How does immigration affect native internal mobility? New evidence from Italy

Sauro Mocetti and Carmine Porello [§]

PRELIMINARY DRAFT

November 2008

Abstract. This paper investigates the relationship between native internal mobility and immigrant inflows. To address this issue we examine both the interregional migration of natives and the demographic evolution of the local labour markets. Endogeneity issue related to the location choices of immigrants are faced exploiting both the existence of previous enclaves and the proximity to "gateways" as instruments for immigrant geographical distribution. We find evidence of complementarities between immigrants and high-educated natives and (if any) a modest displacement effect on less educated natives.

JEL classification: I2, J61, O15, R23

Keywords: Immigration, native internal mobility, enclaves, distance, complementarities

[§] Bank of Italy. Corresponding author: Sauro Mocetti, Regional Economic Research Staff, Bologna Branch, Piazza Cavour 6, 40124, Bologna. E-mail: sauro.mocetti@bancaditalia.it. We thank Antonio Accetturo for useful comments. The views expressed here do not necessarily reflect those of the Bank of Italy.

1. Introduction

Most of the empirical studies regarding the labour market impact of immigration exploit the geographic clustering of immigrants. In these studies, a measure of natives' outcomes (e.g. wage) in a locality is usually regressed on the stock of immigrants in that locality. These "spatial correlations" have two main drawbacks. First, the local labour markets are assumed to be closed, thus ignoring potential selective out-migration and in-migration of natives.[1] Moreover, immigrants are not randomly distributed across labour markets and this makes difficult to isolate the effect of immigration on natives from other related phenomena.

The aim of the present paper is to examine the labour market adjustment to immigration in Italy or, more precisely, how native internal mobility responses to immigrant inflows. Natives are distinguished by educational level, thus allowing for a differential impact depending on the skill group they belong. The empirical findings can be interpreted in both negative and positive terms: on one side, they cast doubts on spatial correlation exercises for Italy if the selective migration of natives is observed to be at work; on the other side, they enable us to test whether natives and immigrants are complements or substitutes looking at what natives express "with their feet".

Previous works on the relationship between natives' internal mobility and immigration inflows find mixed results (see more on this below). Our empirical work adds to the existing literature in several dimensions. First, Italy represents an interesting case of study from an institutional point of view. Most of the existing empirical literature concerns the U.S. whereas we provide evidence on a country that is traditionally characterized by the presence of powerful trade unions, centralised bargaining, and a strong regulation of the labour market in general. Therefore it is reasonable to expect that adjustments to labour shocks occur more on the quantity side rather than on wages. Second, Italy is an interesting case of study also from a geographical point of view. During the last decades it has gradually become one of the main destinations of immigration in Europe; moreover, immigrants are characterized by a high level of ethnic fractionalization. These facts are partly related to its great geographical exposure to the main international migration flows. Third,

[1] In case of natives' mobility, the labour market impact of immigration will not be observed locally but the effects will be spread across the entire country.

contrarily to most of the earlier studies, we distinguish the effect of immigrant inflows on native mobility by educational attainment. Most of foreign immigrants in Italy are low educated and, almost independently from the educational attainment, employed in low qualified jobs. Our aim is to analyze their impact on low- and high-educated natives' location choices. Finally, the use of native internal migration across regions enables us to distinguish between push and pull factors.

The impact of immigration on natives' location choices is examined with two different empirical exercises. In the first, we use data on interregional migration flows of natives drawn from the General Population Register (GPR) and we examine how they respond to immigration. In the second, we analyze the demographic evolution of Local Labour Markets (LLMs) using census data, in order to strengthen our previous results when a finer partition of the territory is adopted.

As mentioned above, identifying the effects of immigration on labour market conditions of natives is challenging since the location of immigrants is itself the outcome of an economic decision. To address the endogeneity issue we rely on two instrumental variables. The first exploits the tendency of newly arriving immigrants to settle in places where previous immigrants of the same country already live. The second uses the distance between local labour markets and the gateway through which immigrants enter Italy to exogenously distribute foreigners over the territory.

Our results suggest (if any) a modest negative displacement effect of immigration on less educated natives. As far as the high skilled natives are concerned, we find that they are complement to foreign workers. A further implication of these findings is that the impact of immigration on natives' wages and employment opportunities is not fully observed locally but it is spread over the territory. OLS estimates are partially reversed by IV estimates thus suggesting that the empirical findings on the labour market of immigrants can be severely biased if endogeneity is not taken into account.

The paper is structured as follows. In section 2 we present some theoretical arguments as well as empirical literature on the relationship between native mobility and immigrant inflows. In section 3 we provide a descriptive evidence of immigration in Italy. In section 4 we examine the impact of immigrant inflows on interregional mobility of natives. In section 5 we analyze the effect of immigration growth on the demographic evolution of LLMs. In section 6 we briefly report our conclusions.

2. The economic impact of immigration

The literature that has examined the labour market effect of immigration flows has been focused on the impact on wages or employment opportunities. The usual assumption is that the entry of immigrants into the labour market of a certain area should lower the wage of competing workers (workers who have the same type of skills), and increase the wage of complementary workers (workers whose skills become more valuable due to immigration). But the empirical evidence showed that these effects are small and often not significant.[2]

A problem related to this literature is that selective out-migration by natives may cancel out the immigrant inflows: if the arrival of one unskilled immigrant leads one unskilled native to leave, then immigrant inflows will have no detectable impact on local wages. As a result, a comparison of the wages of native workers across regions with different incidence of immigrants might show little or no difference because the effects of immigration are diffused throughout the national economy, and not because immigration had no economic effects (Borjas, 2003).

The aim of the our analysis is to verify if this kind of labour market adjustment is at work in Italy. If so, this casts some doubts on spatial correlation evidences. However, such results can still provide evidence interpretable in terms of complementarities and substitution effects. If we observe that a larger fraction of foreign-born people in a labour market is associated to higher out-flows (and/or lower in-flows) of natives, this means that immigrants compete with natives. If the opposite is true then foreigner workers and natives are complements.

Empirical evidence of native-born mobility responses to immigrant inflows is controversial. Filer (1992) finds that metropolitan areas where immigrants cluster have higher rates of native-born out-migration. Similarly, Frey (1996) reports a strong correlation between immigrant inflows and native outflows, and he argues that this behaviour is leading to a "demographic balkanization". Wright *et al.* (1997) show that immigrant inflows are unrelated to natives outflows from large metropolitan areas.[3] In the last decade, Card and Borjas have been the two most prominent researchers working on this issue. Card and DiNardo (2000) find that an increases in immigrant population in specific skill groups lead to small increases in

[2] Card (1990) is probably the most known article on this topic. See Okkerse (2008) and the works cited therein for a review of the literature.
[3] See also Walker *et al.* (1992), White and Liang (1998), and Kritz and Gurak (2001).

the population of native-born individuals of the same skill group; Card (2001) shows that inflows of new immigrants to individual cities did not generate large offsetting mobility flows by natives. On the contrary, Borjas *et al.* (1997) report a strong negative correlation between native net migration and immigration by states; Borjas (2006) finds that immigration is associated with lower in-migration rates, higher out-migration rates, and a decline in the growth rate of the native workforce. Hatton and Tani (2005) is the only work, among those cited above, to examine the relationship between immigration and internal mobility in a country different from the U.S. They analyze the labour adjustments in Britain and find a negative displacement effect.[4]

Whether immigration harms, improves or has no effect on natives' labour opportunities is a complex issue also from a theoretical point of view. Potential complementarities between immigrants and more skilled native are intuitive. The higher supply of immigrants is likely to require workers in jobs where they perform supervisory, managerial, training, and coordinating tasks. Furthermore, high skilled natives can pay less for the services that unskilled foreign-born provide – painting the house, caring old parents and/or young children, etc. – and can specialize in producing the goods and services that better suit their competencies. The relationship between immigrants and low skilled natives seems also obvious. They belong to the same skill group and compete for the same jobs; therefore, negative displacement effects are expected. However, in a segmented labour market the substitutability between immigrants and less skilled natives can be far from perfect. If immigrants undertake jobs which natives refuse, and if these jobs address specific labour shortages, new employment opportunities for natives can become available.[5] The high supply of immigrants might also convince firms to not outsource abroad, thus increasing local employment opportunities. Finally, the presence of new foreign workers also implies higher demand for consumption, so that immigration might simply increase total production and labour demand without any displacement effects.[6]

[4] With respect to Hatton and Tani (2005), our empirical exercise has some advantages. We distinguish native flows by educational level thus allowing for a differential impact of immigration depending on the human capital content of natives. Furthermore, we strengthen our results with a finer partition of the territory, looking at the demographic evolution of LLMs.

[5] According to Gavosto *et al.* (1999), immigrants can contribute to remove potential mismatch in the labour markets. This occurs when firms are unable to expand their output because they do not find workers who are willing to undertake certain (low-skilled) jobs.

[6] See Ottaviano and Peri (2005) for an analysis of complementarities between immigrants and natives.

3. Immigrants in Italy

The scope of this section is to provide a descriptive assessment of immigration in Italy. In particular, we focus on where immigrants come from and how migration patterns are changed across time; where immigrants live in Italy; and what they do, also in comparison with natives.

Italy has been for a long time a country of emigration. However, in the last decades the flow has reversed and it has reached a positive migration balance. Starting from the second half of the 90s Italy and Spain has become prime destinations in the EU (see Table 1). The percentage of foreign born individuals in Italy increased from 0,6 percent in 1991 to 2,3 percent in 2001 to 5,0 percent in 2006. In all areas of the country, most of the population growth results from immigration (see Table 2). In the South, the population growth was much weaker and immigration inflows were not able to counterbalance native outflows towards northern regions.

Looking at the distribution of immigrants by source country two main features arise. First, Italy is characterized by a high degree of fractionalization, though it weakens across time. In 2006 the first 10 countries represent slightly more than half of the total immigrants (see Table 3). This is partly related to the great exposure of Italy towards main international migration routes: the northern-east border is directly connected with the Balkans; the 7,600 kilometres of coastline were exposed to migration flows from Albania (especially Apulia), from Northern Africa and also from Indian subcontinent. Second, the composition of immigrants by source countries is strongly changed. With the exception of Morocco, the ranking of the first 5 countries is different from that of 1991. The incidence of immigrants from Middle and East Europe increases from around 10 percent in the beginning of the 90s to more than 40 percent in 2005; during the same period, the share of immigrants from Africa decreases from 35 to 23 percent (see Figure 1). Generally speaking, in the past the international migration flows were mainly in the direction South-North. During the 90s, with the fall of communist regimes in Central and Eastern Europe, the dissolution of former Yugoslavia and the Soviet Union, and the EU enlargement process, intra-European East-West migrations become predominant.

