient condition for identification in this setting is that the the population shares of emigrants across LLMs are uncorrelated with the error term and hence are exogenous.[12] To test whether this is the case in our setting, we scrutinize the cross-sectional components of the IV. We first calculate the weights that the instrument attributes to each share (the so-called Rotemberg weights). Higher weights correspond to greater relevance in the identifying variation. We then test whether the population share of emigrants for each of the main destination countries (those receiving higher weights) correlates with pre-2008 observable characteristics of the LLM-of-origin.

Tables 4 and 5 report the main results of diagnostic tests as suggested in Goldsmith-Pinkham *et al.* (2018). Table 4 is organized in three panels and show three sets of tests. First, in Panel A we show the share of Rotemberg weights ($\hat{\alpha_c}$) that are positive and negative. Almost all of them are positive indicating that the individual shares are positively correlated with the IV, suggesting our instrument is a convex combination of the country-specific estimated β coefficients and does not show signs of mis-specification. Panel B reports correlations among the components of the IV (G_c and $NTWK_c$), the Rotemberg weights ($\hat{\alpha_c}$), the power of the IV ($\hat{F_c}$) and the estimated coefficients of equation (1) with per-capita stock of firms as the dependent variable ($\hat{\beta_c}$).

[12] Recent work by Borusyak et al. (2018) suggests that a necessary and sufficient condition for identication is that the interaction between the shares and the shift components is asymptotically uncorrelated with the error term. This can be satisfied by uncorrelated "shift" terms as long as they are numerous and idiosyncratic. In our setting this is unlikely, as there is only a dozen of important destination countries and their growth rates are likely correlated. However, Borusyak et al. (2018) point out that the condition they propose is also satisfied by the exogeneity of shares as proposed by Goldsmith-Pinkham *et al.* (2018).

An informative statistic is the correlation between each component of the IV (G_c and $NTWK_c$) and the Rotemberg weights. A larger correlation implies higher relevance of that component of the IV in generating the identifying variation. We see that while the share component $NTWK_c$ has a correlation of 0.84 with the weights, the "shift" component G_c has very low and even negative correlation (-0.05). This confirms that it is mostly the share variation generating identification in our setting, and therefore it is important to check that those emigration shares receiving the highest weights are associated with estimates of β similar to our main estimate and are not correlated with pre-2008 local characteristics. This is what we do in Panel C of Table 4 and in Table 5. Panel C reports the five "country-of-emigration shares" receiving the highest weight and hence driving most of the identifying variation. "Share of emigrants to Germany" generates about 45 percent of the total instrument variation, and "share of emigrants to Switzerland" generates an additional 28 percent. This would be concerning if those shares are correlated with other variables, which we test below. The table shows that emigrants shares to France, Australia and Belgium also have a non-negligible weight. A reassuring feature of our IV is that the estimates of the main coefficient of interest (β in Equation (1)) obtained using any of the top five just-identified instrument are all negative, are close to each other and are not far from the main estimate (-0.007). Estimates obtained using the German or Swiss share only, which exhibit a reasonably high F-statistic above 10, are both significantly negative. Specifically the first equals -0.006 and the second equals -0.003.

Table 5 shows the correlation of the population share of emigrants to the five most important destination countries (according to their Rotemberg weights) with observable characteristics of the origin LLMs measured in the period, 2005-2008. Germany and Switzerland are particularly important and a strong correlation of those shares with pre-existing economic trends would cast doubts on the validity of identification.

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

As an additional exercise, which can potentially increase the power of the instrument, we split the emigrants' destinations into smaller geographical units corresponding to European regions (Eurostat NUTS-2 classification) rather than countries, whenever this information is available in our data (i.e., for Germany, Switzerland, Belgium and the UK). The instrument constructed with this richer set of destinations, but otherwise identical to the one used so far, does not show significantly higher power, and has similar properties when subject to the Goldsmith-Pinkham *et al.*'s (2018) tests (reported in Appendix A.IV). A large share of the variation is driven by three German regions (Stuttgart/Friburg, Dortmund/Koln, and Frankfurt) and two Swiss regions (Zurich and Lugano). As for Germany and Switzerland, using the most important regions-of-destination to estimate $\hat{\beta}_c$ we obtain values that are extremely close to our main estimate and to each other.

Overall, these diagnostic tests indicate a prominent role of Germany and Switzerland in driving most of the variation in the IV. However, they reassure that there is no systematic reason to believe those shares violate the identifying assumptions. Rather, the sufficient conditions for identification outlined by Goldsmith-Pinkham *et al.* (2018) seem to hold in our setting.

3.4 Instrument Power: First Stage Results

In Table 6, we report the first stage results where we predict the emigration rate with the instrument, $Pull_l$. In the regressions we control for GDP per capita and the unemployment rate in 2005, and we include region fixed effects in column (2) and

province fixed effects in column (3). These controls capture pre-determined economic conditions in the LLMs-of-origin. The estimates in the first row of Table 6 show that the *Pull IV* has a significant predictive power for actual emigration and the size of the coefficient is stable across specifications. The first stage F-statistics lie between 14.9 and 29.7, 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 LLMs 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 plots the variation in emigration rates (panel (a)) from LLMs within each province (in bold) as well as the variation in emigration as predicted by the IV pull factor (panel (b)) and provide a visualization of the first stage. Based on historical emigration patterns, the IV predicts more emigration from the South, while the actual emigration in the period 2008-2015 was prominent from Central and Northern regions which are also richer and more dynamic in terms of business creation. This evidence will help us interpret the main results on firm creation we find below.

4 Main Results

4.1 Effects on Firm Creation

Panel A of Table 7 shows the main results of the paper. The coefficients reported are from 2SLS regressions where the endogenous migration flow is instrumented with the pull factor IV. The dependent variables are the change in the stock of firms in column (1), cumulative firm births in column (2), and cumulative firm deaths in column (3)

over the period 2008-2015. All the outcomes are standardized by the LLM population aged 25-64 before the emigration episode (average 2005-2008). The emigration rate is expressed in percentage points, so that the coefficient can be interpreted as the change in the number of firms per person (25-64 years old) in response to an emigration rate of one percent of the population. Standard errors are clustered at the province level.

The estimates indicate that in areas with larger emigration flows in the period 2008-2015 the number of firms declined. This effect is driven by fewer firm births (that is, lower firm creation) rather than more firm deaths: on average, for a one percentage point increase in the emigration rate there has been a decline of 0.007 firms created per person (0.7 firms created per 100 people). A simple calculation, accounting for the under-reporting of Italian migrants discussed in Appendix A.II, implies that for the average LLM, which experienced an outflow of 1,187 individuals in the 2008-2015 period, firm-creation decreased by 303 units. Considering that 3,735 firms were created in the average local labor market in this period, this estimate is consistent with emigration outflows reducing in a significant way the local entrepreneurial activity in the area-of-origin (a decrease in firm creation of roughly 8%). The small and non-significant coefficient of emigration on the number of firm deaths is also reassuring. First, emigration is more logically associated with a decline in potential firm-creation. Second, a correlation of emigration and firm-failure could suggest a reverse channel of causation, namely people left LLMs where firms were closing.

We check the robustness of these results by controlling for lagged values of the outcomes in panel B of Table 7. Similar to the pre-trends check shown in the previous section, this specification aims at testing whether the main coefficients of interest are wrongly capturing long-term trends that may bias the causal estimates of emigration on business dynamism. The results do not support such a hypothesis: the coefficients in panel B are not statistically different from those in panel A.

4.2 Subtraction, Selection and Spillover Effects

How much of this decline is simply due to a *subtraction effect*—namely the fact that with fewer people it is natural to expect fewer firms created in the location-of-origin? And to what extent is it instead due to emigrants being more likely to start a firm than stayers (*selection effect*) and to the fact that other local people may be less induced, by existing entrepreneurs, to create firms (*spillover effect*)? In order to separate these two sources of decline in entrepreneurial activity, we perform a simple accounting exercise. We estimate the *subtraction effect* multiplying the average number of young (25-44 years old) and old (45-64 years old) emigrants by the age-specific average probability of creating a firm over the pre-period 2005-2008, r_{pre}. This would correspond to the total effect if emigrants had the same entrepreneurship rate as stayers and if their departure did not affect the probability of a stayer of becoming an entrepreneur. The remaining part is due partly to *selection* of emigrants among those individuals with higher than average entrepreneurship rates, and to *spillover* effects—namely the potential externality that emigrants might exert on stayers, affecting their probability of starting a firm. For instance, stayers might be less likely to become entrepreneurs if their potential business partners left or perhaps because the lower firm creation has drained potential supply-chain relationships that were needed to start a new firm. Moreover, since emigration reduces the resident population in the sending LLM, the local demand of goods and service might decline and in turn decrease business opportunities for local entrepreneurs. Hence, in equation notation:

$$\Delta \widehat{FirmsBirth} = \overbrace{\underbrace{Emig^{25-44}}_{-907} \times \underbrace{r_{pre}^{25-44}}_{0.018*7}}^{\text{average effect young}} + \overbrace{\underbrace{Emig^{45-64}}_{-282} \times \underbrace{r_{pre}^{45-64}}_{0.008*7}}^{\text{average effect old}} + \underbrace{Residual}_{\text{Selection and Spillover}}$$

$$-313 = -116 - 15 - 182$$

$$100\% = 37\% + 5\% + 58\%$$

The left hand side, $\Delta \widehat{FirmsBirth}$, is the estimated effect on firm creation of the average outflow from the LLM, obtained from Column 2 of Table 7, namely 313 fewer firms. The first two terms on the right hand side are the *subtraction* effects due to emigration of young and old individuals respectively. Emigration drained 907 young individuals 25-44 years old from the average LLM between 2008 and 2015. As their average yearly entrepreneurship rate[13] before the shock (2005-08) was 1.8%, these younger emigrants would have created 116 firms during a seven years period, assuming they had the average entrepreneurship rate. Similarly, the 282 individuals aged 45-64 who left the average LLM between 2008 and 2015 would have created 15 additional firms if they had the average entrepreneurship rate of that group (0.8%).

