# Disentangling the effects of internal high-skilled migrations on regional economic performance: the case of Italy

Giulio Castellano[1], Gaetano Vecchione[2], and Viktor A. Venhorst[3]

[1,2]Department of Political Science, University of Naples Federico II, Italy
[3]Faculty of Spatial Sciences, University of Groningen, The Netherlands

Draft, July 2025

**Abstract**

This work intends to contribute to the literature on the consequences of regional skilled migration on origin regions by investigating the impact of regional skilled migrations in Italy on human capital accumulation and, in the last resort, on economic growth of Southern Italian provinces over the period 2004-2022. Skilled migration reduces the human capital endowment in the origin regions, with a negative effect on labor productivity and economic performance. Moreover, skilled migration might also influence economic performance by influencing individual agents' educational choices. To empirically test the existence of these mechanisms, preliminary OLS estimates with fixed effects are run, including models that allow for spatially heterogeneous effects across Centre-North and South, and across different levels of peripherality defined upon the concept of "inner areas". Finally, an IV 2SLS strategy is adopted in order to control for the endogeneity of migration. This empirical strategy is run using regional data from official statistical sources (ISTAT, Eurostat and the Italian Ministry of University and Research).

**Keywords:** Regional skilled migration; Human capital; Italy; Regional development.

**JEL classification:** J61; R23; O15; I25.

# 1 Introduction

Regional convergence and migration are both central topics in regional science. Nevertheless, there are some aspects of their interaction that still appear unclear. So far, researchers managed to get a good understanding of the determinants of migration (i.e. the individual, collective, and environmental factors that explain migration behaviors), while less has been said about the consequences of migration, especially for sending regions (Faggian et al., 2019). Theoretical approaches like equilibrium and disequilibrium models have been traditionally adopted to identify the determinants of migration, mainly focusing on regional characteristics (labor market and amenities). Those neoclassical approaches looked at migration as an equilibrating mechanism for wages and unemployment across different regions, and so as a converging force. Moreover, equilibrium and disequilibrium models assumed migrants to be homogeneous in terms of characteristics and preferences.

Starting from the Nineties, thanks to the increasing availability of data, the issue of migrants' self-selection into migration became more central, and the main results of neoclassical models have been questioned. A relevant contribution in this sense is represented by the empirical work of Borjas et al. (1992), which proved how migration patterns are determined not only by regional differentials in the level of wages, but also by regional differentials in the returns of skills, which contribute to determine the skill composition of migration flows. These findings are consistent with the work of Krugman (1991) which, with his New Economic Geography (NEG), provided theoretical support towards migration patterns following agglomeration dynamics that widens regional divides. More attractive regions benefit from inflows of talented workers, investments and infrastructure, thus strengthening their ability to further attract skilled migrants, while less attractive, peripheral regions keep losing human capital to more attractive regions, entering in a negative spiral that hinders growth.

Further advances in the literature on the consequences of regional migration have been achieved at the beginning of the twenty-first century in the wake of the endogenous growth theory, which emphasizes the role of human capital as a main driver of economic growth. Although as pointed out by Faggian et al. (2019) the application of growth models *à la* Lucas to the regional context is not straightforward, this strand of literature shows how the human capital content of migration shapes the effects of migration on regional economies: receiving regions benefit from high-skilled immigration as it contributes to create and spread knowledge, while sending regions see their development potential hindered by the loss of human capital caused by the outflow of high-skilled individuals. An insightful theoretical contribution on these heterogeneous effects of skill-selective flows on regional growth and convergence is provided by Fratesi and Riggi (2007). Moreover, high-skilled migration can generate negative spillovers that contribute to worsen its effects on sending regions. Contributions regarding both international (Boeri et al., 2012) and regional migration (Mocetti and Porello, 2010; Nifo et al., 2020) show that migration prospects can act as a disincentive for individuals to acquire tertiary education, reducing investments in human capital in places where there is a structural outflow of high-skilled individuals.

Regional inequalities are a structural feature of many advanced economies, and Italy stands out as one of the countries in which such inequalities are most persistent. In the Italian context, regional migration patterns are deeply linked with the structural divide between North and South (but also between rural and urban areas), and became particularly relevant after the Second World War. Between the Fifties and the Seventies, a large flow of low-skilled workers left the rural areas in Southern regions toward the industrialized areas in the North, especially in the so-called 'industrial triangle', formed by Milan, Turin, and Genoa. There, the labor demand was very high, driven by the presence of big factories. During the Seventies, this migration flow slowed down, due to economic crisis and changes in the productive fabric. During the Eighties and the Nineties, however, there was a resurgence of strong internal migration, this time directed not only towards the main industrial cities, but also towards medium-sized cities in Central Italy, especially in Tuscany and Emilia-Romagna. During these decades, an increasing self-selection of skilled individuals in migration started to emerge, due to the rising service economy. Finally, in the last two decades, especially after the Great Recession, there has been a significant change in the characteristics of migration: in the first place, there was a rise in emigration towards other European countries and the US; in the second place, internal migration was characterized by an increasing skill-selection. Indeed, while in 2002 graduates accounted for only the 10% of the migrant population, by 2022 this share has grown up to 35%.[1] These figures suggest an increased skill-selection of migrants, which is confirmed if we compare the above-mentioned share of graduates in the migrant population with the same share referred to the resident population: as showed in Figure 1.0.1, the difference between the share of graduates in the migrant population the share of graduates in the resident population has been positive and increasing over the last two decades.

This relatively swift change in the skill composition of migrants is likely to affect the regional divide between Southern and Northern regions, yet the economic literature on the consequences of this regional 'brain drain' is still limited. This contribution aims to improve the knowledge and understanding of this phenomenon by trying to estimate and disentangle the direct and indirect effects of high-skilled migration on economic growth in Southern Italy. The first one consists in high-skilled migration reducing the human capital level in the origin region(s) (i.e. the share of highly skilled in the population or workforce). This in turn lowers the productivity of labor and negatively affects economic performance. The second one consists in migration prospects affecting the propensity of individuals to invest in human capital (i.e. whether to acquire tertiary education or not). In the case of international migrations, a relevant strand of the literature has proved this effect to be positive, leading to a mitigation of the 'brain drain' hypothesis. However, in the case of internal migration, the sign of this effect is rather unclear, with some empirical evidence of it being negative, at least in the case of Italy (Nifo et al., 2020).

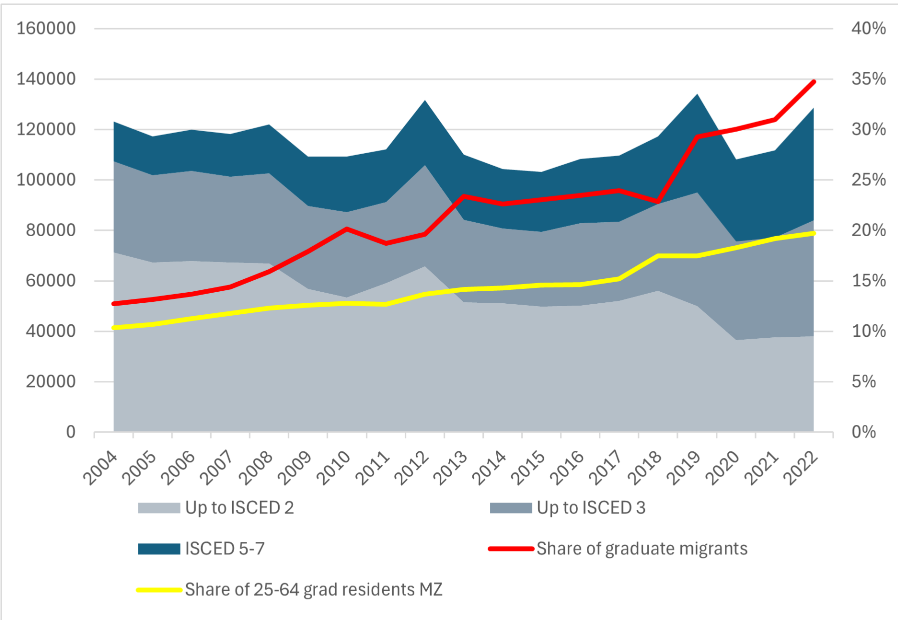



Figure 1.0.1: Internal migrations from Southern to Northern Italy by educational attainment (areas in shades of blue, left axis) and share of graduates in the migrant population from Southern to Northern Italy (red) and in the resident population (yellow). Own elaboration on Istat data.

## 2 Theoretical framework

Assessing the impact of skilled migration on economic performance means observing the effect of those migrations on the variables that contribute to determine the regional GDP (per capita). So, as in Boeri et al. (2012), we start from a GDP equation in the form of a Cobb-Douglas production function. For region $i$ (with $i = S, N$), we can express regional GDP at time $t$ as:

$$Y_{i,t} = A_{i,t} K_{i,t}^{1-\alpha_i} E_{i,t}^{\alpha_i} \tag{2.0.1}$$

[1] *Source:* Istat.

where $A_{i,t}$ is a time-varying scale factor (total factor productivity of region $i$, $K_{i,t}$ is the amount of capital in the economy of region $i$, and $\alpha_i$ is the share of labor in the income of region $i$.

$E_{i,t}$ is labor expressed in terms of efficiency units, and it is a function of highly skilled and low-skilled labor ($H_{i,t}$ and $L_{i,}$, respectively), with constant elasticity of substitution between the two:

$$E_{i,t}^{\alpha_i} = [q_{i,t}\theta_{i,t}H_{i,t}^{\rho} + (1-\theta_{i,t})L_{i,t}^{\rho}]^{1/\rho} \tag{2.0.2}$$

where $q_{i,t}$ is the average productivity of educated workers employed domestically (which in turn depends on the quality of education and on the distribution of innate abilities), $\theta_{i,t}$ is the share of highly skilled labor in total labor of region $i$, $\rho$ is a parameter determining the elasticity of substitution between highly skilled and low-skilled workers, which is given by $1/(1-\rho)$.

High skilled migration has, as a direct consequence, a reduction in human capital accumulation (the share of highly skilled in the workforce of the origin region, $\theta_{S,t}$ reduces). Nevertheless, the literature on international migrations has proven the case for several feedback effects that could contribute to mitigate (or even reverse) the 'brain drain' effect. If so, it could be possible to talk about 'brain gain'.

In line with the brain drain literature, Boeri et al. (2012) consider the investment in education to be endogenous and to be affected by two sets of variables:

- *The proportion of highly skilled in the workforce*: theoretically, the wage premium for education is higher if there are fewer graduates (i.e. if the demand for skilled workers is higher than the supply);
- *Migration prospects to developed countries*: as discussed above.