The distribution of immigrants across Italian provinces largely reflects the economic opportunities available at the local level and the North-South divide is

noticeable (see Figure 2). However, from our point of view, it is much more interesting to look at the distribution of immigrants over the territory by source country. We build a concentration index obtained as the ratio between the share of immigrants of nationality *n* who live in province *i* and the share of all immigrants living in that province. For a sample of the main source countries from each continent, we report in Figure 3 a graphical representation of this index. The incidence of Albanians is relatively higher in Apulia, the closest region from a geographical point of view. Serbians (and other people from the Balkans) are more concentrated in the northern east provinces, close to the frontier. Chinese are relatively more present around Naples and, in particular, in the province of Prato where they work in the local textile industry. Migrants from South America (see Ecuador and Peru) are relatively more distributed in Liguria, and close to the metropolitan areas of Rome and Milan. Finally, the incidence of migrants from Africa is relatively higher in the southern regions. This sketched representation of immigrant distribution over the territory clearly shows that their location choices are not only driven by local economic conditions. It is likely that previous enclaves and proximity to the frontiers (that, in turn, are differentiated by countries and migration trajectories) play a key role. We will exploit these features to address the endogeneity issue in the empirical section.

In table 4, we report occupation and sector distribution of natives and immigrants by educational level. Nearly half of foreign workers are low educated – with at most lower secondary education. However, the occupation and sector distribution of immigrants is not markedly different by educational level, contrarily to what happens for natives. Four immigrants out of five are blue-collar workers, and they are usually employed in occupations that are lower ranked, in terms of skill content and wages, than native born workers with the same level of education. Immigrants usually take jobs avoided by natives (e.g. low paid household and other service jobs) and jobs in sectors that are more affected by seasonal fluctuations (e.g. construction, tourism).[7] Therefore immigrants, almost independently from their educational level, stay at the bottom of the employment ladder. Moreover, their occupation and sector distribution is quite different from that of natives, even from that of low skilled natives.

[7] See also Brandolini *et al.* (2005) and Münz (2007).

4. Analysis of the interregional flows

In this section we examine the impact of immigration on native internal mobility. Our identifying strategy relies on the geographical diversity in immigrants' concentrations and on the changes in immigration patterns over time.

4.1 Data

Data on internal migration of natives are drawn from the GPR.[8] With internal migration of natives we mean the residential migration that occurs when a native changes its place of residence (about 2 percent of the population each year). In particular, we refer to moves across Italian regions. We restrict our analysis throughout to Italian citizens thus excluding mobility of previous immigrants. We select individuals in the age bracket 15-44 to consider only working-age population and to exclude individuals nearing retirement. This choice is dictated by the fact that migration is presumably more common among younger workers since older ones have a smaller expected lifetime gain from moving. Furthermore, young workers are supposed to compete most heavily with foreign-born individuals. We consider the period 1996-2003 for reasons of data availability.

Perhaps the most striking feature of internal mobility in Italy is the persistent net outflow from the South to the Centre-North. This flow was significant during the 1960's, when a considerable number of people were leaving the southern regions in favour of the northern (more developed) regions. The phenomenon loosed strength in the 70's and in the 80's. In the middle of the 90s the migration flows from the South started to grow again, attracting new attention from researchers. With respect to the past, the (human capital) composition of native migrants is changed since the fraction of those who are graduated has increased substantially.[9]

Data on immigrants are drawn from the Ministry of Interior and they refer to the number of residence permits. Evidence on immigrants' presence in Italy is reported in the previous section and we do not provide further details in the following. From a merely descriptive point of view, migration flows of natives and

[8] This measure of residential mobility should be accompanied by two main caveats. First, there may be a time lag between the actual migration and its registration. Second, it does not take into account all the possible types of regional mobility. For example, some people may transfer to another municipality without formalizing it at the register offices.

[9] See Mocetti and Porello (2008) for a detailed description of the phenomenon and for a review of the literature on internal migration in Italy.

of immigrants seem to go in the same directions (see Figure 4). However the cross-region migration patterns are much more complex: they vary considerably over time, and across destination and origin regions, and their composition by educational level is also changed.

Other explanatory variables are used to control for other relevant factors that may affect our outcomes of interest.[10] Unemployment rate (*UNR*) and GDP per worker (*GDP*) are the covariates traditionally used in the literature as main determinants of migrations flows. They measure the job opportunities in a region and clearly affect the expected income (Harris and Todaro, 1970). The share of export of goods with high and increasing international demand (*EXPORT*) should identify the booming economies and approximate future income prospects in a region. Finally, the cost of houses (*HOUSE*) is introduced since it reasonably "deflates" the income prospect in a region; this variable also enable us to control for further ways through which immigrants might affect natives' residential choice, for example through their impact on the housing markets.[11] See Table 5 (panel A) for descriptive statistics.

4.2 Empirical strategy

Our empirical specification is the following:

$$m_{ijet} = \beta_1 IMM_{it} + \beta_2 IMM_{jt} + \beta_3 X_{iet} + \beta_4 X_{jet} + \beta_5 X_{it} + \beta_6 X_{jt} + D_t + \mu_{ijet} \qquad (1)$$

where the dependent variable represents the migration flow of individuals endowed with human capital e, from region i to region j, at time t; migration flow is normalized by population size, with the corresponding educational level, of both sending and destination region. We consider three types of educational level (e): 1 = at most compulsory school; 2 = upper secondary school; and 3 = university degree. IMM_{it} and IMM_{jt} are the incidences of immigrants in regions i and j, at time t. X_{iet} and X_{jet} are control variables varying by educational level (e), region (i or j) and time (t); for example, the unemployment rate of graduated in Lombardy in 2000. X_{it} and X_{jt} are control variables varying by region (i or j) and time (t) and not by educational level; for example, the GDP per worker in Lombardy in 2000. Panel analysis allows

[10] Main determinants of the decision to migrate can be grouped in economic, cultural and institutional factors. We will focus on economic determinants because cultural differences and institutional barriers that play a significant role in explaining international migration are far less important in determining internal migration.

[11] Saiz (2007) find that immigration pushes up rents and housing values in U.S. destination cities and that this impact is an order of magnitude bigger than that found in labour markets.

us to control for fixed effects for education and origin-destination pairs. To avoid simultaneity effect, we relate *current* migration flows to *lagged* values for all the explanatory variables. Finally, we include year dummies (D_t) to take out the effects of economy-wide conditions on internal mobility.

4.3 Endogeneity

There are a number of possible omitted variables that could affect both immigration inflows and native mobility: local demand shocks, expectations about local occupational opportunities, changes in the amenities. If this was the case, the estimates of the relation between immigration and natives' mobility are biased. To address the endogeneity issue, we should use variation in immigration inflows that is plausibly exogenous to the evolution of native internal migration. To this scope we rely on two instrumental variables.

The first instrument exploits the supply-push component of the immigrant inflows and the tendency of newly arriving immigrants to settle in places where previous immigrants from the same country already live (Card and DiNardo, 2000; Card, 2001; Saiz, 2007).[12] For each source country of immigration, we calculate the share of immigrants living in region *i* in 1990, the first year for which this information is available. We apply these weights to "distribute" new immigration inflows from each country into regions for our period of interest. Formally:

$$\overline{IMM}_{it} = \sum_{n=1}^{N} \delta_{ni} \cdot IMM_{nt} \qquad (2)$$

where δ_{ni} measures the share of immigrants from country *n* that are settled in region *i* in 1990, and IMM_{nt} represents the number of immigrants from country *n* at time *t* in Italy.[13] The validity of this instrument relies on the assumption that first settlements

[12] The instrument is motivated by a study of Bartel (1989) who shows that settlement patterns of previous immigrants are a main determinant of immigrants' location choices. Existing networks of previous migrants might ease the arrival of newcomers in several ways: they provide information about the job opportunities; they may help them in finding jobs and/or an apartment; they also provide an existing social network, with individuals sharing the same cultural and linguistic background. See also Zavodny (1999) and Belot and Ermisch (2006).

[13] We consider the first 30 countries in terms of residence permits (excluding countries from Western Europe and North America). They are in alphabetical order: Albania, Algeria, Argentina, Bangladesh, Bosnia, Brazil, Bulgaria, China, Colombia, Croatia, Dominican Republic, Ecuador, Egypt, Philippines, Ghana, India, Macedonia, Morocco, Moldavia, Nigeria, Pakistan, Peru, Poland, Romania, Russia, Senegal, Serbia, Sri Lanka, Tunisia, and Ukraine. Shares for Bosnia, Croatia and Ukraine refer to 1992, for Macedonia to 1993.

are observed with a sufficient lag and local economic shocks are not too persistent over time.

We also exploit the great geographical exposure of Italy to international migration flows and we build an alternative instrumental variable. Specifically, we use the distance between each province and the gateway through which immigrants enter Italy. Angrist and Kugler (2003) and Ottaviano and Peri (2005 and 2006) use a similar approach. Unlike these papers, we consider all the main countries of immigration in Italy and we differentiate them by entry. Assumptions on migration trajectories and the gateways used by immigrants to enter Italy are based on: *i*) geographical reasons, especially for those countries that are close to the Italian borders; *ii*) survey among immigrants in which they declare the frontier used to enter Italy[14]; *iii*) information on migration routes gathered from official reports by the Ministry of Interior and field studies like Monzini *et al.* (2004) and European Migration Network (2005). To not put excessive burden on this section, we refer to the box "Migrants' trajectories and gateways to Italy" in the Appendix for a complete description. In the following we describe how we build the instrumental variable. We regress:

$$\log(IMM_{pnt}) = \alpha_1 \log(DISTANCE_{pn}) + \alpha_2 \log(IMM_{nt}) + D_p + D_n + D_t + D_p \times D_t + \varepsilon_{pnt}$$

where IMM_{pnt} is the number of immigrants in province p from country n at time t; $DISTANCE_{pn}$ is the distance between province p and gateway used by immigrants from country n to enter Italy; IMM_{nt} is the number of immigrants from country n at time t at the national level. We also include dummies for provinces (D_p), year (D_t), source country of immigration (D_n), and the interaction between provinces and year ($D_p \times D_t$), to isolate the "exogenous" component of immigrants' location choices from any nationality and province-year specific effects.[15] Therefore, our second instrument

[14] Data on self-declared region of entry are drawn by ISMU. ISMU is an autonomous and independent organization promoting studies, research and projects on multi-ethnic and multi-cultural society, and focusing in particular on the phenomenon of international migrations. It conducts a survey every year on a sample of immigrants living in Lombardy. In the survey 2002 there was a question in which immigrants were asked to declare the region of entry in Italy. See the Box in the Appendix for further details.