Such an accounting exercise, albeit simple, reveals at least two important patterns. First, the large share of young people in the emigration wave, coupled with the higher propensity of younger individuals to start a business (in Italy as in many other countries), suggests a substantial role of the pure subtraction channel in reducing firm creation. For this reason, in Section 4.3 we focus on the loss of younger individuals as an important channel for the drain in entrepreneurship potential. Second, and even more relevant, the remaining 58% of the total effect, which is unexplained by these mechanical subtraction effects, can be attributed to the *selection* of emigrants in terms of entrepreneurial potential combined with the plausibly negative *spillovers* that they

[13] More precisely, these are yearly entry rates of firms whose owners are under-45 (over-45 respectively) years old between 2005-2008 as percentage of the average 25-44 (45-64 resp.) years old resident population.

exerted on the population of stayers.

Selection of international migrants has been widely documented in the literature on the educational dimension (Grogger and Hanson 2011), but also on pre-migration earnings (Parey *et al.* 2017), occupational skills (Patt *et al.* 2020) non-cognitive characteristics (Jaeger *et al.* 2010, Bütikofer and Peri 2020) and unobservable characteristics (Borjas *et al.* 2018), which likely correlate with higher entrepreneurship potential. While it is hard to quantify the degree of selection of emigrants, in terms of education they are usually two to three times more likely to have a college degree relative to stayers (Grogger and Hanson 2011). Indeed, Italians emigrating during the Great Recession were about 1.8 times more likely to hold a college degree than those who remained in the country according to our data. If the whole residual effect in our decomposition exercise was due to selection, this would imply a share of entrepreneurs among emigrants 2.5 times larger than among stayers. If instead the degree of selection based on entrepreneurship skills was comparable to the selection based on tertiary education observed for Italian emigrants (1.8), then roughly 34% (105 firms) of the whole effect would be attributable to selection and the remaining 24% (77 firms) to spillover. Furthermore, migrants may help create agglomeration of innovation in destination areas (e.g. Kerr *et al.* 2017), thus it is plausible that their departure exerts negative spillover effects on the local economies-of-origin. While in this paper we cannot really identify the channels for the residual effect, we show that the loss of young individuals is an important component of the story, that emigration drained innovative start-ups with higher potential spillover effects, and that our findings are consistent a substantial selection of emigrants among those with higher propensity to be entrepreneurs.

4.3 The Loss of Young People as Potential Entrepreneurs

As shown in the decomposition above, the facts that young people have a relatively high entrepreneurship rate in Italy and that large numbers of this group have emigrated are together responsible for a significant part of the effect. In this section we provide additional evidence that emigration of young people was a significant channel of entrepreneurship loss.

In Table 8 we look at the creation and destruction of firms whose owners and executives are younger than 45. The age of owners and executives is reported in the data from the Chambers of Commerce and 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-executives are under 45. We then look at the effects of pull-driven emigration on the number, creation of, and destruction of this subset of firms. The results in Table 8, which mirror those of Table 7, indicate that, absent emigration (as induced by our pull instrument), there would have been 172 more firms created by young individuals, compared to an average of 2,470 firms created in the average LLM: a 7% reduction in the number of firms created.

Related to this, we see that the demographic characteristics of the LLMs are important determinants of the strength of the instrument: LLMs with lower average age of the population had large migration flows due to strong external pull factors, and would have had low emigration absent such strong pull factors (i.e. younger LLMs are stronger "compliers"). Column (1) of Table 9 reports the main first stage result (from column 3, Table 6), while columns (2) and (3) limit the sample to those LLMs with a low median age and to those with a high median age (measured in the period prior to the emigration wave) respectively.[14] Confirming the descriptive facts about emigration

[14] 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).

presented in Table 1 and Figure 1, the LLMs most affected by the pull factor are those with a relatively low median age. The F-statistic is much larger in relatively younger areas. The estimated first stage coefficient in relatively younger LLMs is also slightly larger than in relatively older ones. Thus, following a LATE interpretation of the instrumental variable identification, the estimated effects are identified more strongly for LLMs with higher shares of young workers—who are potentially also those individuals with higher entrepreneurial skills who are more likely to start a new firm.

4.4 Effects on Innovative Start-ups

As entrepreneurship and firm creation are engines to introduce new technologies and to create new jobs, the loss of entrepreneurial capital due to emigration may be particularly damaging for economic growth if it is also associated with less innovation and slower technological and productivity growth. We analyze the potential impact on economic activity in innovative sectors by focusing on newly created firms that operate in technology-intensive sectors and are not spin-offs of larger established firms. We call this group of firms "innovative start-ups" as they are those more likely to embody genuinely new technologies and ideas.[15] Table 10 shows the results of our baseline model estimated using the net cumulative entry of innovative start-ups in each LLM in the post period as dependent variable.[16]

The estimated coefficient is statistically significant and indicates that, the larger

[15] 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 to at least 15% of revenues (or costs, if higher); or iii) they hold at least one patent of innovative nature. These firms benefit 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.

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

are migrant outflows from Italian LLMs, the less likely those LLMs are to birth innovative start-ups. While on average there were 0.01 additional innovative start-ups per 100 people in a LLM (or 1 per 10,000), a one percentage point higher emigration rate induced a lower creation of about 0.006 start-ups per 100 people (or 0.6 per 10,000). Emigration seems associated with a worrying decline (60%) in the creation of innovative firms, which are responsible for job creation and growth. Such a large effect can be explained by the fact that young start-up entrepreneurs are a rather small group in the population and it is reasonable to expect that they are also the most attracted by pull factors. Considering the well known tendency of STEM (Science, Technology, Engineering and Math) professionals to dominate the group of highly educated migrants to countries such as the US (see Peri *et al.* 2015) or the UK, and considering their significant contribution to innovation in their destination countries (see Kerr and Lincoln 2010), there could be a symmetric slowdown of innovation in their countries-of-origin.

5 Labor Demand Effects, Skill Composition and Wages

The evidence presented so far highlights two important facts. First, emigration produced a loss of entrepreneurship, reducing firm creation by a significant amount. Second, this loss was larger than what the simple "subtraction" of average individuals would imply, suggesting that emigrants were more likely to be entrepreneurs than the average individual. A mechanical consequence of this higher propensity to be entrepreneurs is a lower propensity to be employees. Emigration is traditionally exemplified as a loss of labor supply, and symmetrically immigration is modeled as an increase in the labor supply. However, if emigrants are significantly more likely to be entrepreneurs (relative to non-migrants) and if the firms they start create additional

jobs, then emigration may actually reduce local labor demand together with labor supply. That immigrants are more likely to be entrepreneurs relative to natives is a well-established fact, especially in the US (as shown by extensive evidence summarized, for instance, in Fairlie and Lofstrom 2015). Our paper is the first, to our knowledge, to suggest that emigrants are selected among highly entrepreneurial individuals relative to non-migrants in the country-of-origin. If entrepreneurship (including human capital and know-how to start a firm) is a factor complementary to labor and it is needed in production, then the loss of one person can be thought of as a loss of a fraction of one worker and a fraction of one entrepreneur. As emigrants seem more likely to be entrepreneurs, then they represent a loss of a larger proportion of entrepreneurs than of workers, with a potential effect of reducing the demand for local workers more than they reduce its labor supply, ultimately depressing employment rates and possibly wages. If, instead, emigrants were mostly selected among workers, and thus their departure did not affect job creation but only labor supply, then larger emigration would be associated with larger employment rates, tighter labor markets, and higher wages, as a drop in labor supply and constant labor demand would predict in the canonical model.

