Investigaciones Regionales – Journal of Regional Research
https://doi.org/10.38191/iirr-jorr.24.035

Accepted Article

The effect of internal migration on regional growth in Italy: a dynamic spatial panel data analysis

Francesca Centofanti, Roberto Basile**, Francesca Licari***, Jacopo Pitari*****

Received: 01 de junio de 2023
Accepted: 19 de febrero de 2024

Abstract:

This study assesses the effect of internal migration on regional growth in Italy at the NUTS-3 level over the period 2002-2019. The composition of the internal migration flows of the working-age population in Italy during the sample period appears substantially heterogeneous in nationality and labor skills. The analysis considers this heterogeneity, estimating various specifications of the dynamic spatial model and controlling for the endogeneity of migration variables through a control function approach. The evidence suggests that the internal migration of Italian citizens has a positive direct and spillover impact on regional growth, slowing down the convergence process. On the contrary, there is no evidence of a significant effect of internal migration of foreign citizens. Taking the skill composition of internal movements of Italian citizens into account, the adverse impact on convergence is magnified, thus corroborating the skill-selective hypothesis. Finally, the diverging impact of internal migration increases with the distance of migration flows.

Keywords: Regional growth; convergence; migration; spatial dynamic models.
JEL Classification: F22; J61; R23; C14; C21.

El efecto de la migración interna en el crecimiento regional en Italia: un análisis dinámico de datos de panel espacial

Resumen:

Este estudio evalúa el efecto de la migración interna en el crecimiento regional en Italia a nivel NUTS-3 durante el período 2002-2019. La composición de los flujos de migración interna de la población en edad laboral en Italia durante el período de muestra parece sustancialmente heterogénea en cuanto a nacionalidad y habilidades laborales. El análisis considera esta heterogeneidad, estimando varias especificaciones del modelo espacial dinámico y controlando la endogeneidad de las variables de migración a través de un enfoque de función de control. La evidencia sugiere que la migración interna de ciudadanos italianos tiene un impacto directo y de desbordamiento positivo en el crecimiento regional, ralentizando el proceso de convergencia. Por el contrario, no hay evidencia de un efecto significativo de la migración interna de ciudadanos extranjeros. Teniendo en cuenta la composición de habilidades de los movimientos internos de ciudadanos italianos, el impacto adverso en la convergencia se magnifica, corroborando así la hipótesis de selección por habilidades. Finalmente, el impacto divergente de la migración interna aumenta con la distancia de los flujos migratorios.

[*] Department of Economics and Finance, Tor Vergata University of Rome, Rome. Italy. francesca.centofanti@students.uniroma2.eu
[**] University of L'Aquila, Italy. roberto.basile@univaq.it
[***] National Institute of Statistics (ISTAT) and La Sapienza University of Rome, Italy. licari@istat.it
[****] Tor Vergata University of Rome, Italy. jacopo.pitari@students.uniroma2.eu
Corresponding author: francesca.centofanti@students.uniroma2.eu

Palabras clave: Crecimiento regional; convergencia; migración; modelos espaciales dinámicos.
Clasificación JEL: F22; J61; R23; C14; C21.

1. Introduction

Identifying the effect of internal migration on regional growth has always attracted considerable attention in regional economics. Yet, there still needs to be a clear-cut answer to whether migration is one of the factors fostering or undermining regional convergence. The answer differs by study, showing that the structural spatial-economic impact induced by labor mobility is still an essential and appealing research topic (Brunow, Nijkamp, and Poot, 2015). In the present study, we offer further empirical evidence on this subject by focusing on the case of Italy over the period 2002-2019.

The North-South dualism in Italy has always been characterized by significant income differences and substantial migration flows from the Southern part of the country (the so-called Mezzogiorno) to more industrialized Northern regions. Massive interregional migration flows were observed in the post-II war period until the early 1970s; long-distance flows were quite exclusively movements of low-skilled workers (with primary-school educational attainment) from southern to northern cities. Interestingly, during the same period, income disparities between the North and the South decreased (Daniele, Malanima, et al., 2007), confirming the hypothesis that migration flows boost regional convergence.

With the oil crisis after 1973, internal mobility declined significantly until the mid-1990s, when a new wave of internal migration began again from the South to the North (Bonifazi, Heins, Strozza, and Vitiello, 2009; Istat, 2021). Differently from the 1950s and the 1960s, more recent flows are characterized by a strong component of highly qualified people involving a large and increasing number of workers with tertiary education and many university students (Impicciatore and Strozza, 2016; Piras, 2021).

Over the past twenty years, Italy has also become one of the leading destinations for international migrants, with a growing share of the foreign population over the total population (from 2.5% in 2002 to 8.5% in 2021). The propensity of foreign immigrants to change residence within the country is much higher than that of natives (the so-called established immigrants' secondary migrations); in 2020, respectively, 4.5 per thousand and 2.3 per thousand. They often move from their original destination to a more attractive region in terms of job opportunities and quality of life. The reasons behind the greater internal mobility of foreigners can be traced back to the dynamics of the immigration process and the greater precariousness of their living conditions, especially housing and employment (Bonifazi and Crisci, 2020).

Overall, the composition of the current internal migration flows of the working-age population in Italy appears to be characterized by substantial heterogeneity in nationality and labor skills. The importance of considering such heterogeneity of the labor force is recognized in numerous empirical studies focused on the effect of internal migration on regional economic growth and convergence (Etzo, 2008).

According to the neoclassical theory (Barro and Sala-i Martin, 2004; Badinger, Müller, and Tondl, 2004; Barro, 2015; Bouayad-Agha and Vedrine, 2010), if labor is homogeneous, the migration of workers from poor to rich areas speeds up inter-regional convergence in capital intensity and labor productivity (quantity effect). On the contrary, New Economic Geography models suggest that labor migration enforces a cumulative causation mechanism, which favors the agglomeration of economic activities (Krugman, 1991) and leads to faster growth in regions with higher initial levels of income and larger internal markets (Baldwin, 1999).

Alternative theories point out that, if labor is heterogeneous, the composition effect of migration on the convergence process is ambiguous (Romer, 1986; Stark and Lucas, 1988). In particular, if emigrants from low to highly-developed regions are more productive than the workers left behind, the loss in human capital may outweigh the increase in capital-labor ratio in the source region, causing a slowdown in growth. Moreover, if immigrants are more skilled-selective than host-region workers, the qualitative effect may dominate the quantitative impact on destination areas, despite the reduced capital-labor ratio. This skill-selectivity of migrants may "drain" people with a high human capital endowment from the poorer origin areas leading to divergence in income per capita (Reichlin and Rustichini, 1998; Shioji, 2001). However, some authors have claimed that the opportunity of migrating encourages the accumulation of human capital in the poorer regions; coupled with the return migration of skilled workers, a brain gain process may occur, reinforcing the convergence process (Maza, 2006).