[15] Distance between each province and each gateway is in (log of) kilometres. When immigrants from country *n* are assumed to enter Italy from different gateways, the minimum distance between the province and each gateway is considered. The regression is a pooled OLS at the province level (103) for the period 1996-2002 and, again, we consider the first 30 countries in terms of residence permits.

is obtained collapsing the predicted number of immigrants from country *n* in province *p* at time *t* by region and year.[16] Formally:

$$\overline{\overline{IMM}}_{it} = \sum_{p \in i} \left(\sum_{n=1}^{N} \exp\left(\left(\hat{\alpha}_1 \log\left(DISTANCE_{pn} \right) + \hat{\alpha}_2 \log\left(IMM_{nt} \right) \right) \right) \right) \quad (3)$$

4.4 Results

As a general strategy, we run the regressions on all individuals and after we operate a sample split to consider possible differential effects of immigration on low- and high-educated natives. First we report OLS estimates and then IV estimates are examined and discussed.

Most of the control variables are correctly signed (see Table 7). Native outflows are larger in regions with higher unemployment rate. GDP per worker in the sending region has the expected negative sign. In both cases, less educated response more than graduated to economic factors. Natives (especially graduated) are attracted by those regions which have an export orientation towards those good with a high and increasing international demand. House prices enters, as expected, with negative sign for destination region and positive sign for sending region; the results are confirmed if we look at the subsample of graduated natives whereas are opposite for unskilled ones. Endogeneity of house prices (more inflows lead to a greater demand pressure on the destination housing markets) might explain this striking result.

Turning to our key variable, a larger incidence of immigrants is associated to lower inflows and larger outflows of natives. According to these estimates, a 1 percent increase of immigrant incidence is associated with about 0.2 percent increase of native outflows and about 0.2 percent decrease of native inflows. The results are qualitatively similar for both low- and high-educated natives even if push factors are more relevant for the less educated whereas the unattractiveness of region destination weights more for Italian graduated. Therefore, these results seem to suggest that immigrants push away (and repel) both unskilled and skilled natives. However, some caution is needed because there might be several sources of endogeneity that can bias the OLS estimates. First, there might be a number of potentially relevant omitted

[16] We also repeat the same exercise splitting the sample by area of origin of immigrants. In particular we divide the 30 countries in five main groups: Africa, Asia, South America, Balkans and East-Europe. The coefficient of distance continues to be highly significant even if its magnitude varies by area of origin (see Table 6). The results are fairly consistent with the ones obtained without sample split and are not reported in the text. However, they are available from the authors upon request.

variables that might affect our variables of interest. For example, there might be local demand shocks for jobs that attract immigrants and are avoided by natives (e.g. domestic services, construction) and, together, economic slowdown in sectors traditionally filled by natives. The bias deriving from the omission of these variables should be more severe for natives whose degree of substitutability is lower, i.e. who work in sectors markedly different from those of immigrants. It is also possible, although less likely, that a reverse relationship is at work. That is, immigrants go where natives' outflows are larger. We rely on an instrumental variable strategy to address these potential biases.

IV estimates are partly reversed with respect to OLS estimates. Broadly speaking, the displacement effect on less skilled native is partially confirmed whereas immigrants and high-skilled natives are complements rather than substitutes.[17]

As far as the low-educated natives are concerned, IV estimates based on immigrant enclaves confirm the displacement effect (see Table 8). The coefficients imply that a 1 percent increase of immigrant incidence is associated with about 0.4 percent decrease of native inflows whereas the effect on sending region is not significant. However, when IV estimates based on distances from gateways are considered, the coefficients on immigrant incidence are negative and significant both in the sending and in the destination region, thus leading to an uncertain final outcome (see Table 9). A possible explanation of this contrasting result is that within unskilled native group, some workers are in direct competition with immigrants whereas some others might benefit from them because they are not perfect substitutes.

Results on migration flows of graduated natives are opposite with respect to the OLS estimates. This is likely due to an omitted variable that would affect the employment opportunities of immigrants and natives in the opposite direction, which in turn would bias OLS estimates downward. According to IV estimates, a higher incidence of immigrants is associated to a larger inflow of graduated. The magnitude of elasticity varies between 0.9 to 1.1 percent, depending on the IV adopted. Therefore, the impact seems to be substantial. No detectable impact is found in the sending region.

[17] The instruments proposed are strongly correlated with our endogenous variables. The first stage F-statistics are well above the rule-of-thumb benchmark of 10 in all the specifications adopted.

5. Analysis of the demographic evolution of LLMs

The empirical analysis presented in the previous section has several advantages: it allows use to control for fixed effects for origin-destination pairs in natives' interregional flows; furthermore, it allows us to distinguish between push and pull factors. However, it also has a main drawback. Immigrants tend to be highly geographically concentrated and to examine the effect on the local labour market one may want to look at a finer partition of the territory.[18]

In this section we control the robustness of our previous results trying to answer to the limitation mentioned above. In particular, we examine the natives' demographic evolution of the LLMs in response to the immigration growth.[19]

5.1 Data and empirical approach

LLMs are defined as clusters of municipalities that can be considered as self-contained labour markets on the basis of the degree of work-day commuting by the residents. These territorial units result from the organization of social and economic relations and do not reflect geographical particularities or historical events such as regions or other administrative partitions of the territory. Therefore, they represent the best territorial configuration in terms of labour market features and probably are the most appropriate units to analyze externalities in production. Unfortunately, data on demographic evolution by educational level of the LLMs are available only through census data (1991 and 2001).

Our empirical strategy consists in running the following regression:

$$ITA\_GR_{le} = \beta_1 IMM\_GR_l + \beta_2 X_l + \beta_3 X_{le} + \mu_{le} \qquad (4)$$

where *ITA_GR* is population growth of Italians in LLM *l*. As in the previous section we consider the overall population growth, and that referred to low- and high-educated natives (subscript *e*). *IMM_GR* represent the immigrant growth in the same LLM. As before, if $\beta_1>0$ ($\beta_1<0$) then immigrants and natives are considered as complements (substitutes). X_l and X_{le} are covariates that vary by LLM and by LLM

[18] Immigrants are strongly concentrated towards the major cities. In 1991, the 29 percent of all foreign born individuals lived in cities with more 500,000 inhabitants (Rome, Milan, Turin, Genoa, Naples and Palermo). However, in the last year the concentration of immigrants decreased; the respective figure for 2004 was 17 percent.

[19] See Card and DiNardo (2000), Card (2001) and Card (2007) for a similar exercise.

and educational level, respectively. See Table 5 (panel B) for descriptive statistics of both the dependent variable and the covariates.

The obvious problem with this specification is that the unobserved determinants of population are likely to be correlated with immigrant inflows, leading to an upward-biased estimate of β_1. To isolate the causal effect of immigrant inflows on native population growth, we have to find an instrumental variable that induces more immigrants to move to a certain LLM but is not directly related to its population growth. We rely, as in the previous section, on two instrumental variables based on the existence of previous enclaves and proximity to the frontiers. They are built in a slightly different way due to data availability and the new territorial partition.

The predicted number of immigrants obtained by enclaves' assumption is the following:

$$\overline{IMM}_{lt} = \sum_{a=1}^{A} \delta_{al} \cdot IMM_{at} \tag{5}$$

where δ_{al} represents the share of immigrants from area a living in LLM l in 1991. Unlike section 4, at the LLM level we only know the distribution of immigrants by geographical area (Western Europe, Middle and East Europe, North Africa, Other countries of Africa, Asia, Oceania, North America, and South America).

As far as the IV based on distance from the gateways is concerned, we use the parameter α_1 and α_2 estimated in section 4 and we predict the number of immigrants as in (3) with the only differences that here distances are calculated between each LLM and each gateway. Formally:

$$\overline{\overline{IMM}}_{lt} = \sum_{n=1}^{N} \exp\left[\left(\hat{\alpha}_1 \log\left(DISTANCE_{nl}\right) + \hat{\alpha}_2 \log\left(IMM_{nt}\right)\right)\right] \tag{6}$$

5.2 Results

For each dependent variable (overall native population growth, low- and high-educated native population growth) we present two econometric specifications: in the first we consider only demographic controls, in the second we add variables capturing the economic features of the LLMs.

The OLS estimates are reported in Table 10. *LNPOP* and the share of low- and high-educated natives (*LOWEDUC* and *HIGHEDUC*) are introduced to account for heterogeneity in initial conditions. The population growth in the previous decade

(*POP_GR_1*, *LOWEDUC_GR_1* and *HIGHEDUC_GR_1*) controls for trend effects in the growth pattern. The share of older people in 1991 (*SHARE_65+*) is, as expected, negatively associated to population growth because of the lower natality and the higher risk of mortality. The population density of LLMs in 1991 (*DENSITY*) enters positively, thus suggesting that agglomeration effects prevails on potential congestion effects. Finally, we add variables capturing the economic features of the LLMs. The unemployment rate (*UNR*) is negatively associated with overall population growth, thus suggesting that LLMs with better employment prospects are those who experienced a positive demographic evolution. However, *UNR* enters positively as far as high-educated population growth is concerned. To explain this apparently striking result one may argue that youngsters living in a depressed area would be more inclined to acquire further education rather than quit school and endure a spell of unemployment. The occupation growth in the service sector (*SER_EMP_GR*) is positively associated to population growth and the impact is stronger for high-educated natives. Finally, we include dummies for the production specialization of the LLM.

According to these estimates, immigration growth is positively associated to native population growth. The positive correlation is only partially confirmed when we distinguish between low- and high-educated natives. However, due to paucity of data at the LLM level, we cannot exclude that the relationship we find is driven by some unobserved omitted variables (e.g. LLMs with thriving economies). Again, we rely on IV estimates to try to take account for endogeneity issues.

The predictive power of previous enclaves is weak and the first-stage F-statistics is below the lower bound of 10 suggested by the literature on the strength of instruments (see Table 11). This is in contrast with what we find in the previous section where a more sophisticated enclave instrument is adopted. In some LLMs, the presence of immigrants in 1991 was negligible and therefore the weights can perform badly in capturing the network effects.[20] According to these estimates, immigration growth does not have any significant effect on native population growth.