The dynamics of human capital can be considered to be governed by this process:

$$h_{S,t+1} = h_{S,t}(1-\delta_S) + n(h_{S,t})\Theta_{S,t}(1-m_{S,t}) \tag{2.0.3}$$

where $\delta_S$ is a rate of depreciation of human capital capturing the level and differential in mortality or retirement of existing cohorts in Southern Italy. The variable $n(h_{S,t})$ captures the growth rate of the labor force or the replacement of old by young workers; it is decreasing in $h_{i,t}$ since more educated parents tend to have fewer children. $\Theta_{S,t}$ reflects enrollment rates in tertiary education and captures the proportion of educated within the new cohort of workers. Finally, $m_{S,t}$ is the emigration rate of young educated adults.

The steady state, if it exists, is the following:

$$h_S^* = \frac{n(h_S^*)\Theta_S^*(1-m_S^*)}{\delta_S} \tag{2.0.4}$$

Now we endogenize $\Theta_{S,t}$, that is the investment in education (human capital) in Southern Italy. As previously said, it depends on the share of high skilled ($h_{S,t}$), on the proportion of $h_{S,t}$ employed in the production of human capital ($\hat{h}_{S,t}$), and on migration prospects. The magnitude of the latter depends on the high-skilled wage gap between Southern and Northern Italy, $[1-(w_{S,t}^H/w_{N,t}^H)]$. Therefore, we have

$$\Theta_{S,t} = \Theta\left[h_{S,t}, \hat{h}_{S,t}, m_{S,t}\left(1-\frac{w_{S,t}^H}{w_{N,t}^H}\right)\right] \tag{2.0.5}$$

Plugging (2.0.5) into (2.0.3) and deriving $h_{S,t+1}$ w.r.t. $m_{S,t}$, a marginal increase in $m_{S,t}$ stimulates human capital accumulation if and only if

$$\Theta_3'(1-m_{S,t})\left(1-\frac{w_{S,t}^H}{w_{N,t}^H}\right) > \Theta_{S,t} \tag{2.0.6}$$

where $\Theta_3'$ is the partial derivative of $\Theta_{S,t}$ w.r.t. $(m_{S,t}[1-(w_{S,t}^H/w_{N,t}^H)])$. [2] For condition (2.0.6) to hold, the emigration rate $(m_{i,t})$ must not be too high, the wage differential must be large enough (i.e. $w_{S,t}^H/w_{N,t}^H$ must be small), and $\Theta_3'$ must be large.

In the literature on international brain gain, the prospect of migrating to a richer country could act as an incentive to invest more in human capital for people living in developing countries. Thus, in addition to the direct negative effect of skilled migration on human capital, there is this possible feedback effect through which skilled migration can stimulate human capital accumulation. In our model, this translates to the function $\Theta$ having positive partial derivatives $(\Theta_1', \Theta_2', \Theta_3' > 0)$.

With reference to the internal migrations in Italy, the existence of this education incentive generated by skilled out-migration seems unrealistic for at least two reasons. In first place, international migration flows are heavily influenced by the immigration policies of receiving countries, while regional migrations take place within a single country, where citizens can move without legal constraints. In second place, the differences between Northern and Southern regions in terms of economic performance, quality of life, opportunities, and institutional quality are rather limited compared to the differences between developing and developed countries. Therefore, in the case of internal migrations in Italy, the incentive effect of migration prospects on investment in human capital is unlikely to take place. Moreover, there is some empirical evidence about the existence of a reverse effect, according to which internal migration prospects play as a disincentive to invest in education (Nifo et al., 2020).

Since $\Theta_{S,t}$ is greater than zero (enrollment rates cannot be negative), a necessary but not sufficient condition for (2.0.6) to hold is that its LHS must be greater than zero. The term $(1-m_{S,t})$ is necessarily greater than (or equal to) zero because $0 < m_{S,t} < 1$. So, for the condition (2.0.6) to be satisfied, $\Theta_3'$ and $[1-(w_{S,t}^H/w_{N,t}^H)]$ must be either both positive or both negative. Since we are assuming that in the Italian case the former is negative, then the latter must be negative too. But this would be possible only if the wage ratio were greater than one (i.e. if high-skilled wages were higher in the South Italy than in the North), which is not the case. Therefore, in the Italian scenario the condition (2.0.6) for a brain gain through the human capital channel is not met, and the feedback effect of skilled migration on human capital accumulation contributes to magnify the human capital loss driven by the direct effect.

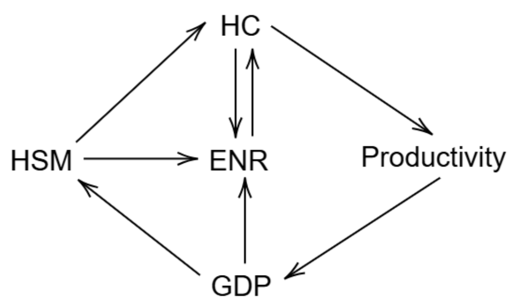



Figure 2.0.1: Mechanics of high skill migration (HSM) impact on GDP

# 3 Data and Methodology

## 3.1 Data

The spatial units of analysis are Italian provinces, which correspond to NUTS3 regions. The period covered by our analysis spans from 2004 to 2022.

[2]Similarly, $\Theta_1'$ and $\Theta_2'$ are the partial derivatives of $\Theta_{S,t}$ w.r.t. $h_{S,t}$ and $\hat{h}_{S,t}$, respectively.

Eurostat provides NUTS-3-level per capita GDP data at current market prices. Given the absence of regional GDP deflators, we used the national GDP deflator to compute GDP per capita at constant 2020 prices.

Our key regressors are the in-migration rate, the out-migration rate, and the net migration rate. Each of these measures are computed both for the overall population and for the graduate population. For province $i$ and year $t$, the in-migration rate in year $t$ for the overall population and for the graduate (high-skilled) population are defined respectively as

$$inmr_{it} = \frac{in_{it}}{pop_{it}} \tag{3.1.1}$$

$$grads\_inmr_{it} = \frac{grads\_in_{it}}{grads_{it}} \tag{3.1.2}$$

where $in_{it}$ ($grads\_in_{it}$) is the number of (graduate) individuals that, during year $t$, transfer their residence in province $i$, and $pop_{it}$ ($grads\_inmr_{it}$) is the (graduate) population of province $i$ on January 1$^{\text{st}}$ of year $t$.

Similarly, the out-migration rates are defined as

$$outmr_{it} = \frac{out_{it}}{pop_{it}} \tag{3.1.3}$$

$$grads\_outmr_{i}t = \frac{grads\_out_{it}}{grads_{it}} \tag{3.1.4}$$

Finally, the net migration rate is defined as the difference between the in-migration rate and the out-migration rate:

$$nmr_{it} = \frac{in_{it} - out_{it}}{pop_{it}} \tag{3.1.5}$$

$$grads\_nmr_{it} = \frac{grads\_in_{it} - grads\_out_{it}}{grads_{it}} \tag{3.1.6}$$

Istat provides data about population by educational attainment at NUTS-3 level only for census years. Therefore, for our period of interest, we have data on graduate residents only for 2011 and from 2018 on.[3] For non-census years, only data at NUTS-2 are available. Therefore, in order to reconstruct the missing data, we leverage the overlapping of NUTS-2 and NUTS-3 data for years 2011 and 2018, to observe the graduates distribution across provinces of each region. These shares do not change significantly between 2011 and 2018, so we assume that they are constant over time and we use them to reconstruct NUTS-3 data starting from NUTS-2 data. For clarity, an example is provided. Consider a NUTS-2 region that we will call $R$. For years 2011 and 2018, we can compute the share of graduates of region $R$ for each of its provinces. For province $i$:

$$\rho_{i2011} = \frac{grads_{i2011}}{grads_{R2011}} \quad ; \quad \rho_{i2018} = \frac{grads_{i2018}}{grads_{R2018}} \tag{3.1.7}$$

By assuming this share $\rho$ constant, we reconstruct the NUTS-3 graduate population for non-census years as:

$$grads_{it} = \rho_i \cdot grads_{Rt} \tag{3.1.8}$$

Descriptive statistics are provided in Table 3.1.1. For an easier interpretation, all the migration rates computed as showed in Equations (3.1.1) to (3.1.6) are multiplied by one thousand. Therefore, they can be interpreted as the number of in-/out-/net-flows per 1000 residents/resident graduates.

[3] Until 2011, censuses were held every ten years. In 2018, Istat introduced the 'permanent census of population and households' (*Censimento permanente della popolazione e delle abitazioni*), so starting from that year census data are available for every subsequent year.