Mirroring such differentiated theoretical predictions, the empirical results reported in the literature are mixed. The neoclassical predictions were confirmed to prevail across US states over the period 1940–1980 (Ganong and Shoag, 2017), across NUTS-2 Spanish regions from 1972 to 1998 (Larramona and Sanso, 2014) and from 1995 to 2002 (Maza, 2006), and across Russian regions from 1995 to 2010 (Vakulenko, 2016). For the case of Spain, Hierro and Maza (2010) also found that internal migration of foreign-born people from 1996 to 2005 did not significantly affect the convergence/divergence process across NUTS-3 regions. Østbye and Westerlund (2007) analyzed the case of Norway and Sweden from 1980–2000. They conclude that the composition effect dominated the quantity effect for Norway, so migration tends to have a divergence effect, whereas in Sweden the opposite is true. Shioji (2001) found no significant impact of migration on the convergence across the Japanese prefectures. He argues that such a migration puzzle could be explained if migrants have higher human capital than non-migrants and if the composition effect of migration overwhelms its quantity effect. Borjas (2019) surveyed some studies on the relationship between immigration and growth, showing that, despite the methodological disagreements about how to measure all the possible effects, there is a consensus on the fact that immigration has a more beneficial impact on growth when the immigrant flow is composed of high-skill workers. For the case of Italy, Piras (2013) found that the composition effect of emigration is stronger for the Southern regions, meaning that the South has not taken advantage of the quantity effect and it experienced a brain drain in favour of the Centre-Northern regions. Fratesi and Percoco (2014) found evidence in favor of the hypothesis of the detrimental effect of the skilled-selective migration flows on convergence from 1980 to 2001 between Italian regions. More generally, the results of previous studies scantly confirm the neoclassical convergence hypothesis and suggest that internal net migration contributes marginally to divergence (Ozgen, Nijkamp, and Poot, 2010).

The above mentioned studies have neglected the distance of migration flows. The migration literature clearly distinguishes between short-distance moves, mainly linked to family and housing adjustments (Thomas, Gillespie, and Lomax, 2019), and long-distance moves, dominated by employment or educationally led motives (Coulter, Ham, and Findlay, 2016). In the case of Italy, the dynamics of long-distance migration between South and North seem quite different from the dynamics regulating shorter distance migration patterns between relatively close provinces.

Spatial spillover effects are also widely neglected in the empirical literature on internal regional migration and growth. Very recently, however, some authors have applied spatial econometric models to explore the role played by migration in spatial-economic development. Spatial dynamic models have been proposed by Incaltarau, Pascariu, Duarte, and Nijkamp (2021) for the case of Romania and Kubis and Schneider (2016) for Germany. In both cases, a System-GMM method has been used to control for the endogeneity of the regressors (especially migration and human capital variables).

As mentioned above, this study aims to unveil the effect of internal mobility on regional growth and convergence in Italy at the NUTS-3 level over the period 2002-2019. It addresses the following questions related to the effects of migration on regional growth: How relevant are the skill composition and the nationality of migrants for explaining migration impacts? Does the impact of internal migration on growth increase with the distance of migration flows? Are there spatial spillover effects from internal migration in the short and the long run?

We answer these questions by estimating several specifications of the spatial dynamic model. Unlike previous studies, we use Quasi-maximum likelihood (QML) estimators instead of System GMM. We also adopt a control function approach to address the endogeneity of migration variables and propose a new instrument of the "shift-share" type.

Section 2 describes the econometric methodology, Section 3 presents the demographic and institutional data used, Section 4 reports the results of the analysis, and Section 5 concludes.

2. Methodology

2.1. A dynamic spatial model

Several specifications of dynamic spatial models can be explored. Among the alternatives, a very general one includes time lags of both the dependent and independent variables and both contemporaneous and time-lagged spatial lags. However, as Elhorst et al. (2014) points out, this generalized model suffers from identification problems and is not helpful for empirical research.

A more parsimonious model can be expressed as:

$$y_t = \tau y_{t-1} + \delta W y_t + X_t \beta + \mu + \xi_t \iota_N + \mathcal{E}_t \quad (1)$$

where y_t denotes an $N \times 1$ column vector consisting of one observation of the dependent variable (i.e., our measure of log GDP per working-age population) for every spatial unit ($i = 1, \ldots, N$) in the sample at time t ($t = 1, \ldots, T$). X_t is an $N \times K$ matrix of the explanatory variables, which here are: *i)* the working-age population (WP) growth rate, *ii)* the shares of employment in agriculture, and *iii)* the net (or gross) migration rates, computed using migration flows and working age population. $K \times 1$ vector β includes the parameters of the explanatory variables.

Coefficients τ and δ are the parameters of the dependent variable lagged in time, y_{t-1}, and in space, Wy_t.[1] The $N \times N$ matrix W is a non-negative matrix of known constants describing the spatial arrangement of the spatial units in the sample. The $N \times 1$ vector μ contains spatial specific effects, μ_i, aimed at controlling for all spatial-specific, time-invariant variables, the omission of which could bias the estimates in a typical cross-sectional study. Similarly, ξ_t denotes timeperiod specific effects, where ι_N is an $N \times 1$ vector of ones, controlling for all time-specific unit-invariant variables, the omission of which could also bias the estimates. These spatial and time-period specific effects are treated as fixed effects in the analysis because unobserved effects might be correlated with the regressors already included in the model. Finally, the disturbance term $\mathcal{E}_t$ is assumed to be i.i.d. across i and t.

As in Lee and Yu (2010), the parameters of model 1 are estimated using bias-corrected quasi-maximum likelihood (QML) estimators. The stationarity conditions on both the spatial and temporal parameters in a model like 1 are stricter than the standard condition $|\tau| < 1$ in serial models, and the standard condition $\frac{1}{\omega_{min}} < \delta < \frac{1}{\omega_{min}}$ in spatial models (with ω_{min} and ω_{max} indicating the minimum and maximum eigenvalues of the W matrix). Specifically, to achieve stationarity in the dynamic spatial panel data model 1, the characteristic roots of the matrix $(I_N - \delta W) - 1_{\tau I_N}$ should lie within the unit circle (Elhorst, 2001; Debarsy, Ertur, and LeSage, 2012), which is the case when

$$\begin{gathered} \tau + \delta\omega_{max} < 1 \; if \; \delta \geq 0 \\ \tau + \delta\omega_{max} < 1 \; if \; \delta < 0 \\ \tau - \delta\omega_{max} > -1 \; if \; \delta \geq 0 \\ \tau - \delta\omega_{max} > -1 \; if \; \delta < 0 \end{gathered} \quad (2)$$

It is also worth noticing that a value of $\tau < 1$ indicates conditional convergence: the lower the absolute value of τ, the greater the estimated rate of convergence (measured as $-ln(\tau)/t$). Moreover, if τ decreases when the migration variables are excluded from the model, we conclude that migration speeds up convergence. On the contrary, if τ increases when the migration variables are excluded from the model, we can say that migration slows down convergence (or even stimulates divergence). As mentioned in Section 1, the neoclassical theory predicts that higher net-internal migration hurts the per capita growth rate, as will gross in-migration. In contrast, gross out-migration will have a positive impact on growth. In

[1] After first estimation attempts, we decided to exclude from the econometric specification the time lag of the spatial lag, Wy_{t-1}, due to the short time series available.

this sense, higher net migration is expected to impact positively on convergence. However, if migration increases human capital in the destination regions at the expense of the regions of origin, migration could have a negative net effect on convergence.