[20] For the validity of the instrumental variable there should be a large enough number of immigrants from a certain area to influence the location choices of future immigrants. Notice also that the distribution of immigrants over the Italian territory is country specific and, between 1991 and 2001, there was a huge change in composition of immigrants by country of origin.

The instrument based on distances between each LLM and the gateways through which immigrants enter Italy perform much better (see Table 12). The IV estimates strengthen the composition effect of the demographic evolution of LLMs. The effect on overall native population growth is negative and it is not significant in the first (more parsimonious) specification. The overall negative effect is driven by that on low-educated natives. On the contrary, the impact of immigration growth is positive and of an higher order of magnitude when high-educated natives are considered. The coefficients vary between 0.6 and 0.9. Taken literally, estimates of this magnitude imply that an increase by 10 percent in immigrant population growth in a LLM lead to an increase between 6 and 9 percent in the Italian graduated growth in that LLM.

6. Conclusion

The possible negative impact of immigration on labour market conditions for natives is one of the core concerns in the public debate on immigration. The present paper investigates the labour market adjustment to the arrivals of foreign workers in Italy.

Italy is an interesting case to study both because of its labour markets institutions and of its great geographical exposure to the main recent international migration flows. Endogeneity issue related to the location choices of the immigrants are faced exploiting both the existence of previous enclaves of immigrants and proximity to gateways as instruments for immigrant geographical distribution.

We find quite different results between OLS and IV estimates, likely due to the omission of some relevant explanatory variable. If we accept the validity of our instrumental variables, our findings show that there is (if any) a modest displacement effect of immigration on less skilled natives. On the contrary, immigrants are found to be complement with more educated natives.

References

Angrist, Joshua D. and Adriana D. Kugler (2003), Protective or counter-productive? Labour market institutions and the effect of immigration on EU natives, *Economic Journal*, 113: F302-F331.

Bartel, Ann P. (1989), Where do the New United-States immigrants live, *Journal of Labour Economics*, 7: 371-391.

Belot, Michele and John Ermisch (2006), Friendship ties and geographical mobility, evidence from BHPS, ISER working paper no. 2006-33.

Borjas, George J. (2003), The labor demand curve is downward sloping: re-examining the impacts of immigration on the labor market, *Quarterly Journal of Economics*, 118: 1335-1374.

Borjas, George J. (2006), Native internal migration and the labour market impact of immigration, *Journal of Human Resources*, 41: 221-258.

Borjas, George J., Richard B. Freeman, Lawrence F. Katz and John DiNardo (1997), How much do immigration and trade affect labour market outcomes?, *Brookings Papers on Economic Activity*, 1997: 1-67.

Brandolini, Andrea, Piero Cipollone and Alfonso Rosolia (2005), Le condizioni di lavoro degli immigrati in Italia, in Massimo Livi Bacci (a cura di), *L'incidenza economica dell'immigrazione*, Giapichelli Editore, Torino.

Card, David (1990), The impact of the Mariel Boatlift on the Miami labour market, *Industrial and Labour Relations Review*, 43: 245-257.

Card, David (2001), Immigrant inflows, native outflows, and the local labour market impacts of higher immigration, *Journal of Labour Economics*, 19: 22-64.

Card, David (2007), How immigration affects U.S. cities, mimeo.

Card, David and John DiNardo (2000), Do immigrant inflows lead to native outflows?, *American Economic Review*, 90: 360-367.

European Migration Network (2005), Irregular migration in Italy.

Filer, Randall K. (1992), The impact of immigrant arrivals on migratory patterns of native workers, in George J. Borjas and Richard B. Freeman (eds.), *Immigration and the workforce: economic consequences for the United States and source areas*, Chicago, University of Chicago Press.

Frey, William H. (1996), Immigration, domestic migration, and demographic Balkanization in America: new evidence for the 1990s, *Population and Development Review*, 22: 741-763.

Gavosto, Andrea, Alessandra Venturini and Claudia Villosio (1999), Do immigrants compete with natives?, *Labour*, 3: 603-621.

Harris, John R. and Michael P. Todaro (1970), Migration, unemployment and development: a two-sector analysis, *American Economic Review*, 60: 126-142.

Hatton, Timothy J. and Massimiliano Tani (2005), Immigration and interregional mobility in the UK, 1982-2000, *Economic Journal*, 115: F342-F358.

Kritz, Mary and Douglas Gurak (2001), The impact of immigration on the internal migration of natives and immigrants, *Demography*, 38: 133-145.

Mocetti, Sauro and Carmine Porello (2008), La mobilità del lavoro in Italia: un'interpretazione dei pattern migratori, mimeo.

Monzini, Paola, Ferruccio Pastore and Giuseppe Sciortino (2004), L'Italia promessa: geopolitica e dinamiche organizzative del traffico di migranti verso l'Italia, CeSPI working paper 9/2004.

Münz, Rainer (2007), Migration, labour markets, and integration of migrants: an overview for Europe, Hamburg Institute of International Economics, working paper 3-6.

Okkerse, Liesbet (2008), How to measure labour market effects of immigration: a review, *Journal of Economic Survey*, 22: 1-30.

Ottaviano, Gianmarco I. P. and Giovanni Peri (2005), Cities and cultures, *Journal of Urban Economics*, 58: 304–337.

Ottaviano, Gianmarco I. P. and Giovanni Peri (2006), The economic value of cultural diversity: evidence from US cities, *Journal of Economic Geography*, 6: 9-44.

Saiz, Albert (2007), Immigration and housing rents in American cities, *Journal of Urban Economics*, 61: 345-371.

Walker, Robert, Mark Ellis and Richard Barff (1992), Linked migration systems: immigration and internal labor flows in the United States, *Economic Geography*, 68: 234-238.

White, Michael and Zai Liang (1998), The effect of immigration on the internal migration of the native born population, 1981-1999, *Population and Policy Review*, 17: 141-166.

Wright, Richard A., Mark Ellis and Michael Reibel (1997), The linkage between immigration and internal labour migration in large metropolitan areas in the United States, *Economic Geography*, 73: 234-254.

Zavodny, Madeline (1999), Determinants of recent immigrants' locational choices, *International Migration Review*, 33: 1014-1030.

Appendix – tables

Table 1: Migration balance into the EU

	1950-1960	1961-1970	1971-1980	1981-1990	1991-2000	2001-2005
France	973 (2.1)	2,033 (4.2)	605 (1.2)	494 (0.9)	227 (0.4)	718 (2.4)
Germany	1.011 (1.4)	1,488 (2.0)	1,505 (1.9)	2,022 (2.6)	3,347 (4.1)	799 (2.0)
Italy	-1,014 (-2.0)	-972 (-1.9)	-84 (-0.2)	-132 (-0.2)	410 (0.7)	1,889 (6.6)
Spain	-796 (-2.6)	-608 (-1.9)	144 (0.4)	-227 (-0.6)	1,302 (3.3)	2,967 (14.2)
Sweden	85 (1.1)	223 (2.9)	84 (1.0)	172 (2.1)	200 (2.3)	140 (3.2)
UK	-539 (-1.0)	-49 (-0.1)	-235 (-0.4)	-2 (0.0)	634 (1.2)	906 (3.0)
EU 25	-2,284 (-0.6)	148 (0.0)	3,078 (0.7)	2,926 (0.7)	7,343 (1.7)	8,786 (3.8)

Cumulative net flows (inflows – outflow) in thousands; annual rate ‰ in parenthesis.
Source: Münz (2007).

Table 2: Demographic evolution by area

	Population growth rate 2001-2006	Contributions to overall population growth (a)		Immigrants incidence (2006)
		Internal migration	Immigrant inflows	
North West	4,6	0,7	3,5	6,8
North East	5,3	1,7	3,7	7,2
Centre	5,8	1,0	3,3	6,3
South and Islands	1,2	-1,2	0,9	1,6

(a) The two columns decompose the overall growth rate into the changes in the native and immigrant populations. More precisely, the first column consider the internal net migration and the second column the immigration balance, each divided by the initial population stock. The residual component of population growth includes the balance between natality and mortality.
Source: Istat

Table 3: Immigrants by source country

1991			2001			2006		
Morocco	63,806	11.6	Morocco	162,254	11.8	Romania	271,491	11.9
Tunisia	31,881	5.8	Albania	146,321	10.6	Albania	256,916	11.2
Philippines	26,166	4.8	Romania	69,999	5.1	Morocco	239,728	10.5
Yugoslavia	22,335	4.1	Philippines	65,073	4.7	Ukraine	115,087	5.0
Senegal	21,073	3.8	China	60,143	4.4	China	114,165	5.0
Q5		*30.1*			*36.5*			*43,6*
Q10		*40.4*			*50.3*			*57,4*

Residence permits by source country. Q5 (Q10) represents the share of the first 5 (10) countries with respect to the total number of the residence permits.
Source: Ministry of Interior.

Table 4: Occupation and sector distribution of natives and immigrants by educational level

	Natives			Immigrants	
	All sample	Low educated	Graduated	All sample	Low educated
Percentage:	*100,0*	*39,7*	*15,2*	*100,0*	*48,5*
Occupation:					
Executive employee	7,9	0,8	31,0	1,5	0,1
White collar	31,8	12,3	37,5	5,6	2,5
Blue collar	32,1	55,7	1,6	76,5	81,3
Self-employee	22,9	26,1	24,1	12,5	12,0
Other	5,4	5,0	5,8	4,0	4,1
of which:					
Unskilled jobs	16,4	29,4	0,7	42,8	48,4
Sector of activity:					
Agriculture	4,3	7,9	0,8	3,9	5,0
Industry	21,8	26,3	9,8	23,7	24,5
Construction	7,7	13,1	1,1	17,2	21,2
Commerce, restaurants, etc.	25,7	29,7	9,7	22,4	22,1
Other private services	13,9	6,0	27,3	8,0	5,1
Public administration	21,1	9,6	47,6	4,3	2,0
Other social & family services	5,6	7,3	3,9	20,5	20,2

Source: authors' elaboration on data from Istat (LFS, year 2006).

