The Impact of Immigration on the Internal mobility of Natives and Foreign-born Residents: Evidence from Italy

Roberto Basile* Luca De Benedictis† María Durban‡ Alessandra Faggian§
Román Mínguez¶

June 5, 2018

Abstract

In this paper we investigate the relationship between immigration flows and internal mobility in Italy during the period 2003-2011. Using semiparametric negative binomial gravity models with smooth spatio-temporal trends, and dealing with endogeneity issues through a control function approach, we provide evidence of a significant negative (or displacement) effect of new foreign immigrants on the internal mobility of foreign-born residents and of Italian citizens with a low education level, as well as a significant positive (or complementarity) effect of new foreign immigrants on the internal mobility of Italian citizens with a high education level.

Keywords: Immigration, Internal mobility, Gravity models.
Jel codes: F22, J61, R23, C14, C21

*Department of Industrial and Information Engineering and Economics, University of L'Aquila, Italy, and Luiss. roberto.basile@univaq.it

†Department of Economics and Law, University of Macerata, Rossi-Doria Centre, University Roma Tre, and Luiss, Italy. luca.debenedictis@unimc.it

‡Statistics Department, University Carlos III. Madrid, Spain. mdurban@est-econ.uc3m.es

§Gran Sasso Science Institute, L'Aquila, Italy. alessandra.faggian@gssi.it

¶Department of Statistics, University of Castilla-La Mancha, Cuenca, Spain. roman.minguez@uclm.es

1 Introduction

During the last decades, Western European countries have become a major destination of international migrants (Münz, 2007), and simulations indicate that the phenomenon will characterize the decades to come (Hanson and McIntosh, 2016; Docquier and Machado, 2017). This evidence has raised an overwhelming debate on the effects of immigrants on domestic labor markets. One of the main results of the current literature is that these effects go beyond their impact on local wages and local unemployment, and could mainly involve the location decision of internal movers (both natives and foreign-born residents). In other words, the adjustment mechanism to immigrant-induced labor supply shocks takes place mainly through interregional migration rather than through changes in local wages (Peri, 2014) and in local unemployment rates (Constant, 2014).[1]

Various studies have indeed analyzed the link between immigration and internal mobility to test whether internal movers and immigrants are complements or substitutes (Filer, 1992; Frey, 1995; Card and DiNardo, 2000; Card, 2001; Kritz and Gurak, 2001; Borjas, 2006; Hatton and Tani, 2005). Overall, the empirical evidence is mixed and a number of research questions remain open. For the case of Italy, Mocetti and Porello (2010) find that immigration is positively associated with inflows of highly-educated natives (thus indicating a *complementarity* effect), and negatively associated with inflows of low-educated natives (thus indicating a *displacement* effect). However, this study - as it is the general case in the literature - does not consider the *effect of foreign immigrants on internal movements of previous immigrant cohorts*, a phenomenon gaining momentum in recent years. Indeed, Figure 1 shows that, during the period 2002-2011, the internal movement of foreign-born residents steeply increased passing from about 115,000 to about 230,000 movers, while the internal movement of Italian citizens slightly decreased from 340,000 to 310,000 movers. This evidence suggests that a growing number of foreign-born

[1]For example, Basile, Girardi, Mantuano, and Russo (2018) find no significant effect of foreign immigrants on the regional unemployment rate in Italy. Thus, a key issue is whether the effects of immigration on unemployment (and on wages) may be masked by interregional labor mobility.

workers immigrated in Italy over the most recent period is taking into account the possibility of moving along the regional gradient existing within the Country in terms of job opportunity differentials, income differentials and spatial differences in the quality of life. Thus, the proper analysis of the labor market effect of immigrants requires to consider how the location decision of these foreign-born internal movers is affected by the new immigration of foreign workers. In particular, since recent and earlier immigrants are likely "perceived" to have similar (low) skills by local employers (because of the lack of "signals"), we expect to find that the location decision of foreign-born internal movers is negatively affected by the new immigration of foreign workers (see next section for a deeper discussion on this point).

Insert Figure 1 about here.

In order to address this issue, we estimate gravity models of internal migration across Italian provinces and test the effect of new foreign immigrants on internal movements of both natives with different education levels (primary school, lower-secondary school, upper-secondary school, and higher education level) and foreign-born residents in Italy by using official data on internal mobility at the province level (NUTS-3 level of territorial aggregation) collected by ISTAT (the Italian National Institute of Statistics) for the period 2002-2011. Only from 2002, indeed, ISTAT provides reliable data on the education level of internal mobility of Italian citizens.

Specifically, within the framework of an extended gravity model, we propose a semiparametric negative binomial model for count data allowing us to control for the unobserved heterogeneity through the inclusion of a smooth spatio-temporal trend both for origin and destination regions. Finally, following Beine and Coulombe (2018), we take seriously into account the problem of endogeneity of foreign immigration rates, by using a control function approach based on the exploitation of the information on the country of origin of the immigrants.

The estimation results of gravity models, including the relative incidence of foreign immigration at destination and origin provinces, show that foreign immigrants have a net displacement effect on the internal movement of low-skilled Italian citizens (those with a primary and a lower-secondary school level of education), and a positive effect on the internal movement of high-skilled natives (those with an upper secondary school level and a higher education level). These preliminary results broadly confirm those, at the regional (NUTS-2) level, reported in Mocetti and Porello (2010). We also show that, beyond and above the classical displacement effect on low-skill natives, the inflow of new waves of immigrants generates a significant negative net effect on the internal movement of foreign-born workers.

The estimation of gravity models including separately the incidence of foreign immigrants at the province of destination of internal movers (as a pull factor) and at the province of origin of internal movers (as a push factor) provides further relevant information on the displacement and on the complementarity effects of foreigners. Starting from the displacement effects, both a push and a pull effect on the internal mobility of natives with a primary school level are in action, and the same appears to be at work in the case of foreign-born internal mobility, while foreign immigration acts only as a push factor for the internal mobility of lower-secondary school level workers. As for the complementarity impact of foreign immigrants, both a push and a pull effect on the internal mobility of natives with a higher level of education are evident. Moreover, the results reveal a significant pull effect of foreign immigrants on the internal mobility of upper-secondary school natives. These results are robust to issues of endogeneity assessed through a control function approach.

These empirical results add to the existing literature in several ways. First, as already mentioned, Mocetti and Porello (2010) only consider the effect of immigrants on the internal movement of natives, while we also quantify the effect on the internal movement of foreign-born workers. Second, we consider the finer NUTS-3 level measuring with higher definition the local effect of immigrants' inflow. Third, their sample period is 1995-2005, while our data cover the years from 2002 and 2011. As mentioned, ISTAT considers the information on the education level of internal movers as reliable only from 2002 onward. Forth, we use count data models in consideration of the very high overdispertion in the distribution of dyadic internal movements, and propose a semiparametric approach with the inclusion of a spatio-temporal trend to control for unobserved heterogeneity, while the literature on gravity count data models lacks the consideration of this control. Finally, as far as the choice of the instrumental variables in the control function approach, we exploit information from the country of origin of the immigrants, while Mocetti and Porello (2010) use the distance from the gateways.

The remainder of the paper is organized as follows. Section 2 provides a conceptual framework on the relationship between immigration and the location choice of both natives and foreign-born residents focusing on the role of formal education as signal of labor skills. Section 3 describes the data and provides some descriptive analysis of migration flows. Section 4 describes the gravity model of migration used. Section 5 presents the econometric results. Section 6 concludes.

2 Are international migrants and internal movers complements or substitutes? The role of formal education as a signal of labor skills

One of the main concerns with international immigration is its effects on local labor markets. What happens when there is a large inflow of immigrants in a city? Since the 1990s empirical evidence from the USA (Filer, 1992; Frey, 1995), Canada (Ley and Tutchener, 2001) and Australia (Sheehan, 1998) shows that when a large number of immigrants move into an area, a large number of natives decide to leave. However, the most recent literature (Card and DiNardo, 2000; Borjas, 2006; Mocetti and Porello, 2010) suggests that the answer to this question depends on the level of substitutability or complementarity between foreign immigrants and natives.

On the one hand, inflows of immigrants into a particular area might lead, because of competition (substitution), to outflows of natives to other locations within the country and/or to a reduction of inflows of natives from other places within the country. Thus, the main effect of foreign immigrants would materialize in a decrease in internal movements of workers to the provinces that receive immigrants (*crowding out* or *displacement* effect). This result is more likely to occur if immigrants have the same skills of internal movers (immigrants and internal movers with low skills may compete directly for jobs). This "skating-rink" model (Card and DiNardo, 2000) may help explain the lack of evidence of a negative effect of immigration on employment and on wages: as immigrants move into a local economy, native workers with similar skills move out, leaving total employment and wages unchanged. Immigrants might still displace native workers by pushing them out of the market, but the wage and the employment (unemployment) effect would not be detectable in the local economy.[2]

On the other hand, inflow of immigrants from abroad and internal native movers may complement each other if their skills are different. In particular, when immigrants fill lower-skill manual-intensive positions, native workers are able to specialize in more cognitive-intensive production tasks. In Europe, immigration over the last decades is indeed associated with job creation and employment upgrading into higher-skills, better paying jobs for native workers. In these cases, inflows of immigrants into a particular area, rather than displacing native workers, might lead to a reduction of outflows of high-skilled natives to other locations within the country. Moreover, if qualified native workers and cognitive-intensive production jobs do not match locally, the effect of foreign immigrants would materialize in an increase in internal movements of skilled workers to the provinces that receive immigrants (*complementarity* effect). "But this complementarity may consign generations of immigrant workers to low-skill, low-paying employment, especially in hierarchically structured labor markets. By reinforcing perceptions that some occupations are immigrant jobs and solidifying stereotypes, this pattern of employment and mobility reduces social cohesion and integration and may prevent immigrants from investing in education and entering higher-skilled occupations" (Constant, 2014). In turn, this pattern

[2] An alternative hypotheses which grew in popularity since the 1990s is that of "cultural avoidance" (Filer, 1992; Frey, 1994, 1996, 1999; Faggian, Partridge, Rickman, et al., 2012). Simply put, natives might be more reluctant to live in areas with a high concentration of immigrants, especially if very different from themselves in terms of individual characteristics, such as ethnicity and religion. Whatever the reasons behind the displacement of natives, one thing is clear: an increase in the internal mobility rates of natives creates a ripple effect to other regions, allowing the effects of international immigration in a particular area to spread nationally (and sometimes even beyond) in a "bathtub" model fashion (Ali, Partridge, and Rickman, 2012; Borjas, 2006).

mainly exacerbates the competition between different cohorts of foreign immigrants.

A crucial issue in this debate is how to measure labor skills of migrants. Although it is widely accepted that skills predict individuals' employbility (e.g., Heckman, Stixrud, and Urzua 2006), it is recognized that employers cannot directly observe the skills of job applicants. Thus, individuals make costly investments to signal skills to potential employers. Formal education is one of these investments. Numerous studies, for example, show that more schooling and higher educational degrees lead to more success on the labor market (Card 1999; Heckman, Lochner, Todd 2006; Tyler, Murnane, and Willett 2000; Clark and Martorell 2014).

Although non-cognitive skills (such as social skills and experience or maturity) may also play a key role in determining the labor market performance of job applicants, we may broadly assume that firms mainly use educational credentials as "*signal*" for skill levels when they allocate native and foreign immigrant workers between manual-intensive and cognitive-intensive positions. It seems indeed intuitive that cognitive skill signals (such as formal education) are more important for labor-market entrants (migrants) than for workers with substantial experience. Based on this considerations, we can make our **first prediction**:

> *In the lack of cognitive-skill signals (for example when there is not equipollence between formal educational attainments of new foreign immigrants and those of natives), firms classify new foreign immigrant as low-skill workers, and allocate them in manual-intensive positions. These workers are therefore considered as substitute of (observed) low-skill natives and complement of high skill natives*

However "this complementarity may consign generations of immigrant workers to low-skill, low-paying employment, especially in hierarchically structured labor markets. By reinforcing perceptions that some occupations are immigrant jobs and solidifying stereotypes, this pattern of employment and mobility reduces social cohesion and integration and may prevent immigrants from investing in education and entering higher-skilled occupations" (Constant, 2014). These further considerations leave us to our **second prediction**:

Immigrants persist in this "poverty trap" even in the subsequent years after their immigration. And thus, when foreign-born movers move within the country, they are considered as substitutes of new foreign immigrants.

These two predictions are based on the hypothesis of asymmetric information about the educational attainment of both new foreign migrants and foreign-born migrants. Obviously, if these predictions are confirmed, matching foreign worker skills with the right vacancies becomes problematic. Moreover, inefficient foreign worker allocation could be long lasting, resulting in lower overall productivity.

3 Internal mobility and immigration in Italy

3.1 Data source

We use official data at province level (NUTS-3 level of territorial aggregation) on the internal mobility of Italian citizens and of foreign-born residents, as well as on the immigration from abroad of foreign people, collected through the "*Indagine sui trasferimenti di residenza*", which is a survey carried out by ISTAT (the Italian Institute of Statistics). In keeping with the methodological standards set by the EU Regulation 862/2007, ISTAT has revised the entire data set from 1995 onward.