Table 3.1.1: Descriptive statistics, 2004-2022

| | Full sample | | | | | Provinces of Centre - North | | | | | Provinces of Mezzogiorno | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Obs. | Mean | St.dev. | Min. | Max. | Obs. | Mean | St.dev. | Min. | Max. | Obs. | Mean | St.dev. | Min. | Max. |
| Per capita real GDP growth (log-difference), 2020 prices | 2033 | -0.000370 | 0.0442 | -0.228 | 0.299 | 1311 | -0.000653 | 0.0446 | -0.228 | 0.142 | 722 | 0.000142 | 0.0435 | -0.202 | 0.299 |
| Per capita real GDP, 2020 prices | 2033 | 27876.6 | 7596.2 | 14946.2 | 63454.2 | 1311 | 32023.4 | 5901.2 | 18913.3 | 63454.2 | 722 | 20346.9 | 3362.7 | 14946.2 | 32603.7 |
| Overall in-migration rate (all moves) | 2002 | 9.782 | 3.299 | 3.975 | 21.13 | 1299 | 11.28 | 2.860 | 4.784 | 21.13 | 703 | 7.005 | 2.000 | 3.975 | 13.79 |
| Overall out-migration rate (all moves) | 2002 | 9.634 | 2.358 | 3.331 | 22.67 | 1299 | 9.771 | 2.444 | 3.331 | 20.03 | 703 | 9.382 | 2.171 | 4.381 | 22.67 |
| Overall net migration rate (all moves) | 2002 | 0.148 | 2.569 | -9.587 | 8.351 | 1299 | 1.513 | 1.636 | -3.977 | 8.351 | 703 | -2.376 | 2.008 | -9.587 | 4.530 |
| Overall in-migration rate (intra-regional moves excl.) | 2002 | 6.025 | 2.332 | 1.959 | 14.43 | 1299 | 6.809 | 2.293 | 2.675 | 14.43 | 703 | 4.576 | 1.592 | 1.959 | 11.80 |
| Overall out-migration rate (intra-regional moves excl.) | 2002 | 6.041 | 2.042 | 2.260 | 15.99 | 1299 | 5.544 | 1.752 | 2.260 | 13.04 | 703 | 6.959 | 2.215 | 2.525 | 15.99 |
| Overall net migration rate (intra-regional moves excl.) | 2002 | -0.0163 | 2.289 | -8.555 | 6.334 | 1299 | 1.265 | 1.332 | -2.303 | 6.334 | 703 | -2.383 | 1.733 | -8.555 | 2.381 |
| Overall in-migration rate (intra-repartition moves excl.) | 2002 | 4.542 | 1.916 | 1.865 | 11.89 | 1299 | 5.041 | 2.069 | 2.218 | 11.89 | 703 | 3.621 | 1.115 | 1.865 | 9.734 |
| Overall out-migration rate (intra-repartition moves excl.) | 2002 | 4.648 | 1.916 | 1.724 | 14.39 | 1299 | 3.870 | 1.448 | 1.724 | 10.06 | 703 | 6.087 | 1.841 | 2.443 | 14.39 |
| Overall net migration rate (intra-repartition moves excl.) | 2002 | -0.106 | 2.166 | -7.995 | 6.106 | 1299 | 1.172 | 1.189 | -3.747 | 6.106 | 703 | -2.466 | 1.472 | -7.995 | 1.538 |
| Overall in-migration rate (moves between CN and MZ) | 2002 | 2.932 | 1.111 | 1.029 | 9.186 | 1299 | 2.748 | 1.096 | 1.029 | 7.373 | 703 | 3.273 | 1.059 | 1.626 | 9.186 |
| Overall out-migration rate (moves between CN and MZ) | 2002 | 3.041 | 2.294 | 0.626 | 13.93 | 1299 | 1.580 | 0.557 | 0.626 | 4.123 | 703 | 5.742 | 1.782 | 2.244 | 13.93 |
| Overall net migration rate (moves between CN and MZ) | 2002 | -0.109 | 2.030 | -8.001 | 4.988 | 1299 | 1.168 | 0.777 | -0.465 | 4.988 | 703 | -2.469 | 1.425 | -8.001 | 1.605 |
| High-skilled in-migration rate (all moves) | 2046 | 16.54 | 5.882 | 4.649 | 47.16 | 1299 | 18.61 | 5.404 | 5.944 | 39.27 | 747 | 12.93 | 4.849 | 4.649 | 47.16 |
| High-skilled out-migration rate (all moves) | 2046 | 19.15 | 6.477 | 5.817 | 49.14 | 1299 | 16.65 | 4.587 | 5.817 | 35.95 | 747 | 23.52 | 6.954 | 8.583 | 49.14 |
| High-skilled net migration rate (all moves) | 2046 | -2.616 | 8.471 | -35.70 | 22.30 | 1299 | 1.967 | 5.045 | -16.52 | 22.30 | 747 | -10.59 | 7.228 | -35.70 | 22.02 |
| High-skilled in-migration rate (intra-regional moves excl.) | 2046 | 10.41 | 4.805 | 2.144 | 35.33 | 1299 | 11.89 | 4.988 | 4.213 | 35.33 | 747 | 7.835 | 3.092 | 2.144 | 24.17 |
| High-skilled out-migration rate (intra-regional moves excl.) | 2046 | 12.60 | 6.291 | 3.430 | 42.61 | 1299 | 9.556 | 3.527 | 3.430 | 26.10 | 747 | 17.88 | 6.542 | 6.003 | 42.61 |
| High-skilled net migration rate (intra-regional moves excl.) | 2046 | -2.188 | 7.693 | -29.25 | 20.68 | 1299 | 2.331 | 4.056 | -7.076 | 20.68 | 747 | -10.05 | 6.018 | -29.25 | 11.62 |
| High-skilled in-migration rate (intra-repartition moves excl.) | 2046 | 7.826 | 3.924 | 1.228 | 29.73 | 1299 | 8.778 | 4.255 | 1.926 | 29.73 | 747 | 6.169 | 2.525 | 1.228 | 23.37 |
| High-skilled out-migration rate (intra-repartition moves excl.) | 2046 | 9.956 | 6.290 | 2.234 | 40.03 | 1299 | 6.295 | 2.465 | 2.234 | 19.35 | 747 | 16.32 | 5.827 | 5.792 | 40.03 |
| High-skilled net migration rate (intra-repartition moves excl.) | 2046 | -2.130 | 7.446 | -28.92 | 19.25 | 1299 | 2.483 | 3.397 | -5.386 | 19.25 | 747 | -10.15 | 5.515 | -28.92 | 11.35 |
| High-skilled in-migration rate (moves between CN and MZ) | 2046 | 5.196 | 2.416 | 0.813 | 19.94 | 1299 | 5.014 | 2.498 | 0.813 | 19.60 | 747 | 5.513 | 2.233 | 1.092 | 19.94 |
| High-skilled out-migration rate (moves between CN and MZ) | 2046 | 7.026 | 7.451 | 0.559 | 39.11 | 1299 | 2.041 | 0.736 | 0.559 | 5.036 | 747 | 15.69 | 5.721 | 5.406 | 39.11 |
| High-skilled net migration rate (moves between CN and MZ) | 2046 | -1.829 | 7.317 | -29.32 | 15.11 | 1299 | 2.973 | 2.190 | -1.596 | 15.11 | 747 | -10.18 | 5.330 | -29.32 | 8.376 |
| Observations | 2077 | | | | | 1311 | | | | | 766 | | | | |

### 3.2 Methodology

When it comes to estimate the effects of migration on growth and convergence, the literature has highlighted four major methodological challenges, as pointed out by Kubis and Schneider (2016):

1. **Endogeneity of migration:** migration and regional growth are characterized by a two-way relation;
2. **Regional heterogeneity:** unobserved regional amenities that are relevant to regional growth and correlate with migration decisions;
3. **Spatial dependence:** Heterogeneity and dependence of growth rates between spatially related units;
4. **Skill selectivity:** The human capital content of migration is typically unobserved.

The most widespread approach in the literature is represented by the IV method, which allows to control for the endogeneity of migration, to mitigate the collinearity issue and to simultaneously estimate the sign of the relation between growth and the migration rate and the human capital gain (Card, 2001; Østbye and Westerlund, 2007; Biagi et al., 2011; Fratesi and Percoco, 2014; Pinate et al., 2023; D'Ingiullo et al., 2023). To account for regional heterogeneity, most empirical research relies on regional fixed effects (Etzo, 2008; Fratesi and Percoco, 2014; Kubis and Schneider, 2016; Basile et al., 2017; Vecchione, 2018a).

## 4 Results

Before developing an IV strategy, we attempt to estimate the following baseline model of $\beta$-convergence:

$$\ln\left(\frac{y_{it}}{y_{it-1}}\right) = \alpha + \sum_i \phi_i D_i + \sum_t \phi_t D_t + \beta \ln(y_{it-1}) + \gamma nmr_{it-2} + \varepsilon_{it} \qquad (4.0.1)$$

where the dependent variable is the logarithmic approximation of per capita GDP growth rate; $D_i$ and $D_t$ are province and year dummies, respectively; $\ln(y_{it-1})$ is the initial level of per capita GDP; and $nmr_{it-n}$ is the (overall or high-skilled) net migration rate at lag $n$.

The results for Equation (4.0.1) are displayed in Table 4.0.1, where the net migration rates are computed over both the overall and the graduate population. For models (1) and (2), the net migration rates are computed by taking into account all inter-provincial migration. In (3) and (4), movements within the same NUTS 2 region are excluded. In (5) and (6), movements within the same NUTS 1 region[4] are excluded. Finally, for (7) and (8), only movements between Center-North and *Mezzogiorno* are considered.[5] Although affected by endogeneity, these results might be a good starting point to see how the model behaves and to decide what improvement are best to be implemented later on.

The results show the presence of conditional convergence: the coefficients attached to the lagged GDP per capita are negative and statistically significant, meaning that the a higher (lower) per capita GDP is associated with a lower (higher) growth rate, holding other variables constant. More precisely, a 1 percent increase (decrease) in GDP per capita is associated with a 0.026-0.032 percentage points reduction (increase) of per capita GDP growth rate. Moving to the key variable of interest, the net migration rates taken with a two-period lag, we observe non significant coefficients for the overall migration rates (models 1-3-5-7).
When taking into account the net migration rates for the high-skilled, the results change: all the coefficients are positive and statistically significant, indicating that a net inflow of graduates boosts local growth. Moreover, they increase in both magnitude and statistical significance as the minimum threshold for migration distance increases. A percentage point increase in the net migration rate of graduates (i.e. + 10 net graduate in-migrants

[4]Italian NUTS 1 regions are five: North-West, North-East, Center, South, Islands.
[5]Center-North consists of North-West, North-East and Center; *Mezzogiorno* consists of South and Islands.

per 1000 graduate residents) that move between the two macro-areas is associated with a 0.01 percentage point increase of the GDP per capita growth after two years.

Table 4.0.1: Net migration rates

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y)_{t-1}$ | -0.260*** | -0.320*** | -0.258*** | -0.321*** | -0.257*** | -0.320*** | -0.257*** | -0.320*** |
| | (0.0192) | (0.0210) | (0.0185) | (0.0209) | (0.0185) | (0.0210) | (0.0185) | (0.0212) |
| Overall net migr rate, $t-2$ | 0.316 | | 0.00255 | | -0.311 | | -0.165 | |
| | (0.568) | | (0.626) | | (0.697) | | (0.925) | |
| HS net migr rate, $t-2$ | | 0.474* | | 0.691** | | 0.763** | | 1.017*** |
| | | (0.187) | | (0.231) | | (0.260) | | (0.290) |
| Constant | 2.649*** | 3.265*** | 2.627*** | 3.270*** | 2.623*** | 3.260*** | 2.625*** | 3.267*** |
| | (0.196) | (0.214) | (0.189) | (0.213) | (0.189) | (0.214) | (0.189) | (0.217) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Adj. $R^2$ | 0.702 | 0.727 | 0.702 | 0.727 | 0.702 | 0.727 | 0.702 | 0.728 |
| N | 1994 | 1788 | 1994 | 1788 | 1994 | 1788 | 1994 | 1788 |

Standard errors clustered by province in parentheses.
$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

As pointed out by Østbye and Westerlund (2007) and Etzo (2008), assessing the impact of migration on growth by using only the net migration rate can be misleading. Indeed, the in-migration and out-migration rate could affect growth independently, especially when they are not negligible but they are of similar size and then they result in a net migration rate which is close to zero. Therefore, we run a model that estimates the gross migration rates separately:

$$\ln\left(\frac{y_{it}}{y_{it-1}}\right) = \alpha + \sum_i \phi_i D_i + \sum_t \phi_t D_t + \beta \ln(y_{it-1}) + \gamma_1 inmr_{it-h} + \gamma_2 outmr_{it-k} + \varepsilon_{it} \quad (4.0.2)$$

Following this reasoning, we try to estimate this model adopting the same lag for the two gross migration rates, so that $h = k = 2$. As showed in Table 4.0.2, all the coefficients attached to gross migration rates for the overall population are not statistically significant, confirming the results obtained by using the net migration rates (Table 4.0.1). Moving to high-skilled migration, the results seem to be more informative. In this case, the coefficients for in-migration rates are positive but not statistically significant, while those for out-migration rates are negative and statistically significant, indicating that outflows of graduates negatively affect growth. The effect is bigger for longer distances of migration.