Table 5: Definition of the variables and descriptive statistics

A – Empirical exercise on natives interregional mobility

Name	Description [source]	Mean	St.dev.
LOWEDUC_FLOW	Log of migration flow of low-educated (at most compulsory schooling) natives across regions, normalized by the corresponding population of both sending and destination region [GPR].	-7.54	0.814
HIGHEDUC_FLOW	Log of migration flow of high-educated (graduated) natives across regions, normalized by the corresponding population of both sending and destination region [GPR].	-5.00	0.859
IMM	Log of the ratio between immigrants (measured by the number of residence permits) and population [Ministry of Interior].	-4.19	0.708
UNR (low-educated)	Unemployment rate of low educated [ISTAT].	0.14	0.101
UNR (high-educated)	Unemployment rate of high educated [ISTAT].	0.11	0.055
GDP	Log of GDP per employee [ISTAT]	3.72	0.090
EXPORT	Ratio between the export of goods with high and increasing international demand and the total export [ISTAT].	30.2	16.24
HOUSE	Log of the housing price [Bank of Italy].	7.45	0.308

B – Empirical exercise on LLMs' demographic evolution

Name	Description [source]	Mean	St.dev.
ITA_GR	Overall Italians growth rate between 1991 and 2001 [ISTAT].	-0.02	0.043
LOWEDUC_ITA_GR	Low-educated (at most compulsory schooling) Italians growth rate between 1991 and 2001 [ISTAT].	-0.15	0.045
HIGHEDUC_ITA_GR	High-educated (graduated) Italians growth rate between 1991 and 2001 [ISTAT].	0.78	0.368
IMM_GR	Immigrant growth rate between 1991 and 2001 [ISTAT].	3.53	2.632
LNPOP	Log of population in 1991 [ISTAT].	10.5	1.147
LOWEDUC	Share of Italians with at most compulsory school in 1991 [ISTAT].	0.82	0.046
HIGHEDUC	Share of Italians with a university degree in 1991 [ISTAT].	0.02	0.011
SHARE_65+	Share of population aged 65 or more in 1991 [ISTAT].	0.15	0.036
POP_GR_1	Overall population growth rate between 1981 and 1991 [ISTAT].	0.01	0.046
LOWEDUC_GR_1	Low-educated growth rate between 1981 and 1991 [ISTAT].	-0.07	0.059
HIGHEDUC_GR_1	High-educated growth rate between 1981 and 1991 [ISTAT].	0.41	0.166
DENSITY	Log of density (inhabitants per squared kilometre) in 1991 [ISTAT].	5.84	1.144
UNR	Unemployment rate in 1991 [ISTAT].	0.19	0.123
SER_EMP_GR	Employment growth in the service sector between 1991 and 2001 [ISTAT].	0.02	0.088
LLM TYPE:			
NO_SPEC	Dummy equal to 1 if the LLM has not a prevailing specialization (reference category) [ISTAT].	0.32	0.467
NON_MANIF	Dummy equal to 1 if the LLM is mainly non-manufacture [ISTAT].	0.19	0.394
MADE_ITALY	Dummy equal to 1 if the LLM is specialized in made-in-Italy production [ISTAT].	0.34	0.473
HEAVY_IND	Dummy equal to 1 if the LLM is specialized in heavy-industry production [ISTAT].	0.08	0.274
URBAN	Dummy equal to 1 if the LLM is urbanized [ISTAT].	0.07	0.250

Table 6: Residence permits and distance from the gateways

	All sample	By source area:				
		Africa	South America	Asia	East Europe	Balkans
IMM $_{(n,t)}$	0.807***	0.929***	0.760***	0.752***	0.748***	0.946***
	(0.029)	*(0.244)*	*(0.077)*	*(0.164)*	*(0.024)*	*(0.100)*
DISTANCE $_{(i,n)}$	-0.329***	-0.412***	-0.152***	-0.137***	-0.084***	-0.502***
	(0.007)	*(0.017)*	*(0.017)*	*(0.021)*	*(0.013)*	*(0.022)*
PROVINCE $_{(i)}$	YES	YES	YES	YES	YES	YES
YEAR $_{(t)}$	YES	YES	YES	YES	YES	YES
COUNTRY $_{(n)}$	YES	YES	YES	YES	YES	YES
PROVINCE $_{(i)}$ x YEAR $_{(t)}$	YES	YES	YES	YES	YES	YES
Observations	20,703	4,814	4,211	4,139	4,004	3,535
R-squared	0.74	0.78	0.82	0.74	0.89	0.84

The dependent variable is IMM(i,n,t) and it is the log of the residence permits in province *i* of nationality *n* in year *t*; the key explanatory variables are IMM(n,t) that measures the log of the residence permits in Italy of nationality *n* in year *t*, and DISTANCE(i,n) that measures the log of kilometres between capital province *i* and the gateway used by immigrants of nationality *n* to enter Italy. The specification includes fixed effect at the province level (PROVINCE), year dummies (YEAR), dummies for the source country of immigration (COUNTRY) and interaction between PROVINCE and YEAR to capture any province-year effects. The sample contains the residence permits in the 103 Italian provinces distinguished by nationality, for the period 1996-2002. We consider only the 30 most important nationalities in terms of number of residence permits. Standard errors in parentheses; *, **, *** significantly different from zero at the 10, 5 and 1 percent level, respectively.

Table 7: Natives migration flows – OLS estimates

	All sample		Low-educated		High-educated	
IMM (d)	-0.189***	-0.224***	-0.088	-0.092	-0.372***	-0.440***
	(0.051)	*(0.052)*	*(0.064)*	*(0.065)*	*(0.100)*	*(0.102)*
IMM (s)	0.165***	0.183***	0.269***	0.282***	0.090	0.128
	(0.050)	*(0.050)*	*(0.069)*	*(0.067)*	*(0.102)*	*(0.101)*
UNR (d)	-0.448	-0.327	-0.293	-0.261	-0.283	-0.214
	(0.314)	*(0.323)*	*(0.466)*	*(0.479)*	*(0.508)*	*(0.525)*
UNR (s)	0.958***	0.890***	1.841***	1.774***	0.251	0.191
	(0.337)	*(0.341)*	*(0.477)*	*(0.472)*	*(0.548)*	*(0.550)*
GDP (d)	0.628*	0.141	0.384	0.278	0.431	-0.442
	(0.359)	*(0.385)*	*(0.512)*	*(0.569)*	*(0.723)*	*(0.748)*
GDP (s)	-1.023***	-0.775**	-2.127***	-1.972***	0.090	0.607
	(0.388)	*(0.380)*	*(0.612)*	*(0.544)*	*(0.770)*	*(0.760)*
EXPORT (d)		0.004***		0.001		0.007***
		(0.001)		*(0.002)*		*(0.003)*
EXPORT (s)		-0.002		-0.001		-0.004*
		(0.001)		*(0.002)*		*(0.002)*
HOUSE (d)		-0.362***		0.355**		-0.886***
		(0.138)		*(0.147)*		*(0.322)*
HOUSE (s)		0.201*		0.236		0.405*
		(0.112)		*(0.159)*		*(0.238)*
YEAR	YES	YES	YES	YES	YES	YES
Obs.	9,120	9,120	3,040	3,040	3,040	3,040

Panel with education and destination-source region pairs fixed effects (*ije*). The dependent variable is the natives' interregional migration flow. The explanatory variables, for both destination (d) and source (s) region, are: the incidence of immigrants (IMM), the unemployment rate (UNR), the GDP per worker (GDP), the propensity to export good with increasing international demand (EXPORT), the house prices (HOUSE). All the explanatory variables are lagged by one year. YEAR are time dummies (period 1996-2003). The table is divided in three panels: the first panel reports OLS estimates for all natives; the second and the third panels report OLS estimates for low- and high-educated natives, respectively. Clustered standard errors in parenthesis. *, **, *** significantly different from zero at the 10, 5 and 1 percent level, respectively.

Table 8: Natives migration flows – IV estimates using previous enclaves

	All sample		Low-educated		High-educated	
IMM (d)	0.243**	0.115	-0.495***	-0.438***	1.109***	0.869***
	(0.110)	*(0.113)*	*(0.155)*	*(0.159)*	*(0.236)*	*(0.241)*
IMM (s)	-0.313***	-0.264**	-0.204	-0.158	-0.249	-0.096
	(0.110)	*(0.113)*	*(0.155)*	*(0.159)*	*(0.236)*	*(0.241)*
UNR (d)	-0.274	-0.252	-0.191	-0.103	1.173**	0.919*
	(0.248)	*(0.247)*	*(0.434)*	*(0.440)*	*(0.507)*	*(0.494)*
UNR (s)	0.768***	0.796***	1.960***	1.974***	-0.083	-0.003
	(0.248)	*(0.247)*	*(0.434)*	*(0.440)*	*(0.507)*	*(0.494)*
GDP (d)	-0.010	-0.171	0.916**	0.525	-2.010***	-1.969***
	(0.317)	*(0.314)*	*(0.446)*	*(0.447)*	*(0.707)*	*(0.692)*
GDP (s)	-0.318	-0.366	-1.510***	-1.659***	0.649	0.868
	(0.317)	*(0.314)*	*(0.446)*	*(0.447)*	*(0.707)*	*(0.692)*
EXPORT (d)		0.002**		0.002*		0.002
		(0.001)		*(0.001)*		*(0.002)*
EXPORT (s)		-0.000		0.001		-0.003
		(0.001)		*(0.001)*		*(0.002)*
HOUSE (d)		-0.342***		0.336**		-0.764***
		(0.103)		*(0.151)*		*(0.223)*
HOUSE (s)		0.173*		0.211		0.384*
		(0.103)		*(0.151)*		*(0.223)*
YEAR	YES	YES	YES	YES	YES	YES
First stage F-statistics	571.5	547.7	209.0	202.4	206.7	192.7
Obs.	9,120	9,120	3,040	3,040	3,040	3,040

Panel with education and destination-source region pairs fixed effects (*ije*). The dependent variable is the natives' interregional migration flow. The explanatory variables, for both destination (d) and source (s) region, are: the incidence of immigrants (IMM), the unemployment rate (UNR), the GDP per worker (GDP), the propensity to export good with increasing international demand (EXPORT), the house prices (HOUSE). All the explanatory variables are lagged by one year. YEAR are time dummies (period 1996-2003). The table is divided in three panels: the first panel reports IV estimates for all natives; the second and the third panels report IV estimates for low- and high-educated natives. Instrumental variable is built using enclaves of previous immigrants. Clustered standard errors in parenthesis. *, **, *** significantly different from zero at the 10, 5 and 1 percent level, respectively.