Regarding the information on the level of schooling of native internal movers, the series are given as reliable by ISTAT only from 2002 onward. Thus, for the sample period 2002-2011 used in this paper, ISTAT provides reliable information on the number of Italian citizens canceled from the municipality of a province and registered to the municipality of another province disaggregated by age class and by level of schooling: primary school, lower-secondary school, upper-secondary school, and higher (i.e. *Laurea* corresponding to the Bachelor's degree, *Laurea magistrale* corresponding to the Master's degree, and *Dottorato di ricerca* corresponding to the PhD). Then, for each level of education, we select movers in the age class 15-64 years (25-64 in the case of higher-education level), i.e. movers belonging to the working-age population.

For the case of foreign-born residents and of immigration from abroad of foreign people, there is no information on the educational attainment of the migrants. Thus, foreign migrants are considered in the analysis as a unique category.

3.2 Overall trends

Interregional migration flows of natives in Italy reduced from the mid-1970s to the mid-1990s. Faini, Galli, Gennari, and Rossi (1997) show that this happened because of several socio-economic factors, like expectations of North-South wage convergence (in line with the "option value of waiting" approach sketched by Burda, 1995), large-scale job creation in the public sector, transaction costs due to mobility and job-matching failures (Alesina, Danninger, and Rostagno, 2001; Attanasio and Schioppa, 1991). The 1992 crisis also caused the fiscal consolidation required to join the Euro area and the end of the "*intervento straordinario*" (extraordinary intervention), namely a special program of public transfers to Southern regions. These factors have stimulated a renewal of migration flows (Basile and Causi, 2007). Starting from the mid-90s, indeed, interregional migration flows of natives have dramatically increased, especially from the South to the North. However, differently from the 50s and the 60s, when long-distance flows were mainly movements of low-skilled workers, more recent flows are characterized by a strong component of human capital (in terms of schooling), involving a large number of workers with a tertiary education (Piras, 2012a,b, 2017; Etzo, 2011).

During the sample period 2003-2011, ISTAT registered about three millions of internal movements of Italian citizens in the working age class (15-64). About 8% have a primary school level of education, 35% a lower-secondary school level, 40% an upper-secondary school level, and 18% a tertiary (or higher) level. However, while the number of movers with a tertiary education registered an upward trend, all other categories showed a slightly downward trend (Figure 2). Moreover, once we normalize by the population stock with the same educational level, it turns out that the internal migration rate of Italian citizens with a tertiary or higher education level is higher than the migration rate of the other categories over the whole sample period (Figure 3). This evidence is in line with the broad migration literature according to which internal mobility in developed countries is mostly movement of high-skilled workers (see, for instance, Antolín and Bover, 1993; Burda and Wyplosz, 1994; Greenwood, 1997; Hughes and McCormick, 1994).

Insert Figure 2 and Figure 3 about here.

The recent upsurge of internal mobility of natives described above has been accompanied by an increase of immigration from abroad of foreign workers, especially from North- and and Sub-Saharan Africa (Morocco, Nigeria, Senegal, Egypt), China, Indian subcontinent (Bangladesh, Pakistan, Sri Lanka and India), and new EU member states (Romania *in primis*, bus also Albania and Ukraine), with a consequent growing share of the foreign-born population over the total population. In fact, starting from the second half of the 90s, Italy has also become a prime destination of foreign immigrants in the EU along with Spain (Münz, 2007), partly because of the great exposure of Italy towards the main international migration routes. Figure 1 shows the total annual gross immigration flows in Italy of foreign people in the age class 18-64.

Once arrived in Italy, several of these foreign immigrants change residence within the Country, moving from the original place of destination to a more attractive one in terms of job opportunity and/or in terms of quality of life (the so-called established immigrants' secondary migrations). Looking at internal migration flows of foreign-born residents (Figure 1), it emerges that in 2003 their total number was equal to the total number of natives with an upper-secondary school. However, during the sample period, the annual number of foreign-born residents moving within the Country rose sharply changing from 115,000 to 230,000. Over the whole sample period, ISTAT registered more than one million and seven hundred-thousands (1,744,419) of internal movements of foreign-born residents in the age class 18-64.

3.3 Spatial distribution of migration flows

There is a clear North-South spatial pattern in the provincial distribution of internal net migration rates (given by the balance between in-migration to minus out-migration from the province divided by the corresponding stock of population in the province). In particular, as expected, the higher the level of educational attainment of internal native movers, the higher the net migration rate towards Northern areas (Figures 4-7). There is also a clear North-South spatial divide in the distribution of new foreign immigration rates (Figure 8), as well as in the distribution of internal net mobility rates of foreign-born residents (Figure 9), across provinces both in 2003 and in 2011.

Insert Figures 4-9 about here.

Finally, the distribution of interregional migration flows across Italian provinces is highly right skewed regardless the type of flow considered (Figure 10). The percentage of zeros in each matrix of internal flows ranges from 18% (in the case of natives with an upper-secondary school level of education) to 52% (in the case of natives with a primary school level of education) (Table 1). The average value of dyadic flows ranges from 2.5 (in the case of natives with primary education) to 18.3 (in the case of foreign-born residents). The maximum number of internal movers is instead rather high (ranging from 621 to 14,676), and the standard deviation is much higher than the mean, thereby indicating the presence of strong overdispersion in the data. All in all, these evidences suggest that dyadic migration flows are better modeled as count data rather than as normally distributed values. We will formally test this statement in section 5.1.

Insert Figure 10 about here.

Insert Table 1 about here.

4 A semiparametric gravity model of interregional mobility flows

4.1 Baseline econometric specification

We use the data described above to assess the impact of immigration on the internal mobility of both natives (distinguished by level of education) and foreign-born residents. For each year, and for each type of internal mover, the square matrix of gross flows of inter-regional movements has 103 rows, corresponding to Italian provinces. For each of the five categories of movers, the panel of inter-regional flows has 95,481 observations (9 annual dates from 2003 to 2011, by 103 $\times$ 103 province pairs).

In our econometric analysis the dependent variable is m^h_{jkt}, i.e. the number of Italian residents aged 15-64 with level of education h (or the number of foreign-born residents, considered as a unique, homogeneous, category) moving at time t from province of origin j to province of destination k. The econometric specification takes into consideration the relationship between internal mobility (m^h_{jkt}) and a measure of the incidence of immigration from abroad on k (Imm_{kt}) and on j (Imm_{jt}), which according to our theoretical framework affects the probability of employment and, thus, the location decision of internal movers. Specifically, Imm_{lt} is the share of foreign immigrants on total population at region l ($l = k, j$).

Given the discrete nature of the dyadic outcome (number of internal movers with presence of zeros and a right-skewed distribution), we apply a gravity negative binomial model in order to explain the count of movers, accounting for overdispersion in the data.[3] The *baseline* negative binomial gravity model specification with variables included in relative (i.e. destination/origin)

[3] As in the case of Ortega and Peri (2013) and Beine, Bertoli, and Fernández-Huertas Moraga (2016), the gravity model of migration flows can be derived from a Random Utility of Maximization (RUM) model that describes the location decision problem that individuals residing in a region j face when deciding to move to another region k at time t.

form is given by:

$$\begin{aligned}\eta_{jkt} &= \log(\mu_{jkt}) = \beta_1 \log\left(\frac{Imm_{kt}}{Imm_{jt}}\right) + \beta_2 \log(\phi_{jk}) + \beta_3 \log\left(\frac{y_{kt}}{y_{jt}}\right) \\ \mu_{jkt} &= E\left(m^h_{jkt}\right) \qquad m^h_{jkt} \sim Negbin(\mu, \theta)\end{aligned} \tag{1}$$

where on the r.h.s. we include, in line with Beine and Coulombe (2018), *i*) a (relative) measure of the incidence of new immigrants, $\log\left(\frac{Imm_{kt}}{Imm_{jt}}\right)$, affecting the probability of employment and thus the relative expected income of the two regions, *ii*) a measure of accessibility of destination k for potential migrants from j ($\log \phi_{jkt}$ i.e. the cost of moving), and *iii*) destination-specific and origin-specific characteristics, $\log\left(\frac{y_{kt}}{y_{jt}}\right)$, which contribute to determine the (relative) expected income of the two regions.

The cost of moving is influenced by the physical distance (D_{jk}) between region j and region k, and we use the great circle distance between the centroids of j and k as proxy for the physical distance. As discussed above, the expected income depends on the average wage level and by the probability of finding a job. As proxy for relative wages, we include a measure of relative disposable income per head (source: Prometeia), $\log\left(\frac{Inc_{kt}}{Inc_{jt}}\right)$, as well the differential (in logs) in the average house price (source: Bank of Italy), $\log\left(\frac{House_{kt}}{House_{jt}}\right)$. A part from the incidence of new immigrants, $\log\left(\frac{Imm_{kt}}{Imm_{jt}}\right)$, the probability of employment is affected by the unemployment rate differential ($u_{kt} - u_{jt}$), and by the relative economic structure of the province. The last variable is measured in terms of differentials (in logs) in the share of agriculture employment, construction employment, and manufacturing employment on total employment, $\log\left(\frac{Agr_{kt}}{Agr_{jt}}\right)$, $\log\left(\frac{Cons_{kt}}{Cons_{jt}}\right)$, $\log\left(\frac{Man_{kt}}{Man_{jt}}\right)$. Figures 11-16 display the spatial distribution of these control variables.

Insert Figures 11-16 about here.

Within the set of regional characteristics, we also include an offset term, i.e. the sum of the log of the population aged 15-64 (25-64 in the case of tertiary education) with education level h residing in region j at time t ($\log Pop_{jt}^{h}$), as proxy for the ability to send out migrants, and of the log of population aged 15-64 (25-64) with education level h residing in region k at time t ($\log Pop_{kt}^{h}$), as proxy for the ability to receive migrants. In the case of foreign-born internal movers, we have used the log of population aged 15-64 with a primary school level of education.

To mitigate simultaneity biases, and to account for information on which natives base their decision to move, we relate current migration flows to lagged values for all the explanatory variables.

4.2 Including spatio-temporal trends

Controlling for unobserved heterogeneity is a fundamental challenge in empirical analysis, as failing to do so can introduce omitted variable bias and preclude causal inference. In our case relevant omitted variables refer to regional business cycle and regional amenities (Beine and Coulombe, 2018). In Italy, indeed, there is no official information at a local scale (province or administrative region) on the dynamics of the business cycle. Moreover, the number of physical characteristics of a province that may enhance the location as a place to live (natural amenities) is very high and, thus, hard to be captured by one or a few proxies. To accomplish this task, scholars using gravity models for longitudinal data usually apply fixed effects estimators including both origin-by-year and destination-by-year dummy variables. In our case this would imply including $2 \times N + T + 2 \times N \times T$ dummy variables and, thus, estimating 2,069 incidental parameters (i.e. α_k, α_j, τ_t, β_{kt}, and β_{jt}). However, with count data models, this approach is usually problematic since negative binomial models with so many parameters can hardly converge.

In order to solve this problem, and thus to account for time-varying and time-invariant unobserved heterogeneity, we propose the inclusion of smooth spatio-temporal trends both for origin

– $f^j(x_{s1}, x_{s2}, x_\tau)$ – and destination – $f^k(x_{s1}, x_{s2}, x_\tau)$ – regions within model 1:

$$\begin{aligned}
\eta_{jkt} &= \log(\mu_{jkt}) = \beta_1 \log\left(\frac{Imm_{kt}}{Imm_{jt}}\right) + \beta_2 \log(\phi_{jk}) + \beta_3 \log\left(\frac{y_{kt}}{y_{jt}}\right) \\
& \quad f^j(x_{s1}, x_{s2}, x_\tau) + f^k(x_{s1}, x_{s2}, x_\tau) \\
\mu_{jkt} &= E\left(m^h_{jkt}\right) \qquad m^h_{jkt} \sim Negbin(\mu, \theta)
\end{aligned} \tag{2}$$

More specifically, if most of the spatial and time unobserved heterogeneity is smoothly distributed across space and time, we can approximate it by a smooth interaction between longitude (x_{s1}), latitude (x_{s2}) and time (x_τ).

Because of the complexity in the estimation of a three-dimensional smooth function, we use the ANOVA-type decomposition of $f(x_{s1}, x_{s2}, x_\tau)$ proposed by Lee and Durbán (2011) where spatial and temporal main effects, and second- and third-order interactions between them can be identified:

$$\begin{aligned}
f(x_{s1}, x_{s2}, x_\tau) &= f_1(x_1) + f_2(x_2) + f_t(x_\tau) + \\
& \quad f_{1,2}(x_1, x_2) + f_{1,\tau}(x_1, x_\tau) + f_{2,\tau}(x_2, x_\tau) + f_{12\tau}(x_1, x_2, x_\tau)
\end{aligned}$$

First, as already pointed out in Basile, Durbán, Mínguez, Montero, and Mur (2014), the geoadditive terms given by $f_1(x_1)$, $f_2(x_2)$ and $f_{1,2}(x_1, x_2)$ work as control functions to filter the spatial trend out of the residuals, and transfer it to the mean response in a model specification. Thus, they make it possible to capture the shape of the spatial distribution of the outcome variable, eventually conditional on the determinants included in the model. These control functions also isolate stochastic spatial dependence in the residuals, that is spatially autocorrelated unobserved heterogeneity. Thus, they can be regarded as an alternative to the use of individual regional dummies (α_k and α_j) to capture unobserved heterogeneity, as long as such heterogeneity is smoothly distributed over space. Regional dummies peak at significantly higher and lower levels of the mean response variable. If these peaks are smoothly distributed over a two-dimensional surface (i.e., if unobserved heterogeneity is spatially autocorrelated), the smooth spatial trend is able to capture them. In our case, we can easily assume that time-invariant omitted variables (such as the physical characteristics or amenities of the provinces) affecting internal migrations are smoothly distributed across space, so as the smooth spatial trend is a good proxy for them.