A possible way to explain the lack of significance of the coefficients attached to the in-migration rates is that in-migration and out-migration could produce their effects at different times. In fact, the impact of out-migration on growth can be expected to manifest after a short time, because the loss in labor force and human capital is immediate. On the contrary, the effect of in-migration might need more time to kick in, as migrants might need some time to adapt to a new context before starting contributing to economic growth through their work. Therefore, we try to estimate the same model with five lags for the in-migration and two lags for the out-migration rates. The results are displayed in Table 4.0.3.

Also in this case the coefficients attached to overall migration rates are not significant, while those referred to high-skilled migration show a relevant improvement: the effect of the in-migration of graduates on growth is positive and both magnitude and significance of the coefficients increases as shorter distance moves are excluded: a 1 percentage point increase in the high-skilled in-migration rate at time $t-5$ is associated with a 0.008 - 0.019 percentage point increase of the growth rate between time $t-1$ and $t$. The coefficients attached to high-skilled out-migration rates are all negative and behave in a similar fashion, as their size and significance increase when longer distance moves are considered. More precisely, these models estimate that a 1 percentage point increase in the graduate out-migration rate is associated with a 0.009 - 0.014 percentage point reduction of the per capita

Table 4.0.2: Gross migration rates at lag 2

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y\_t - 1)$ | -0.260*** | -0.323*** | -0.258*** | -0.323*** | -0.257*** | -0.320*** | -0.257*** | -0.321*** |
| | (0.0192) | (0.0211) | (0.0185) | (0.0208) | (0.0184) | (0.0210) | (0.0187) | (0.0211) |
| Overall in-migr rate, $t-2$ | 0.196 | | -0.0216 | | 0.00962 | | -0.454 | |
| | (0.651) | | (0.720) | | (0.785) | | (1.040) | |
| Overall out-migr rate, $t-2$ | -0.505 | | -0.0374 | | 0.742 | | -0.144 | |
| | (0.772) | | (0.964) | | (1.053) | | (1.280) | |
| HS in-migr rate, $t-2$ | | 0.204 | | 0.361 | | 0.476 | | 0.625 |
| | | (0.230) | | (0.307) | | (0.362) | | (0.575) |
| HS out-migr rate, $t-2$ | | -0.939*** | | -1.122*** | | -1.031** | | -1.177*** |
| | | (0.253) | | (0.302) | | (0.334) | | (0.334) |
| Constant | 2.652*** | 3.310*** | 2.628*** | 3.302*** | 2.620*** | 3.273*** | 2.628*** | 3.279*** |
| | (0.197) | (0.216) | (0.190) | (0.212) | (0.189) | (0.214) | (0.191) | (0.215) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Adj. $R^2$ | 0.702 | 0.728 | 0.702 | 0.728 | 0.702 | 0.727 | 0.702 | 0.728 |
| N | 1994 | 1788 | 1994 | 1788 | 1994 | 1788 | 1994 | 1788 |

Standard errors clustered by province in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.0.3: Gross migration rates: in-migration at lag 5; out-migration at lag 2

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y_{t-1})$ | -0.314*** | -0.411*** | -0.313*** | -0.413*** | -0.313*** | -0.413*** | -0.315*** | -0.416*** |
| | (0.0212) | (0.0291) | (0.0213) | (0.0286) | (0.0214) | (0.0287) | (0.0216) | (0.0286) |
| Overall in-migr rate, $t-5$ | -0.932 | | -0.860 | | -1.057 | | -1.747 | |
| | (0.668) | | (0.992) | | (1.175) | | (1.461) | |
| Overall out-migr rate, $t-2$ | -0.763 | | -0.586 | | 0.0376 | | -1.157 | |
| | (0.835) | | (1.081) | | (1.178) | | (1.330) | |
| HS in-migr rate, $t-5$ | | 0.494 | | 0.821* | | 1.210** | | 1.910** |
| | | (0.277) | | (0.341) | | (0.424) | | (0.661) |
| HS out-migr rate, $t-2$ | | -0.888** | | -1.089** | | -1.138** | | -1.420*** |
| | | (0.314) | | (0.409) | | (0.430) | | (0.412) |
| Constant | 3.217*** | 4.191*** | 3.203*** | 4.207*** | 3.195*** | 4.209*** | 3.216*** | 4.235*** |
| | (0.220) | (0.297) | (0.220) | (0.291) | (0.221) | (0.292) | (0.222) | (0.290) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Adj. $R^2$ | 0.726 | 0.775 | 0.726 | 0.775 | 0.726 | 0.775 | 0.726 | 0.776 |
| N | 1775 | 1466 | 1775 | 1466 | 1775 | 1466 | 1775 | 1466 |

Standard errors clustered by province in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

GDP growth rate after two years.

## 4.1 Heterogeneous effects

### 4.1.1 Heterogeneous effects for Center-North and *Mezzogiorno*

Until now, we estimated models that did not allow the coefficients on our variables of interest (net and gross migration rates) to vary by broader regions. In other words, we did not allow the effect of migration to differ between Northern and Southern provinces. This goes against the empirical intuition derived by observing the features of Italian internal migration, which is fundamentally unidirectional (from South to North), selective by skill level, and takes place between regions with different structural characteristics. Therefore, we add to the previous models (Equations (4.0.1) and (4.0.2)) interaction terms between net and gross migration rates and a dummy equal to one for provinces of *Mezzogiorno*:

$$\ln\left(\frac{y_{it}}{y_{it-1}}\right) = \alpha + \sum_i \phi_i D_i + \sum_t \phi_t D_t + \beta \ln(y_{it-1}) + \gamma_1 net_{it-n} + \gamma_2 (net_{it-n} \times D_{south}) + \varepsilon_{it} \quad (4.1.1)$$

and

$$\begin{aligned}\ln\left(\frac{y_{it}}{y_{it-1}}\right) = {} & \alpha + \sum_i \phi_i D_i + \sum_t \phi_t D_t + \beta \ln(y_{it-1}) + \gamma_1 inmr_{it-h} + \gamma_2 (inmr_{it-h} \times D_{south}) \\ & + \lambda_1 outmr_{it-k} + \lambda_2 (outmr_{it-k} \times D_{south}) + \varepsilon_{it} \quad (4.1.2)\end{aligned}$$

Results from the model in Equation (4.1.1) do not show a significant impact on growth, and the coefficients on the interacted variable are not significant either, suggesting a homogeneous effect of migration across provinces of Center-North and *Mezzogiorno*. To check whether these results are misleading, we run model Equation (4.1.2), which allows to estimate the coefficients on in- and out-migration rates separately. We estimate this model with to different lag settings: the first with both gross migration rates at lag 2, the second with in-migration rates at lag 5 and out-migration rates at lag 2. Even these specifications do not bring evidence towards the existence of heterogeneous effects across Center-North and *Mezzogiorno*.[6]

The results discussed above do not match our theoretical expectations, as the interacted effects fail to reach statistical significance. This might be caused by a number of weaknesses in the models:

- *Spatial heterogeneity across provinces of Mezzogiorno.* Provinces hosting big cities like Naples, Bari or Palermo are different from more rural, inland provinces. Therefore, one could expect different effects across provinces with different characteristics, and an interaction with an aggregate dummy could hide true but opposite or varied effects;
- *Compensation effects across provinces.* Some provinces may be penalized by the loss of human capital, others not (e.g. based on economic structure, universities, connectivity). A single interacted coefficient fails to grasp this complexity;
- *Poor statistical power to estimate interacted effects in the presence of high noise.* The interaction increases the number of parameters and the estimated variance. With high noise and covariance between explanatory variables, even true effects may be insignificant.

### 4.1.2 Heterogeneous effects by degree of peripherality

To address these weaknesses, a classification about the degree of 'centrality' of provinces has been developed starting from the definition of inner areas (*aree interne*). Born in the Fifties, the concept of inner areas has been formalized in a more precise way with the definition of the National Strategy for Inner Areas (*Strategia*

[6] The tables in which the detailed results of those specifications are available in the Appendix.

*Nazionale Aree Interne (SNAI)*, a policy program developed in the context of the 2014-2020 Programming Cycle and later revised. Currently, inner areas are defined as municipalities that are peripheral in terms of accessibility of essential public services (healthcare, education, transport). The identification of such municipalities starts with the definition of hubs, that is municipalities which are provided with a diverse offer in terms of secondary education, at least one level 1 DEA hospital[7], and a 'silver' level train station[8]. The current classification distinguishes between hubs, inter-municipal hubs, belt, intermediate, peripheral.[9]

Since the definition of inner areas is at municipality level while our analysis is at province level, we construct an indicator that expresses the degree of peripherality of a province. This is done by computing for each province the share of resident population in 2011 living in a municipality classified as intermediate, peripheral or ultra-peripheral. Provinces with a share lower or equal to 15 percent are classified as central, those with a share between 15 and 40 percent are considered intermediate, and those with a share higher than 40 percent are considered peripheral. The choice of setting the thresholds at 15 and 40 percent is motivated by two reasons: they preserve the conceptual consistency between tags and values, and create categories of similar dimensions, as shown in Table 4.1.1. This table, together with Figure 4.1.1, also shows that the geographical distribution of these categories is not homogeneous across the two macro-areas: among sixty-nine Northern provinces, thirty-four are classified as central (49%), and only nine fall within the 'peripheral' category (13%). Conversely, among thirty-eight provinces in the *Mezzogiorno*, twenty-three are peripheral (more than 60%), and only three (8%) are classified as central (Barletta-Andria-Trani, Napoli, Cagliari). This picture highlights how a relevant share of Southern Italians suffer from the scarcity of essential public services, which are the pillars of a good quality of life and, therefore, important determinants of migration.

In order to allow for different effects of migration across central, intermediate, and peripheral provinces, we estimate the following model:

$$\ln\left(\frac{y_{it}}{y_{it-1}}\right) = \alpha + \sum_i \phi_i D_i + \sum_t \phi_t D_t + \beta \ln(y_{it-1}) + \gamma_1 net_{it-2} + \gamma_4(net_{it-2} \times D_{1i}) + \gamma_5(net_{it-2} \times D_{2i}) + \varepsilon_{it} \quad (4.1.3)$$

where $D_1$ and $D_2$ are dummies equal to one if province $i$ is intermediate or peripheral, respectively. The coefficient $\gamma_1$ represents the marginal effect of net migration rates in central provinces, while $\gamma_4$ and $\gamma_5$ show how the effect of net migration rates varies in intermediate and peripheral provinces respectively, relative to central provinces.