With model 1 estimated in implicit form, economic interpretations can only be drawn from its reduced form. Assuming the invertibility of the matrix $(I_N - \delta W)^{-1}$ (known as the global interaction multiplier), the reduced form in 1 can be written as follows:

$$y_t = (I_N - \delta W)^{-1}(\tau I_N)y_{t-1} + (I_N - \delta W)^{-1}(X_t\beta + \mu + \xi_t \iota_N + \varepsilon_t) \tag{3}$$

The partial derivatives of the expected value of y for each $k - th$ variable in X in each unit i at each time t give the so-called impact matrices in the short run:

$$\left[\frac{\partial E(y_t)}{\partial \chi_{1k}} \cdots \frac{\partial E(y_t)}{\partial \chi_{Nk}}\right](I_N - \hat{\delta}W)^{-1}\hat{\beta}k \tag{4}$$

and in the long run:

$$\left[\frac{\partial E(y)}{\partial \chi_{1k}} \cdots \frac{\partial E(y)}{\partial \chi_{Nk}}\right]\left[(1-\hat{\tau})I_N - \hat{\delta}W\right]^{-1}\hat{\beta}k \tag{5}$$

These matrices are generally of full rank and not symmetric regardless of the sparsity and structure of the interaction matrix W. For the explanatory variable x_k, the diagonal elements of the matrices 4 and 5 measure the so-called "direct effects", i.e. how much a 1 unit change in the explanatory variable k for the province of origin i would affect the dependent variable for the same province i. This effect is heterogeneous across provinces in spatial autocorrelation due to higher-order feedback effects. They arise due to impact passing through neighboring provinces and back to the provinces themselves. Debarsy and Ertur (2010) call interactive heterogeneity in contrast to standard individual heterogeneity in panel data models. The magnitude of these direct effects depends primarily on the value of $\hat{\beta}k$, which is constant across the sample. Heterogeneity in short- and long-run direct effects thus comes from the diagonal elements of the matrix $(I_N - \hat{\delta}W)^{-1}$ and the matrix $\left[(1-\hat{\tau})I_N - \hat{\delta}W\right]^{-1}$ representing the magnitude of pure feedback effects in the short- and the long-run, respectively. In applied works, heterogeneity in the short-run direct effect is typically negligible compared to the value of $\hat{\beta}k$. On the contrary, in the computation of the long-run direct effect, the heterogeneity is amplified by the cumulative impacts of transitory shocks over time. However, the main question in this type of spatial econometric specification concerns the impact of a variation of an explanatory variable in a province *i* on the dependent variable in other Italian provinces, that is the indirect or spillover effect, i.e. the off-diagonal impact matrices 4 and 5. In contrast to direct effects, the central part is played here by the information content, and the structure of the interaction matrix W, which is the primary source of heterogeneity, with all the parameters kept constant across the whole sample. Again, the long-run spillover effect captures an amplified heterogeneity due to the cumulative impact of transitory shocks over time. Not surprisingly, strongly related areas are more influenced than less connected provinces. The average diagonal elements of 4 and 5 provide summary statistics for the short-run and the long-run direct effect (ADE), while the average row-sum of off-diagonal elements gives a summary indicator of the indirect (spillover) effect (AIE). The significance levels of these spillover effects are obtained via Monte Carlo simulations. Moreover, the sum of the $i - th$ row of the impact matrices (net of the diagonal element) represents the total impact on the dependent variable in province *i* due to a 1 unit change in x_k in each of the Italian provinces. The sum of the $j - th$ column of the impact matrices (net of the diagonal element) quantifies the total impact on the response variable of all provinces of a 1 unit change in x_k in province j, which is of particular interest for the present study.

2.2. A control function approach

One of the main difficulties in estimating equation 1 is to address the endogeneity of the migration variables included among the X_t regressors. In other words, the X_t variables in equation 1 can be split into two parts: $Migr_t$, including all gross or net migration variables considered as endogenous, and Z_{1t}

collecting all exogenous regressors. The primary source of potential endogeneity is the reverse causality between regional growth and migration. As migrant movements are mainly based on income opportunities, changes in regional growth prospects could be the cause rather than the effect of migration flows. As mentioned in Section 1, System-GMM estimators are widely used in the empirical growth literature to account for the endogeneity of the regressors, essentially using internal instruments (i.e. time lags of migration variables). On the contrary, this study uses a control function approach to address the endogeneity issue in the QMLE procedure, which requires the migration variables to be instrumented through an external (and excluded) instrument, say Z_{2t}.

The matrix of included and excluded instruments, $Z_t = (Z_{1t}, Z_{2t})$, is assumed to be strongly correlated with the endogenous variable $Migr_t$, but not correlated with the errors of the structural model 1, i.e. $E(\varepsilon_t|Z_t) = 0$ (orthogonality condition). In addition, we assume the rank condition for identification holds. This condition can be tested by estimating a linear reduced form for $Migr_t$ where the gross or net migration rate ($Migr_t$) is regressed on its contemporaneous spatial lag, the exogenous regressors Z_{1t}, the spatial fixed effects (θ), the time fixed effects (ι_N), and the exogenous external instrument (Z_{2t}):

$$Migr_t = \alpha \mathrm{W}\, Migr_t + \mathrm{Z}_{1t}\gamma + \mathrm{Z}_{2t}\pi + \theta + \lambda_t\iota_N + \nu_t \quad (6)$$

and rejecting the null $H_0: \pi = 0$ at a suitable small significant level.

Moreover, we assume that the i.i.d. errors in equation 6 are not correlated with the instruments, $E(\nu_t|Z_t) = 0$. The CF approach proceeds by noting that the correlation between the structural error, ε_t, and the reduced form error, ν_t, can be captured using a linear relationship: $\varepsilon_t = \nu_t\rho + \eta_t$, with $E(\eta_t|\nu_t) = 0$. Because ε_t and ν_t are uncorrelated with Z_t, it follows that η_t is also uncorrected with Z_t, and then η_t must also be uncorrelated with $Migr_t$. Therefore, we obtain a valid estimating equation by plugging $\varepsilon_t = \nu_t\rho + \eta_t$ into the structural equation 1 to get:

$$\mathrm{y}_t = \tau\mathrm{y}_{t-1} + \delta W y_t + X_t\beta + \mu + \xi_t 1_N + \nu_t\rho + \eta_t \quad (7)$$

By including ν_t as an explanatory variable in Equation 7, we obtain a new error term, η_t, that is uncorrelated with all other right-hand-side variables, including $Migr_t$. Therefore, including ν_t in the equation "controls for" the endogeneity of $Migr_t$. One can think of ν_t as proxying for the factors in ε_t that are correlated with $Migr_t$.