Second, the smooth time trend, $f_\tau(x_\tau)$, and the smooth interactions between space and time - $f_{1,\tau}(x_1,x_\tau)$, $f_{2,\tau}(x_2,x_\tau)$, and $f_{12\tau}(x_1,x_2,x_\tau)$ - work as control functions to capture the heterogeneous effect of common shocks, as well as the effect of the cyclical component of provincial output (i.e. the output gap). In particular, according to the theoretical framework discussed in Section 2, these region-specific but time-varying effects influence the probability of employment along with the unemployment rate. However, there is no statistical information on the business cycle either at province or at regional level in Italy. But, if this heterogeneity is smoothly distributed across time and space, we can approximate it using smooth interactions between time and space, in alternative to origin-by year and destination-by year dummy variables. The smooth terms in the model can be estimated using different methods, and we use the penalized spline approach described in Appendix A.

The inclusion of the spatio-temporal trends may also help accomodate for *multilateral resistance effects*. These effects arise when the independence assumption on the error term is violated because of the existence of *third-region effects* (or spatial spillover effect) and/or because of the presence of *common factors* (i.e. time-specific shocks) affecting simultaneously all region-pairs kj at a specific time t, potentially with a heterogeneous intensity. In general, ignoring the effect of multilateral resistance to migration generates biases in the estimation of the parameters of the observable determinants of migration and, thus, different strategies have been proposed in the literature to relax the strong assumption of independence. First, as firstly suggested by Bertoli and Moraga (2013), when the longitudinal dimension of the data set is large enough, one may apply the CCE estimator proposed by Pesaran (2006), by introducing cross-sectional averages of the variables to control for the effect of unobserved *common factors*. Second, following the literature on international FDI, the effects of multilateral resistance to migration due to shocks in neighboring regions can be controlled by introducing spatial lags of the observed variables or by introducing a spatial autoregressive process in the error term. Third, the origin-year and destination-year dummies may also help control for the effect of multilateral resistance to migration due to a heterogeneity in the preference for migration, as proposed by Ortega and Peri (2013). The first method cannot be adopted in our case because time panel series are too short and because there is no equivalent estimator for negbin models. The inclusion of spatial lags of the explanatory variables will be considered as an option in section **??**. Finally, we suggest that the model with the ANOVA decomposition of the spatio-temporal trends for the origin and destination regions can accommodate various forms of unobserved heterogeneity due to omitted variables including common shocks, and thus, capture most of the residual cross-sectional dependence. In order to empirically verify this statement, we follow Beine, Bertoli, and Fernández-Huertas Moraga (2016)'s suggestion of adapting the Pesaran's cross-sectional dependence test (CD test) to make sure there are no remaining signs of cross-sectional correlation in the residuals.

4.3 Endogeneity issues

Another important econometric issue affecting the estimation of model 2 concerns the problem of endogeneity of the relative immigration rate, Imm_{kt}/Imm_{jt}. The provincial inflow of immigrants might be endogenous with respect to the interprovincial mobility of workers (m^h_{jkt}) because of reverse causality or omitted variables. Reverse causality occurs, for instance, if insufficient inflows of internal migrants in a given province (or excessive outflows of natives toward other provinces) induce local policy makers or local firms to expand the demand of international migrants to compensate the negative effect on the labor supply. In this case, reverse causality would determine a negative correlatio between $\log\left(\frac{Imm_{kt}}{Imm_{jt}}\right)$ and the error term and, thus, a downward bias of the estimated effect of the immigration rate on m^h_{jkt}. Reverse causality bias is partly mitigated by the inclusion of the immigration rate in the model with one-year time lag with respect to the dependent variable. However, if migration variables are persistent over time, this empirical strategy is not sufficient to properly accommodate the endogeneity bias.

Omitted variables correlated with the immigration rate can also determine an endogeneity bias. Suppose, for instance, that there is a positive local demand shock which determines an increase in labor demand for specific skills in a given Italian province. These skills can be found both in other Italian provinces and in the rest of the world. In this case, omitted variables would generate a positive correlation between $\log\left(\frac{Imm_{kt}}{Imm_{jt}}\right)$ and the error term, and thus an upward bias of the estimates of the relation between immigration and internal mobility. It could also be the case, however, that local demand shocks determine an increase in the demand for jobs that attract immigrants and are avoided by natives. In this case, the estimated impact of immigration on native mobility is downward biased.

As discussed above, the introduction of the spatio-temporal trends – $f^j(x_{s1}, x_{s2}, x_\tau)$ and $f^k(x_{s1}, x_{s2}, x_\tau)$ – is aimed at cleaning the error term from these unobserved shocks and, thus, to mitigate the bias. Nevertheless, if these unobserved effects are not fully accounted for by the smooth spatio-temporal trends, some endogeneity bias will remain. To further reduce this bias, we follow Beine and Coulombe (2018) in using a control function approach based on the information on the country of origin of foreign immigrants (see Appendix B for details).

5 Econometric results

In this section we discuss the estimation results of the semiparametric negative binomial gravity model reported in Tables 2-6. The dependent variable is, alternatively, the number of Italian citizens (natives) aged 15-64 with education level h, and the number of foreign-born residents moving at time t from province j to province k. In order to properly assess the effect of immigration on each category of internal movers, we always control for the cost of moving (approximated by the log of physical distance between each pair of provinces), and for the attractiveness of destination and origin regions in terms of per capita disposable income, unemployment rate, real estate price, and sectoral composition of the employment. In Tables 2, 3 and 4, we include all these variables in relative terms, i.e. as destination-origin differentials, while in Tables 5 and 6 we try to assess the asymmetric effects of push and pull factors on internal mobility of workers. Moreover, within the set of regional characteristics, we always include an offset term capturing the scale of the population at "risk" of movement at destination and at origin (i.e. the sum of the log of the population at origin and the log of the population at destination). Finally, a smooth spatio-temporal trend is always included (except for Table 2) to control for both unobserved spatial heterogeneity and for spatially heterogeneous time-varying common effects (Mínguez, Durbán, and Basile, 2016).

Insert Tables 2-6 about here.

The main results can be summarized as follows. Immigration has a significant impact on internal mobility of both natives and foreign-born residents. In particular, the effects of foreign immigrants materialize in *i*) both a decrease in the gross inflow and an increase in the gross outflow of foreign-born residents and of natives with a very low level of education (up to the primary school); *ii*) both an increase in the gross inflow and a decrease in the gross outflow of natives with a high level of education (tertiary or higher); *iii*) an increase in the gross outflow of natives with the lower-secondary school level of education; and *iv*) an increase in the gross inflow of natives with the upper secondary-school level of education. Cases *i*) and *iii*) represent *displacement* effects, while cases *ii*) and *iv*) represent *complementarity* effects. These results are robust to issues of endogeneity.

5.1 Diagnostics and model performance

All regressions are estimated using the Restricted Maximum Likelihood (REML) estimation method, as described in Appendix A. First, overdispersion is detected by estimating the baseline specification using a Poisson regression model (Table 2). The ratio between residual deviance and the degrees of freedom (the so called *overdispersion ratio*) turned out to be much higher than 1 in all cases, indicating severe overdispersion. More formally, the overdispersion test

(Cameron and Trivedi, 1990) always rejects the null hypothesis of equidispersion. This means that the assumptions of the Poisson model are not met, and we cannot trust its results. Definitely, a negative binomial distribution must be adopted to properly model internal migration flows.

Second, for each estimated model we test the presence of cross-sectional dependence using Pesaran's CD test. The results suggest that the introduction of the smooth spatio-temporal trends both at origin and destination allows us to control for residual strong cross-sectional dependence. Without the introduction of these smooth trends (Table 2), the CD statistics turns out to be much higher and significant at 1% in three out five cases. Once we include smooth spatio-temporal trends in the model (Tables 3-6), the evidence of cross-sectional dependence disappears or it is weakly significant (in the cases of lower-secondary and upper-secondary school levels of native movers). Thus, in the discussion below we will focus on the results from Table 3 onward.

The percentage of explained deviance increases with the level of education of native internal movers. However, the explained deviance reaches its highest value in the case of foreign-born internal movers, suggesting that our gravity model better fits the spatio-temporal variability of this specific flow of internal migrants. Comparing Tables 2 and 3, it emerges that about 30% of the explanatory power comes from the spatio-temporal trend in the cases of natives with primary school and lower-secondary school levels. This percentage decreases to about 20% in the cases of natives with upper-secondary school and tertiary levels. Finally, the spatio-temporal trend account only for about 4% of the total deviance in the case of foreign-born residents.

Finally, to mitigate the endogeneity bias due to reverse causality and omitted variables, we also adopt a two-step control function (CF) approach, as mentioned above. The results of the first step are reported in Appendix B. The control functions $f(res)$ (i.e. the smooth functions of the residuals from the first step) are always significant in the second steps (see Tables 4 and 6), thus confirming that the immigration rate represents an endogenous variable.

5.2 Testing the relative incidence of foreign immigrants

Table 3 shows the results of the gravity model using the *relative incidence of foreign immigrants* at destination and at origin, $\log(Imm_{kt}/Imm_{jt})$. All the control variables are also included in relative terms. Table 4 provides the results obtained using the same specification but controlling for the endogeneity of foreign immigrants (second stage results of the CF approach).

The results in Table 3 suggest that immigration has a significant but weak net displacement effect on the internal movement of low-skilled natives (those with a primary and a lower-secondary school level of education). Once we control for endogeneity (Table 4), it also emerges a significant positive (complementarity) effect on the internal movement of medium-skilled natives (those with an upper-secondary school level of education) and high-skilled natives (those with a tertiary or a higher level of education), and a higher negative effect on the internal movement of low-skilled natives. Thus, we conclude that the results without control for endogeneity are upward biased for low-educated natives and downward biased for highly-educated ones.

These preliminary results broadly confirm those reported in Mocetti and Porello (2010). We also complement them by displaying a significant negative effect of foreign immigration on the internal movement of foreign-born residents (column 5). Again the magnitude of the negative coefficient of the second stage of the CF approach is much higher than that obtained without control for the endogeneity bias, indicating the existence of unobserved omitted variables that are positively correlated with immigrants and that also attract foreign-born residents.

The magnitude of the coefficients estimated with the CF approach suggests that a 1% increase in the immigration rate leads to a reduction of 0.30% of net-flows of foreign-born residents. The elasticity of the displacement effect on the natives with primary school and lower-secondary school level of education is -0.40% and -0.14%, respectively. Quite interestingly, the complementarity effect on the net-flow of natives with higher education (0.55%) is higher than any displacement effect.

Most of the control variables enter with the expected sign. Focusing on the results in Table 4, we find that all kinds of internal movers migrate towards provinces with better employment opportunities (i.e. lower unemployment rate) with respect to their province of origin. As expected, the effect of the unemployment rate differential between destination and origin regions $(u_{kt} - u_{jt})$ is higher in the case of natives with a primary education level and of foreign-born residents (-0.03). The effect of the relative disposable income – $\log(Inc_{kt}/Inc_{jt})$ – is positive and significant but in the case of natives with lower secondary school level of education and, curiously, in the case of natives with higher level of education. The effect of relative house prices – $\log(House_{kt}/House_{jt})$ – is never significant in the case of native internal movers while, surprisingly, it is positive and significant in the case of foreign-born residents. As expected, moving costs hamper labor mobility: the elasticity of $\log(\phi_{jkt})$ is a bit higher (about 1%) in the case of foreign-born residents. Finally, the effect of the relative industry composition is rather heterogeneous across the different kinds of internal movers. A relatively higher weight of agriculture – $\log(Agr_{kt}/Agr_{jt})$ – and construction – $\log(Cons_{kt}/Cons_{jt})$ – seems to attract low educated workers, and to discourage highly educated workers. This last category of native movers is definitely more attracted by provinces with a higher share of services (reference category), while foreign-born internal movers are more attracted by a higher weight of manufacturing – $\log(Man_{kt}/Man_{jt})$.

5.3 Testing the impact of foreign immigrants at origin and at destination

Tables 5 and 6 report the results (without and with control for the endogeneity bias) of the gravity model specified with all the explanatory variables included separately for the region of origin (j) and for the region of destination (k) of internal movers. The coefficients of the variables measuring the incidence of foreign immigrants provide relevant information on displacement and complementary of foreign immigrants.

Starting from the displacement effects and focusing on the results of the CF approach (Table 6), it emerges both a push and a pull effect of immigration on the internal mobility of natives with a primary school level and on foreign-born internal mobility. Thus, both natives with a primary education level and foreign-born residents tend to leave, and are discouraged to move towards provinces of destination of the immigrants. However, while in the case of low-educated natives the push effect dominates the pull effect (a 1% increase of immigration leads to 0.42% increase of low-educated outflows and 0.13% decrease of their inflow), in the case of foreign-born movers push and pull effects have a similar magnitude (0.24% and -0.29%, respectively). Foreign immigration acts only as a push factor for the internal mobility of lower-secondary school level workers, that is they leave provinces of destination of the immigrants (the elasticity is 0.21%).