Table 9: Natives migration flows – IV estimates using distance from the gateways

	All sample		Low-educated		High-educated	
IMM (d)	-0.035	-0.055	-0.907**	-0.773**	0.961*	0.857*
	(0.233)	*(0.220)*	*(0.398)*	*(0.368)*	*(0.504)*	*(0.473)*
IMM (s)	-0.222	-0.193	-0.775*	-0.686*	0.400	0.391
	(0.233)	*(0.220)*	*(0.398)*	*(0.368)*	*(0.504)*	*(0.473)*
UNR (d)	-0.378	-0.287	-0.088	0.050	1.028	0.909
	(0.257)	*(0.248)*	*(0.474)*	*(0.488)*	*(0.666)*	*(0.606)*
UNR (s)	0.810***	0.813***	2.103***	2.215***	0.556	0.418
	(0.257)	*(0.248)*	*(0.474)*	*(0.488)*	*(0.666)*	*(0.606)*
GDP (d)	0.399	-0.015	1.454**	0.764	-1.767*	-1.955**
	(0.435)	*(0.356)*	*(0.672)*	*(0.527)*	*(1.015)*	*(0.838)*
GDP (s)	-0.454	-0.432	-0.765	-1.283**	-0.421	0.301
	(0.435)	*(0.356)*	*(0.672)*	*(0.527)*	*(1.015)*	*(0.838)*
EXPORT (d)		0.003**		0.004*		0.002
		(0.001)		*(0.002)*		*(0.002)*
EXPORT (s)		-0.000		0.003		-0.005**
		(0.001)		*(0.002)*		*(0.002)*
HOUSE (d)		-0.352***		0.316**		-0.765***
		(0.103)		*(0.161)*		*(0.226)*
HOUSE (s)		0.178*		0.181		0.429*
		(0.103)		*(0.161)*		*(0.226)*
YEAR	YES	YES	YES	YES	YES	YES
First stage F-statistics	112.9	130.8	32.2	37.7	39.4	45.1
Obs.	9,120	9,120	3,040	3,040	3,040	3,040

Panel with education and destination-source region pairs fixed effects (*ije*). The dependent variable is the natives' interregional migration flow. The explanatory variables, for both destination (d) and source (s) region, are: the incidence of immigrants (IMM), the unemployment rate (UNR), the GDP per worker (GDP), the propensity to export good with increasing international demand (EXPORT), the house prices (HOUSE). All the explanatory variables are lagged by one year. YEAR are time dummies (period 1996-2003). The table is divided in three panels: the first panel reports IV estimates for all natives; the second and the third panels report IV estimates for low- and high-educated natives. Instrumental variable is built using distance between local labour market and gateways through which immigrants enter Italy. Clustered standard errors in parenthesis. *, **, *** significantly different from zero at the 10, 5 and 1 percent level, respectively.

Table 10: Native population growth in LLMs – OLS estimates

	All sample		Low-educated		High-educated	
IMM_GR	0.001***	0.001***	0.001**	0.001	-0.001	0.001
	(0.001)	*(0.000)*	*(0.001)*	*(0.001)*	*(0.006)*	*(0.006)*
Demographic controls:						
LNPOP	-0.002	-0.001	-0.005**	-0.002	-0.000	0.016
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.023)*	*(0.025)*
POP_GR_1	0.573***	0.473***				
	(0.047)	*(0.047)*				
LOWEDUC			0.065	0.140***		
			(0.046)	*(0.050)*		
LOWEDUC_GR_1			0.359***	0.298***		
			(0.047)	*(0.048)*		
HIGHEDUC					-19.731***	-21.405***
					(1.693)	*(2.164)*
HIGHEDUC_GR_1					-0.929***	-0.938***
					(0.139)	*(0.134)*
SHARE_65+	-0.241***	-0.250***	-0.345***	-0.344***	-1.825***	-1.121
	(0.054)	*(0.052)*	*(0.062)*	*(0.059)*	*(0.639)*	*(0.722)*
DENSITY	0.005**	0.005**	0.008***	0.009***	0.050**	0.030
	(0.003)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.022)*	*(0.021)*
Economic controls:						
UNR		-0.079***		-0.124***		0.503**
		(0.028)		*(0.032)*		*(0.219)*
SER_EMP_GR		0.108***		0.067***		0.707***
		(0.019)		*(0.017)*		*(0.153)*
LLM TYPE	-	YES	-	YES	-	YES
AREA fixed effects	YES	YES	YES	YES	YES	YES
Observations	686	686	686	686	686	686

OLS estimates on the determinants of native population growth, distinguishing overall, low- and high-educated growth. The explanatory variables are: immigrants population growth (IMM_GR), log of population (LNPOP), population growth in the previous period (POP_GR_1 and, for low- and high-educated, LOWEDUC_GR_1 and HIGHEDUC_GR_1), share of low- and high-educated (LOWEDUC and HIGHEDUC) share of older people (SHARE_65+), density (DENSITY), unemployment rate (UNR), economic features of the LLMs (NON_MANIF, MADE_ITALY, HEAVY_IND, NO_SPEC and URBAN) and employment growth in service sector (SER_EMP_GR). Area dummies are Centre-North and South. Robust standard errors, to control for heteroschedasticity, in parenthesis. *, **, *** significantly different from zero at the 10, 5 and 1 percent level, respectively.

Table 11: Native population growth in LLMs – IV estimates using enclaves

	All sample		Low-educated		High-educated	
IMM_GR	0.006	0.005	0.008	0.008	0.009	0.001
	(0.005)	*(0.004)*	*(0.005)*	*(0.005)*	*(0.045)*	*(0.041)*
Demographic controls:						
LNPOP	-0.002	-0.001	-0.005*	-0.002	-0.001	0.016
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.023)*	*(0.025)*
POP_GR_1	0.568***	0.473***				
	(0.047)	*(0.047)*				
LOWEDUC			0.062	0.152***		
			(0.050)	*(0.055)*		
LOWEDUC_GR_1			0.341***	0.286***		
			(0.049)	*(0.049)*		
HIGHEDUC					-19.599***	-21.406***
					(1.774)	*(2.141)*
HIGHEDUC_GR_1					-0.931***	-0.938***
					(0.138)	*(0.135)*
SHARE_65+	-0.238***	-0.238***	-0.349***	-0.326***	-1.820***	-1.119
	(0.056)	*(0.057)*	*(0.066)*	*(0.065)*	*(0.640)*	*(0.739)*
DENSITY	0.004*	0.005**	0.006**	0.008***	0.048**	0.030
	(0.003)	*(0.002)*	*(0.003)*	*(0.003)*	*(0.023)*	*(0.021)*
Economic controls:				-0.108***		0.504**
UNR		-0.071**		(0.034)		(0.219)
		(0.029)		*0.070***		*0.707****
SER_EMP_GR		0.109***		(0.019)		(0.153)
		(0.019)		*-0.108****		*0.504***
LLM TYPE	-	YES	-	YES	-	YES
AREA fixed effects	YES	YES	YES	YES	YES	YES
First stage F-statistics	5.1	4.4	5.4	4.6	5.4	4.2
Observations	686	686	686	686	686	686

IV estimates on the determinants of native population growth, distinguishing overall, low- and high-educated growth. The explanatory variables are: immigrants population growth (IMM_GR), log of population (LNPOP), population growth in the previous period (POP_GR_1 and, for low- and high-educated, LOWEDUC_GR_1 and HIGHEDUC_GR_1), share of low- and high-educated (LOWEDUC and HIGHEDUC) share of older people (SHARE_65+), density (DENSITY), unemployment rate (UNR), economic features of the LLMs (NON_MANIF, MADE_ITALY, HEAVY_IND, NO_SPEC and URBAN) and employment growth in service sector (SER_EMP_GR). Area dummies are Centre-North and South. Instrumental variable is built using enclaves of previous immigrants. Robust standard errors, to control for heteroschedasticity, in parenthesis. *, **, *** significantly different from zero at the 10, 5 and 1 percent level, respectively.

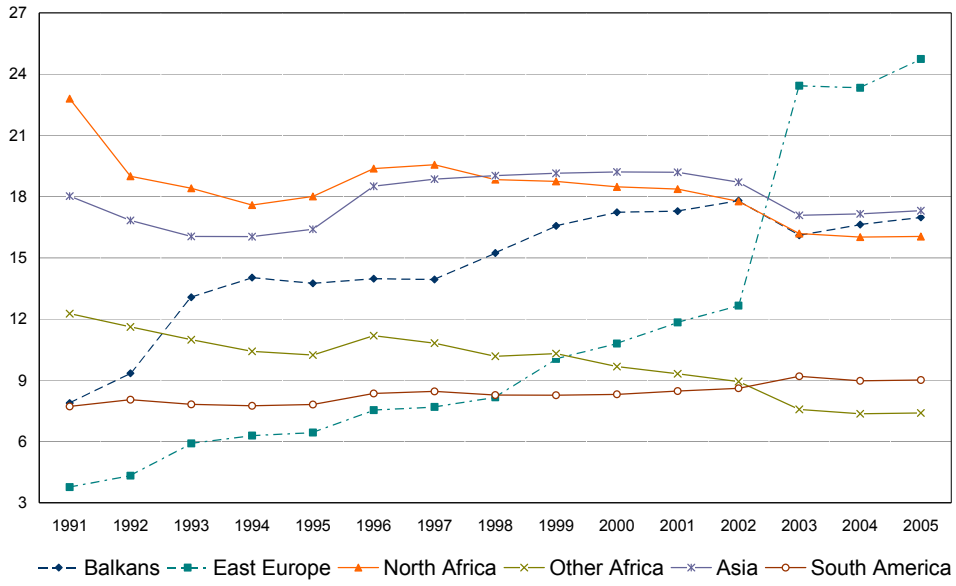
Table 12: Native population growth in LLMs – IV estimates using distance

	All sample		Low-educated		High-educated	
IMM_GR	-0.001	-0.008**	-0.006**	-0.016***	0.062***	0.092***
	(0.002)	*(0.003)*	*(0.003)*	*(0.005)*	*(0.022)*	*(0.032)*
Demographic controls:						
LNPOP	-0.002	-0.001	-0.005*	-0.003	-0.004	0.019
	(0.002)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.024)*	*(0.029)*
POP_GR_1	0.576***	0.472***				
	(0.047)	*(0.053)*				
LOWEDUC			0.067	0.110		
			(0.048)	*(0.067)*		
LOWEDUC_GR_1			0.377***	0.329***		
			(0.052)	*(0.064)*		
HIGHEDUC					-18.880***	-21.587***
					(1.769)	*(2.361)*
HIGHEDUC_GR_1					-0.941***	-0.911***
					(0.142)	*(0.136)*
SHARE_65+	-0.243***	-0.284***	-0.341***	-0.388***	-1.792***	-0.738
	(0.054)	*(0.058)*	*(0.066)*	*(0.078)*	*(0.659)*	*(0.782)*
DENSITY	0.006**	0.006**	0.009***	0.011***	0.036	0.018
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*	*(0.023)*	*(0.024)*
Economic controls:						
UNR		-0.101***		-0.163***		0.731***
		(0.031)		*(0.041)*		*(0.263)*
SER_EMP_GR		0.105***		0.060**		0.726***
		(0.021)		*(0.024)*		*(0.181)*
LLM TYPE	-	YES	-	YES	-	YES
AREA fixed effects	YES	YES	YES	YES	YES	YES
First stage F-statistics	19.0	13.6	20.2	13.8	19.8	13.7
Observations	686	686	686	686	686	686

IV estimates on the determinants of native population growth, distinguishing overall, low- and high-educated growth. The explanatory variables are: immigrants population growth (IMM_GR), log of population (LNPOP), population growth in the previous period (POP_GR_1 and, for low- and high-educated, LOWEDUC_GR_1 and HIGHEDUC_GR_1), share of low- and high-educated (LOWEDUC and HIGHEDUC) share of older people (SHARE_65+), density (DENSITY), unemployment rate (UNR), economic features of the LLMs (NON_MANIF, MADE_ITALY, HEAVY_IND, NO_SPEC and URBAN) and employment growth in service sector (SER_EMP_GR). Area dummies are Centre-North and South. Instrumental variable is built using distance between LLMs and gateways trough which immigrants enter Italy. Instrumental variable is built using enclaves of previous immigrants. Robust standard errors, to control for heteroschedasticity, in parenthesis. *, **, *** significantly different from zero at the 10, 5 and 1 percent level, respectively.