Table 4.1.1: Table of frequencies for degree of peripherality (rows) and macro-area (columns) of Italian provinces.

| | Center-North | *Mezzogiorno* | Total |
|---|---|---|---|
| Central | 34 | 3 | 37 |
| Intermediate | 26 | 12 | 38 |
| Peripheral | 9 | 23 | 32 |
| Total | 69 | 38 | 107 |

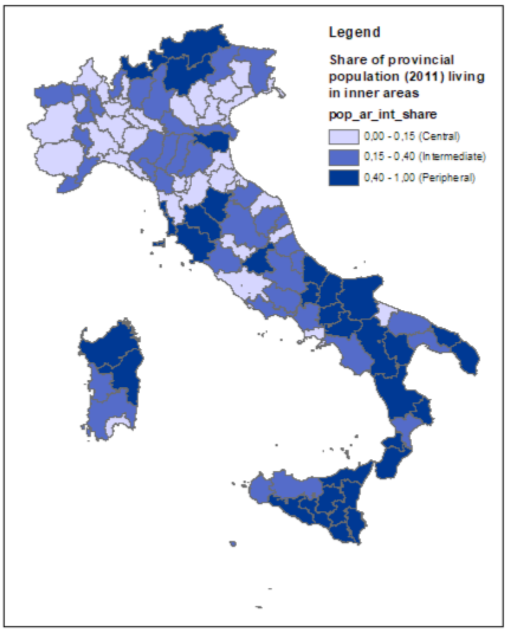
Results for the model in Equation (4.1.3) are reported in Table 4.1.2. When considering the overall net migration rates, coefficients are not significant across all categories of provinces, thus confirming the findings in Table 4.0.1. More insightful results emerge when taking into account graduate net migration rates: although no significant effect is detected for central provinces, the interacted coefficients are all positive, and their signifi-

[7]In Italy, a Level I Emergency Department (DEA di I livello) is the highest classification for an emergency and admission department within a hospital. It signifies a facility equipped to handle the most complex and critical medical emergencies, acting as a regional or supra-regional reference point for severe trauma, acute medical conditions, and other high-acuity cases.

[8]According to the Italian Railway Network (*Rete Ferroviaria Italiana*, RFI) classification, a 'silver' railway station has the following characteristics: medium-to-low passenger traffic (mostly at metropolitan and regional level); essential passenger services (basic information, ticketing, etc.); meeting of basic accessibility standards; intermodal connections; regional connectivity role.

[9]Belts include municipalities that are within 27.7 minutes away from a hub. Intermediate municipalities lie between 27.7 and 40.9 minutes away, and peripheral ones between 40.9 and 66.9. Municipalities that are more than 66.9 away from a hub are considered ultra-peripheral.

Figure 4.1.1: Share of provincial population (2011) living in municipalities classified as inner areas.




cance is stronger when all movements across provinces are considered (specification 2). Therefore, the marginal effect of an increase in the inflow of graduates is stronger in intermediate and peripheral provinces relative to central ones.

Again, following the reasoning of Østbye and Westerlund (2007) and Etzo (2008) on the drawbacks of net migration rates, the same model with gross migration rates is estimated:

$$\begin{aligned} \ln\left(\frac{y_{it}}{y_{it-1}}\right) &= \alpha + \sum_i \phi_i D_i + \sum_t \phi_t D_t + \beta \ln(y_{it-1}) + \gamma_1 inmr_{it-h} + \lambda_1 outmr_{it-k} \\ &\quad + \gamma_2 (inmr_{it-h} \times D_{1i}) + \gamma_3 (inmr_{it-h} \times D_{2i}) \\ &\quad + \lambda_2 (outmr_{it-k} \times D_{1i}) + \lambda_3 (outmr_{it-k} \times D_{2i}) + \varepsilon_{it} \end{aligned} \tag{4.1.4}$$

This model is estimated twice: the first time with both in- and -out-migration rates at two lags; the second time with the in-migration rates at five lags and the out-migration rates at two lags. The results are displayed in Tables 4.1.3 and 4.1.4, respectively.

The coefficients reported in Table 4.1.3 show no significant effect of both in- and out-migration rates on growth when overall migration is considered. The estimates for high-skilled migration rates are once again more insightful. A 1 percentage point increase in the in-migration rate is associated, for intermediate provinces, with a 0.009-0.019 percentage point increase in the growth rate of GDP per capita. Instead, for central and peripheral provinces, the effect of high-skilled in-migration is not statistically significant. As for out-migration, the effect on growth is particularly negative for intermediate (-0.019) and peripheral provinces (-0.01) when only movements towards a different macro-area are considered.

When considering in-migration rates at lag 5, the results slightly change, as shown in Table 4.1.4. Gross migration rates computed over the overall population fail to produce a statistically significant impact on growth, while some specifications for graduate migration show some significant coefficients. More in detail, when movements across all provinces are considered (model 2), peripheral provinces benefit more (0.0147 percentage point increase in growth rate) from a 1 percentage point increase in graduate inflows than intermediate and central provinces do. This result might suggest the existence of decreasing marginal returns from migration. This is especially true for short-distance moves. Conversely, the negative effect of graduate out-migration on growth is more pronounced in non-central provinces when considering migration between the two macro-areas: a 1

Table 4.1.2: Heterogeneous effect of net migration rates, by degree of peripherality.

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y_{t-1})$ | -0.260*** | -0.324*** | -0.257*** | -0.324*** | -0.257*** | -0.321*** | -0.257*** | -0.321*** |
| | (0.0192) | (0.0211) | (0.0185) | (0.0212) | (0.0185) | (0.0213) | (0.0185) | (0.0213) |
| Overall net migr rate, $t-2$ | 0.213 | | 0.466 | | 0.657 | | 0.865 | |
| | (0.820) | | (0.952) | | (1.082) | | (1.670) | |
| Intermediate × Overall net migr rate, $t-2$ | 0.270 | | -0.502 | | -1.288 | | -0.774 | |
| | (1.106) | | (1.285) | | (1.429) | | (1.970) | |
| Peripheral × Overall net migr rate, $t-2$ | 0.0639 | | -0.936 | | -1.650 | | -1.930 | |
| | (1.186) | | (1.398) | | (1.656) | | (2.271) | |
| HS net migr rate, $t-2$ | | -0.485 | | -0.451 | | -0.291 | | -0.491 |
| | | (0.366) | | (0.573) | | (0.690) | | (0.813) |
| Intermediate × HS net migr rate, $t-2$ | | 1.424** | | 1.663* | | 1.490 | | 2.358* |
| | | (0.467) | | (0.681) | | (0.829) | | (0.964) |
| Peripheral × HS net migr rate, $t-2$ | | 1.296** | | 1.389* | | 1.272 | | 1.456 |
| | | (0.457) | | (0.683) | | (0.807) | | (0.920) |
| Constant | 2.650*** | 3.308*** | 2.624*** | 3.304*** | 2.620*** | 3.280*** | 2.624*** | 3.279*** |
| | (0.196) | (0.215) | (0.189) | (0.217) | (0.189) | (0.218) | (0.188) | (0.218) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Adj. $R^2$ | 0.701 | 0.728 | 0.701 | 0.728 | 0.702 | 0.728 | 0.701 | 0.728 |
| N | 1994 | 1788 | 1994 | 1788 | 1994 | 1788 | 1994 | 1788 |

Standard errors clustered by province in parentheses.
$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

percentage point increase in graduate out-migration reduces the growth rate of GDP per capita by 0.0175 percentage points in intermediate provinces and by 0.0153% in peripheral ones. These results, together with the absence of significant effects of out-migration on central provinces, suggests that the latter suffer less from the loss of human capital since they manage to balance this loss with comparable graduates inflows. On the contrary, intermediate and peripheral provinces suffer from the loss of human capital and see their economic growth hindered by the outflow of graduates.

## 4.2 IV 2SLS estimates

Although all the models that have been estimated so far include migration rates at two or five lags, the results that they yield are likely to be biased because of endogeneity. The reason behind this concern is that migration patterns are determined by expectations about regional growth, which may be self-fulfilling in the case of selective migration. In other words, there could be a reverse causality mechanism between growth and migration.

To isolate the exogenous variability, an instrumental variable approach has been adopted. Following Fratesi and Percoco (2014) and Card (2001), we construct the following instrument for net migration rates:

$$z_{it} = \frac{in_{i2004} - out_{i2004}}{tot\_moves_{2004}} \cdot \frac{tot\_moves_t}{pop_{it}} \tag{4.2.1}$$

where $in_{i2004} - out_{i2004}$ is the net migration flow for province $i$ in 2004, $tot\_moves_t$ represents the total number of inter-provincial movements in Italy in year $t$, and $pop_{it}$ is the resident population of province $i$ in year $t$. Instruments of the same kind have been computed also for in-migration and out-migration rates, replacing the net migration flows with inflows and outflows, respectively.

Table 4.1.3: Heterogeneous effect of gross migration rates at lag 2, by degree of peripherality.

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y_{t-1})$ | -0.260*** | -0.327*** | -0.257*** | -0.325*** | -0.257*** | -0.322*** | -0.259*** | -0.323*** |
| | (0.0191) | (0.0209) | (0.0185) | (0.0208) | (0.0182) | (0.0210) | (0.0185) | (0.0208) |
| Ovrl in-migr rate, $t-2$ | -0.0209 | | 0.163 | | 0.505 | | 0.331 | |
| | (0.961) | | (1.013) | | (1.176) | | (1.802) | |
| Intermediate × Ovrl in-migr rate, $t-2$ | 0.424 | | -0.173 | | -0.910 | | -1.047 | |
| | (1.177) | | (1.228) | | (1.384) | | (1.950) | |
| Peripheral × Ovrl in-migr rate, $t-2$ | 0.354 | | -0.197 | | -0.238 | | -0.649 | |
| | (1.384) | | (1.607) | | (1.800) | | (2.246) | |
| Ovrl out-migr rate, $t-2$ | -0.740 | | -1.492 | | -1.571 | | -2.653 | |
| | (0.961) | | (1.401) | | (1.500) | | (1.911) | |
| Intermediate × Ovrl out-migr rate, $t-2$ | 0.120 | | 1.600 | | 2.464 | | 1.068 | |
| | (1.509) | | (2.187) | | (2.282) | | (2.955) | |
| Peripheral × Ovrl out-migr rate, $t-2$ | 0.461 | | 2.234 | | 3.424 | | 4.133 | |
| | (1.376) | | (1.892) | | (2.158) | | (2.701) | |
| HS in-migr rate, $t-2$ | | -0.654 | | -0.596 | | -0.351 | | -0.364 |
| | | (0.379) | | (0.564) | | (0.686) | | (0.884) |
| Intermediate × HS in-migr rate, $t-2$ | | 1.521** | | 1.769* | | 1.634* | | 2.242* |
| | | (0.471) | | (0.683) | | (0.822) | | (1.108) |
| Peripheral × HS in-migr rate, $t-2$ | | 0.877 | | 0.956 | | 0.798 | | 0.584 |
| | | (0.546) | | (0.789) | | (0.989) | | (1.383) |
| HS out-migr rate, $t-2$ | | -0.0428 | | -0.307 | | -0.156 | | 1.306 |
| | | (0.495) | | (0.913) | | (1.235) | | (1.058) |
| Intermediate × HS out-migr rate, $t-2$ | | -1.178 | | -1.060 | | -1.022 | | -3.187** |
| | | (0.598) | | (1.000) | | (1.350) | | (1.208) |
| Peripheral × HS out-migr rate, $t-2$ | | -1.151* | | -0.923 | | -0.962 | | -2.332* |
| | | (0.538) | | (0.953) | | (1.282) | | (1.124) |
| Constant | 2.653*** | 3.347*** | 2.627*** | 3.325*** | 2.624*** | 3.289*** | 2.641*** | 3.297*** |
| | (0.196) | (0.214) | (0.189) | (0.212) | (0.187) | (0.214) | (0.189) | (0.212) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Adj. $R^2$ | 0.701 | 0.729 | 0.701 | 0.728 | 0.701 | 0.727 | 0.701 | 0.728 |
| N | 1994 | 1788 | 1994 | 1788 | 1994 | 1788 | 1994 | 1788 |

Standard errors clustered by province in parentheses.