As far as the complementary impact of foreign immigrants is concerned, it emerges both a push and a pull effect on the internal mobility of natives with a higher level of education. In this case, however, the pull effect dominates the push effect: a 1% increase of immigration leads to 0.15% decrease of high-educated outflows and 0.41% increase of their inflow. Moreover, it appears a significant pull effect of foreign immigrants on the internal mobility of upper-secondary school natives (0.14%), while the relative incidence (Table 4) did not show any significant effect.

Results for the control variables as push and pull factors are consistent with those discussed above for their relative incidence. For the sake of brevity we focus on the effect of disposable income and unemployment displayed in Table 6. Income at destination – $\log(Inc_{kt})$ – is a key determinant of the attractiveness of each location only for native wit an upper-secondary and a higher education. Moreover, in line with the credit constraints hypothesis, income at origin – $\log(Inc_{jt})$ – has also a positive impact on the outflows of higher-educated natives. Thus, investments in human capital by Italian residents in poor regions is not a sufficient condition for exploiting better income opportunities; (potential) workers with a tertiary or a higher education must also have enough financial resources to fly away. On the other hand, low-educated natives are discouraged to leave provinces with higher disposable income (the coefficient of $\log(Inc_{jt})$ is negative); surprisingly, they are discouraged to move to provinces with higher income. A higher unemployment rate at origin (u_{jt}) work instead as a push factor for low- and medium-educated natives and for foreign-born movers. The last category of internal movers is also discouraged to move to provinces with highers unemployment rate.

6 Conclusions

We have investigated for Italy the hypothesis that immigration is a determinant of interregional migration flows of both natives and foreign-born residents. Our results indicate a displacement effect of the immigrants on the internal mobility of foreign-born residents and of Italian citizens with a low education level, but also a positive impact on the internal mobility of natives with a high education level. These findings suggest that interregional migration is an important mechanism through which the Italian labor market adjusts to immigration, and are consistent with the null unemployment effect (**any evidence of the effect of immigration on local wages?**) of immigration at the local level reported by (Basile, Durbán, Mínguez, Montero, and Mur, 2014).

Incentivizing human capital accumulation and job creation in human capital intensive sectors looks a policy able to reduce the displacement effect of immigration and, at the same time, to benefit low-skill workers through increases in their productivity.

A The semiparametric negbin gravity model with spatio-temporal trends: specification, identification and estimation technique

As it was mentioned in section 4.1, an alternative to model count data in the presence of overdispersion is the specification of a distribution that permits more flexible modeling of the variance than the Poisson distribution. The standard parametric model to account for overdispersion is the Negative Binomial. The most common way to derive this distribution is through a Poisson-Gamma mixture. This is a two-stage model that assumes that data are Poisson, but there is a heterogeneity that it is not observed. The Negative Binomial has been derived and presented with different reparameterizations. We follow the one derived by letting the mean of the Poisson distribution vary according to a parameter ζ given by the Gamma distribution. The stochastic component is given by

$$\mathbf{y}|\zeta \sim \mathscr{P}\text{oisson}(\zeta\mu), \quad \text{and}$$
$$\zeta \sim \tfrac{1}{\kappa}\mathscr{G}\text{amma}(\kappa).$$

The marginal distribution of $\mathbf{y}$ is, then, the Negative Binomial with mean μ and variance $\mu + \mu^2/\kappa$, where κ is the dispersion parameter. Note that, for a large value of κ, the Negative Binomial model reduces to Poisson. Then, we have the response $\mathbf{y}$ defined as:

$$\mathbf{y} \sim \text{Neg Bin}(\mu, \kappa)$$

with probability function:

$$P(\mathbf{y} = y_i|\mu_i, \kappa) = \binom{y_i + \kappa - 1}{y_i} \left(\frac{\mu_i}{\kappa + \mu_i}\right)^{y_i} \left(\frac{\kappa}{\kappa + \mu_i}\right)^{\kappa},$$
$$\text{for } \kappa \geq 0, \text{ and } y_i = 0, 1, 2, \ldots$$

The fact that the Negative Binomial distribution is not in the exponential family, makes more difficult the use of the standard methodology developed for Generalized Linear Models (GLMs). However, it can be formulated as a GLM, if the parameter κ is assumed constant.

$$\mathscr{L}(\mu_i, \kappa|y_i) = y_i \ln\left(\frac{\mu_i}{\mu_i + \kappa}\right) - \kappa \ln\left(\frac{\mu_i}{\mu_i + \kappa}\right) + \ln\Gamma(y_i + \kappa) - \ln\Gamma(\kappa) - \ln\Gamma(y_i + 1) + \kappa \ln \kappa,$$

from which we can see that the canonical link is $\eta_i = \ln(\mu_i/\mu_i + \kappa)$.

If κ were known, this would be within the exponential family. For a given κ, the log-likelihood for the vector μ is

$$\mathscr{L}(\mu_i; \kappa) = \sum_{i=1}^{n} y_i \ln\left(\frac{\mu_i}{(\mu_i + \kappa)}\right) - \sum_{i=1}^{n} \kappa \ln\left(\frac{1 + \mu_i}{\kappa}\right) + c(y, \kappa), \tag{3}$$

where $c(y,\kappa)$ is a function of the y_i's and κ. For a given μ, the log likelihood for κ is

$$\mathscr{L}(\mu_i,\kappa) = n\{\kappa \ln\kappa - \ln\Gamma(\kappa)\} + \sum_{i=1}^{n}\{\ln\Gamma(y_i,\kappa) - (y_i+\kappa)\ln(\kappa+\mu_i)\} + d(y_i,\mu_i) \tag{4}$$

for some function $d(y_i,\kappa)$.

A.1 Incorporation of covariates and parameters estimation

For simplicity, we will include a single covariate, the extension to a more general case is immediate. As in any GLM, we will assume that a transformation of the mean is linear on the covariates:

$$g(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{1i}$$

In this case, we will use the log-link to obtain the GLM-based Negative Binomial that yields identical parameters estimates to the Poisson-Gamma mixture. However, the linearity assumption is very strong, and we may need to relax it, by assuming that the linear predictor is a non-linear smooth function of the covariate, i.e., $\eta_i = f(x_{1i})$. There are several ways in which $f(x_1)$ can be estimated. One of the most flexible and well know technique is the Penalized Spline, P-spline (Eilers and Marx, 1996).

The method is based on a basis representation of the unknown function, which is combined with a penalty on the likelihood to control the wiggliness of the curve. In particular, we will use cubic B-splines as basis functions, and second-order differences of adjacent coefficients, $\lambda\beta'D'D\beta$ (where D is a matrix that calculates second order differences) as penalty controlled by the so called *smoothing parameter*, λ. Then, the smooth function is represented by:

$$f(x_1) = \sum_{j=1}^{c} B_j(x)\theta_j,\ j = 1,...,c, \tag{5}$$

with B_j a B-spline basis function, and θ_j is a component of a vector of regression coefficients of length c (the number of knots used to construct the basis). So, in this case, the vector μ can be written as $\mu = exp(B\theta)$, and the log-likelihood (3) becomes:

$$\mathscr{L}(\theta,\kappa) = y'(B\theta - \log\{\kappa 1 + \exp(B\theta)\}) - \kappa 1' \log\{\kappa 1 + \exp(B\theta)\} + c(y,\kappa). \qquad (6)$$

And the penalized log-likelihood is

$$\mathscr{L}_p(\theta,\kappa) = \mathscr{L}(\theta,\kappa) - \frac{\lambda}{2}\beta' D' D\beta \ , \qquad (7)$$

The first term of the equation above is the likelihood of y given θ, and the second can be seen as the log of the prior density of θ, and so, we could assume that $\theta \sim N\left(0, \sigma_\theta^2 (D'D)^{-1}\right)$, with $\sigma_\theta^2 = \lambda^{-1}$. Then, our initial model has become a hierarchical model:

$$\begin{aligned}
&\mathbf{y} \sim \text{Neg Bin}(\mu,\kappa) \quad log(\mu) = \eta \\
&\eta = B\theta \\
&\theta \sim N\left(0, \sigma_\theta^2 (D'D)^{-1}\right).
\end{aligned}$$

But, difference (derivative) penalties of order d do not penalize powers of x up to degree $d-1$, in our case, $d = 2$, so $D'D$ will be singular with to eigenvalues equal to zero, therefore, θ has a degenerate distribution. The standard solution is to reparameterize the linear predictor:

$$B\theta = X\beta + Z\alpha \quad \alpha \sim N(0, G),$$

where the 2 columns of X span the polynomial null space of $D'D$, and the $(n-2)$ columns of Z expand its complement. The reparametrization is not unique, we use the one based on the singular value decomposition of the penalty matrix $D'D = U\Lambda U'$ (Lee and Durbán, 2011). Then,

$$X = BU_0 \quad Z = BU_+ \quad G = \sigma_\alpha^2 \Lambda_+$$

where U_0 and U_+ are the eigenvectors corresponding to the zero and no zero eigenvalues of the penalty, and $\sigma_\alpha^2 = \lambda^{-1}$.

Now we are in the context of a Generalized Linear Mixed Model (GLMM), however, the maxium likelihood solution for the hierarchical GLMM requires integrating over the random effects. Breslow and Clayton (1993) popularized the use of Penalized Quasi-likelihood (PQL) for estimation and inference in these models. PQL is a simple method for estimation of GLMMs that can be implemented by iterative fitting a linear mixed model to a modified dependent variable. We rewrite the penalized likelihood in (7) as:

$$\mathscr{L}_p(\beta,\alpha,\kappa,\sigma_\alpha^2) = \mathscr{L}(\beta,\alpha,\kappa,\sigma_\alpha^2) - \alpha' G^{-1}\alpha \; . \tag{8}$$

Taking the derivative of (8) with respect to β and α, yields:

$$X'\left(\frac{y-\exp(X\beta+Z\alpha)}{\kappa 1+\exp(X\beta+Z\alpha)}\right) = 0 \tag{9}$$

$$\kappa Z'\left(\frac{y-\exp(X\beta+Z\alpha)}{\kappa 1+\exp(X\beta+Z\alpha)}\right) = G^{-1}\alpha \; . \tag{10}$$

The system of equations in (9) and (10) can be solved using a Fisher's scoring algorithm with working vector $z = \eta + W^{-1}(y-\mu)$, and matrix of weights given by:

$$W = \kappa \text{ diag}\left(\frac{\exp(X\beta+Z\alpha)}{\kappa 1+\exp(X\beta+Z\alpha)}\right). \tag{11}$$

Then, the estimated fixed and random effects are,

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}z \tag{12}$$

$$\hat{\alpha} = GZ'V^{-1}(z - X\hat{\beta}), \tag{13}$$

with $V = V^{-1} + ZGZ'$. Then, conditional on the estimates obtained in (12) and (13), $(\sigma_\alpha^2, \kappa)$ are estimated by the approximate REML quasi-likelihood:

$$-\tfrac{1}{2}\log|V| - \tfrac{1}{2}\log|X'V^{-1}X| - \tfrac{1}{2}z'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})z \tag{14}$$

The PQL solution is obtained by iteration between (12), (13) and (14) until convergence.

A.2 Smooth interaction terms

The linear predictor described above may be a linear combination over several variables, that might interact with each other. In particular, as proposed in section 4.2, if most of the spatial and time unobserved heterogeneity is smoothly distributed across space and time, we can approximate it by a smooth interaction between longitude (x_{s1}), latitude (x_{s2}) and time (x_τ), i.e., $\eta = f(x_{s1}, x_{s2}, x_\tau)$. In correspondence to the ANOVA decomposition in ordinary regression models, we can perform a similar decomposition and write,

$$\begin{aligned} f(x_{s1}, x_{s2}, x_\tau) &= f_1(x_1) + f_2(x_2) + f_t(x_\tau) + \\ &\quad f_{1,2}(x_1, x_2) + f_{1,\tau}(x_1, x_\tau) + f_{2,\tau}(x_2, x_\tau) + f_{12\tau}(x_1, x_2, x_\tau), \end{aligned}$$

where $f_j(x_j),\ j = 1,2,t$ are main effects and $f_{jk}(x_j, x_k),\ j,k = 1,2,t$ and $f_{12\tau}(x_1, x_2, x_\tau)$ are interaction effects. One of the advantajes of this decomposition is that it allows to determine if trasformed mean response is sufficiently described by the simple sum of the main effects or if, in addition, the interaction effects are needed as well.

Tensor product smooths are the natural way of representing smooth interaction terms. They are constructed based of smooths of single covariates, i.e., using the basis and penalties for each covariate described above, and combining them appropriately. To show the construction of tensor product basis we will consider a smooth of two covariates (and the extension to 3 or more covariates is immediate). To construct the basis we start by representing, for example, a smooth function of longitude using the basis expansion given in (5),

$$f(x_{s1}) = \sum_{k=1}^{c} B_k(x_{s1}) a_k.$$

Then, a smooth function of longitude and latitude, $f_{1,2}(x_1, x_2)$, can be constructed by allowing the coefficients a_k to vary smoothly with latitude,

$$a_k(x_{s2}) = \sum_{l=1}^{\breve{c}} \breve{B}_l(x_{s2}) a_{kl},$$

where $\breve{B}_l(x_{s2})$ are also B-spline basis functions. Therefore,

$$f_{1,2}(x_1, x_2) = \sum_k \sum_l B_k(x_{s1}) B_l(x_{s2}) a_{kl}. \tag{15}$$

and the matrix of coefficients $A = [a_{kl}]$ summarizes the fit. The roughness of the elements of A determines how smooth the surface will be. To tune roughness, each column and each row of A is penalized, using a different smoothing parameter in each direction. In this case the penalty term to be added to the log-likelihood (as in (7)), would be:

$$a' \left(\lambda_{s1}(D'_{s1} D_{s1} \otimes I_{s2}) + \lambda_{s2}(I_{s1} \otimes D'_{s2} D_{s2}) \right) a,$$

where D_{s1} and D_{s1} are difference matrices of order 2 applied to the rows and columns of A. A similar approach is taken to represent all interactions in terms of tensor products of B-spline basis, and appropriate penalties are added to the likelihood. Finally, as in the univariate case, each interaction term is reparameterized within the mixed model framework by using an appropriate transformation (see Lee and Durbán, 2011, for details), and estimation proceeds as previously.