Since we cannot observe ν_t, we proceed with a two-step procedure. First, we run the MLE regression of equation 6 to obtain the residuals $\widehat{\nu_t}$. Then, we include these residuals in place of ν_t in equation 7. The inclusion of $\widehat{\nu_t}$ has also the advantage of producing a heteroskedasticityrobust Hausman test of the null hypothesis $H_0: \rho = 0$, which means $Migr_t$ is actually exogenous.

2.3. A new "shift-share" instrument

The application of the control function procedure to the analysis of the effect of internal migration on regional growth requires the use of a valid external instrument. As mentioned above, the central identification issue is the selection problem: immigrants do not randomly sort into locations, but rather are attracted to areas with favorable demand conditions (Jaeger, Ruist, and Stuhler, 2018). A simple comparison between high- and low-immigration areas may therefore yield a biased estimate of the impact of immigration. To address the selection problem, most studies exploit the observation that immigrants tend to settle into existing areas with large immigrant populations. Bartel (1989), for example, observed that immigrants locate near previous immigrants from the same country of origin. Card (2001), trying to identify the causal impact of immigration on natives' labor market outcomes, created for each labor market a predicted inflow based on the previous share of the immigrant population from each country of origin combined with the current inflow of immigrants from those countries of origin at the national level. The predicted inflow can be seen as a weighted average of the national immigration rates from each country of origin (the "shift"), with weights that depend on the distribution of earlier immigrants at a reference period in the past (the "shares"). The potential advantage of this specification arises from the considerable variation in the geographic clustering of immigrants from different countries of origin. Here, we propose a new instrument, adapting the logic behind the shift-share approach to the internal migration context.

Exploiting data on the internal bilateral flows of migrants from origin-province o to destinationprovince d at time t, $M_{od,t}$, we estimate the origin-time fixed effects $\phi_{0,t}$ and the destination-time fixed effects $\gamma_{d,t}$:

$$M_{od,t} = \phi_{o,t} + \gamma_{d,t} + \varepsilon_{od,t} \tag{8}$$

where these fixed effects represent the *shifts*.

Our instruments will be:

$$\widehat{Emi}_{i,t=} \sum_{d=1}^{N} w_{id,0}, \hat{\gamma}_{d,t} \tag{9}$$

$$\widehat{Imm}_{i,t=} \sum_{o=1}^{N} w_{io}, 0\hat{\varphi}_{o,t}$$

where $w_{id,0}$, is the *share* of emigrants from province i to destination d at the beginning of the period (2002) and $w_{id,0}$ is the *share* of immigrants to province i from origin o at the beginning. The instruments for the net flows are simply $\widehat{Migr_t} = \widehat{Imm}_{it} - \widehat{Emi}_{it}$ and the migration rates are obtained by dividing the instruments by the working-age population.

The sources of endogeneity of migration variables are the adjustments in local labor markets to changes in labor supply as imposed by migration flows, which react to variation in both real wages and living standards (as proxied, in this case, by regional GDP). Hence, in this case, reverse causality depends on short-run shocks in local labor markets. The instruments proposed above account for the exogenous spatial distribution of migrants. For the emigration flows, we consider the exogenous short-run shocks that occurred in destination provinces; for the immigration flows, we consider the exogenous short-run shocks that happened in the origin provinces.

This strategy should guarantee the respect of the exclusion restriction assumption.

3. Data and descriptive analysis

We built a panel database using Italian National Institute of Statistics (Istat) data. The spatial units of analysis are the 107 NUTS-3 level Italian regions (provinces). The sample period is 2002-2019, with a 3-year average frequency. Data are averaged over time to eliminates shortrun business cycle dynamics, following one of the preferred approaches in convergence studies.

GDP per working-age population is computed using total gross value added (GVA) data and the total population of ages 15-64. We also calculate the growth rate for the workingage population (WP). The employment shares in the agriculture, manufacturing, services, and construction sectors are computed using employment data in the different macro-sectors and total employment. Migration flows are calculated using data on changes of residence between provinces. Net migration rates are computed by dividing the migration flows of Italian (foreign) citizens in the age class 15-64 by the working-age Italian (foreign) population for each province:

$$migr_{it} = \frac{(immi_{it} - emi_{it})}{wp_{it}}$$

We also capture the possible asymmetrical impact of migration by differentiating between inflows and outflows in a province.

Table 1 shows summary statistics for the variables mentioned above, including the two first moments, the extreme values, and their interquartile distribution.