Table 11 tests some of these implications by regressing employment outcomes on emigration, instrumented with the Pull IV. First, we test the impact on employment in Column (1). The estimate shows a negative and significant effect of emigration on employment. The magnitude of the coefficient is 6.8% fewer employees per one percentage point of emigration. A back-of-the-envelope calculation shows that for the average LLM, with 1,187 emigrants, this would imply about 1,160 fewer employees. Such an impact is much larger than subtracting the average number of employed people among those who left. Based on the average employment to population ratio in 2005 (equal to 0.57), the number of workers for 1,187 people (lost) would have been only

676, rather than 1,160. Therefore this implies the loss of additional jobs on top of those subtracted by a simple loss in labor supply. Column (2) shows, consistently, that the employment-population ratio—a measure capturing the number of jobs per capita in a local economy—declines in response to emigration, albeit not significantly. Column (3) shows that the average firm size does not change by a significant amount in response to emigration, again suggesting that this phenomenon was not simply a subtraction of workers to a fixed number of existing firms, which would have implied a significant decline in that size. Finally, Column (4) shows that the overall wage bill in the LLM experienced a non-significant negative change in response to emigration, signaling a decline in labor income in the local economy. Taken together, these results do not suggest that the departure of emigrants was associated with a tightening of the labor market, which would have implied an increase in jobs per capita (the employment-population ratio) and possibly an increase in wages.[17] Therefore, the overall picture is more consistent with the idea that emigration leaves the labor demand depressed relative to the impact on labor supply, or at least does not generate a tighter labor market.

Furthermore, in Table 12 we explore whether emigration has altered the relative skill composition of employment in the economy. In particular, we analyze whether emigration rates affected employment of specific skill groups more than others. The effect on each group's employment is a combination of the loss of potential workers, with different skills, and the loss of potential firms creating jobs for those types of workers. We distinguish between blue collars, white collars, and managerial jobs (the only breakdown in the Inps data available to us). We find that, while there is a small non-significant negative effect on the number of blue collar workers in the labor market,

[17] We do not show the effect on average wages, which is hardly significant in most specifications, as its interpretation is less clear. Indeed, its effect combines a change in employment composition (as shown below) together with the relative demand and supply effects.

there is a larger, negative and significant effect on white collar workers. Emigration is also associated with a negative and large change in managers, but is imprecisely estimated and, because of the large size of the standard error, not statistically significant.

This finding is consistent both with the selection of emigrants among the high-skilled and with the notion that the loss of new firms depressed demand for skilled labor more than that for unskilled labor. Overall, a local economy that lost emigrants experienced lower firm creation, fewer innovative start-ups, a (non-significant) decline in employment-population ratio, and a decline in skilled employment. Taken together, these effects appear consistent with a loss in local entrepreneurship generating a drop in labor demand together with a decline in labor supply—rather than just a drop in labor supply that would have increased labor market tightness and the employment-population ratio. These results point out that a simple representation of migration as a change in the supply of workers, with all else being equal, is likely to miss important features and major effect of this phenomenon.

6 Robustness Checks: Other Forms of Mobility and Trade

Emigration abroad is only one of the potential flows of individuals from and to a local area. Local economies also experienced internal flows of Italian citizens (who moved within the country) and inflow of foreign immigrants. Those flows may be correlated with local economic conditions and hence with the flows of Italians moving abroad. Moreover, they can partially compensate for the impact of emigration on firm creation. If the IV is not totally uncorrelated with other migration flows into or out of the local area, their presence may generate spurious results. To address the potential confounding effect of other migration flows, we perform several robustness

checks. First, in column (1) of Table 13, we augment the main analysis by adding, as a control, the immigration rate to each LLM. The estimated effect of the emigration rate is quantitatively unchanged and still significant, implying a small correlation between the emigration rate, as triggered by our pull factor, and immigration flows.[18]

As a second check of our results, we exclude those areas which are more likely to be strongly affected by cross-country commuting and trade, which are also potentially correlated with emigration. The map in Figure 4(a) shows that migration outflows are more intense in border regions, which are also strongly connected with foreign countries in terms of commuting patterns and local trade. As trade relations and migration flows may be correlated (Rauch 1999, 2001) and both are correlated with past economic conditions, in a robustness check we exclude those Italian LLMs at the border with other countries, for which this correlation may be stronger. The results omitting those border areas are presented in Table 13 Column 2. The point estimates of the coefficient of interest barely change, offering reassurance that our main conclusions are not biased by the presence of specific channels in border regions. A more direct way of controlling for potential trade flows is presented in column (3). There, we show estimates of the standard regression when we add a control for the share of firms in the tradeable sector as of 2005 (column 1). The introduction of this control does not change the coefficient on the emigration rate suggesting that the effects identified in this paper is not due to a spurious correlation with trade.[19] In Appendix Table A8 we also split firms between the tradeable and the non-tradeable sectors. The largest impact of emigrants on firm creation is for non-tradeable sector firms. This indicates that the emigration flows we are analyzing do not seem to be particularly linked to

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

[19] We also perform a regression directly controlling for the value of exports per capita as of 2005 and a control for the change in export per capita in 2008-2015. The coefficient of the emigration rate remains virtually unchanged. We are unable to show these results due to a confidentiality agreement on the use of those trade data.

international trade activity.

7 Conclusions

In this paper, we provide empirical evidence on an important question about which we know very little: what happens to firm creation when emigration increases. We shed light on this question by taking advantage of a sudden and large emigration wave from Italy, occurred between 2008 and 2015, 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 is large and independent of pre-2008 local trends in firm creation and economic outcomes. These features provide support to causal interpretations of our estimates.

Our results indicate that Italian LLMs that lost more people due 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 employment and in the share of highly qualified workers. We then provide a decomposition of our quantitative effect into three parts. The first two are due to the simple subtraction of people, young and old, which would decrease firm creation in line with their average entrepreneurship rate of these groups. This component indicates that 37% of the effect is due to loss of young people and 5% to a loss of older people. The estimated causal effect, however, is much larger than these two components, indicating that the remaining 58% is due to a combination of selection of emigrants among those who may have higher entrepreneurship rate, and of potential spillovers effects on those who remain and become less likely to start businesses.

The primary role of the youth channel in explaining the loss of entrepreneurship is consistent with ideas put forth by Liang *et al.* (2018), Acemoglu *et al.* (2017) and Engbom (2019), namely that the lack of young managers and young entrepreneurs, may negatively affect firm creation, innovation and ultimately labor demand. However, the loss of firm creation due to emigration is much more than a simple subtraction of average young individuals, suggesting that emigrants are a highly selected group.

The findings in this paper have two main implications. First, we show that international migration implies much more than simple "labor flows". In some cases, migrants' roles as job-creators can be larger than their roles as employees, so that traditional models of migration constituting changes in labor supply may be missing a crucial part of the story. Second, our results suggest that emigrants are a highly selected group with high entrepreneurial abilities. This is in line with recent research showing that emigrants have a higher propensity to take risk (Jaeger *et al.* 2010) and higher intensity of traits such as "adaptability to new circumstances" (Bütikofer and Peri 2020). This positive selection of migrants on non-cognitive traits may be very important for understanding their economic impact and potential, and we hope to stimulate more research in this area.

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] 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.

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

[4] Borjas, George J., Ilpo Kauppinen and Panu Poutvaara. 2018. "Self-selection of emigrants: theory and evidence on the stochastic dominance in observable and unobservable characteristics," *The Economic Journal*, vol. 129, pages 143-171.

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

[6] Bütikofer, Aline and Giovanni Peri. 2020. "How Cognitive Ability and Personality Traits Affect Geographic Mobility," Journal of Labor Economics, forthcoming.

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

[8] 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.

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

[10] 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.

[11] 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.

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

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

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

[15] Fairlie, Robert W. and Lofstrom, Magnus. 2015. "Immigration and Entrepreneurship," CESifo Working Paper Series No. 5298.

[16] Fouka, Vasiliki, Soumyajit Mazumder, and Marco Tabellini. "From Immigrants to Americans: Race and Assimilation During the Great Migration." Harvard Business School Working Paper, No. 19-018, August 2018.

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

[18] 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.

[19] 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.

[20] Hunt, Jennifer and Marjolaine Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation?," *American Economic Journal: Macroeconomics* vol. 2, pages 31-56.

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

[22] Jaeger, David A., Thomas Dohmen, Armin Falk, David. Huffman, Uwe Sunde, and Holger Bonin. 2010. "Direct Evidence on Risk Attitudes and Migration," *The Review of Economics and Statistics*, vol. 92, pages 684-689.

[23] Karahan, Fatih, Benjamin Pugsley and Aysegül Sahin. 2019. "Demographic Origins of the start-up Deficit," NBER Working Papers 25874, National Bureau of Economic Research, Inc.

[24] Kerr, Sari Pekkala, William Kerr and Caglar Ozden and Christopher Parsons. 2017. "High-Skilled Migration and Agglomeration," *Annual Review of Economics* vol. 9, pages 201-34.

[25] 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.

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

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

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

[29] 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.

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

[31] Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German Jewish Émigrés and US Invention," *American Economic Review*, vol. 104(10), pages 3222-55.