It is important to note that the partition of $f(x_{s1}, x_{s2}, x_\tau)$ given above is is non-identifiable since, for example, $f_1(x_{s1})$ and $f_2(x_{s2})$ are confounded with $f_{12}(x_{s1}, x_{s2})$ (as in a three-way anova, lower order effects are confounded with higher order interactions). Identifiability constraints need to be impose to make the model identifiable. Lee and Durbán (2011) shows that it is enough to impose the constraints used in a factorial design to the coefficients of the B-splines basis of main effects and lower order interactions.

B The construction of the instrumental variable and the control function approach

As discussed in section 4.3, endogeneity issues can be generated by the occurrence of reverse causality between foreign immigration and internal mobility, and by the existence of omitted factors of internal migration that are correlated with foreign immigration. To address these issues, we extend the REML methodology discussed in Appendix A to estimate the parameters of model (2) in a 2-stage control function (CF) approach (Blundell and Powell, 2003), that is an alternative to standard instrumental variable/two-stage least square methods. In the first stage, the endogenous variables – i.e. $\log(Imm_{kt}/Imm_{jt})$ in Table 3, and $\log(Imm_{kt})$ and$\log(Imm_{jt})$ in Table 5 – are regressed on a set of conformable instrumental variables, using a semiparametric negbin model. The residuals from the first stages are then included in the original model (2) to control for the endogeneity of $\log(Imm_{kt}/Imm_{jt})$, $\log(Imm_{kt})$, and $\log(Imm_{kt})$. Since the second stage contains generated regressors (i.e., the first-step residuals), a bootstrap procedure is used to compute p-values (see Basile, Durbán, Mínguez, Montero, and Mur, 2014, for details on the bootstrap procedure).

To apply the CF approach, we need external instruments (i.e. variables not correlated with the error term but with the observed immigration flows) whose values need to vary across provinces and over time. To this end we follow an approach previously proposed in the literature of international trade and in the migration literature by Beine and Coulombe (2018), which consists on exploiting the information on the country of origin of foreign immigrants.

More specifically, the instrumental variable is constructed into two steps. First, we estimate a gravity negative binomial model that explains the magnitude of the flows of immigrants from each origin country c of the world to each province k for each year t (Imm_{ckt}):

$$\begin{aligned} \eta &= \log(\mu) = \alpha \log(X_{kt}) + \delta \log(\phi_{ckt}) + \\ & \quad f^k(x_{s1}, x_{s2}, x_\tau) + \gamma_c \times \tau_t \\ \mu &= E(Imm_{ckt}) \qquad Imm_{ckt} \sim Negbin(\mu, \theta) \end{aligned} \tag{16}$$

The covariates of this model include exogenous time-varying variables (X_{kt}) that are specific to the province of destination of immigrants (namely the unemployment rate, the disposable income per capita, the average house price, and the share of agriculture, manufacturing and construction employment on total employment), a smooth spatio-temporal trend for the provinces of destination – $f^k(x_{s1}, x_{s2}, x_\tau)$ – using the ANOVA decomposition explained above, the bilateral distance between each country of origin and the province of destination (ϕ_{ck}), a dummy variable indicating whether the country of origin and the province are contiguous ($contig_{ck}$), and country-of-origin-by-year fixed effects ($\gamma_c \times \tau_t$). These fixed effects capture the effect of origin-specific time invariant and time-varying variables that are supposed to be uncorrelated to the unobserved provincial shocks (and thus with the error term of model 2). Thus, they allow us to use variation in immigration that is exogenous to the evolution of internal migration.

The estimation results of model (16) are in line with our expectations (Table 7). Foreign immigrants to a given province from a given origin country increase with the average wage rate of the province, and decreases with its unemployment rate. A higher share of services on total employment (the reference category) seems to attract foreign immigrants. Average house prices of the province do not exert any effect. Finally, flows of immigrants decrease with distance and increase with contiguity between the country of origin and the province of destination.

Insert Table 7 about here.

Based on the estimation of this gravity model, we recover the predicted bilateral flows for each country-province-year triplet ($\widehat{Imm}_{ckt} = exp(\widehat{\mu}_{ckt})$). Then, we sum up these predicted flows across origin country to get a total predicted flow of immigrants by year and province ($\widehat{Imm}_{kt} = \sum_c \widehat{Imm}_{ckt}$), which is supposed to be generated by exogenous factors, i.e. variables that are uncorrelated with the unobservable province-time shock (and the error term). This total predicted value (once divided by the total province population) can therefore be used as an instrument of the observed immigration rate in model 2. More precisely, the observed immigration rate is regressed against the predicted immigration rate in a regression model including all other exogenous variables included in model 2. Finally, a smooth function of the residuals from this model ($f(res)$) is included in model 2 as a control function to correct the endogeneity bias. The overall significance of these control functions can be used as a Wu-Hausman test and confirms the endogeneity of the immigration rate.

The validity of the instruments has to fulfill the usual two conditions. First, the instruments must be strong predictors of the observed immigration rates. The percentage of explained deviance (69%) reported in Table 7 suggests that this is the case at the bilateral level. Furthermore, at the aggregate level, i.e., after summing up across origins, this can be evaluated by the F-stat of the first stage of the final CF procedure. The value of the F-stat in the case of relative incidence of immigrants (Tables 4) is 24165 with a p-value = 0.000, while the value of the F-stat in the case of absolute incidence (Table 6) is 91148 with a p-value = 0.000. The second condition is that the instrument must be uncorrelated with the error term of the final regression. In this case, the error term contains the influence of unobserved provincial shocks on the net interprovincial immigration flows. The covariates used for the prediction of Imm_{ckt} are obviously uncorrelated with the contemporaneous shocks.

References

ALESINA, A., S. DANNINGER, AND M. ROSTAGNO (2001): "Redistribution through public employment: the case of Italy," *IMF Staff Papers*, 48(3), 447–473.

ALI, K., M. D. PARTRIDGE, AND D. S. RICKMAN (2012): "International immigration and domestic out-migrants: are domestic migrants moving to new jobs or away from immigrants?," *The Annals of Regional Science*, 49(2), 397–415.

ANTOLÍN, P., AND O. BOVER (1993): *Regional migration in Spain: the effect of personal characteristics and of unemployment, wage and house price differentials using pooled cross-sections*. Banco de España, Servicio de estudios.

ATTANASIO, O. P., AND F. P. SCHIOPPA (1991): "Regional inequalities, migration and mismatch in Italy, 1960-86," *Mismatch and labour mobility*, pp. 237–324.

BASILE, R., AND M. CAUSI (2007): "Le determinanti dei flussi migratori nelle province italiane: 1991-2001," *Economia e Lavoro*, 41, 139–159.

BASILE, R., M. DURBÁN, R. MÍNGUEZ, J. M. MONTERO, AND J. MUR (2014): "Modeling regional economic dynamics: Spatial dependence, spatial heterogeneity and nonlinearities," *Journal of Economic Dynamics and Control*, 48, 229 – 245.

BASILE, R., A. GIRARDI, M. MANTUANO, AND G. RUSSO (2018): "Interregional Migration of Human Capital and Unemployment Dynamics: Evidence from Italian Provinces," *mimeo*.

BEINE, M., S. BERTOLI, AND J. FERNÁNDEZ-HUERTAS MORAGA (2016): "A practitioners' guide to gravity models of international migration," *The World Economy*, 39(4), 496–512.

BEINE, M., AND S. COULOMBE (2018): "Immigration and internal mobility in Canada," *Journal of Population Economics*, 31(1), 69–106.

BERTOLI, S., AND J. F.-H. MORAGA (2013): "Multilateral resistance to migration," *Journal of Development Economics*, 102, 79–100.

BLUNDELL, R., AND J. POWELL (2003): *Advances in Economics and Econometrics Theory and Application*chap. Endogeneity in Nonparametric and Semiparametric Regression Models. Cambridge University Press.

BORJAS, G. J. (2006): "Native internal migration and the labor market impact of immigration," *Journal of Human resources*, 41(2), 221–258.

BRESLOW, N., AND D. CLAYTON (1993): "Aproximated inference in generalised linear mixed models," *Journal of the American Statistical Association*, 88(421), 9–25.

BURDA, M., AND C. WYPLOSZ (1994): "Gross worker and job flows in Europe," *European Economic Review*, 38(6), 1287–1315.

BURDA, M. C. (1995): "Migration and the option value of waiting," *Economic and Social Review*, 27(1), 1.

CAMERON, A. C., AND P. K. TRIVEDI (1990): "Regression-based tests for overdispersion in the Poisson model," *Journal of econometrics*, 46(3), 347–364.

CARD, D. (2001): "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 19(1), 22–64.

CARD, D., AND J. DINARDO (2000): "Do immigrant inflows lead to native outflows?," *American Economic Review*, 90(2), 360–367.

CONSTANT, A. F. (2014): "Do migrants take the jobs of native workers?," *IZA World of Labor*.

DOCQUIER, F., AND J. MACHADO (2017): "Income Disparities, Population and Migration Flows Over the Twenty First Century," *Italian Economic Journal*, 3(2), 125–149.

EILERS, P., AND B. MARX (1996): "Flexible Smoothing with B-Splines and Penalties," *Statistical Science*, 11, 89–121.

ETZO, I. (2011): "The determinants of the recent interregional migration flows in Italy: a panel data analysis," *Journal of Regional Science*, 51(5), 948–966.

FAGGIAN, A., P. MCCANN, AND S. SHEPPARD (2007): "Some evidence that women are more mobile than men: Gender differences in UK graduate migration behavior," *Journal of Regional Science*, 47(3), 517–539.

FAGGIAN, A., M. D. PARTRIDGE, D. S. RICKMAN, ET AL. (2012): "Cultural avoidance and internal migration in the USA: Do the Source Countries Matter?," *Migration impact analysis*, pp. 203–224.

FAINI, R., G. GALLI, P. GENNARI, AND F. ROSSI (1997): "An empirical puzzle: falling migration and growing unemployment differentials among Italian regions," *European Economic Review*, 41(3-5), 571–579.

FILER, R. (1992): "The effect of immigrant arrivals on migratory patterns of native workers," in *Immigration and the workforce: Economic consequences for the United States and source areas*, ed. by G. J. Borjas, and R. B. Freeman, pp. 245–270. University of Chicago Press, Chicago.

FREY, W. (1999): "Immigration and demographic balkanization: Toward one America or two?," in *America's demographic tapestry: baseline for the new millennium*, ed. by J. W. Hughes, and J. J. Seneca, pp. 78–97. Rutgers University Press, New Brunswick, NJ.

FREY, W. H. (1994): "The new white flight.," *American Demographics*, 16(4), 40–46.

——— (1995): "Immigration and internal migration 'flight' from US metropolitan areas: Toward a new demographic Balkanisation," *Urban Studies*, 32(4-5), 733–757.

——— (1996): "Immigration, domestic migration, and demographic balkanization in America: new evidence for the 1990s," *Population and Development Review*, pp. 741–763.

GREENWOOD, M. J. (1997): "Internal migration in developed countries," *Handbook of population and family economics*, 1, 647–720.

HANSON, G., AND C. MCINTOSH (2016): "Is the Mediterranean the New Rio Grande? US and EU immigration pressures in the long run," *Journal of Economic Perspectives*, 30(4), 57–82.

HATTON, T. J., AND M. TANI (2005): "Immigration and Inter-Regional Mobility in the UK, 1982–2000," *The Economic Journal*, 115(507).

HUGHES, G., AND B. MCCORMICK (1994): "Did migration in the 1980s narrow the North-South divide?," *Economica*, pp. 509–527.

KRITZ, M. M., AND D. T. GURAK (2001): "The impact of immigration on the internal migration of natives and immigrants," *Demography*, 38(1), 133–145.

LEE, D., AND M. DURBÁN (2011): "P-Spline ANOVA Type Interaction Models for Spatio-Temporal Smoothing," *Statistical Modelling*, 11, 49–69.

LEY, D., AND J. TUTCHENER (2001): "Immigration, globalisation and house prices in Canada's gateway cities," *Housing Studies*, 16(2), 199–223.

MÍNGUEZ, R., M. L. DURBÁN, AND R. BASILE (2016): "Spatio-Temporal Autoregressive Semiparametric Model for the analysis of regional economic data," Discussion paper, Dipartimento di Economia e Finanza, LUISS Guido Carli.

MOCETTI, S., AND C. PORELLO (2010): "How does immigration affect native internal mobility? new evidence from Italy," *Regional Science and Urban Economics*, 40(6), 427–439.

MÜNZ, R. (2007): "Migration, labor markets, and integration of migrants: An overview for Europe," Discussion paper, HWWI policy paper.