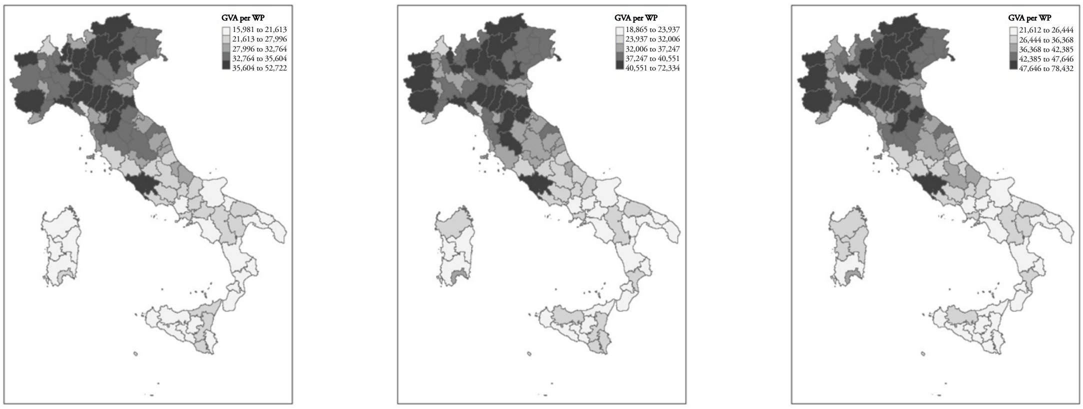
TABLE 1.
Summary statistics

Variable:	Mean	Std Dev	Min	1^{st} quartile	Median	3^{rd} quartile	Max
GVA per WP	35,126	9,962	17,026	26,130	356,99	41,446	78,432
WP growth rate	-0.415	3.394	-52.547	-0.512	-0.112	0.308	2.450
Agriculture employment share	0.054	0.042	0.002	0.023	0.042	0.075	0.212
Manufacturing employment share	0.171	0.082	0.041	0.103	0.161	0.231	0.383
Services employment share	0.686	0.069	0.503	0.639	0.688	0.733	0.889
Construction employment share	0.074	0.014	0.046	0.064	0.073	0.083	0.116
Net-migration rate (*Italians*)	$9.75e^{-05}$	0.003	-0.010	-0.002	$3.73e^{-04}$	0.002	0.10
Net-migration rate (*Italians-HK*)	$-3.12e^{-05}$	0.001	-0.004	-0.001	$9.95e^{-05}$	0.001	0.004
Net-migration rate (*Foreigners*)	-0.002	0.015	-0.089	-0.008	$-4.65e^{-04}$	0.006	0.057
In-migration rate (*Italians*)	0.024	0.008	0.006	0.018	0.023	0.031	0.043
In-migration rate (*Italians-HK*)	0.007	0.002	0.001	0.005	0.007	0.009	0.015
Out-migration rate (*Italians*)	0.024	0.006	0.005	0.019	0.023	0.029	0.043
Out-migration rate (*Italians-HK*)	0.007	0.002	0.001	0.006	0.007	0.009	0.013
In-migration rate (*Foreigners*)	0.070	0.028	0.020	0.051	0.065	0.082	0.181
Out-migration rate (*Foreigners*)	0.072	0.022	0.019	0.055	0.069	0.085	0.164
Work. age pop. (*Italians*)	333,720	354,086	50,608	144,643	229,828	373,120	2,418,270
Work. age pop. (*Foreigners*)	27,661	42,486	733	7,819	16,528	30,237	405,891

Notes: Number of observations: 642. Migration variables with the suffix *-HK* are the human-capital weighted migration rates computed as defined in Section 3.

We restrict our attention to migrants in the working-age class since they are typically considered more mobile. The decision to move is often motivated primarily by working purpose, due to the strong attraction role played by higher wages and better employment opportunities. Indeed, descriptive analysis shows that net population movements are oriented towards prosperous areas with higher real income prospects. Looking at the distribution of GVA per working-age population (Figure 1), we clearly see that Northern provinces have always experienced higher levels than Southern provinces during the years, confirming their attraction potential for migrants from the rest of Italy. As a matter of facts, all Southern provinces have negative net migration rates and almost all Northern provinces have positive rates for Italian and foreign citizens (Figure 2).

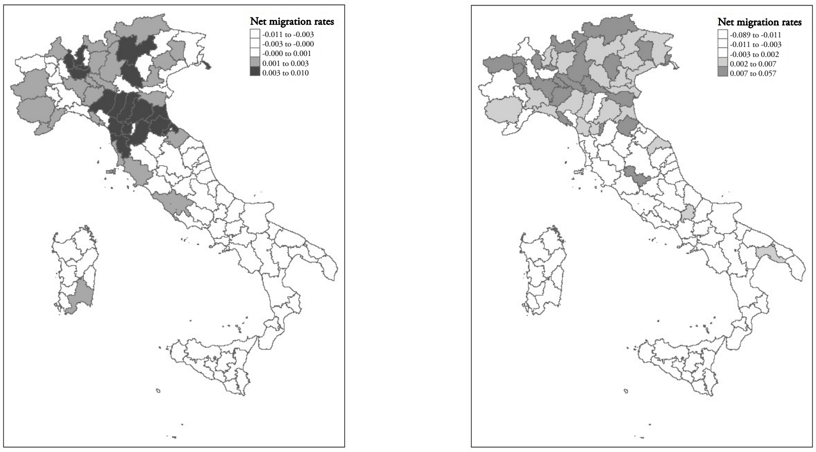
FIGURE 1.
Gross value added (GVA) per working-age population (2002, 2010, 2019)



(a) GVA per WP (2002) (b) GVA per WP (2010) (c) GVA per WP (2019)

Source: personal elaboration on ISTAT data.

FIGURE 2.
Net (average) migration rates of Italian and foreign citizens (2002-2019)



(a) Italian citizens (b) Foreign citizens (of any nationality)

Source: personal elaboration on ISTAT data.

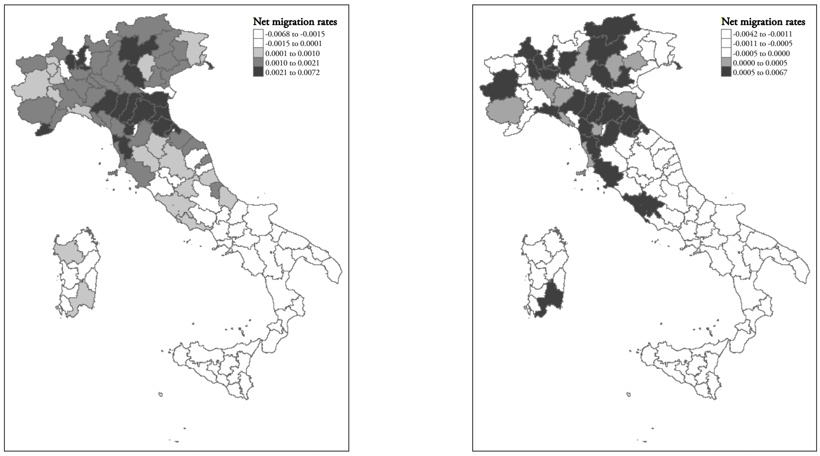
The South-North migration (long-distance movements) is generally mainly driven by differences in economic and labor market conditions, so its impact on growth is expected to be greater than the impact of shorter-distance movements. We use data on migration flows distinguished by distance thresholds to correct the measure of migration, excluding migration flows below gradually increasing cut-offs.[2]

As discussed in Section 1, one of the main objectives of our analysis is to capture the composition effects of migrants. One of the problems in assessing the role of skill selectivity of migrants in fostering or reducing regional imbalances is the measurement of migrants' skills or competencies. However, there is a general consensus in using the information on educational attainment, as employers generally use formal education as a "signal" for labor skills to reduce asymmetric information problems (Spence, 1978). In the present paper, we use the information on educational attainment for Italian migrants. As a driver of bilateral migration flows, human capital is expected to have opposite effects in the sending and the receiving regions whenever they are structural different, namely whenever regions are heterogeneous (Piras,

[2] The results are discussed in Section 4.2.

2017). Descriptive evidence shows no substantial differences in the direction of movements of Italian citizens with medium-low education level (up to the high school graduation) and Italian citizens with high education level (University degree or beyond). Indeed, regardless of education level, all the Southern provinces register negative net migration rates, while almost all of the Northern provinces see positive net migration rates.

FIGURE 3.
Net (average) migration rates of Italian citizens by the level of education (2002-2019)



(a) Low-medium educational level (b) High educational level

Source: personal elaboration on ISTAT data.