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

[33] Patt, Alexander, Jens Ruhose, Simon Wiederhold and Miguel Flores. 2020. "International Emigrant Selection on Occupational Skills," *Journal of European Economic Association*, forthcoming.

[34] 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.

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

[36] Rauch, James E.. 1999. "Networks versus markets in international trade," *Journal of International Economics*, vol. 48(1), pages 7-35.

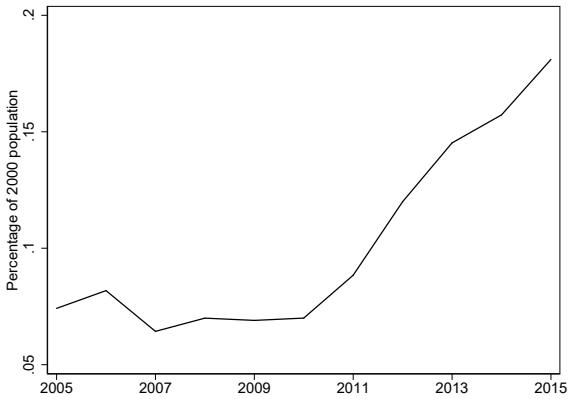
[37] Rauch, James E.. 2001. "Business and social networks in international trade," *Journal of Economic Literature*, vol. 39(1), pages 1177-1203.

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

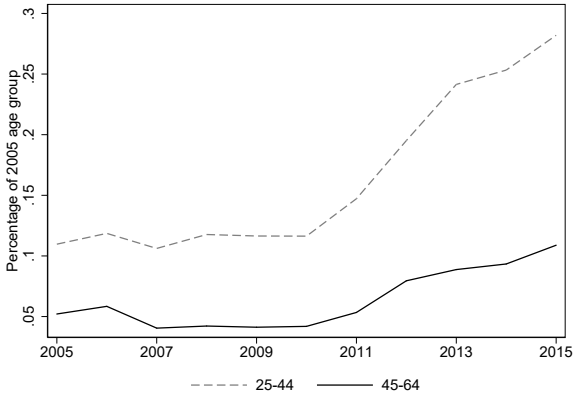
[39] Waldinger, Fabian. 2010. "Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany," *Journal of Political Economy*, vol. 118(4), pages 787-831.

Figures

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

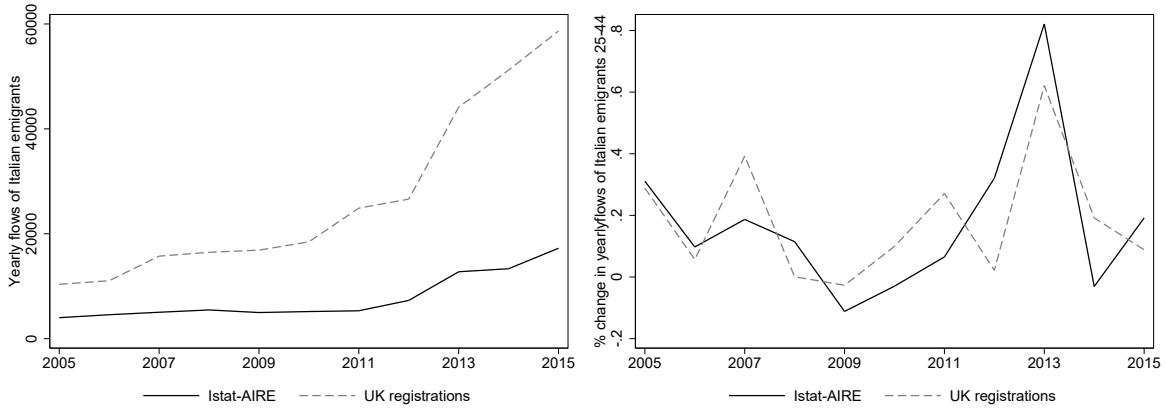


(a) Annual overall outflows from AIRE-Istat



(b) Annual outflows, by age group

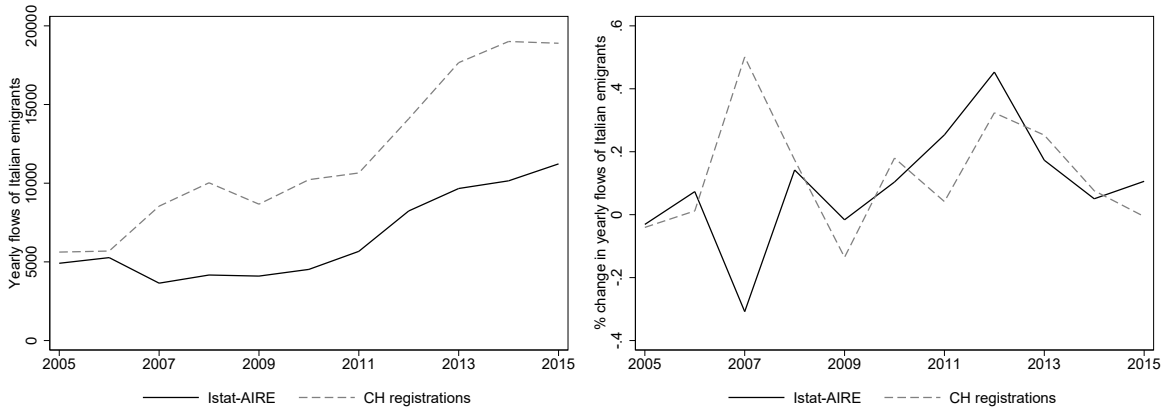
Figure 2: Recorded Emigration and UK Registered Italian Inflows



(a) Annual outflows from AIRE-Istat and inflows from UK Social Security Registry

(b) Percent changes in annual flows from AIRE-Istat and UK Social Security Registry

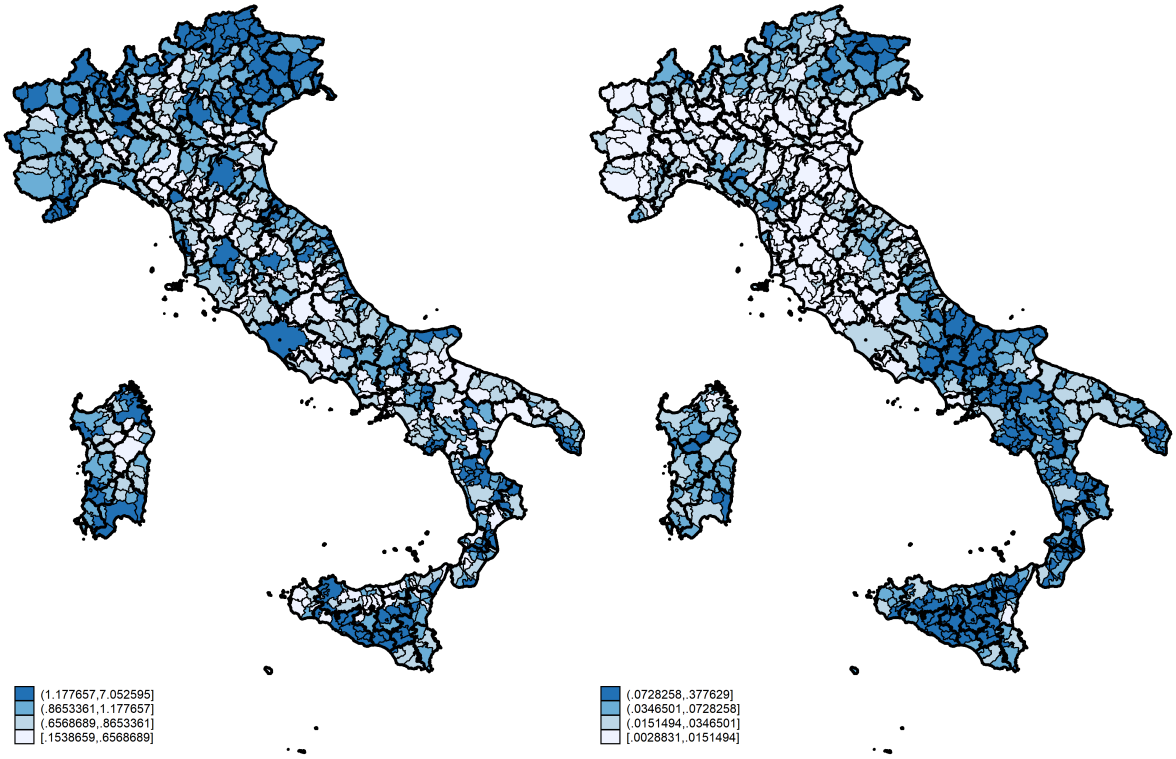
Figure 3: Recorded Emigration and Switzerland Registered Italian Inflows