ORRENIUS, P. M., AND M. ZAVODNY (2007): "Does immigration affect wages? A look at occupation-level evidence," *Labour Economics*, 14(5), 757–773.

ORTEGA, F., AND G. PERI (2013): "The effect of income and immigration policies on international migration," *Migration Studies*, 1(1), 47–74.

PERI, G. (2014): "Do immigrant workers depress the wages of native workers?," *IZA world of Labor*.

PIRAS, R. (2012a): "Disentangling the role of migrants' educational level in the long-run Italian internal migration trends," *Studies in Regional Science*, 42(2), 377–396.

——— (2012b): "Internal migration across Italian regions: macroeconomic determinants and accommodating potential for a dualistic economy," *The Manchester School*, 80(4), 499–524.

——— (2017): "A long-run analysis of push and pull factors of internal migration in Italy. Estimation of a gravity model with human capital using homogeneous and heterogeneous approaches," *Papers in Regional Science*, 96(3), 571–602.

SHEEHAN, P. (1998): *Among the Barbarians: The Dividing of Australia*. Sydney: Random House.

FIGURE 1
Annual total flows of foreign immigrants and of internal movements of Italian citizens (natives) and foreign-born residents

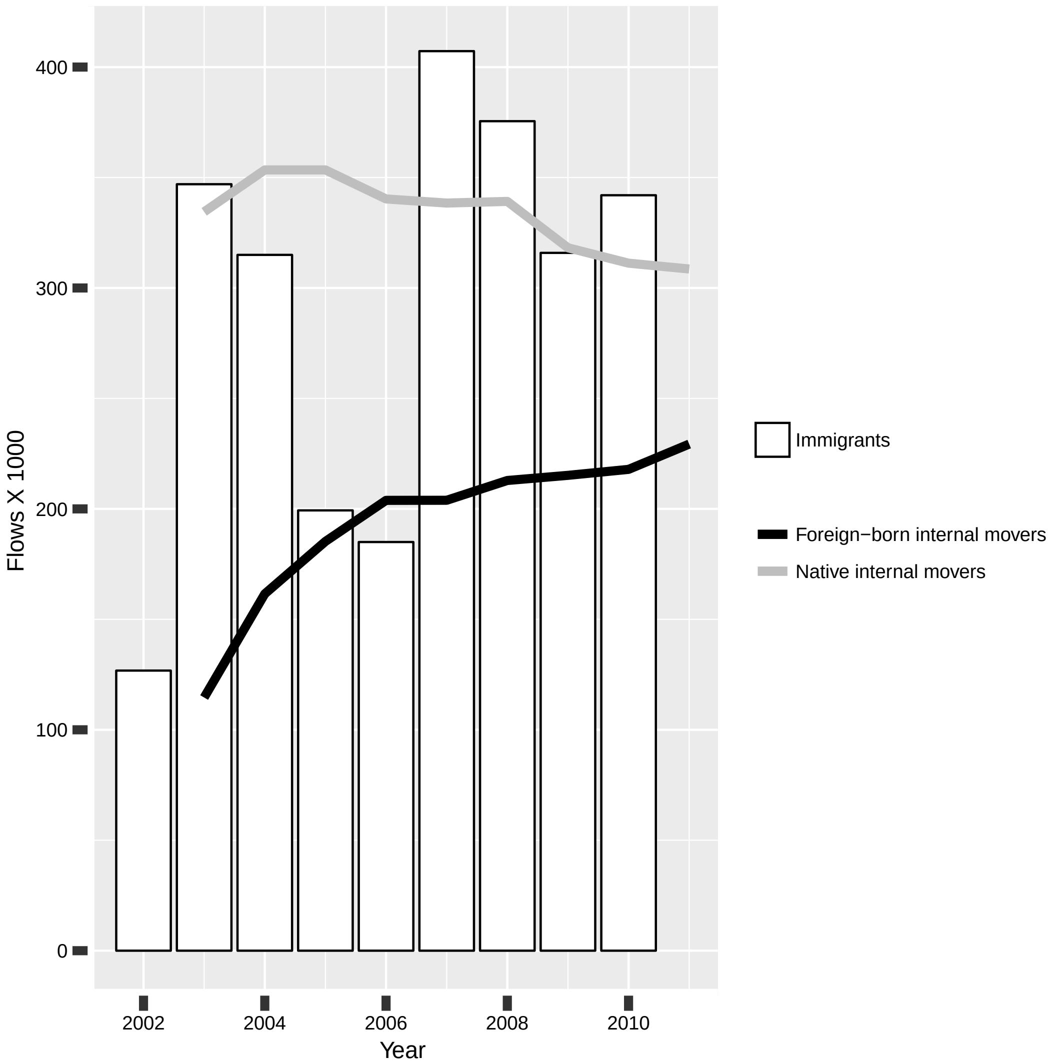


FIGURE 2
Annual total internal flows of Italian citizens by level of education

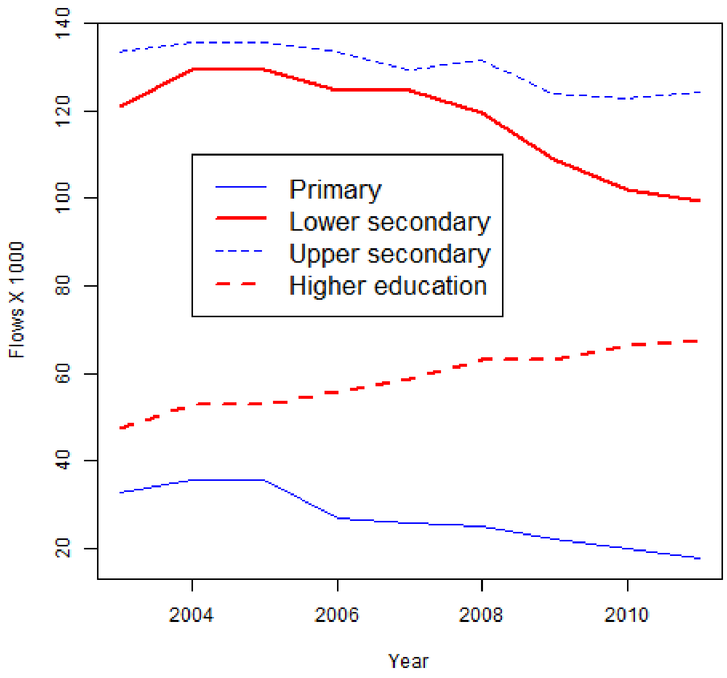


FIGURE 3
Annual migration rates of natives by level of education. Migration rates are computed as gross annual total flows divided by the stock of population with the same level of educational attainment

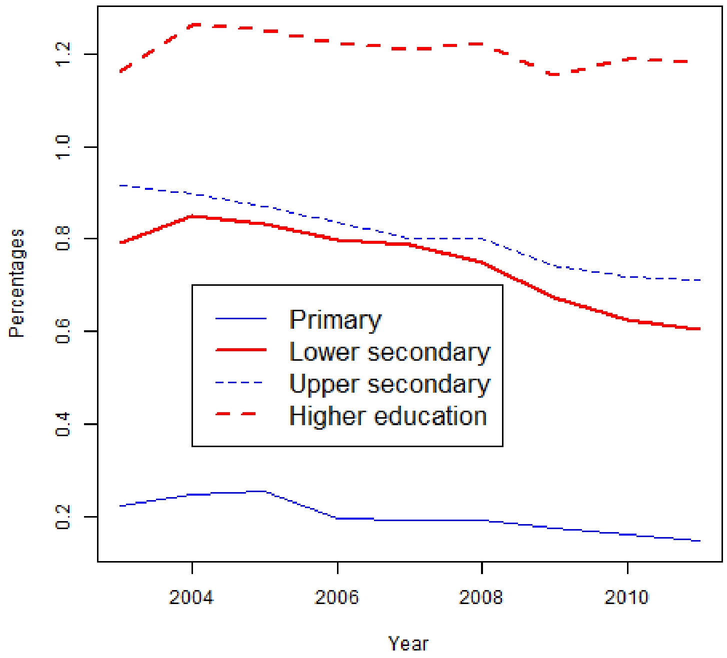
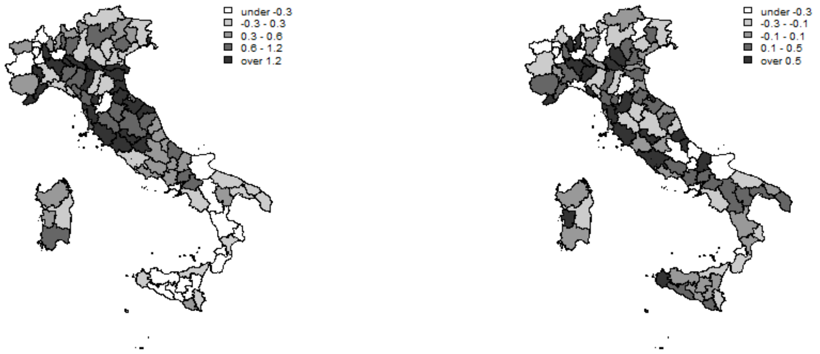
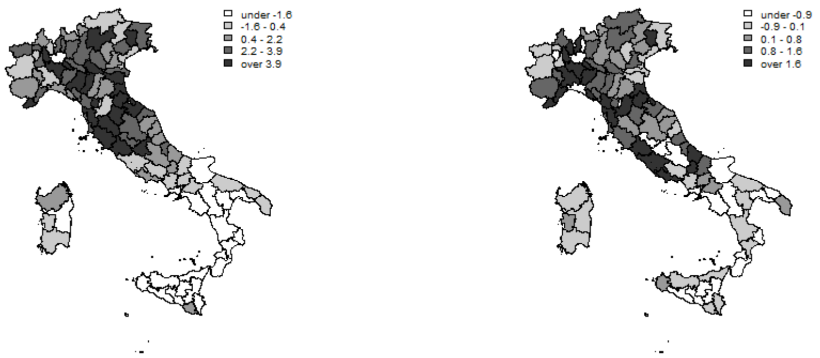


FIGURE 4
Map of internal mobility rates of Italian citizens with a primary education level



(a) 2003 (b) 2011

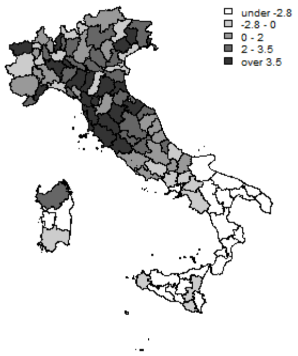
FIGURE 5
Map of internal mobility rates of Italian citizens with a lower secondary education level



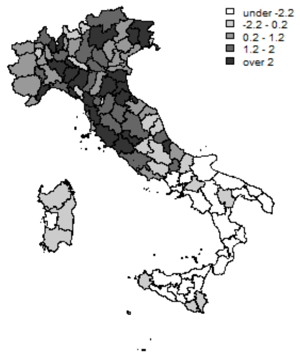
(a) 2003 (b) 2011

FIGURE 6

Map of internal mobility rates of Italian citizens with a upper secondary education level



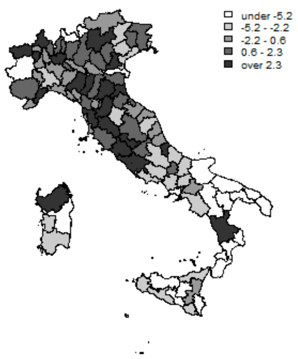
(a) 2003



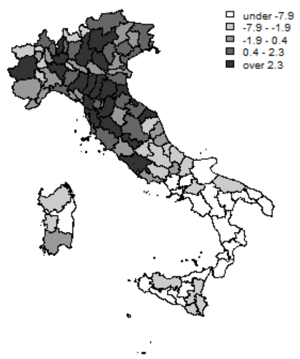
(b) 2011

FIGURE 7

Map of internal mobility rates of Italian citizens with a higher education level

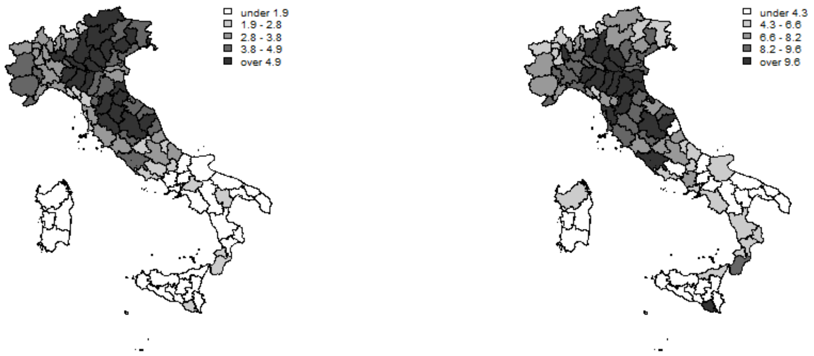


(a) 2003



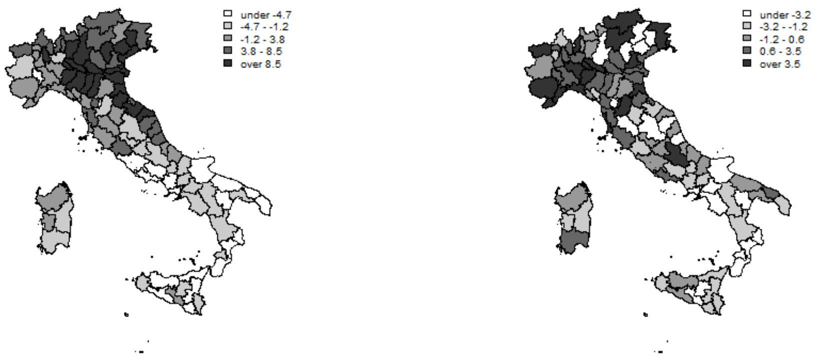
(b) 2011

FIGURE 8
Map of immigration rates of foreign workers



(a) 2003 (b) 2011

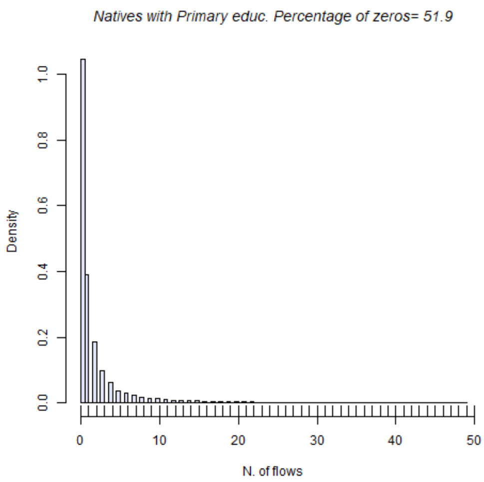
FIGURE 9
Map of internal mobility rates of foreign-born residents