We cannot capture skill heterogeneity in migration flows of foreign citizens due to the lack of reliable measures of the education level. Nevertheless, it is widely recognized that, despite their educational attainment, foreign immigrants are largely "perceived" as low-skilled workers by local employers and are mainly concentrated in low-wage, low-productivity occupations. Therefore, while net internal migration rates of natives are expected to positively impact on growth because of their higher skill intensity, net internal migration rates of foreign workers are expected to hurt growth.

Following Barro and Lee (2010), we capture the human-capital endowment of migrants computing a human-capital weighted migration rate (only for Italian citizens):

$$migr_{it}^{hk} = \frac{\left(\sum_{k=1}^{3} immi_{it,k} D_k - \sum_{k=1}^{3} emi_{it,k} D_k\right) / \sum_{k=1}^{3} D_k}{wp_{it}}$$

where $immi_{it,k}$ and $emi_{it,k}$ are the number of immigrants and emigrants of school level k. Duration D_k is the number of years spent in schooling necessary to achieve a particular educational level k: $Dk = 8$ for lower-secondary education, $D_k = 13$ for higher-secondary education, and $D_k = 18$ for tertiary education.

Among the covariates in the growth regression model, the sectoral employment shares were also considered.[3] In the context of regional growth dynamics, the industrial structure plays a pivotal role, as extensively discussed in the development economics literature. Overall, the industrial structure serves as a foundation for economic growth, shaping the trajectory of development and influencing various aspects of the economy, from innovation and productivity to employment and trade. In the early stages of economic development, agriculture typically plays a significant role in driving growth, especially in

[3] After preliminary estimates, we opted to omit the shares of employment in manufacturing, services, and construction sectors from the econometric specification. The decision was driven by the observation that the associated coefficients failed to reach statistical significance across almost all model specifications.

agrarian economies (Johnston and Mellor, 1961). As economies develop, the relative importance of agriculture tends to decline, and the sector's contribution to overall GDP growth diminishes, giving way to manufacturing, often considered a driver of economic growth due to its potential for productivity gains and value addition (Syverson, 2011). The transition towards manufacturing is associated with structural transformation, characterized by increases in manufacturing productivity and employment share (Cantore, Clara, Lavopa, and Soare, 2017). More recently, services have become the largest sector in most modern economies, contributing significantly to GDP growth and employment generation (Griliches, 2008). Services tend to concentrate in urban areas, where economies of scale and agglomeration effects support service provision and consumption. This stylized narrative of structural transformation embodies a recurring theme within the literature, delineating the interplay between GDP per capita growth and shifts in employment patterns across sectors (Herrendorf, Rogerson, and Valentinyi, 2014). Specifically, it underscores the empirical association linking rising GDP per capita with declining employment shares in agriculture, juxtaposed with increasing shares in manufacturing and services.

In the specific case of Italy, the industrial landscape has historically been a cornerstone of economic development, notably in regions such as Lombardy, Emilia-Romagna, and Veneto. These regions have fostered manufacturing excellence, driving innovation, productivity, and export competitiveness. However, Italy's industrial structure also faces challenges, particularly concerning regional disparities. Northern regions, characterized by industrial clusters and innovation hubs, tend to have higher productivity levels and economic performance compared to southern regions, which lag behind in terms of infrastructure, human capital, and economic diversification. Indeed, despite the importance of manufacturing, Italy's agricultural sector remains relevant, particularly in rural regions, although its share of GDP has declined over time. In light of these considerations, incorporating the Italian provinces' industrial structure into the model allows for a more accurate analysis of the country's growth dynamics.

Finally, spatial lags of dependent and independent variables have been computed using an inverse-distance matrix[4], whose general term is defined as:

$$w1_{ij} = \begin{cases} 0 & if\ i = j\ and\ if\ d_{ij} > \bar{d} \\ d_{ij}^{-1} / \sum_{j \neq i} d_{ij}^{-1} & otherwise \end{cases} \tag{10}$$

where d_{ij} is the great circle distance between the centroids of the provinces and d is a cutoff value, corresponding to the minimum distance allowing all provinces to have at least one neighbor.

4. Econometric results

4.1. Baseline results

In Table 2, we report the estimation results of the dynamic spatial lag model 1, obtained using a Quasi-Maximum Likelihood (QML) estimator, without control for the endogeneity of the migration variables.[5]

[4] We also use a contiguity matrix for robustness check. The results (available upon request) are robust to this alternative spatial weight matrix.

[5] The Likelihood-ratio test for the spatial model specifications shows that the SAR (spatial autoregressive) model outperforms the Spatial Durbin model.

TABLE 2.
QMLE results - Net rates: no control for endogeneity of migration

Dependent variable: log of GDP per WP	(1) total migration	(2) human capital	(3) human capital no spatial	(4) without migration	(5) without migration no spatial
$Y_{(t-1)}$	0.530***	0.524***	0.615***	0.609***	0.733***
	(0.044)	(0.044)	(0.069)	(0.043)	(0.090)
WY_t	0.243***	0.238***		0.257***	
	(0.035)	(0.034)		(0.034)	
WP growth rate	-0.018**	-0.019***	-0.021***	-0.011*	-0.014**
	(0.007)	(0.007)	(0.008)	(0.006)	(0.009)
Agriculture employment share	-1.374***	-1.413***	-1.277***	-1.261***	-1.109***
	(0.266)	(0.267)	(0.256)	(0.261)	(0.253)
Net-migration rate (*Italians*)	4.949***	15.162***	15.814***		
	(1.172)	(3.354)	(3.835)		
Net-migration rate (*Foreigners*)	0.059	0.134	0.061		
	(0.186)	(0.181)	(0.178)		
Implied convergence speed (%)	10.573	10.761	8.097	8.255	5.178
BIC	-2331.45	-2334.41	-1872.61	-2305.69	-1902.73

Notes: The spatial weights matrix adopted is an inverse distance matrix. Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels, respectively. All models include spatial and time fixed effects. Variables are defined in Section 3 of the main text.

Different model specifications were employed. For all of them, the time-lag parameter of the dependent variable (τ) is significant but lower than one, indicating conditional convergence. The parameter of the spatial lag of the dependent variable (δ) is also positive and significant, indicating the existence of spatial contagion effects. The sum of the two parameters (τ and δ) is lower than 1, suggesting that controlling for spatiotemporal persistence is sufficient to satisfy the stationarity condition. In the model without the spatial lag (Column 3), τ is higher, confirming the prediction of Ertur and Koch (2007), according to which spatial interdependence reduces interregional convergence (the implied convergence speed decreases substantially) because of local technological interactions between economies which favor the proximity to other economies with high initial conditions (i.e., economies closer to the technological frontier). Across all model specifications, the employment share in agriculture exhibits statistical significance and, as expected, a negative effect on GDP per capita growth. The adverse impact of the agricultural sector on provincial economic growth can be attributed to the comparatively lower levels of innovation and knowledge intensity associated with agriculture (Di Berardino, D'Ingiullo, Pozzi, Quaglione, and Sarra, 2020).