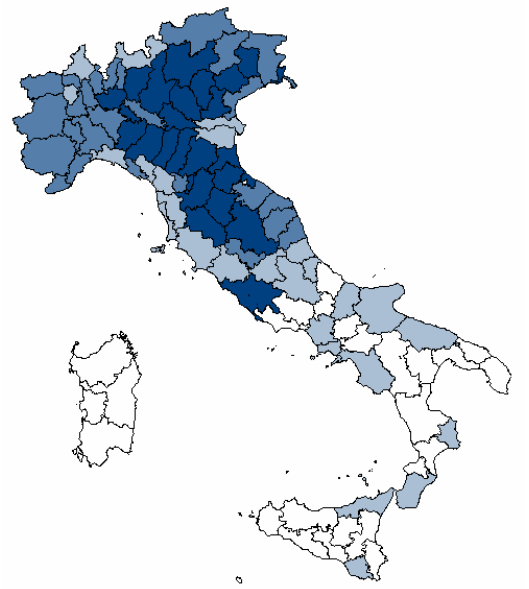
Appendix – Figures

Figure 1: Composition of immigrants by area of origin



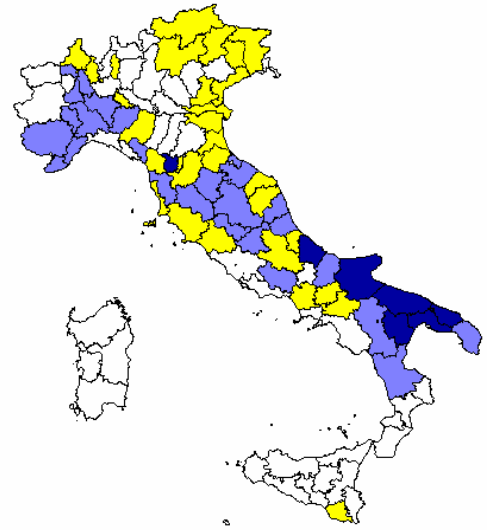
Source: Authors' elabouration on data drawn from Istat.

Figure 2: Distribution of immigrants across provinces

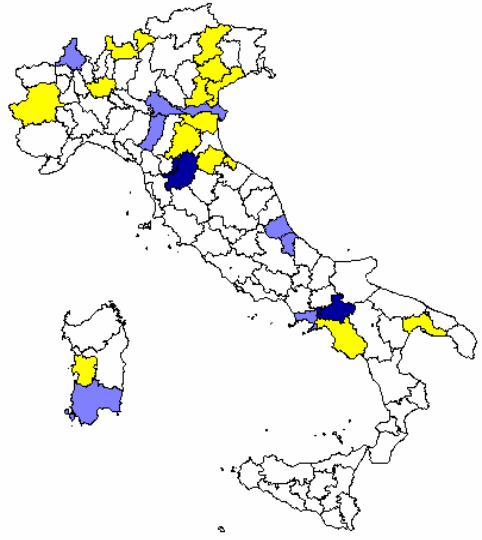


Italian provinces are divided in quartiles according to the percentage of immigrants (year 2006). Those with a darker blue have higher incidence of immigrants, and viceversa.
Source: Authors' elabouration on data drawn from Istat.

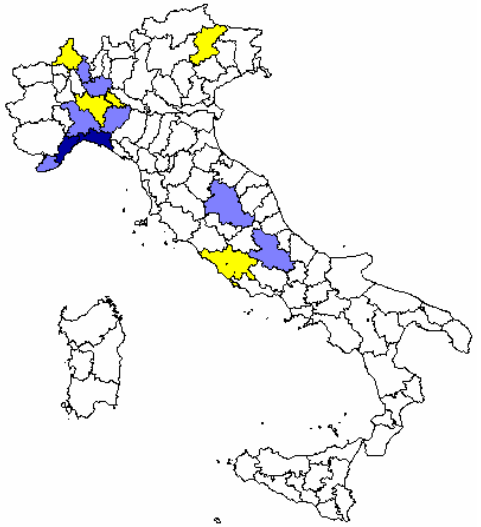
Figure 3: Distribution of immigrant across Italian provinces