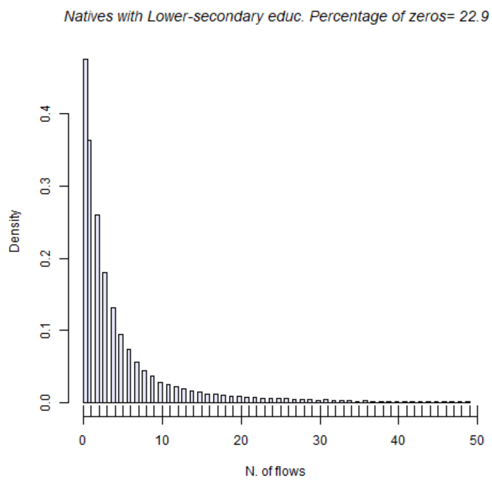
(a) 2003 (b) 2011

FIGURE 10

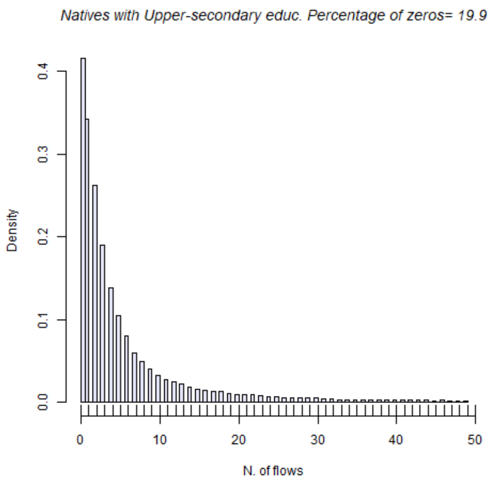
Histograms of dyadic internal flows of migrants. Distribution truncated at the threshold of 50 dyadic flows.



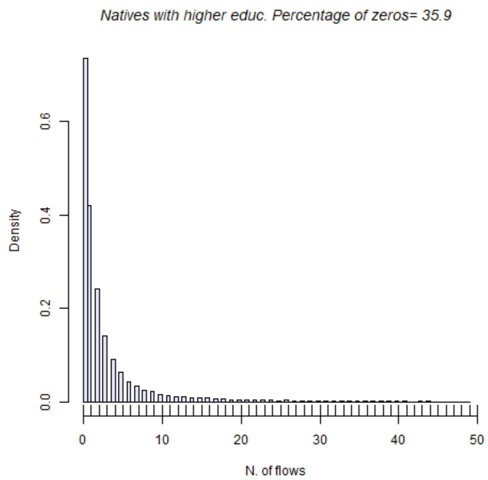
(a) Natives. Primary



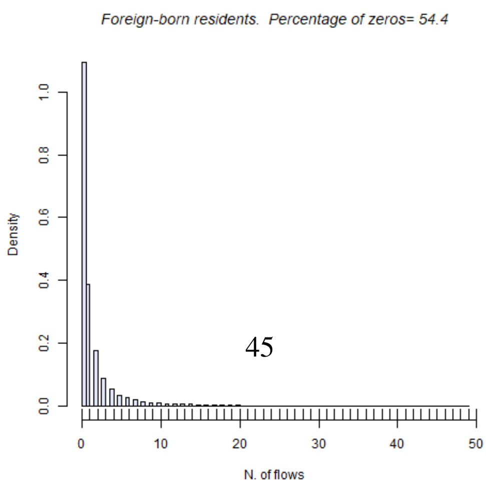
(b) Natives. Lower-sec



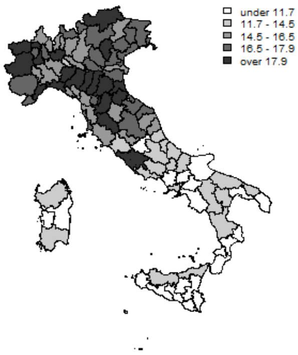
(c) Natives. Upper-sec



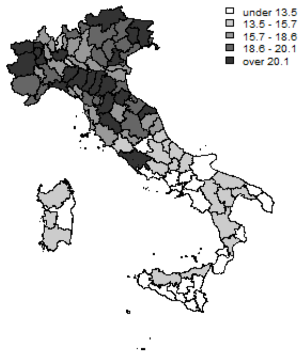
(d) Natives. Higher



(e) Foreign-born residents



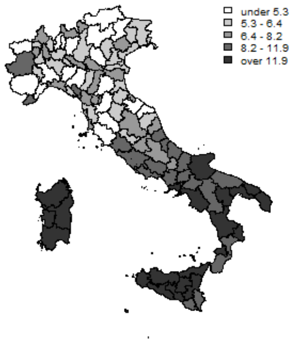
(f) 2003



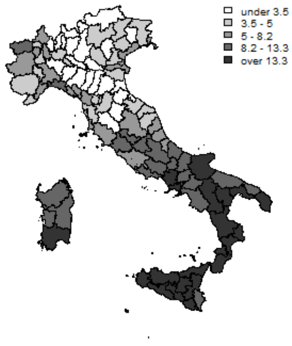
(g) 2011

FIGURE 11
Disposable income per capita

FIGURE 12
Unemployment rates

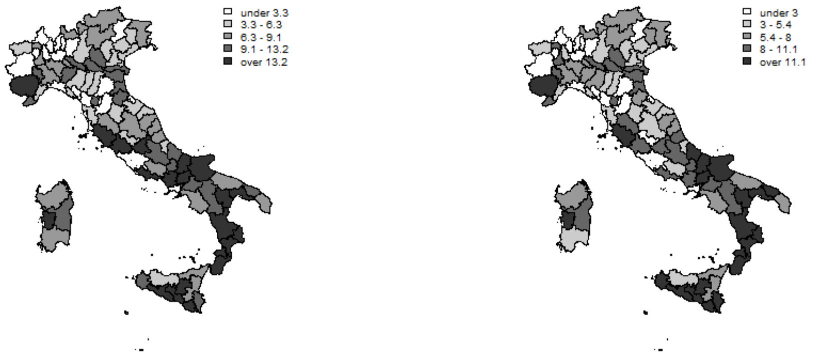


(a) 2003



(b) 2011

FIGURE 13
Share of agriculture employment on total employment

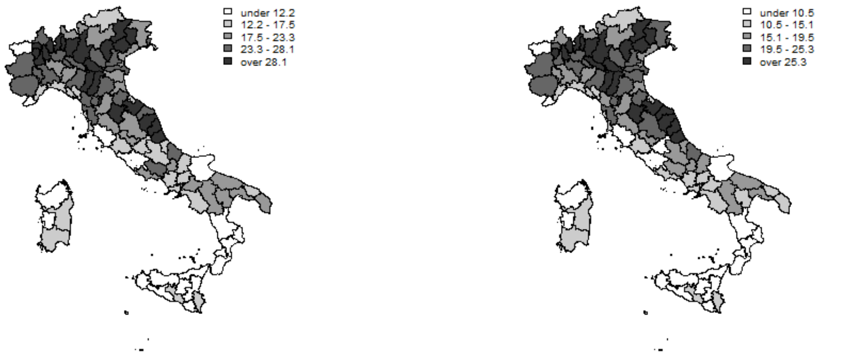


(a) 2003 (b) 2011

FIGURE 14
Share of construction employment on total employment

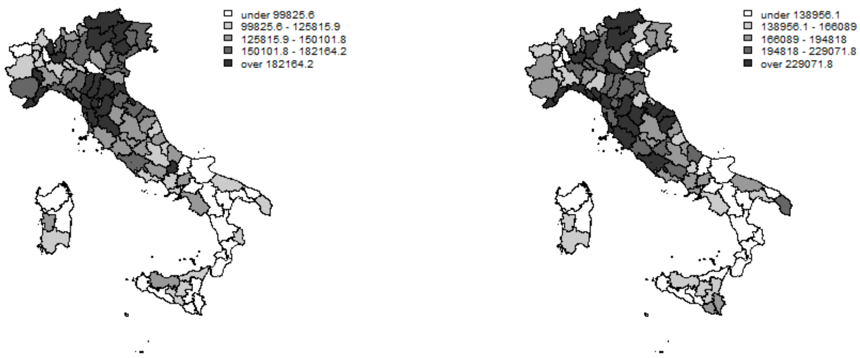
(a) 2003 (b) 2011

FIGURE 15
Share of manufacturing employment on total employment



(a) 2003 (b) 2011

FIGURE 16
House prices



(a) 2003 (b) 2011

TABLE 1

Distribution of dyadic internal migration flows in Italy (Period: 2003-2011)

Type	Perc. of zeros	Mean	Max	Std.Dev.	Total
Foreign-born residents	33%	18.27	14,676	203.92	1,744,419
Italian citizens					2,997,327
- Primary school	52%	2.53	621	10.32	241,781
- Lower secondary	20%	11.09	1,925	41.44	1,058,623
- Upper secondary	18%	12.25	1,764	43.76	1,169,271
- Higher education	36%	5.53	644	20.15	527,652

Source: ISTAT "*Indagine sui trasferimenti di residenza*".

TABLE 2

Gravity models of internal mobility in Italy. Testing the relative incidence of foreign immigrants. Negative Binomial Estimates. Baseline specification (1)

	Italian citizens				Foreign-born
	Primary	Lower sec.	Upper sec.	Higher	residents
Parametric terms	*Coefficients (standard errors)*				
Intercept	-5.675***	-4.056***	-3.959***	-2.323***	-1.582***
	(0.031)	(0.024)	(0.022)	(0.025)	(0.023)
$\log(Imm_{kt}/Imm_{jt})$	-0.037***	0.001	0.064***	0.077***	-0.113***
	(0.014)	(0.010)	(0.009)	(0.011)	(0.011)
$\log(Inc_{kt}/Inc_{jt})$	-0.090**	0.067**	0.300***	0.661***	0.526***
	(0.040)	(0.028)	(0.027)	(0.031)	(0.029)
$u_{kt} - u_{jt}$	-0.020***	-0.019***	-0.015***	-0.012***	-0.038***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
$\log(House_{kt}/House_{jt})$	0.054***	0.082***	0.082***	0.088***	0.071***
	(0.015)	(0.011)	(0.010)	(0.012)	(0.011)
$\log(Agr_{kt}/Agr_{jt})$	0.062***	0.034***	0.015***	-0.027***	0.027***
	(0.006)	(0.004)	(0.004)	(0.005)	(0.004)
$\log(Cons_{kt}/Cons_{jt})$	0.024	-0.031*	-0.027	-0.044*	-0.025
	(0.027)	(0.019)	(0.018)	(0.021)	(0.020)
$\log(Man_{kt}/Man_{jt})$	-0.047***	-0.026***	0.004	-0.010	0.021**
	(0.023)	(0.008)	(0.008)	(0.009)	(0.008)
$\log(\phi_{jk})$	-0.574***	-0.619***	-0.638***	-0.665***	-1.100***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Diagnostics and performance					
Explained deviance (%)	41.8	51.2	57.2	58.2	79.7
Log-lik.	-144,990	-270,245	-273,707	-198,938	-220,966
AIC	289,940	540,510	547,434	397,896	441,952
CD stat.	2.421	3.512	2.548	1.172	-0.942
	[0.015]	[0.000]	[0.010]	[0.241]	[0.346]
Overdispersion ratio	10.174	46.692	43.136	16.166	41.235
Overdispersion test	29.467	29.252	30.311	35.753	25.616
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

Notes: the dependent variables (m^h_{jkt}) are the number of Italian citizens aged 15-64 with level of education h, and the number of foreign-born residents, moving at time t from province of origin j to province of destination k. REML estimates. Standard errors in parenthesis and P-values in brackets. $*$, $**$ and $***$ denote significance at the 1, 5 and 10 per cent levels respectively. CD is Pesaran's statistics for cross-sectional dependence test. Number of obs.: 95.481.