The estimation results provide evidence of a positive and significant effect of the net-internal migration of Italian citizens on regional growth (Column 1). This corroborates the conclusion of much of the literature that, on average, migration has a positive effect on GDP per capita (Campo, Forte, and Portes, 2018; Ozgen, Nijkamp, and Poot, 2010), in contrast to the neoclassical prediction. On the contrary, the net-internal migration rate of foreign citizens does not enter significantly, in line with the evidence for Spain (Hierro and Maza, 2010). Moreover, the relevance of the skill-selection effects of migration emerges by comparing the coefficients of the net-internal migration of Italians in Columns (1) and (2). The estimated parameter of HK-weighted net migration is higher than that of the unweighted variables, in line with Fratesi and Percoco (2014), who showed that internal migration flows in Italy are selective and biased toward the most qualified people moving from the South to the North. The relevance of migration of human capital is widely recognized in growth studies, even if there is no agreement on the nature of the effects of brain drain in the sending regions (Milio, Lattanzi, Casadio, Crosta, Raviglione, Ricci, and Scano, 2012). Beine, Docquier, and Rapoport (2008) provided evidence that areas with high skilled emigration rates are the most vulnerable, as they are the least likely to benefit from brain gain out of brain drain. Indeed, the losers of the brain drain are often those provinces that have already experienced large outflows of skilled workers (Marchiori, Shen, and Docquier, 2013), as this is the case in the southern Italian provinces. The Bayesian Information Criterion (BIC) reaches its lowest value with the HK-weighted migration SAR model (Column 2), confirming the better performance of the skill-selectivity specification in explaining the migration effects. Our findings reinforce the divergence effect of migration, as the convergence speed increases when the human capital measure is accounted for in the migration rates, showing that the composition effect prevails over the neoclassical quantitative effect. This is also in line with some studies (e.g. Haque and Kim, 1995) arguing that brain drain causes a reduction in the rate of growth of per capita income and, thus, in levels of prosperity. Excluding the net migration rates (Column 4), τ increases and thus the implied speed of convergence decreases, indicating a diverging impact of migration of Italian citizens, in line with the evidence for Romania (Incaltarau, Pascariu, Duarte, and Nijkamp, 2021).

Removing migration variables and the spatial lag term (Column 5), the implied speed of convergence decreases even more, showing that spatial interdependences contribute to migration amplifying regional disparities. Even if the role of spatial spillovers in the convergence process is still little addressed in empirical studies, the existing literature confirms that a region grows faster when its neighboring regions exhibit high productivity levels (Kubis and Schneider, 2016) and the clustering tendency of regions with a similar development level undermines convergence.

We repeat the estimates using the control function approach to address the endogeneity problem. Before going into detail about the coefficient estimates, we need to establish whether our instruments are significantly related to migration variables. A valid instrument must meet two conditions (Baum, 2006): *i)* not be correlated with the error term in the main performance equation, except through control variables included in the regression (instrument exogeneity), and *ii)* be correlated with the endogenous variable (instrument relevance).

To assess the relevance of these external instruments from a statistical viewpoint, we report in Table 3 the *F-test* results of the first stage of the control function approach (model 6), where each potentially endogenous variable included in the model is considered alternatively as the dependent variable and regressed on the exogenous variables and the external instruments. The *F-statistic* shows that the coefficients associated with the external instrumental variables are jointly significant. While control function estimates do not provide weak-instruments tests, we can trust the results of these estimates to conclude that our external instruments are strongly correlated with the potentially endogenous variables included in the model.

As emphasized by Wooldridge (2014), because instrument exogeneity involves the covariance between the instrument and the unobserved error in equation 1, we cannot generally hope to test this assumption. Even if we do not have a formal test as to the exogeneity of the chosen instrument, such as the Hansen overidentification test[6], in the vast majority of cases, we maintain the assumption of instrument exogeneity by appealing to economic behavior. Indeed, a shift-share instrument uses earlier migrants' settlement patterns to identify information (Card, 2009). As a result of the tendency for new migrants to move to the same provinces as earlier migrants from the same province of origin, the number of migrants arriving in a province over a given interval of time is fairly predictable. If the migration rates from each source province are exogenous to conditions in a specific province, then the predicted migration flows will be exogenous.

[6] This test cannot be implemented here because we have only just-identification, that is, one instrument for one endogenous explanatory variable.

Table 3 reports the estimation results of the second stage of the control function procedure (model 7), obtained using QMLE.

TABLE 3.
QMLE results: Control function - Second stage. HK-weighted migration only

Dependent variable: log of GDP per WP	(6) No spatial		(7) SAR
$Y(t—1)$	0.468****		0.525***
	(0.044))		(0.044)
WY_t			0.236***
			(0.034)
WP growth rate	-0.008		-0.018**
	(0.007)		(0.007)
Agriculture employment share	-1.361***		-1.400***
	(0.259)		(0.264)
Net-Migration rate (*Italians*)	10.657**		11.649**
	(3.909)		(4.109)
Net-Migration rate (*Foreigners*)	-0.205		0.195
	(0.234)		(0.237)
First-stage residuals (*Italians*)	12.832*		4.381
	(4.995)		(5.152)
F-test		337.47***	
		[0.000]	
First-stage residuals (*Foreigners*)	0.368		-0.104
	(0.311)		(0.292)
F-test		567.03***	
		[0.000]	

Notes: The spatial weights matrix adopted is an inverse distance matrix. Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels, respectively. All models include spatial and time fixed effects. Variables are defined in Section 3 of the main text. For the first stages, the same model specification has been used.

Again, different model specifications were employed. Given the previous results, we focus on models with *human-capital-augmented* migration measures. We first estimate the model without control for spatial dependence (Column 6). The HK-weighted net migration rate parameter is still positive and significant but lower than that without control for endogeneity, indicating an upper endogeneity bias. This direction of difference is expected, since studies that use instrumental variable (IV) estimation methods yield smaller coefficients of net migration in growth regressions (Ozgen, Nijkamp, and Poot, 2010), highlighting the importance of two-stage estimation techniques to overcome the reverse causality problem in the relationship between migration and growth. The dimension of the difference sizes the importance of accounting for endogeneity. The residuals of the first step enter significantly, confirming the endogeneity of the migration term. The results with the SAR model (Column 7) confirm the positive effect of the net-migration rate of Italians. The residuals of the first step are not significant after controlling for the spatial dependence, suggesting that the SAR model, to some extent, filters the endogeneity bias out. The spatial and temporal fixed effects introduced in Model 1 cannot capture all the unobserved heterogeneity and omitted time and space-varying variables, which could underlie the endogeneity bias. For example, shocks could simultaneously affect migration choices and growth dynamics, generating endogeneity in the net migration variable inserted in the model. As a space and time-varying variable, the spatial lag term may have partially captured the effect of unobserved heterogeneity and these omitted time and space varying variables.