(a) Annual outflows from AIRE-Istat and inflows from Switzerland BFS Registry

(b) Percent changes in annual flows from AIRE-Istat and Switzerland BFS Registry

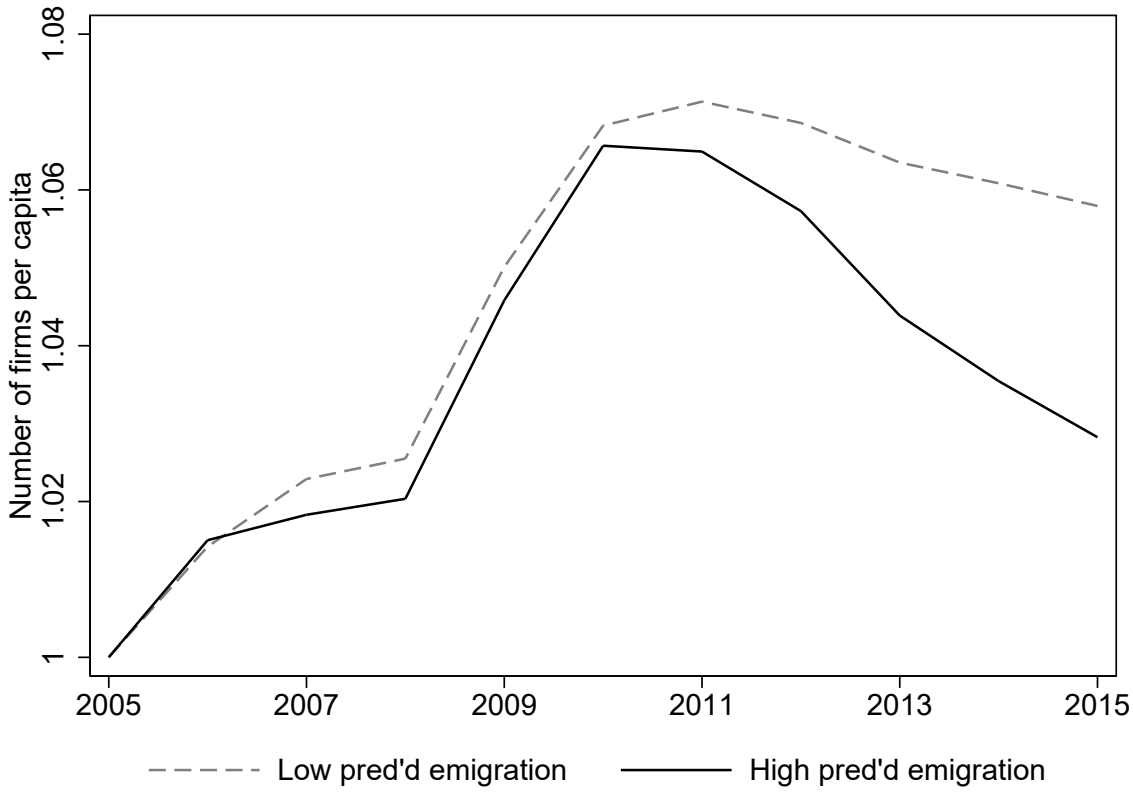
Figure 4: Actual and Predicted Emigration from Italian LLMs



(a) Emigration

(b) Predicted Emigration

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



Notes: The graph reports the series of the number of firms per person in LLMs predicted to have low and high emigration. The two series are normalized by their values in 2005: the two levels in 2005 are 0.19 and 0.18 for low and high-predicted emigration LLMs, respectively.

Tables

Table 1: Emigration by country-of-destination, top 5 countries: 2000 stock, 2008-2015 flows and 2008-2015 GDP performance

Panel A Top countries in 2000	Stock of Emigrants	GDP 2015/2008
Germany	286,570	1.07
Switzerland	228,725	1.09
France	165,244	1.04
Belgium	117,935	1.06
Argentina	99,506	1.11
Panel B		
Top Countries in $2008-15$	Flows	$\%\ of\ 25-44-y.o.$
Germany	70,104	48.6
U.K.	66,094	61.2
Switzerland	53,567	52.3
France	45,046	46.8
United States	27,563	54.9

Notes: Panel A reports the top 5 countries in terms of size of the emigration network as of 2000 as measured in the AIRE data, and the GDP growth between 2008 and 2015 calculated on IMF data (out of a total of 176 countries considered). For reference, GDP growth in both UK and US growth was 1.19 and in Italy was 0.93. Panel B reports the cumulated flows of emigrants to the top destination countries in the period 2008-2015 and the share of 25-44 years old measured in the Istat data. Stocks, flows, and the denominator of the share of young individuals include emigrants of all age groups.

Table 2: OLS 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
Emig Rate	0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.001)
Unemp Rate 2005	0.061	-0.000	-0.061
	(0.040)	(0.039)	(0.047)
GDP PC 2005	0.073	0.041***	-0.032
	(0.045)	(0.007)	(0.044)
Observations	686	686	686
R-squared	0.185	0.566	0.241
Avg. Outcome	-0.001	0.079	0.080
Avg. Treatment	1.020	1.020	1.020
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 (Column 1), cumulative firm entry (Column 2) and exit (Column 3) between 2008 and 2015 as a fraction of population 25-64 years old in the LLM (average 2005-2008). The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. We control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level and we include 110 province FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 3: Instrument validity check: Effect of emigration rates on pre-shock change in stock and flows of firms (2005-08)

VARIABLES	(1) All Firms Δ Stock 2005-08	(2) All Firms $\sum$ Births 2005-08	(3) All Firms $\sum$ Deaths 2005-08
Emig IV	-0.010	-0.002	0.008
	(0.011)	(0.011)	(0.010)
Observations	686	686	686
R-squared	0.161	0.626	0.181
Avg. Outcome	0.003	0.039	0.036
Avg. Treatment	1.020	1.020	1.020
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 population 25-64 years old in the LLM (average 2005-2008) 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 4: Emigration pull factor IV diagnostics

Panel A: Negative and positive weights

	Sum	Mean	Share
$\hat{\alpha}_c \leq 0$	-0.007	0	0.007
$\hat{\alpha}_c > 0$	1.007	0.01	0.993

Panel B: Correlations

	$\hat{\alpha}_c$	G_c	$\hat{\beta}_c$	$\hat{F}_c$	$Var(NTWK_c)$
$\hat{\alpha}_c$	1.0000				
G_c	-0.0523	1.0000			
$\hat{\beta}_c$	-0.0142	-0.0407	1.0000		
$\hat{F}_c$	0.0161	0.0704	-0.0149	1.0000	
$Var(NTWK_c)$	0.8419	-0.1015	-0.0281	0.0075	1.0000

Panel C: Top 5 destination countries

	$\hat{\alpha}_c$	G_c	$\hat{\beta}_c$	$\hat{F}_c$	95% C.I.
Germany	0.455	1.075	-0.006	12.94	(-0.02, -0.01)
Switzerland	0.278	1.01	-0.003	16.40	(-0.01, 0.00)
France	0.075	1.007	-0.005	3.53	(-0.20, 0.20)
Australia	0.039	1.064	-0.003	0.58	(-0.20, 0.20)
Belgium	0.029	1.005	-0.002	0.83	(-0.20, 0.20)

Notes: The table reports the Pull IV diagnostics as suggested by Goldsmith-Pinkham *et al.* (2018). Panel A reports the sum, the mean and the share of negative and positive Rotemberg weights $\hat{\alpha}_c$. Panel B reports correlations between the weights ($\hat{\alpha}_c$), the 2008-2015 destination country GDP growth (G_c), the just-identified coefficients ($\hat{\beta}_c$), the first stage F-statistic of destination countries emigrant networks ($\hat{F}_c$) and the variance in the emigrant networks across destination countries ($Var(NTWK_c)$). Panel C reports the top five destination countries according to the Rotemberg weights. The coefficients $\hat{\beta}_c$ are based on the regression of Table 7, Panel A, Column 2, where the outcome is the change 2008-2015 in the stock of firms per capita, and control variables include GDP per capita and unemployment rate in 2005 as well as 110 province FEs. We computed the Rotemberg decomposition using Goldsmith-Pinkham *et al.*'s Stata package.