TABLE 3

Gravity models of internal mobility in Italy. Testing the relative incidence of foreign immigrants. Negative Binomial Estimates. Model specification (2) with spatio-temporal trends

	Italian citizens				Foreign-born
	Primary	Lower sec.	Upper sec.	Higher	residents
Parametric terms	*Coefficients (standard errors)*				
Intercept	-2.462***	-0.830**	-1.269***	-0.914***	-2.552***
	(0.486)	(0.399)	(0.323)	(0.284)	(0.229)
$\log(Imm_{kt}/Imm_{jt})$	-0.041**	-0.032**	0.006	-0.006	-0.130***
	(0.019)	(0.014)	(0.013)	(0.016)	(0.014)
$\log(Inc_{kt}/Inc_{jt})$	-0.013	-0.008	0.121***	0.325***	0.020
	(0.058)	(0.043)	(0.039)	(0.047)	(0.039)
$u_{kt} - u_{jt}$	-0.012***	-0.011***	-0.010***	-0.011***	-0.023***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
$\log(House_{kt}/House_{jt})$	-0.012	-0.004	-0.002	0.042***	0.043***
	(0.019)	(0.013)	(0.012)	(0.015)	(0.013)
$\log(Agr_{kt}/Agr_{jt})$	0.044***	0.028***	0.001	-0.041***	0.004
	(0.009)	(0.007)	(0.006)	(0.007)	(0.006)
$\log(Cons_{kt}/Cons_{jt})$	0.125***	0.010	-0.027	-0.067**	-0.061**
	(0.038)	(0.027)	(0.025)	(0.030)	(0.025)
$\log(Man_{kt}/Man_{jt})$	0.038*	0.010	0.020	-0.072***	0.072***
	(0.022)	(0.016)	(0.015)	(0.017)	(0.015)
$\log(\phi_{jk})$	-0.722***	-0.790***	-0.815***	-0.816***	-0.981***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
Nonparametric terms					
Spatio-temporal trends (ANOVA specification)	Yes	Yes	Yes	Yes	Yes
Diagnostics and performance					
Explained deviance (%)	55.8	65.3	69.7	68.2	82.8
Log-lik.	-149,233	-263,599	-266,547	-193,548	-210,076
AIC	298,757	527,504	533,392	387,452	420,529
CD stat.	0.800	1.830	1.822	0.469	-0.851
	[0.423]	[0.067]	[0.069]	[0.638]	[0.394]

Notes: see notes in Table 2.

TABLE 4

Gravity models of internal mobility in Italy. Testing the relative incidence of foreign immigrants. Negative Binomial Estimates. Model specification (2) with spatio-temporal trends, accounting for endogeneity (second step CF approach)

	Italian citizens				Foreign-born
	Primary	Lower sec.	Upper sec.	Higher	residents
Parametric terms	*Coefficients (standard errors)*				
Intercept	-2.486***	-0.861**	-1.328***	-0.954***	-2.557***
	(0.484)	(0.398)	(0.322)	(0.304)	(0.221)
$\log(Imm_{kt}/Imm_{jt})$	-0.402***	-0.141***	0.096**	0.548***	-0.298***
	(0.073)	(0.053)	(0.049)	(0.059)	(0.051)
$\log(Inc_{kt}/Inc_{jt})$	0.165**	0.044	0.079*	0.057	0.099**
	(0.068)	(0.049)	(0.046)	(0.055)	(0.047)
$u_{kt} - u_{jt}$	-0.027***	-0.016***	-0.006**	0.011***	-0.030***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
$\log(House_{kt}/House_{jt})$	0.007	0.002	-0.007	0.010	0.051***
	(0.019)	(0.014)	(0.013)	(0.015)	(0.013)
$\log(Agr_{kt}/Agr_{jt})$	0.045***	0.029***	0.001	-0.044***	0.004
	(0.009)	(0.007)	(0.006)	(0.007)	(0.006)
$\log(Cons_{kt}/Cons_{jt})$	0.151***	0.017	-0.031	-0.097***	-0.047*
	(0.038)	(0.028)	(0.026)	(0.030)	(0.026)
$\log(Man_{kt}/Man_{jt})$	0.017	0.003	0.025*	-0.036**	0.062***
	(0.022)	(0.017)	(0.015)	(0.017)	(0.015)
$\log(\phi_{jk})$	-0.723***	-0.789***	-0.814***	-0.814***	-0.978***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
Nonparametric terms	*χ^2-test [EDF]*				
$f(res)$	28.10***	12.28**	14.03***	104.39***	135.27***
	[2.163]	[3.096]	[3.040]	[3.215]	[3.956]
Spatio-temporal trends (ANOVA specification)	Yes	Yes	Yes	Yes	Yes
Diagnostics and performance					
Explained deviance (%)	55.8	65.4	69.7	68.4	82.9
Log-lik.	-149,219	-264,040	-266,541	-193,970	-210,010
AIC	298,742	527,497	533,386	387,344	420,405
CD stat.	0.782	1.841	2.461	0.481	-0.863
	[0.436]	[0.066]	[0.073]	[0.631]	[0.388]

Notes: see notes in Table 2. $f(res)$ is a smooth function of the residuals from the first step model; χ^2-test is the test to assess the overall significance of this control function, while EDF indicates the corresponding estimated degrees of freedom.

TABLE 5

Gravity models of internal mobility in Italy. Testing the impact of foreign immigrants at origin and at destination. Negative Binomial Estimates

	Italian citizens				Foreign-born
	Primary	Lower sec.	Upper sec.	Higher	residents
Parametric terms	*Coefficients (standard errors)*				
Intercept	-0.649	-0.919*	-1.533***	-4.314***	-0.537
	(0.697)	(0.535)	(0.471)	(0.507)	(0.472)
$\log(Imm_{kt})$	0.042**	0.010	0.030**	0.039**	0.025*
	(0.021)	(0.016)	(0.015)	(0.017)	(0.015)
$\log(Imm_{jt})$	0.113***	0.071***	0.013	0.039**	0.249***
	(0.021)	(0.016)	(0.015)	(0.017)	(0.015)
$\log(Inc_{kt})$	-0.670***	-0.231***	0.226***	1.109***	-0.189***
	(0.082)	(0.059)	(0.053)	(0.062)	(0.055)
$\log(Inc_{jt})$	-0.644***	-0.218***	-0.028	0.530***	-0.191***
	(0.081)	(0.060)	(0.055)	(0.063)	(0.055)
u_{kt}	0.001	-0.002	-0.005**	-0.014***	-0.031***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
u_{jt}	0.024***	0.019***	-0.013***	0.009***	0.014***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)
$\log(House_{kt})$	0.093***	0.074***	0.063***	0.100***	-0.016***
	(0.027)	(0.019)	(0.018)	(0.021)	(0.018)
$\log(House_{jt})$	0.115***	0.075***	0.060	0.023	-0.102***
	(0.026)	(0.019)	(0.018)	(0.021)	(0.018)
$\log(Agr_{kt})$	0.096***	0.116***	0.041***	-0.034***	0.045 ***
	(0.013)	(0.009)	(0.009)	(0.010)	(0.009)
$\log(Agr_{jt})$	0.004	0.057***	0.037***	0.050***	0.040***
	(0.013)	(0.010)	(0.009)	(0.010)	(0.009)
$\log(Cons_{kt})$	-0.329***	-0.292***	-0.348***	-0.385***	-0.049
	(0.053)	(0.040)	(0.035)	(0.040)	(0.036)
$\log(Cons_{jt})$	-0.575***	-0.315***	-0.292***	-0.214***	0.072**
	(0.052)	(0.039)	(0.036)	(0.040)	(0.036)
$\log(Man_{kt})$	-0.155***	-0.149***	-0.252***	-0.441***	0.047**
	(0.031)	(0.022)	(0.020)	(0.022)	(0.020)
$\log(Man_{jt})$	-0.216***	-0.156***	-0.275***	-0.308***	-0.089***
	(0.030)	(0.023)	(0.021)	(0.023)	(0.020)
$\log(\phi_{jk})$	-0.732***	-0.791***	-0.819***	-0.828***	-0.978***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
Nonparametric terms					
Spatio-temporal trends (ANOVA specification)	Yes	Yes	Yes	Yes	Yes
Diagnostics and performance					
Explained deviance (%)	55.9	65.2	71.2	73.1	82.6
Log-lik.	-149,017	-263,371	-266,258	-192,685	-209,742
AIC	298,345	527,063	532,831	385,733	419,968
CD stat.	0.830	1.883	1.253	0.374	-0.872
	[0.407]	[0.060]	[0.070]	[0.708]	[0.382]

Notes: see notes in Table 2.

TABLE 6

Gravity models of internal mobility in Italy. Testing the impact of foreign immigrants at origin and at destination. Negative Binomial Estimates accounting for endogeneity (second step CF approach)

	Italian citizens				Foreign-born
	Primary	Lower sec.	Upper sec.	Higher	residents
Parametric terms	*Coefficients (standard errors)*				
Intercept	-0.640	-1.119**	-1.532***	-4.457***	-0.367
	(0.687)	(0.529)	(0.470)	(0.513)	(0.467)
$\log(Imm_{kt})$	-0.125**	0.014	0.139***	0.414***	-0.291***
	(0.051)	(0.036)	(0.034)	(0.040)	(0.035)
$\log(Imm_{jt})$	0.415***	0.205***	0.016	-0.154***	0.238***
	(0.050)	(0.037)	(0.035)	(0.040)	(0.036)
$\log(Inc_{kt})$	-0.652***	-0.276***	0.140**	0.916***	0.031
	(0.085)	(0.061)	(0.056)	(0.065)	(0.057)
$\log(Inc_{jt})$	-0.806***	-0.305***	-0.067	0.570***	-0.073
	(0.083)	(0.062)	(0.058)	(0.066)	(0.058)
u_{kt}	-0.004	-0.001	-0.001	0.002	-0.049***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
u_{jt}	0.034***	0.025***	0.014***	0.004	0.011***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)
$\log(House_{kt})$	0.102***	0.081***	0.064***	0.102***	-0.031*
	(0.027)	(0.019)	(0.018)	(0.021)	(0.018)
$\log(House_{jt})$	0.117***	0.081***	0.064***	0.040*	-0.112***
	(0.026)	(0.019)	(0.018)	(0.021)	(0.018)
$\log(Agr_{kt})$	0.092***	0.114***	0.039***	-0.031***	0.038***
	(0.013)	(0.009)	(0.009)	(0.010)	(0.009)
$\log(Agr_{jt})$	0.008	0.059***	0.036***	0.050***	0.043***
	(0.013)	(0.010)	(0.009)	(0.010)	(0.009)
$\log(Cons_{kt})$	-0.328***	-0.315***	-0.377***	-0.435***	0.028
	(0.054)	(0.039)	(0.035)	(0.041)	(0.037)
$\log(Cons_{jt})$	-0.630***	-0.349***	-0.310***	-0.221***	0.107***
	(0.052)	(0.039)	(0.036)	(0.041)	(0.036)
$\log(Man_{kt})$	-0.179***	-0.151***	-0.244***	-0.404***	0.003
	(0.030)	(0.022)	(0.020)	(0.023)	(0.021)
$\log(Man_{jt})$	-0.202***	-0.145***	-0.278***	-0.332***	-0.086***
	(0.029)	(0.023)	(0.021)	(0.023)	(0.020)
$\log(\phi_{jk})$	-0.720***	-0.788***	-0.816***	-0.822***	-0.992***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Nonparametric terms	*χ^2-test [EDF]*				
$f(res_a)$	16.50***	25.38***	50.65***	178.88***	262.52***
	[2.882]	[3.642]	[3.717]	[3.823]	[3.759]
$f(res_b)$	56.13***	43.59***	14.94***	48.08***	77.93***
	[2.620	[3.543]	[3.588]	[3.748]	[3.715]
Spatio-temporal trends (ANOVA specification)	Yes	Yes	Yes	Yes	Yes
Diagnostics and performance					
Explained deviance (%)	56.0	65.2	71.1	73.0	82.8
Log-lik.	-149,400	-263,790	-266,220	-192,573	-209,457
AIC	298,298	526,994	532,780	385,523	419,315
CD stat.	0.848	1.882	1.748	0.390	-0.871
	[0.396]	[0.060]	[0.071]	[0.692]	[0.383]

Notes: see notes in Table 4. $f(res_a)$ and $f(res_b)$ are smooth functions of the residuals from the first step models; χ^2-test is the test to assess the overall significance of these control functions, while EDF indicates the corresponding estimated degrees of freedom.

TABLE 7

First step results of the control function approach

Parametric terms	*Coefficients (standard errors)*
Intercept	3.496^{***}(0.530)
$\log(Inc_{kt})$	0.316^{***} (0.083)
u_{kt}	-0.032^{***}(0.004)
$\log(House_{kt})$	0.006 (0.026)
$\log(Agr_{kt})$	-0.138^{***} (0.012)
$\log(Cons_{kt})$	-0.032 (0.052)
$\log(Man_{kt})$	-0.224^{***} (0.030)
$\log(\phi_{ck})$	-0.649^{***} (0.034)
$Contig_{ck}$	1.019^{***} (0.170)
Country by time fixed effects	Yes
Nonparametric terms	
Spatio-temporal trend	Yes
Diagnostics and performance	
Explained deviance (%)	68.5
Log-lik.	-170,250

Notes: the dependent variables (imm_{ckt}) are the number of foreign immigrants moving at time t from country c to province k. REML estimates. Standard errors in parenthesis. $*$, $**$ and $***$ denote significance at the 1, 5 and 10 per cent levels respectively. Number of obs.: 41200.