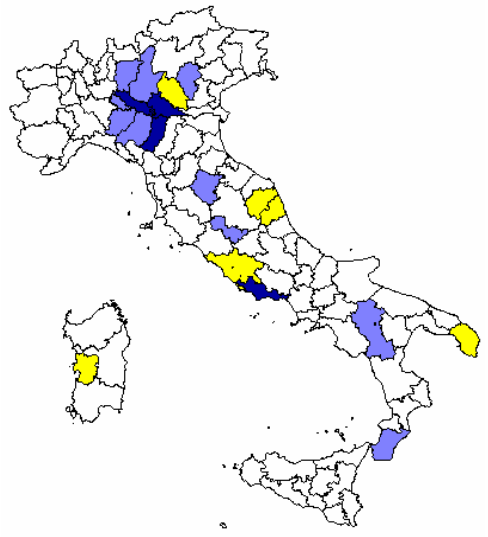
Albania



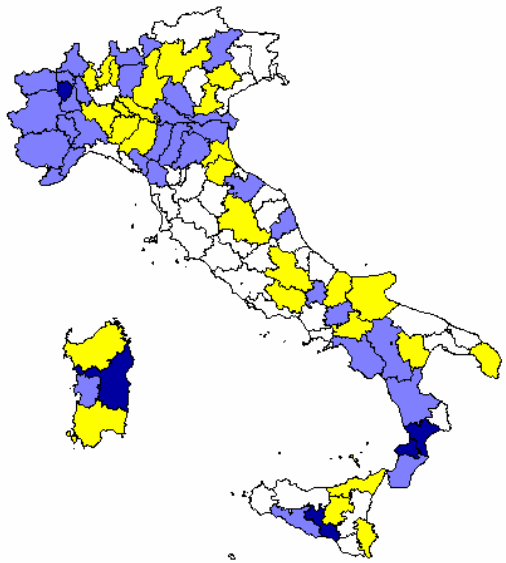
China



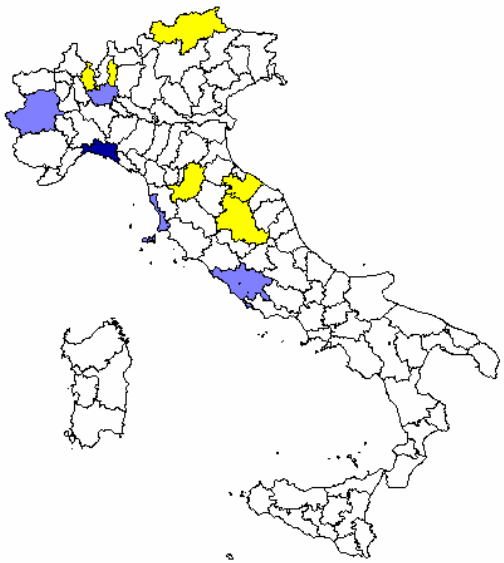
Ecuador



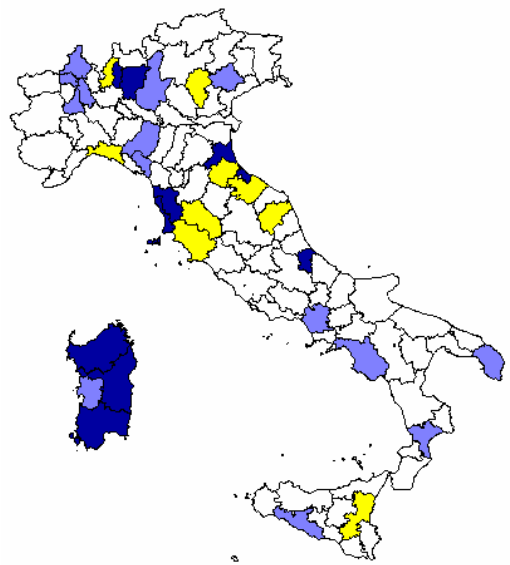
India



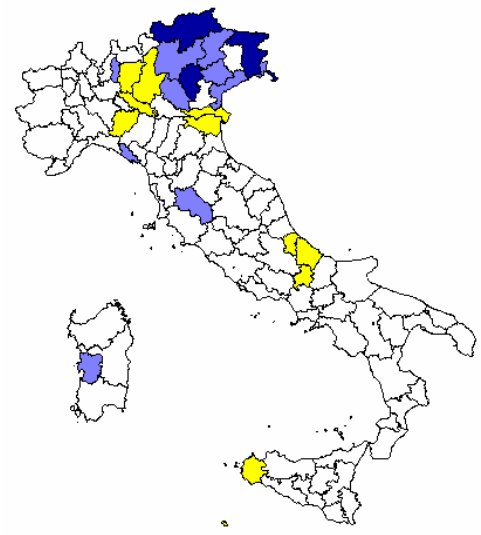
Morocco



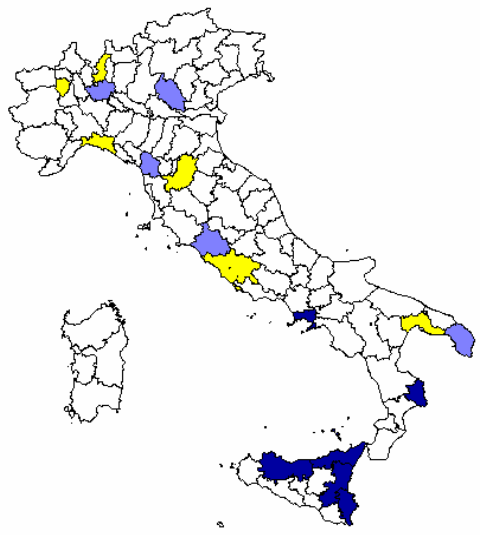
Peru



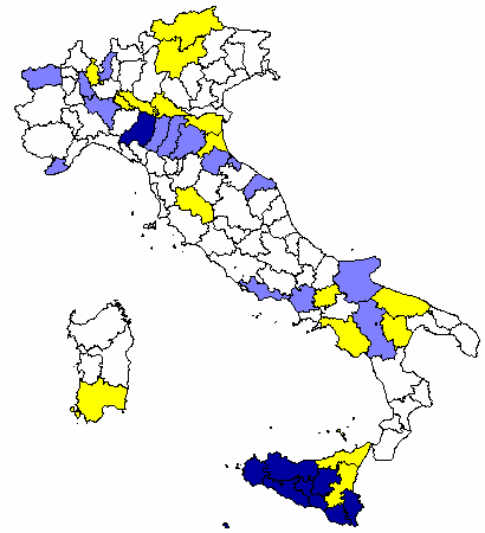
Senegal



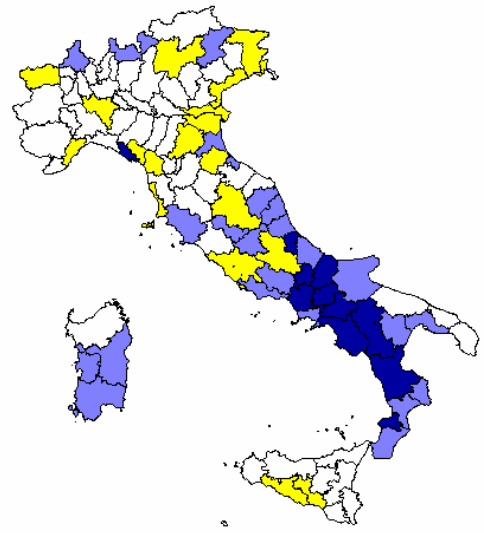
Serbia



Sri Lanka



Tunisia



Ukraine

$I \leq 1$

$1 < I \leq 1.5$

$1.5 < I \leq 3$

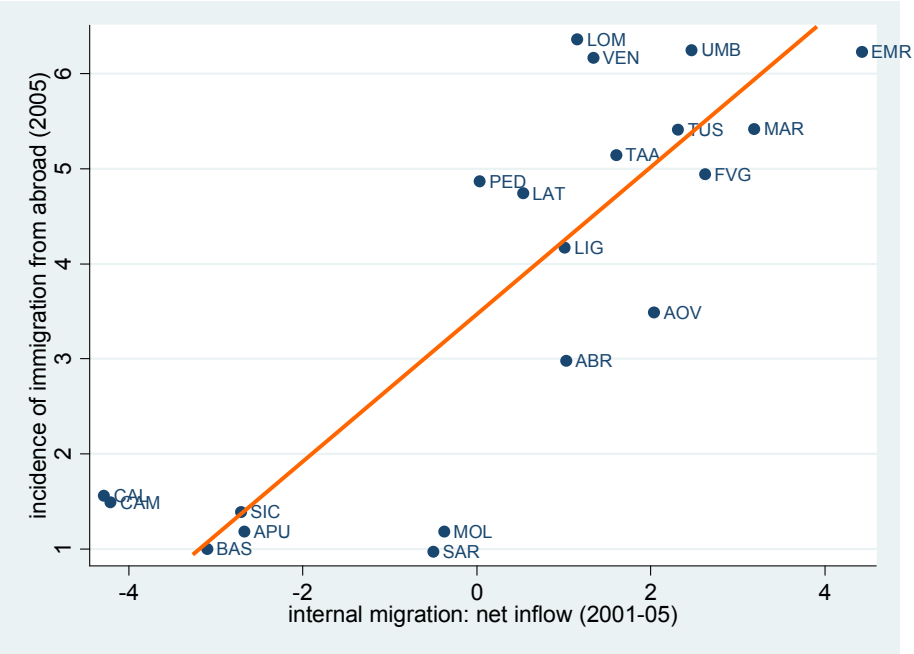
$I > 3$

Where the geographical concentration index *I* for each province *i* and each nationality *n* is built as follows:

$$I_{in} = \frac{IMM_{in} \Big/ \sum_i IMM_{in}}{\sum_n IMM_{in} \Big/ \sum_i \sum_n IMM_{in}}$$

Source: Residence permits provided by Ministry of Interior (year 2002).

Figure 4: Immigrant inflow and native internal mobility



Source: authors' elaborations from ISTAT.

Migrants' trajectories and gateways to Italy

The scope of this box is twofold. First, we identify the main gateways to which immigrants enter Italy. Second, we associate to each gateway one (or more) source country of immigration, replicating the migration trajectories.

Italy is geographically located at the crossroads of the main international migration flows that have interested Europe in the last decades. We identify the following main gateways: the northern-east border (in particular, Friuli Venezia-Giulia), the northern-west border (Piedmont and Liguria), the two most important international airports Rome and Milan, and the coastline in the southern regions (especially Apulia and Sicily).

The prominence of these gateways is confirmed by the results of a survey conducted among immigrants living in Lombardy and by data of the Ministry of Interior on illegal immigration at the frontiers.

Gateways to Italy

	Self-declared gateway		Refused entry at the border	
Piedmont	3,8	(6)	12,9	(4)
Aosta Valley	0,7	(13)	0,7	(13)
Lombardy	41,9	(1)	34,2	(1)
Trentino Alto-Adige	2,1	(8)	0,8	(10)
Veneto	1,8	(10)	1,9	(6)
Friuli Venezia-Giulia	3,2	(7)	25,5	(2)
Liguria	10,8	(3)	0,8	(11)
Emilia-Romagna	0,9	(12)	1,1	(8)
Tuscany	0,5	(15)	0,3	(14)
Umbria	0,4	(17)	0,0	(18)
Marche	0,6	(14)	1,0	(9)
Latium	13,5	(2)	13,4	(3)
Abruzzi	0,2	(19)	0,2	(17)
Molise	0,1	(20)	0,0	(18)
Campania	1,5	(11)	1,8	(7)
Apulia	8,8	(4)	4,4	(5)
Basilicata	0,3	(18)	0,0	(18)
Calabria	2,1	(8)	0,2	(15)
Sicily	6,4	(5)	0,8	(12)
Sardinia	0,5	(15)	0,2	(16)

Self-declarations of the gateways used to enter Italy are drawn from a survey conducted among about 8,000 immigrants living in Lombardy (ISMU, 2002). Data on refused entry at the borders – foreigners turned away at the border due to their non-fulfilment of the requirements prescribed by the law – are drawn from the Ministry of Interior (2005). For each region, we report the percentages (and rank positions in parenthesis).

The information drawn from ISMU should be interpreted with some caution since they refer to a selected sample of immigrants, those living in Lombardy. Therefore there is plausibly a "northern bias" in the declaration of the region of entry. However, two main evidences arise. First, immigrants show a high geographical mobility across regions. Only 40 percent of immigrants living in Lombardy arrived directly in this region whereas about one out of five comes from the southern regions. Second, apart the expected result from Lombardy, the other main gateways result Latium, Liguria, Apulia and Sicily. The northern-east border appears relatively less important whereas the role of other regions is almost negligible.

The data by the Ministry of Interior refer to illegal immigration intercepted at the border. They enable use to have a clearer geographical pattern of immigration inflows. The main gateways are Lombardy, Friuli Venezia-Giulia, Latium and Piedmont. More precisely, in 2005 the terrestrial frontiers with the highest number of refusal of immigrants were Malpensa and Fiumicino – the airports of Milan and Rome, respectively – and the frontier of Trieste (in the North-East) and Verbania-Domodossola (in Piedmont). In the same year, the CPTs (centres of temporary detection) with the larger number of immigrants were in Rome and in Lampedusa (in the Sicily Canal).

The next step consists in assigning to each gateway one (or more) source country of immigration, depending on the typical migration trajectories used by immigrants to enter Italy. These assignments are based on several factors. All the countries close to the Italian border are assigned to the gateways in terms of geographical proximity: the Albanians are assumed to enter from Apulia (through the Otranto Canal); immigrants from the Balkans and from other East-Europe countries are assumed to enter from the Italian-Slovenian border (Trieste); Tunisians are assumed to enter from West-Sicily. For countries that are far away Italians border, we use other sources in order to identify the main migration trajectories. We refer to the survey by ISMU in which immigrants self-declare their region of entry, and to field studies (Monzini *et al.*, 2004; European Migration Network, 2005) that analyzed international migration flows.[21] The available evidence supports the

[21] ISMU data have been used to build a specialization index in order to identify the main gateway used to enter in Italy. For each source country-region pairs we calculated the ratio between the percentage of immigrants coming from country *n* entering in region *i* and the corresponding percentage calculated for all immigrants (independently from country of origin).

following assumptions. The northern-west border is crossed by immigrants coming from the Maghreb and the Latin America. Foreigners from the Maghreb cross the Strait of Gibraltar that represents the first gateway to other countries of the EU, and then continue by land to Ventimiglia (in Liguria). On the other side, Portugal and Spain are the preferred destinations for migrants from the Latin America, and Liguria is the closer region to the Iberian Peninsula both by land (again Ventimiglia) and by sea (Genoa). Milan and Rome are the first destination for all immigrants coming from more distant countries (e.g., South America and Asia) due to the presence of the two main international airports. The Mediterranean coastline (Sicily, Calabria and Apulia) is the first destination of immigrants coming from North Africa but also of those coming from the Indian subcontinent (through the Suez Canal). Finally, in some cases, the relationship between a gateway and a source country of immigration is dictated by other "exogenous" reasons.[22]

[22] To give a couple of examples: the historical linkage between the seaport of Naples and Odessa (in Ukraine), and the preferences of Polish for Rome probably due to religious motives.