Table 5: Relationship between destination countries' emigration networks and pre-period LLM characteristics

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Germany	Switzerland	France	Australia	Belgium
Δ Stock	-0.002	-0.002	0.004	-0.000	0.002
	(0.007)	(0.004)	(0.002)	(0.002)	(0.004)
$\sum$ Births	0.174	-0.083	-0.076	0.010	0.004
	(0.157)	(0.081)	(0.084)	(0.063)	(0.094)
Unemp Rate 2005	0.058	-0.050	-0.004	-0.054	0.004
	(0.075)	(0.047)	(0.027)	(0.037)	(0.047)
GDP PC 2005	-0.017	-0.016	-0.012	-0.003*	-0.005
	(0.011)	(0.011)	(0.009)	(0.002)	(0.004)
Observations	683	683	683	628	660
R-squared	0.485	0.454	0.416	0.497	0.360
Province FE	X	X	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the share of emigrants in each of the 5 top destination countries described in each column. The independent variables are the main LLMs observable characteristics, namely the change in stock and cumulative entry of firms between 2005 and 2008, unemployment rate and value added per capita in 100,000 euros in 2005, as well as 110 province FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 6: First stage regressions

VARIABLES	(1) Emig Rate	(2) Emig Rate	(3) Emig Rate
Pull IV	5.678***	5.837***	5.739***
	(1.067)	(1.072)	(1.489)
Unemp Rate 2005	-2.504**	0.641	2.561
	(1.120)	(1.469)	(2.177)
GDP PC 2005	0.664***	0.769***	0.873***
	(0.177)	(0.129)	(0.240)
Observations	686	686	686
R-squared	0.137	0.244	0.398
F-excluded instrument	28.310	29.667	14.865
Avg. Outcome	1.020	1.020	1.020
FE	-	Region	Province

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level. Column 1 include no fixed effects while Columns 2 and 3 include region (20) and province (110) FEs respectively. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table 7: Effect of emigration rates on change in stock and flows of firms

VARIABLES	(1) All Firms Δ Stock 2008-15	(2) All Firms $\sum$ Births 2008-15	(3) All Firms $\sum$ Deaths 2008-15
	Panel A: main results		
Emig Rate	-0.007***	-0.007**	-0.000
	(0.002)	(0.003)	(0.003)
R-squared	0.174	0.522	0.241
F-excl. instr.	14.865	14.865	14.865
Back of Env Firms	-303	-313	-11
	Panel B: controlling for lagged outcome		
Emig Rate	-0.005*	-0.006***	-0.002
	(0.003)	(0.002)	(0.002)
Δ Stock	1.129***		
	(0.007)		
$\sum$ Births		1.662***	
		(0.078)	
$\sum$ Deaths			1.157***
			(0.018)
R-squared	0.963	0.837	0.965
F-excl. instr.	14.864	14.863	14.855
Back of Env Firms	-217	-287	-80
Observations	686	686	686
Avg. Outcome	-0.001	0.079	0.080
Avg. Treatment	1.020	1.020	1.020
Back of Env Avg. Emig.	1187	1187	1187
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 population 25-64 years old in the LLM (average 2005-2008). The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 LLM over the period adjusted by the misreporting factor of 2.6 and the estimated effect on each outcome in an average LLM. In Panel B, we further control for the lagged outcome, i.e. for the change in firm stock, cumulative firm entry and exit between 2005 and 2008 as a fraction of population in each column respectively. ***, **, * 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 young-owned firms

VARIABLES	(1) Young Firms Δ Stock 2008-15	(2) Young Firms $\sum$ Births 2008-15	(3) Young Firms $\sum$ Deaths 2008-15
Emig Rate	-0.004**	-0.004*	0.000
	(0.002)	(0.002)	(0.003)
Observations	686	686	686
R-squared	0.340	0.472	0.470
F-excl. instr.	14.865	14.865	14.865
Avg. Outcome	-0.018	0.053	0.072
Avg. Treatment	1.020	1.020	1.020
Back of Env Avg. Emig.	1187	1187	1187
Back of Env Firms	-173	-172	1
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 stock, cumulative entry and exit of firms owned and managed by under 45 between 2008 and 2015 as a fraction of population 25-64 years old in the LLM (average 2005-2008). The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 LLM 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: First stage regression by LLM median age

VARIABLES	(1) All	(2) Rel. Younger	(3) Rel. Older
Emig IV	5.739***	6.128***	4.912*
	(1.489)	(1.753)	(2.777)
Observations	686	343	343
R-squared	0.398	0.477	0.442
F-excluded instrument	14.865	12.227	3.129
Avg. Outcome	1.020	1.058	0.982
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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 10: Effect of emigration rates on innovation

VARIABLES	(1) Start-Ups $\sum$ Births 2008-15
Emig Rate	-0.006**
	(0.002)
Observations	686
R-squared	0.325
F-excl. instr.	14.865
Avg. Outcome	0.010
Avg. Treatment	1.020
Back of Env Avg. Emig.	1187
Back of Env Firms	-3
Province FE	X

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 population 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 LLM 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: Effect of emigration rates on change in LLM employment

VARIABLES	(1) Δ Employees 2008-15	(2) Δ Emp/Pop 2008-15	(3) Δ Avg. Size 2008-15	(4) Δ Wage Bill 2008-15
Emig Rate	-0.068**	-0.048	-0.018	-0.027
	(0.030)	(0.030)	(0.037)	(0.034)
Observations	686	685	686	686
R-squared	0.198	0.205	0.244	0.264
F-excl. instr.	14.865	14.852	14.865	14.865
Avg. Outcome	-0.110	-0.072	-0.014	-0.113
Avg. Outcome 2005	16709.0	0.3	5.5	348.6
Avg. Treatment	1.020	1.020	1.020	1.020
Back of Env Avg. Emig.	1187	1188	1187	1187
Back of Env. Effect	-1160	-0	-0	-10
Province FE	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, employment to population (25-64) ratio, average firm size and total wage bill between 2008 and 2015, as a fraction of each outcome in 2005. The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 LLM 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 12: Effect of emigration rates on change in LLM skills

VARIABLES	(1) Δ Blue Coll 2008-15	(2) Δ White Coll 2008-15	(3) Δ Managers 2008-15
Emig Rate	-0.026	-0.086**	-1.697
	(0.041)	(0.043)	(1.621)
Observations	686	686	584
R-squared	0.199	0.233	0.187
F-excl. instr.	14.865	14.865	6.287
Avg. Outcome	-0.117	-0.013	0.248
Avg. Outcome 2005	8950	6737	192
Avg. Treatment	1.020	1.020	0.980
Back of Env Avg. Emig.	1187	1187	1141
Back of Env. Effect	-234	-590	-319
Province FE	X	X	X

2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variables are the change 2008-2015 in LLM employment by skill level as a share of 2005 employment by skills in each column respectively. The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 LLM over the period adjusted by the misreporting factor of 2.6 and the estimated effect on each outcome in an average LLM.

Table 13: Robustness checks

VARIABLES	(1) Controlling for Immigration $\sum$ Births	(2) Excluding border provinces $\sum$ Births	(3) Controlling for tradable share $\sum$ Births
Emig Rate	-0.006**	-0.006**	-0.008**
	(0.003)	(0.003)	(0.003)
Immig Rate 05-08	0.004***		
	(0.000)		
Tradable sh. 2005			-0.041
			(0.028)
Observations	686	590	686
R-squared	0.613	0.503	0.511
F-excl. instr.	15.652	15.652	13.646
Avg. Outcome	0.079	0.080	0.079
Avg. Treatment	1.020	0.945	1.020
Back of Env Avg. Emig.	1187	1100	1187
Back of Env Firms	-255	-253	-358
Province FE	X	X	X

Notes: 2SLS estimates. In Columns 1 and 3, 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. The dependent variable is the change in cumulative firm entry between 2008 and 2015 as a fraction of population 25-64 years old in the LLM (average 2005-2008). The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. In Column 1, we also include the cumulative immigration rate between 2005 and 2008 as a percentage of LLM population in 2000. In Column 3 we also control for the share of LLM firms in tradable sectors in 2005. In all columns, we further control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level (not reported) as well as 110 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