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 4.1.4: Heterogeneous effect of gross migration rates, by degree of peripherality. In-migr at lag 5, out-migr at lag 2.

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y_{t-1})$ | -0.314*** | -0.411*** | -0.313*** | -0.412*** | -0.314*** | -0.412*** | -0.315*** | -0.414*** |
| | (0.0213) | (0.0286) | (0.0212) | (0.0286) | (0.0210) | (0.0288) | (0.0212) | (0.0284) |
| Ovrl in-migr rate, $t-5$ | -0.202 | | 0.466 | | 0.493 | | 0.00758 | |
| | (0.905) | | (1.292) | | (1.569) | | (2.287) | |
| Intermediate × Ovrl in-migr rate, $t-5$ | -1.136 | | -2.306 | | -3.192* | | -2.254 | |
| | (1.168) | | (1.458) | | (1.594) | | (2.405) | |
| Peripheral × Ovrl in-migr rate, $t-5$ | -1.862 | | -2.303 | | -1.917 | | -2.537 | |
| | (1.629) | | (2.365) | | (2.855) | | (3.430) | |
| Ovrl out-migr rate, $t-2$ | -0.850 | | -1.897 | | -1.934 | | -3.497 | |
| | (1.125) | | (1.795) | | (1.923) | | (2.315) | |
| Intermediate × Ovrl out-migr rate, $t-2$ | 0.646 | | 1.965 | | 2.702 | | 1.257 | |
| | (1.673) | | (2.530) | | (2.642) | | (3.179) | |
| Peripheral × Ovrl out-migr rate, $t-2$ | -0.135 | | 1.699 | | 2.582 | | 3.415 | |
| | (1.498) | | (2.164) | | (2.455) | | (2.852) | |
| HS in-migr rate, $t-5$ | | 0.0458 | | 0.427 | | 0.593 | | 0.653 |
| | | (0.356) | | (0.473) | | (0.564) | | (0.852) |
| Intermediate × HS in-migr rate, $t-5$ | | 0.146 | | -0.0157 | | 0.144 | | 0.556 |
| | | (0.517) | | (0.731) | | (0.837) | | (1.201) |
| Peripheral × HS in-migr rate, $t-5$ | | 1.516* | | 1.364 | | 1.917 | | 2.606 |
| | | (0.703) | | (0.824) | | (1.095) | | (1.532) |
| HS out-migr rate, $t-2$ | | -0.223 | | -0.331 | | -0.0437 | | 1.943 |
| | | (0.471) | | (0.786) | | (1.040) | | (1.485) |
| Intermediate × HS out-migr rate, $t-2$ | | -0.399 | | -0.572 | | -0.955 | | -3.702* |
| | | (0.550) | | (0.868) | | (1.124) | | (1.592) |
| Peripheral × HS out-migr rate, $t-2$ | | -1.031* | | -0.951 | | -1.309 | | -3.472* |
| | | (0.470) | | (0.793) | | (1.062) | | (1.516) |
| Constant | 3.217*** | 4.189*** | 3.202*** | 4.200*** | 3.203*** | 4.197*** | 3.222*** | 4.217*** |
| | (0.221) | (0.292) | (0.220) | (0.291) | (0.217) | (0.292) | (0.218) | (0.289) |
| Province FE | | | | | | | | |
| Year FE | | | | | | | | |
| Adj. $R^2$ | 0.726 | 0.776 | 0.726 | 0.775 | 0.726 | 0.775 | 0.726 | 0.776 |
| N | 1775 | 1466 | 1775 | 1466 | 1775 | 1466 | 1775 | 1466 |

Standard errors clustered by province in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

When building the instrument for high-skilled migration rates, the same reasoning been followed:

$$z_{it}^{HS} = \frac{grads\_in_{i2004} - grads\_out_{i2004}}{grad\_moves_{2004}} \cdot \frac{grad\_moves_t}{grads_{it}} \tag{4.2.2}$$

where $grads\_in_{i2004} - grads\_out_{i2004}$ represents the net flow of graduates for province $i$ in 2004, $grad\_moves_t$ is the total number of people with a university degree moving from one province to another at time $t$, and $grads_{it}$ is the number of graduates living in province $i$ in year $t$.

To test for the relevance of the instrument, a Stock-Yogo weak ID F test is run: for each model specification, we check whether the first-stage Kleibergen-Paap F statistic is greater than 16.38. The F statistics are reported in Table 4.2.2, and they are all greater than 16.38, with the sole exception of model 2, whose instrument is severely weak. The outcomes of F-tests are consistent with the significance of the coefficients in the first-stage regressions, which are all significant at 0.1% except the one for model 2, which fails to reach any level of significance.

Since the 2SLS model that we are running is exactly identified (one endogenous variable and one instrument), it is not possible to formally test the exclusion restriction (i.e. that the instruments are related to provincial growth only through migration rates). However, the rationale of Card (2001) and Fratesi and Percoco (2014) in favor of this kind of instrument is quite convincing, as they represent the trends in inter-provincial migration flows as if the attractiveness of destination regions – as measured by the first terms of Equations (4.2.1) and (4.2.2) – remained unchanged at 2004 level. In other words, they can be interpreted as counterfactual migration rates in a world where provincial economies don't adjust in response to migration trends. This interpretation provides a theoretical argument in favor of the instruments exclusion restriction being satisfied. Finally, an additional point in support of this claim is represented by the results of the reduced-form estimates (Table 4.2.1), which are consistent with our theoretical expectations.

The second-stage results reported in Table 4.2.2 show a positive effect of net migration on growth only when movements across the five repartitions or the two macro-areas are considered. A 1 percentage point increase in the net migration rate computed over those long-range movements produces an increase of about 0.027 percentage point in the GDP per capita growth rate after two years. This result is bigger in size than the OLS estimate, which returned a 0.01 percentage point increase in the growth rate (Table 4.0.1), yet its statistical significance is weaker.

Table 4.2.1: Reduced-form estimates. Instruments for net migration rates at lag 2.

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y_{t-1})$ | -0.313*** | -0.315*** | -0.312*** | -0.315*** | -0.312*** | -0.316*** | -0.313*** | -0.316*** |
| | (0.0210) | (0.0207) | (0.0210) | (0.0207) | (0.0210) | (0.0208) | (0.0208) | (0.0208) |
| $z_{it-2}$ | 4.324* | | 2.880 | | 2.041 | | -1.126 | |
| | (2.128) | | (2.693) | | (2.826) | | (3.229) | |
| $z_{it-2}^{HS}$ | | 0.921* | | 1.104* | | 1.254** | | 1.483** |
| | | (0.459) | | (0.478) | | (0.468) | | (0.490) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Adj. $R^2$ | | | | | | | | |
| N | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 |

Standard errors clustered by province in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As with OLS estimates, we now estimate a 2SLS model with gross migration rates, which are instrumented with variables that follow the same intuition previously discussed for Equations (4.2.1) and (4.2.2). Therefore,

Table 4.2.2: Instrumental variables' estimates. Net migration rates at lag 2.

| | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *First-stage regressions* | | | | | | | | |
| $\ln(y_{t-1})$ | 0.00695*** | 0.0166*** | 0.00329** | 0.0112*** | 0.00247** | 0.00870** | 0.00251*** | 0.00689** |
| | (0.00156) | (0.00332) | (0.00100) | (0.00306) | (0.000841) | (0.00280) | (0.000660) | (0.00253) |
| $z_{it-2}$ | 0.785*** | | 1.744*** | | 1.997*** | | 1.836*** | |
| | (0.134) | | (0.128) | | (0.106) | | (0.0885) | |
| $z^{HS}_{it-2}$ | | 0.198 | | 0.439*** | | 0.473*** | | 0.548*** |
| | | (0.108) | | (0.106) | | (0.111) | | (0.119) |
| *Second-stage regressions.* *Dep. variable:* $\ln(y_t) - \ln(y_{t-1})$ | | | | | | | | |
| $\ln(y\_t - 1)$ | -0.351*** | -0.393*** | -0.317*** | -0.343*** | -0.314*** | -0.339*** | -0.311*** | -0.334*** |
| | (0.0297) | (0.0578) | (0.0212) | (0.0237) | (0.0211) | (0.0226) | (0.0216) | (0.0220) |
| Net migr rate, $t-2$ | 5.510 | | 1.651 | | 1.022 | | -0.613 | |
| | (2.830) | | (1.566) | | (1.430) | | (1.753) | |
| HS net migr rate, $t-2$ | | 4.665 | | 2.515 | | 2.650* | | 2.708* |
| | | (3.373) | | (1.348) | | (1.202) | | (1.052) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Kleibergen-Paap rk Wald F stat. | 34.27 | 3.368 | 186.1 | 17.12 | 355.1 | 18.28 | 430.3 | 21.24 |
| Adj. $R^2$ | 0.113 | -0.0651 | 0.155 | 0.134 | 0.156 | 0.138 | 0.156 | 0.149 |
| N | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 |

Standard errors clustered by province in parentheses.
$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

overall and high-skilled in-migration rates are instrumented with the following variables, respectively:

$$z_{it} = \frac{in_{i2004}}{tot\_moves_{2004}} \cdot \frac{tot\_moves_t}{pop_{it}} \tag{4.2.3}$$

$$z^{HS}_{it} = \frac{grads\_in_{i2004}}{grad\_moves_{2004}} \cdot \frac{grad\_moves_t}{grads_{it}} \tag{4.2.4}$$

Similarly, these are the instruments for overall and high-skilled out-migration rates, respectively:

$$z_{it} = \frac{out_{i2004}}{tot\_moves_{2004}} \cdot \frac{tot\_moves_t}{pop_{it}} \tag{4.2.5}$$

$$z^{HS}_{it} = \frac{grads\_out_{i2004}}{grad\_moves_{2004}} \cdot \frac{grad\_moves_t}{grads_{it}} \tag{4.2.6}$$

In the 2SLS models we are going to estimate, there are two endogenous variables ($inmr_{it-2}, outmr_{it-2}$) and two instruments ($z\_in_{it-2}, z\_out_{t-2}$. Therefore, a multi-dimensional F-test must be run in order to assess the joint relevance of the instruments. In the case of two endogenous regressors and two instruments, the Stock-Yogo threshold for the weak instruments test is 7.03. The Kleibergen-Paap rk Wald F statistics reported in Table 4.2.5 refer to the joint relevance of the two instruments and are above the threshold for all models from 3 to 8. With respect to the exclusion restriction, the theoretical rationale for using the instruments in Equations (4.2.3) to (4.2.6) is essentially the same as for instruments for net migration rates (Equations (4.2.1) and (4.2.2)). However, the reduced-form estimates reported in Table 4.2.3 show a significant association with growth only for high-skilled out-migration rates. The second stage results (Table 4.2.5) suggest the presence of a negative and significant causal relationship between out-migration and provincial growth, while there is no clear effect of in-migration on growth. More in detail, a 1 percentage point increase in the high-skilled out-migration rate causes a 0.025-0.039 percentage point decrease in the growth rate of GDP per capita after two years.