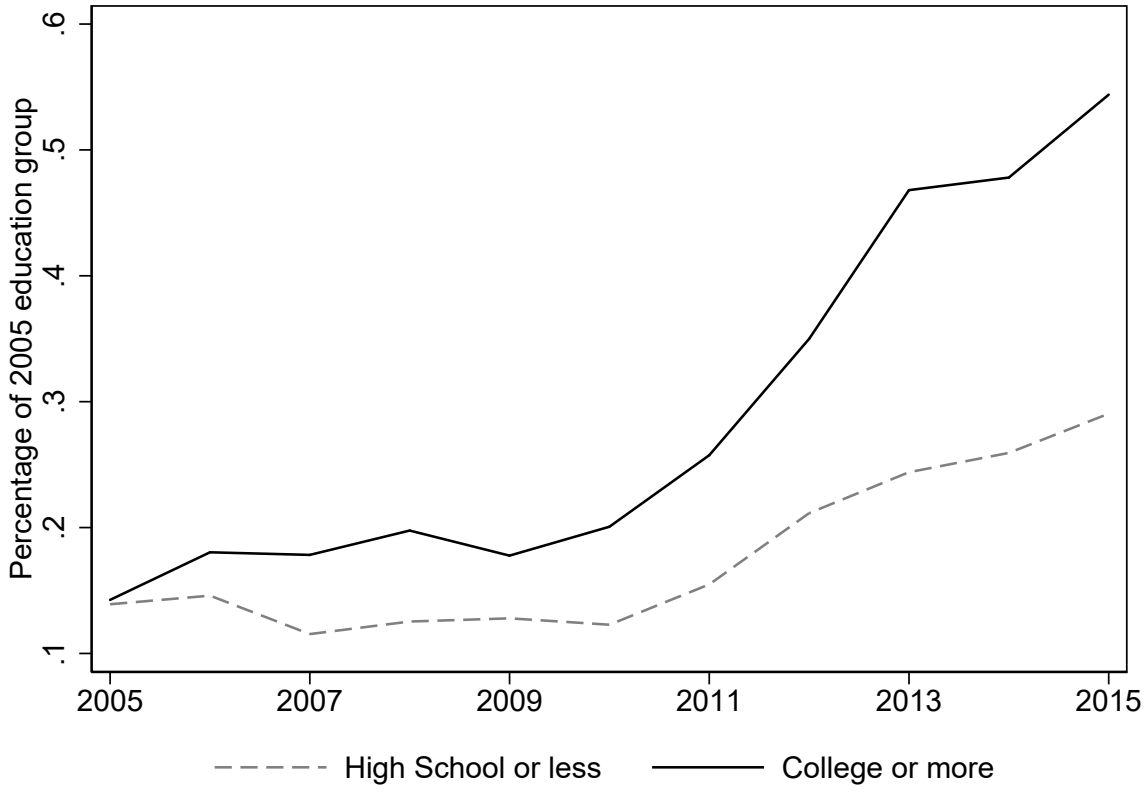
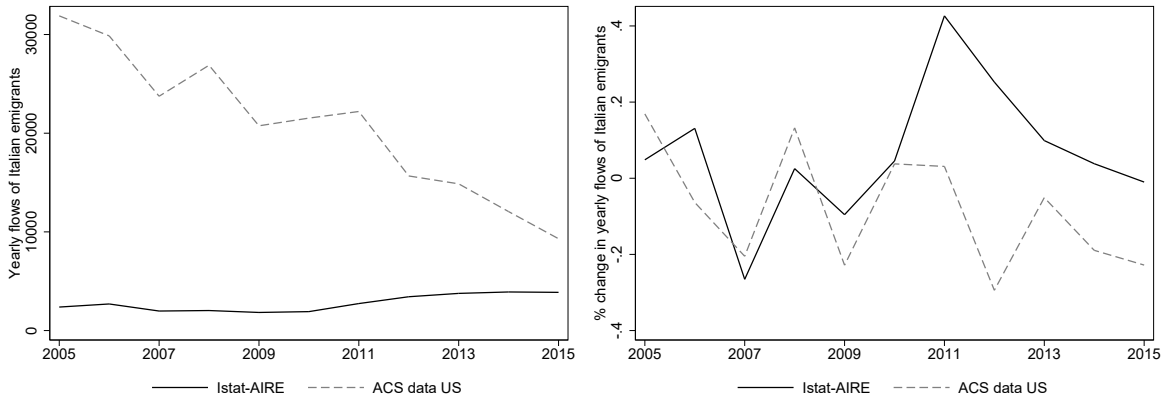


Figure A2: Recorded Emigration and US Italian Inflows



(a) Annual outflows from AIRE-Istat and inflows from US Census ACS

(b) Percent changes in annual flows from AIRE-Istat and US Census ACS

A.II Accounting for Under-registration in AIRE-Istat Emigration Data

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

Table A1: Correction factor based on destination country data

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

Table A1 compares the emigration flows registered in the Italian AIRE-Istat data to the immigration flows registered by each foreign source respectively. The variable *Factor* shows the ratio between the immigration and emigration flow in each year. The data shows that emigration flows are systematically under-reported in the AIRE-Istat data, in almost every year and for all the three countries considered. In the last two columns, we construct a weighted average of the correction factors (weighted by the emigration flows). If we include the US (penultimate column), the average correction factor for the period 2009-15 is about 3.48. However, as the ACS data are survey-based and thus less reliable than the administrative sources from UK and Switzerland, in the last column we only consider the two latter countries, for which the average correction factor ranges between 2.62 and 3.23 over the period 2009-15 and is 2.87 on average. Based on these results, we use the minimum value of the average correction factor, 2.6, to adjust upwards the emigration flows between 2009-15 throughout our empirical analysis.

A.III Instrument Validity Checks: Additional Pre-trends

Table A2: Instrument validity check: Effect of emigration rates on pre-shock change in stock and flows of Young-owned firms (2005-08)

	(1)	(2)	(3)
	Young Firms	Young Firms	Young Firms
	Δ Stock	$\sum$ Births	$\sum$ Deaths
VARIABLES	2005-08	2005-08	2005-08
Emig IV	0.003	-0.001	-0.004
	(0.010)	(0.010)	(0.006)
Observations	686	686	686
R-squared	0.222	0.529	0.340
Avg. Outcome	-0.001	0.028	0.029
Avg. Treatment	1.020	1.020	1.020
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 stock, cumulative entry and exit of firms owned and managed by individuals under 45 years old between 2005 and 2008 as a fraction of population 25-64 years old in the LLM (average 2005-2008) 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 A3: Instrument validity check: Effect of emigration rates on pre-shock change in LLM employment (2005-08)

VARIABLES	(1) Δ Employees 2005-08	(2) Δ Emp/Pop 2005-08	(3) Δ Avg. Size 2005-08	(4) Δ Wage Bill 2005-08
Emig IV	-0.148	-0.024	-0.160	-0.192
	(0.198)	(0.201)	(0.188)	(0.221)
Observations	686	686	686	686
R-squared	0.282	0.260	0.177	0.194
Avg. Outcome	0.119	0.126	0.025	0.127
Avg. Outcome 2005	16709.0	0.3	5.5	348.6
Avg. Treatment	1.020	1.020	1.020	1.020
Province FE	X	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variables are the change in employment, employment to population (25-64) ratio, average firm size and total wage bill between 2005 and 2008, as a fraction of each outcome in 2005 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 A4: Instrument validity check: Effect of emigration rates on pre-shock change in LLM skills (2005-08)

VARIABLES	(1) Δ Blue Coll 2005-08	(2) Δ White Coll 2005-08	(3) Δ Managers 2005-08
Emig IV	0.027	-0.222	5.863
	(0.219)	(0.237)	(5.799)
Observations	686	686	584
R-squared	0.323	0.137	0.134
Avg. Outcome	0.129	0.133	0.247
Avg. Outcome 2005	8950	6737	163
Avg. Treatment	1.020	1.020	0.980
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variables are the change 2005-2008 in LLM employment by skill level as a share of 2005 employment by skills 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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.

A.IV IV Diagnostics for the Consulate-based IV

The following tables, A5 and A6, replicate the main tests proposed by Goldsmith-Pinkham *et al.* (2018) for the IV based on the consulates rather than on the countries. Similarly to what shown in Table 4, Table A5 shows that the cross-sectional components of the pull emigration instrumental variable is driven by networks of Italian emigrants towards German and Swiss regions. Interestingly, the estimated coefficients of Stuttgart/Friburg and Dortmund/Koln, that alone make up about 40 percent of the IV variation, are close to each other (0.044 and 0.045) and close to the main estimate at the consulate and country level (about 0.05). Table A6 shows the correlations between the share of emigrants towards the most relevant regions and the main labor market characteristics: we fail to find statistically significant correlations with observable LLM characteristics, similarly to what shown in the country level analysis in the main text, although we acknowledge that the estimated coefficients are large in magnitude.

Table A5: Emigration pull factor IV diagnostics

Panel A: Negative and positive weights

	Sum	Mean	Share
$\hat{\alpha}_c \leq 0$	-0.007	0	0.007
$\hat{\alpha}_c > 0$	1.007	0.01	0.993

Panel B: Correlations

	$\hat{\alpha}_c$	G_c	$\hat{\beta}_c$	$\hat{F}_c$	$Var(NTWK_c)$
$\hat{\alpha}_c$	1.0000				
G_c	-0.0804	1.0000			
$\hat{\beta}_c$	-0.0217	-0.0365	1.0000		
$\hat{F}_c$	0.0179	0.0688	-0.0137	1.0000	
$Var(NTWK_c)$	0.7535	-0.1360	-0.0386	-0.0005	1.0000

Panel C: Top 5 destination regions

	$\hat{\alpha}_c$	G_c	$\hat{\beta}_c$	$\hat{F}_c$	95% C.I.
Stuttgart/Friburg	0.250	1.075	-0.008	7.33	(-0.04, -0.01)
Zurich	0.106	1.01	-0.002	12.59	(-0.01, 0.00)
Dortmund/Koln	0.090	1.075	-0.005	2.64	(-0.20, 0.20)
Lugano	0.077	1.010	-0.000	6.51	(-0.12, 0.00)
Frankfurt	0.075	1.075	-0.003	2.61	(-0.20, 0.20)

Notes: The table reports the Pull IV diagnostics as suggested by Goldsmith-Pinkham *et al.* (2018). Panel A reports the sum, the mean and the share of negative and positive Rotemberg weights $\hat{\alpha}_c$. Panel B reports correlations between the weights ($\hat{\alpha}_c$), the 2008-2015 destination region/country GDP growth (G_c), the just-identified coefficients ($\hat{\beta}_c$), the first stage F-statistic of destination regions/countries emigrant networks ($\hat{F}_c$) and the variance in the emigrant networks across destination regions/countries ($Var(NTWK_c)$). Panel C reports the top five destination regions according to the Rotemberg weights. The coefficients $\hat{\beta}_c$ are based on the regression of Table 7, Panel A, Column 2, where the outcome is the change 2008-2015 in the stock of firms per capita, and control variables include GDP per capita and unemployment rate in 2005 as well as 110 province FEs. We computed the Rotemberg decomposition using Goldsmith-Pinkham *et al.*'s Stata package.

Table A6: Relationship between country-of-destination emigration networks and LLMs' characteristics

VARIABLES	(1) Stuttgart/Friburg	(2) Zurich	(3) Dortmund/Koln	(4) Lugano	(5) Frankfurt
Δ Stock	-0.002	-0.001	0.001	0.000	-0.001
	(0.006)	(0.001)	(0.002)	(0.001)	(0.001)
$\sum$ Births	0.142	-0.035	0.042	0.013	-0.053**
	(0.099)	(0.032)	(0.056)	(0.024)	(0.023)
Unemp Rate 2005	0.017	-0.010	0.034	0.009	-0.010
	(0.048)	(0.021)	(0.036)	(0.016)	(0.009)
GDP PC 2005	-0.007	-0.004	-0.004	-0.003	-0.004
	(0.005)	(0.003)	(0.003)	(0.002)	(0.003)
Observations	653	666	651	645	648
R-squared	0.301	0.416	0.401	0.457	0.388
Province FE	X	X	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variable is the share of emigrants in each of the 5 top destination regions described in each column. The independent variables are the main LLMs observable characteristics, namely the change in stock and cumulative entry of firms between 2005 and 2008, unemployment rate and value added per capita in 100,000 euros in 2005, as well as 110 province FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

A.V Additional Robustness Checks

One may be concerned that our instrument is correlated with internal migration flows. While these should not be correlated with pull factors from abroad, the network of emigrants may be correlated with the internal flows and with local push factors. For instance, LLMs with high emigration rates to foreign countries could also exhibit substantial emigration to other Italian LLMs, and the latter may reduce firm creation, violating the exclusion restriction. We thus test whether our estimates are robust to this potential threat. In Table A7, Columns 1 and 2, we report the results of our first stage regression where internal migration outflows and inflows are the outcome variables respectively. The effect is not statistically significant: this indicates that the instrument does not predict emigration to or immigration flows from other LLMs in Italy. In Column 3 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 13).

In Table A8 we report the results of regressing firm entry by considering firms operating in tradable and non-tradable sectors separately, as we discuss in Section 6.

In Tables A9, A10 and A11 we include the lag of the outcome variables among the set of controls. In all cases the main results continue to hold.

Table A7: First stage regression on internal migration flows and immigration

VARIABLES	(1) Internal Emig	(2) Internal Immig	(3) Immig Rate 05-08
Emig IV	8.104	-7.464	-2.028
	(5.931)	(5.498)	(1.794)
Observations	686	686	686
R-squared	0.410	0.462	0.754
F-excluded instrument	1.867	1.843	1.277
Avg. Outcome	15.860	14.694	4.352
Province FE	X	X	X

Notes: OLS estimates. The sample is composed of 686 local labor markets (LLMs). In Columns 1 and 2, the dependent variable is the cumulative emigration and immigration of Italians to and from different LLMs in Italy between 2008 and 2015 respectively, while in Column 3 the dependent variable is the cumulative inflow of foreign-born immigrants from abroad between 2005 and 2008. All the outcomes are as a fraction of population 25-64 years old in the LLM (average 2005-2008). 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level, as well as 110 province FEs. ***, **, * indicate significance at 1-percent, 5-percent and 10-percent level, respectively.

Table A8: Effect of emigration rates on firms growth, in tradable and non tradable sectors

VARIABLES	(1) Tradable $\sum$ Births 2008-15	(2) Non Tradable $\sum$ Births 2008-15
Emig Rate	-0.001**	-0.006**
	(0.000)	(0.003)
Observations	686	686
R-squared	0.545	0.508
F-excl. instr.	14.865	14.865
Avg. Outcome	0.006	0.073
Avg. Treatment	1.020	1.020
Back of Env Avg. Emig.	1187	1187
Back of Env Firms	-49	-264
Province FE	X	X

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, as a fraction of population 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 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 LLM 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 A9: Robustness check - Effect of emigration rates on change in stock and flows of firms owned by under 45 (2008-15) controlling for lagged outcomes (2005-08)

	(1)	(2)	(3)
	Young Firms	Young Firms	Young Firms
	Δ Stock	$\sum$ Births	$\sum$ Deaths
VARIABLES	2008-15	2008-15	2008-15
Emig Rate	-0.004*	-0.004*	0.001
	(0.002)	(0.002)	(0.002)
Δ Stock	0.736***		
	(0.066)		
$\sum$ Births		1.497***	
		(0.083)	
$\sum$ Deaths			1.068***
			(0.195)
Observations	686	686	686
R-squared	0.656	0.808	0.791
F-excl. instr.	15.027	14.929	14.869
Avg. Outcome	-0.018	0.053	0.072
Avg. Treatment	1.020	1.020	1.020
Back of Env Avg. Emig.	1187	1187	1187
Back of Env Firms	-190	-162	32
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 stock, cumulative entry and exit of firms owned and managed by under 45 between 2008 and 2015 as a fraction of population 25-64 years old in the LLM (average 2005-2008). The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level (not reported) as well as 110 province FEs. Additionally, we control for the change in stock, cumulative entry and exit of firms owned and managed by under 45 between 2005 and 2008 as a fraction of population 25-64 years old in the LLM (average 2005-2008). The back of the envelope report the number of emigrants from the average LLM 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 A10: Robustness check - Effect of emigration rates on LLM employment (2008-15) controlling for lagged outcome (2005-08)

VARIABLES	(1) Δ Employees 2008-15	(2) Δ Emp/Pop 2008-15	(3) Δ Avg. Size 2008-15	(4) Δ Wage Bill 2008-15
Emig Rate	-0.067**	-0.048	-0.017	-0.027
	(0.030)	(0.031)	(0.037)	(0.034)
Δ Employees	0.060			
	(0.087)			
Δ Emp/Pop		0.133		
		(0.098)		
Δ Avg. Size			0.046	
			(0.115)	
Δ Wage Bill				0.017
				(0.063)
Observations	686	685	686	686
R-squared	0.200	0.210	0.245	0.265
F-excl. instr.	14.948	15.174	13.939	14.557
Avg. Outcome	-0.110	-0.072	-0.014	-0.113
Avg. Outcome 2005	16709.0	0.3	5.5	348.6
Avg. Treatment	1.020	1.020	1.020	1.020
Back of Env Avg. Emig.	1187	1188	1187	1187
Back of Env. Effect	-1133	-0	-0	-9
Province FE	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, employment to population (25-64) ratio, average firm size and workers by qualification between 2008 and 2015, as a fraction of each outcome in 2005. The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level (not reported) as well as 110 province FEs. Additionally, we control for the percentage change in each outcome between 2005 and 2008. The back of the envelope report the number of emigrants from the average LLM 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 A11: Robustness check - Effect of emigration rates on LLM skills (2008-15) controlling for lagged outcome (2005-08)

VARIABLES	(1) Δ Blue Coll 2008-15	(2) Δ White Coll 2008-15	(3) Δ Managers 2008-15
Emig Rate	-0.026	-0.080*	-1.816
	(0.041)	(0.043)	(1.673)
Δ Blue Coll	0.001		
	(0.079)		
Δ White Coll		0.155	
		(0.103)	
Δ Managers			0.049
			(0.250)
Observations	686	686	584
R-squared	0.199	0.242	0.182
F-excl. instr.	15.208	14.761	6.234
Avg. Outcome	-0.117	-0.013	0.248
Avg. Outcome 2005	8950	6737	192
Avg. Treatment	1.020	1.020	0.980
Back of Env Avg. Emig.	1187	1187	1141
Back of Env. Effect	-234	-549	-341
Province FE	X	X	X

2SLS estimates. The sample is composed of 686 local labor markets (LLMs). The dependent variables are the change 2008-2015 in LLM employment by skill level as a share of 2005 employment by skills in each column respectively. The independent variable is the cumulative emigration rate between 2008 and 2015, i.e. the cumulative flow of Italians aged 25-64 emigrating abroad as a fraction of the population of 25-64 years old in the LLM (average 2005-2008) and multiplied by 100. The instrument 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, $Pull_l = \sum_c NTWK_{l,c} * G_c$. We further control for unemployment rate and value added per capita in 100,000 euros in 2005 at the LLM level (not reported) as well as 110 province FEs. Additionally, we control for the percentage change in each outcome between 2005 and 2008. The back of the envelope report the number of emigrants from the average LLM over the period adjusted by the misreporting factor of 2.6 and the estimated effect on each outcome in an average LLM.