Table 4.2.3: Reduced-form estimates. Instruments for gross migration rates at lag 2.

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y_{t-1})$ | -0.313*** | -0.317*** | -0.312*** | -0.316*** | -0.312*** | -0.315*** | -0.313*** | -0.317*** |
| | (0.0210) | (0.0204) | (0.0210) | (0.0204) | (0.0210) | (0.0205) | (0.0209) | (0.0204) |
| $z\_in_{it-2}$ | 4.313 | | 1.792 | | 0.955 | | -2.318 | |
| | (2.535) | | (3.226) | | (3.522) | | (4.828) | |
| $z\_out_{it-2}$ | -4.345 | | -4.625 | | -3.381 | | 0.762 | |
| | (2.356) | | (2.991) | | (3.149) | | (3.325) | |
| $z\_in^{HS}_{it-2}$ | | -0.0782 | | 0.0290 | | 0.0272 | | -1.878 |
| | | (0.647) | | (0.792) | | (1.103) | | (1.614) |
| $z\_out^{HS}_{it-2}$ | | -1.226** | | -1.346** | | -1.359** | | -1.224* |
| | | (0.451) | | (0.450) | | (0.458) | | (0.520) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| $R^2$ | | | | | | | | |
| N | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 |

Standard errors clustered by province in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As already done for the OLS estimates, we run the same model with *inmr* at five lags. In this case, the Stock-Yogo F-test for weak instruments supports the joint relevance of the instruments across all specifications with the only exception of model 1. Kleibergen-Paap F statistics are reported in Table 4.2.8. The reduced-form estimates (Table 4.2.6) are consistent with and stronger than the ones for the previous specification (Table 4.2.3). The second-stage estimates reported in Table 4.2.8 indicate a significant impact of out-migration on provincial growth across all specifications. A 1 percentage point increase in out-migration rates produces a 0.027-0.031 drop in the growth rate of GDP per capita after two years.

# 5 Conclusion

Although further IV 2SLS specifications need to be run for a comprehensive analysis, especially to verify the existence of heterogeneous causal effects for Center-North and South and for different degrees of peripherality, the results obtained so far allow to reach some preliminary conclusions. First of all, conditional convergence is stronger when controlling for endogeneity: the coefficients attached to the lagged GDP per capita range from -0.26 to -0.32 in the OLS setting, while in the IV 2SLS they range from -0.31 to -0.39. Therefore, the OLS might underestimate the speed of convergence. Secondly, and consistently with the literature on regional migration, there is no evidence of a relevant effect of non selective net migration on growth: the coefficients fail to reach statistical significance in both OLS and IV settings. Conversely, high-skilled migration produces an effect on growth. The OLS coefficients are positive and statistically significant, and they increase in size as longer distance movements are considered, until reaching the value of 1.02 in model 8, in which only high-skilled migration between Center-North and South is considered, indicating a 0.01 percentage point increase in growth after a 1 percent point increase in high-skilled net migration. However, because of endogeneity bias, OLS might underestimate the true causal effect of skilled migration on growth, as the IV coefficient for model 8 is almost three times bigger and estimates a 0.027 percentage point increase in growth as the effect of a 1 percentage point increase in high-skilled net migration.

This pattern holds even when considering models with gross migration rates. Gross migration rates for overall migration are not significant in both OLS and IV specifications. As for skilled migration, both estimation methods fail to detect a significant effect of in-migration rates on provincial growth, while coefficients attached

Table 4.2.4: First-stage regressions. Gross migration rates at lag 2.

| | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Dependent variable:* $inmr_{it-2}$ | | | | | | | | |
| $\ln(y_{t-1})$ | 0.00435*** | 0.00703* | 0.00194* | 0.00412 | 0.00159* | 0.00286 | 0.00103 | 0.000862 |
| | (0.00119) | (0.00310) | (0.000961) | (0.00224) | (0.000788) | (0.00187) | (0.000544) | (0.00139) |
| $z\_in_{it-2}$ | 0.766*** | | 1.224*** | | 1.344*** | | 1.731*** | |
| | (0.189) | | (0.135) | | (0.138) | | (0.181) | |
| $z\_out_{it-2}$ | -0.0216 | | -0.559*** | | -0.630*** | | -0.397*** | |
| | (0.173) | | (0.131) | | (0.102) | | (0.0722) | |
| $z\_in^{HS}_{it-2}$ | | 0.686*** | | 0.780*** | | 0.793*** | | 0.690*** |
| | | (0.114) | | (0.105) | | (0.143) | | (0.195) |
| $z\_out^{HS}_{it-2}$ | | -0.0218 | | -0.0873 | | -0.0885 | | -0.0397 |
| | | (0.0860) | | (0.0562) | | (0.0508) | | (0.0491) |
| *Dependent variable:* $outmr_{it-2}$ | | | | | | | | |
| $\ln(y\_t-1)$ | -0.00261* | -0.00973** | -0.00135 | -0.00719* | -0.000886 | -0.00579* | -0.00147** | -0.00610* |
| | (0.00117) | (0.00335) | (0.000879) | (0.00284) | (0.000681) | (0.00270) | (0.000549) | (0.00274) |
| $z\_in_{it-2}$ | -0.200 | | -0.557*** | | -0.678*** | | -0.0312 | |
| | (0.158) | | (0.109) | | (0.103) | | (0.134) | |
| $z\_out_{it-2}$ | 0.420 | | 1.125*** | | 1.337*** | | 1.461*** | |
| | (0.226) | | (0.204) | | (0.170) | | (0.0942) | |
| $z\_in^{HS}_{it-2}$ | | 0.401*** | | 0.211* | | 0.115 | | -0.183 |
| | | (0.106) | | (0.102) | | (0.103) | | (0.168) |
| $z\_out^{HS}_{it-2}$ | | 0.149 | | 0.323** | | 0.367** | | 0.533*** |
| | | (0.121) | | (0.115) | | (0.116) | | (0.128) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Adj. $R^2$ | | | | | | | | |
| N | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 |

Standard errors clustered by province in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.2.5: Instrumental variables' estimates. Gross migration rates at lag 2.

| *Dependent variable:* $\ln(y_t) - \ln(y_{t-1})$ | NUTS 3 (1) | (2) | NUTS 2 (3) | (4) | NUTS 1 (5) | (6) | CN-MZ (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_{t-1})$ | -0.352*** | -0.421*** | -0.317*** | -0.349*** | -0.313*** | -0.338*** | -0.311*** | -0.329*** |
| | (0.0341) | (0.0801) | (0.0219) | (0.0268) | (0.0217) | (0.0240) | (0.0217) | (0.0242) |
| Overall in-migr rate, $t-2$ | 2.978 | | -0.527 | | -0.741 | | -1.336 | |
| | (3.562) | | (2.996) | | (3.036) | | (2.793) | |
| Overall out-migr rate, $t-2$ | -10.19 | | -4.373 | | -2.878 | | 0.158 | |
| | (6.699) | | (2.926) | | (2.668) | | (2.002) | |
| HS in-migr rate, $t-2$ | | 4.324 | | 1.085 | | 0.550 | | -3.400 |
| | | (3.869) | | (1.385) | | (1.549) | | (2.783) |
| HS out-migr rate, $t-2$ | | -7.595 | | -3.880* | | -3.569* | | -2.549* |
| | | (5.193) | | (1.759) | | (1.546) | | (1.138) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Kleibergen-Paap rk Wald F stat. | 1.798 | 1.336 | 7.373 | 7.941 | 11.77 | 10.22 | 43.53 | 7.115 |
| Adj. $R^2$ | 0.0894 | -0.198 | 0.150 | 0.129 | 0.153 | 0.135 | 0.155 | 0.121 |
| N | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 | 1751 |

Kleibergen-Paap rk Wald F statistics refer to the joint relevance of the system. Standard errors clustered by province in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.2.6: Reduced-form estimates. Instruments for in-migration rates at lag 5 and out-migration rates at lag 2.

| *Dependent variable:* $\ln(y_t) - \ln(y_{t-1})$ | NUTS 3 (1) | (2) | NUTS 2 (3) | (4) | NUTS 1 (5) | (6) | CN-MZ (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_{t-1})$ | -0.405*** | -0.415*** | -0.406*** | -0.412*** | -0.408*** | -0.412*** | -0.405*** | -0.416*** |
| | (0.0295) | (0.0293) | (0.0296) | (0.0292) | (0.0296) | (0.0289) | (0.0294) | (0.0285) |
| $z\_in_{it-5}$ | -3.200 | | -5.368 | | -8.906* | | -7.715 | |
| | (2.298) | | (3.696) | | (4.037) | | (6.162) | |
| $z\_out_{it-2}$ | -0.0637 | | -3.500 | | -1.380 | | -1.280 | |
| | (2.895) | | (3.005) | | (3.101) | | (2.919) | |
| $z\_in^{HS}_{it-5}$ | | -0.686 | | -0.420 | | -0.541 | | -1.595 |
| | | (0.859) | | (0.968) | | (1.504) | | (2.238) |
| $z\_out^{HS}_{it-2}$ | | -1.504** | | -1.483** | | -1.690** | | -1.800*** |
| | | (0.466) | | (0.484) | | (0.513) | | (0.513) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| $R^2$ | | | | | | | | |
| N | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 |

Standard errors clustered by province in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.2.7: First-stage regressions. In-migration rates at lag 5, out-migration rates at lag 2.

| | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Dependent variable:* $inmr_{it-5}$ | | | | | | | | |
| $\ln(y\_t-1)$ | -0.000885 | 0.00510 | -0.000406 | 0.00456 | -0.000407 | 0.00359 | -0.000844 | 0.00190 |
| | (0.00129) | (0.00401) | (0.000937) | (0.00292) | (0.000725) | (0.00233) | (0.000539) | (0.00163) |
| $z\_in_{it-5}$ | 1.265*** | | 1.536*** | | 1.479*** | | 1.893*** | |
| | (0.211) | | (0.169) | | (0.176) | | (0.194) | |
| $z\_out_{it-2}$ | -0.273 | | -0.269* | | -0.144 | | 0.102 | |
| | (0.154) | | (0.112) | | (0.0874) | | (0.0731) | |
| $z\_in^{HS}_{it-5}$ | | 0.725*** | | 0.879*** | | 0.905*** | | 0.798*** |
| | | (0.135) | | (0.149) | | (0.201) | | (0.173) |
| $z\_out^{HS}_{it-2}$ | | 0.00548 | | 0.00328 | | -0.00186 | | -0.0110 |
| | | (0.0635) | | (0.0550) | | (0.0442) | | (0.0248) |
| *Dependent variable:* $outmr_{it-2}$ | | | | | | | | |
| $\ln(y\_t-1)$ | -0.00186 | -0.00357 | -0.00125 | -0.00376 | -0.00103 | -0.00325 | -0.00206*** | -0.00438* |
| | (0.00103) | (0.00327) | (0.000908) | (0.00250) | (0.000689) | (0.00230) | (0.000595) | (0.00212) |
| $z\_in_{it-5}$ | -0.118 | | 0.166 | | 0.0725 | | 0.437* | |
| | (0.117) | | (0.174) | | (0.142) | | (0.200) | |
| $z\_out_{it-2}$ | 0.663*** | | 0.971*** | | 1.153*** | | 1.398*** | |
| | (0.189) | | (0.184) | | (0.159) | | (0.0938) | |
| $z\_in^{HS}_{it-5}$ | | -0.0360 | | -0.160 | | -0.160 | | -0.234 |
| | | (0.0990) | | (0.0969) | | (0.148) | | (0.235) |
| $z\_out^{HS}_{it-2}$ | | 0.475*** | | 0.540*** | | 0.539*** | | 0.675*** |
| | | (0.0901) | | (0.0752) | | (0.0766) | | (0.0858) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Adj. $R^2$ | | | | | | | | |
| N | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 |

Standard errors clustered by province in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.2.8: Instrumental variables' estimates. In-migration rates at lag 5, out-migration rates at lag 2.

| *Dependent variable:* | NUTS 3 | | NUTS 2 | | NUTS 1 | | CN-MZ | |
|---|---|---|---|---|---|---|---|---|
| $\ln(y_t) - \ln(y_{t-1})$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(y\_t - 1)$ | -0.410*** | -0.420*** | -0.413*** | -0.418*** | -0.413*** | -0.418*** | -0.409*** | -0.422*** |
| | (0.0307) | (0.0367) | (0.0296) | (0.0330) | (0.0298) | (0.0338) | (0.0299) | (0.0337) |
| Overall in-migr rate, $t-5$ | -2.640 | | -3.017 | | -5.925* | | -3.931 | |
| | (1.804) | | (2.458) | | (2.565) | | (3.228) | |
| Overall out-migr rate, $t-2$ | -1.181 | | -4.441 | | -1.937 | | -0.628 | |
| | (4.378) | | (3.046) | | (2.650) | | (2.111) | |
| HS in-migr rate, $t-5$ | | -1.103 | | -0.977 | | -1.152 | | -2.796 |
| | | (1.379) | | (1.291) | | (1.917) | | (2.938) |
| HS out-migr rate, $t-2$ | | -3.154* | | -2.740* | | -3.140** | | -2.712** |
| | | (1.279) | | (1.103) | | (1.105) | | (0.806) |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Kleibergen-Paap rk Wald F stat. | 5.871 | 14.13 | 15.89 | 16.95 | 41.77 | 9.989 | 63.43 | 9.694 |
| Adj. $R^2$ | 0.206 | 0.144 | 0.198 | 0.181 | 0.200 | 0.177 | 0.208 | 0.168 |
| N | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 | 1442 |

Kleibergen-Paap rk Wald F statistics refer to the joint relevance of the system. Standard errors clustered by province in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

to out-migration rates are significant across all specifications, except model 2 for the IV setting. Again, the IV coefficients are much larger than OLS ones: when considering only significant results, the OLS estimates that the effect of 10 additional skilled out-migrants per 1000 resident graduates on provincial growth ranges from -0.009 to -0.014 percentage points, while the IV returns a much bigger effect (from -0.026 to -0.039 percentage points).

The upward shift in the coefficient values under IV estimation can be interpreted as a correction for endogeneity bias in the OLS model. Several sources of endogeneity may be at play. First, reverse causality is a central concern: regions experiencing faster income growth are likely to attract more migrants, especially those with higher qualifications, thus inflating the correlation between migration and growth in the absence of a causal mechanism. Second, unobserved confounding factors—such as institutional quality, infrastructure investment, or local governance capacity—may simultaneously influence both migration patterns and economic outcomes. Third, the self-selection of migrants based on expected future opportunities could result in migration decisions that are endogenous to anticipated, but not yet realized, income changes. Collectively, these mechanisms induce a downward bias in OLS estimates, obscuring the full economic benefits of migration. Overall, the IV estimates highlight the significant role of skilled migration in fostering regional economic growth. Most importantly, the discussed results highlight the strong, negative effect of human capital loss through skilled out-migration on the local economies of sending regions. These findings suggest that policies aimed at attracting and retaining highly educated individuals may yield substantial economic returns, particularly in lagging regions that suffer from structural human capital loss, which in the Italian context correspond to inner areas and many Southern provinces.

## References


AA.VV. (2018). *New Frontiers in Interregional Migration Research*. Springer.

Affuso, S. and Vecchione, G. (2012). *Migrazioni intellettuali e Mezzogiorno d'Italia*. McGraw-Hill.

Barro, R. J. and Sala-I-Martin, X. (1992). Convergence. *Journal of Political Economy*, pages 223–251.

Barro, R. J. and Sala-i Martin, X. (2004). *Economic growth*. MIT Press, Cambridge, Massachusetts; London, England, second edition edition.

Basile, R., Girardi, A., Mantuano, M., and Russo, G. (2017). Interregional Migration, Human Capital Externalities and Unemployment Dynamics: Evidence from Italian Provinces.

Biagi, B., Faggian, A., and McCann, P. (2011). Long and Short Distance Migration in Italy: The Role of Economic, Social and Environmental Characteristics. *Spatial Economic Analysis*, 6(1):111–131.

Boeri, T., Brücker, H., Docquier, F., and Rapoport, H. (2012). *Brain drain and brain gain: The global competition to attract high-skilled migrants*. Oxford University Press, United Kingdom.

Bonifazi, C. and Heins, F. (2000). Long-term trends of internal migration in italy. *International Journal of Population Geography*, 6:111–131.

Borjas, G. J., Bronars, S. G., and Trejo, S. J. (1992). Self-selection and internal migration in the United States. *Journal of Urban Economics*, 32(2):159–185.

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

Di Liberto, A. (2008). Education and Italian regional development. *Economics of Education Review*, 27(1):94–107.

D'Ingiullo, D., Odoardi, I., and Quaglione, D. (2023). Stay or emigrate? How social capital influences selective migration in Italy. *Regional Studies, Regional Science*, 10(1):529–548.

Etzo, I. (2008). Internal migration and growth in Italy.

Faggian, A., Modrego, F., and McCann, P. (2019). Human capital and regional development. In Capello, R. and Nijkamp, P., editors, *Handbook of regional growth and development theories*, chapter 8, pages 149–171. Elgar, 2nd edition.

Faggian, A., Rajbhandari, I., and Dotzel, K. R. (2017). The interregional migration of human capital and its regional consequences: a review. *Regional Studies*, 51:128–143.

Faini, R. (2003). The brain drain: an unmitigated blessing ?

Fratesi, U. (2014). Editorial: The mobility of high-skilled workers – causes and consequences.

Fratesi, U. and Percoco, M. (2014). Selective Migration, Regional Growth and Convergence: Evidence from Italy. *Regional Studies*, 48(10):1650–1668.

Fratesi, U. and Riggi, M. R. (2007). Does migration reduce regional disparities? The role of skill-selective flows. *Review of Urban and Regional Development Studies*, 19(1):78–102.

Greenwood, M. J. (1975). Research on internal migration in the united states: A survey. *Source: Journal of Economic Literature*, 13:397–433.

i Martin, X. X. S. and Barro, R. J. (1995). Technological diffusion, convergence, and growth.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, 99(3):483–499.

Kubis, A. and Schneider, L. (2016). Regional Migration, Growth and Convergence – A Spatial Dynamic Panel Model of Germany. *Regional Studies*, 50(11):1789–1803.

Mocetti, S. and Porello, C. (2010). La mobilità del lavoro in Italia: nuove evidenze sulle dinamiche migratorie.

Niebuhr, A., Granato, N., Haas, A., and Hamann, S. (2012). Does labour mobility reduce disparities between regional labour markets in germany? *Regional Studies*, 46:841–858.

Nifo, A., Scalera, D., and Vecchione, G. (2020). Does skilled migration reduce investment in human capital? An investigation on educational choices in Italian regions (2001–2016). *Metroeconomica*, 71(4):781–802.

Nifo, A. and Vecchione, G. (2014). Do institutions play a role in skilled migration? the case of italy. *Regional Studies*, 48:1628–1649.

Østbye, S. and Westerlund, O. (2007). Is migration important for regional convergence? Comparative evidence for Norwegian and Swedish counties, 1980-2000. *Regional Studies*, 41(7):901–915.

Ozgen, C., Nijkamp, P., and Poot, J. (2010). The effect of migration on income growth and convergence: Meta-analytic evidence. *Papers in Regional Science*, 89:537–561.

Pinate, A. C., Faggian, A., Di Berardino, C., and Castaldi, C. (2023). The heterogenous relationship between migration and innovation: Evidence from Italy. *Industry and Innovation*, 30(3):336–360.

Puga, D. (1999). The rise and fall of regional inequalities. *European Economic Review*, 43:303–334.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3:135–146.

Vecchione, G. (2017). *Brain drain o brain gain? Un tentativo di misurazione per il Mezzogiorno*, pages 297–312. Editoriale scientifica.

Vecchione, G. (2018a). Skilled Migration and Human Capital Accumulation in Southern Italy. In Bellandi, M., Biagi, B., Faggian, A., Marrocu, E., and Usai, S., editors, *Regional Development Trajectories Beyond the Crisis*, volume 55, pages 199–210. Franco Angeli, Milano.

Vecchione, G. (2018b). *Skilled Migration and Human Capital Accumulation in Southern Italy*, volume 55, pages 199–210. Franco Angeli.