Finally, we report in Table 4 the direct, indirect, and total marginal effects of our main migration variable (the HK-weighted net migration of Italian citizens).

Table 4.
Marginal effects of HK-weighted Net-Migration rate of Italians

	No control for endogeneity			Control function		
Short run	DIRECT	INDIRECT	TOTAL	DIRECT	INDIRECT	TOTAL
	15.437***	4.533***	19.970***	11.856***	3.399**	15.254**
	(3.393)	(1.287)	(4.475)	(4.151)	(1.346)	(5.404)
Long run						
	35.134***	30.490**	65.623***	26.899***	22.467**	49.366**
	(7.804)	(11.326)	(17.696)	(9.466)	(9.944)	(18.678)

Notes: Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels, respectively.

We first look at the effects computed from the SAR model specification 3, without control for endogeneity (Column 2 in Table 2). The marginal effects are all significant and have the expected (positive) sign. Furthermore, the long-run effects are larger than their short-run counterparts. The results are robust to controlling for endogeneity: the estimated elasticities are still positive and significant, but the marginal effects are lower in the short and in long run, confirming the upper endogeneity bias. Focusing on the results obtained after applying the control function procedure, we observe that a 1% increase in the net migration rate in a province generates, on average, an increase in the growth rate of that province (average direct effect, ADE) of about 11.9% in the short run and of 27% in the long run. Moreover, a 1% increase in the net migration rate in a province generates an increase in the growth rate of the other Italian provinces (average indirect effect, AIE) of about 3.4% in the short run and of approximately 22.5% in the long run. Thus, the average total effect (ATE) of the human-capital augmented net migration rate is 15.3% in the short run and 49.4% in the long run.

Overall, spatial spillovers (indirect effects) highlight the interdependence of local economies and the importance of considering regional dynamics in economic analysis and policymaking. By understanding the mechanisms through which idiosyncratic shocks (such as net immigration) propagate across space, policymakers can better address regional disparities, promote inclusive growth, and enhance the resilience of interconnected labor markets to external shocks.

In principle, any type of network linkages can drive spatial productivity growth spillovers. Labor mobility in a broad sense (not only migration) plays a central role in driving spatial spillovers. Workers may commute between neighboring areas in response to economic opportunities, creating linkages between local labor markets (Ray, Haqiqi, Hill, Taylor, and Hertel, 2023). The stronger these linkages, the more interconnected the labor markets become, favouring other spillover channels. Indeed, the social networks established by migrants often span across geographical boundaries, facilitating the flow of information, resources, and entrepreneurial opportunities between regions. These social connections can serve as conduits for knowledge exchange, trade collaboration, and resource mobilization, fostering regional development beyond the immediate impact of migration.

Indeed, one key mechanism is the diffusion of knowledge, skills, and technologies. When migrants bring their human capital and expertise to a new region, they often contribute to the local labor force and innovation ecosystem. This influx of talent can enhance the productivity and competitiveness of local industries, generating positive spillover effects that extend to neighboring regions (Rosenthal and Strange, 2004). For instance, the adoption of innovative practices by firms in the recipient region can inspire similar advancements in neighboring areas, leading to increased economic activity and growth.

Another important transmission channel of spatial spillover is trade linkages. Migration can stimulate investment and consumption patterns, which in turn, can generate economic linkages across regions. As migrants settle and integrate into their new communities, they contribute to local demand for goods and services, creating opportunities for businesses in neighboring regions to expand their customer base and supply chains. This interregional trade and investment can amplify the economic benefits of migration, fostering growth in both the recipient region and its neighbors. Additionally, enhanced infrastructure and transportation networks can augment connectivity between regions, reducing spatial barriers and facilitating the flow of goods, services, workers and ideas. This can lead to spillovers in productivity, economic development, and innovation.

4.2. Distance thresholds

Different migration movements may respond differently to economic incentives. Short-distance migration flows (e.g. within the same region) are notoriously greater than long-distance ones (e.g. interregional or South-North migrations). In the case of Italy, for instance, Bonifazi and Heins (2000) detected differences in the features of short-distance and long-distance interprovincial migration for the time span 1955-1995. Biagi, Faggian, and McCann (2011) found that economic/labor market variables play a dominant role in long-distance migration decisions. Moreover, provinces with a local university, a better-educated population (human capital), and more affordable houses are preferred. The results differ for short-distance migration, primarily towards relatively smaller provinces, with people giving more weight to differences in quality of life and amenities.

It is, therefore, appropriate to verify whether the above results are robust to eliminating migration flows below a certain distance threshold. One way to correct the migration flow measure would be to exclude intra-regional migrations. However, this criterion is not without criticisms, as some inter-regional movements between municipalities close to the borders between two regions would cover a smaller distance than some intra-regional flows. Therefore, we prefer to correct the measure of migration by excluding migration flows below gradually increasing thresholds. In particular, we use the thresholds from 100 to 500 kilometers (km), with steps of 100 km.[7] The marginal effects of the HK-weighted migration of Italian citizens, computed using the SAR model 3, are reported in Table 5.

[7] See *Appendix A* for further details on the use of distance cut-offs in the literature.

TABLE 5.
Marginal effects of HK-weighted Net-Migration rate of Italians - Distance thresholds

	No control for endogeneity			**Control function**		
Short run	DIRECT	INDIRECT	TOTAL	DIRECT	INDIRECT	TOTAL
Above 100 km	15.535***	4.153***	19.688***	12.189**	3.357**	15.546**
	(3.445)	(1.204)	(4.425)	(4.133)	(1.459)	(5.462)
Above 200 km	16.879***	4.444***	21.323***	14.608***	3.997**	18.606**
	(3.641)	(1.256)	(4.633)	(4.736)	(1.727)	(6.306)
Above 300 km	16.936***	4.505***	21.441***	16.866****	4.662**	21.529***
	(3.641)	(1.285)	(4.661)	(5.376)	(1.981)	(7.175)
Above 400 km	18.127***	4.786***	22.913***	17.905***	4.921**	22.825***
	(3.791)	(1.334)	(4.821)	(5.309)	(2.009)	(7.120)
Above 500 km	18.363***	4.971***	23.334***	22.842****	6.430**	29.273***
	(4.005)	(1.401)	(5.104)	(6.604)	(2.533)	(8.886)
Long run						
Above 100 km	38.136***	33.099**	71.234***	29.339**	25.750*	55.089**
	(8.527)	(13.046)	(19.765)	(10.263)	(14.898)	(23.944)
Above 200 km	41.220***	34.872**	76.092***	35.051***	30.422*	65.472**
	(8.915)	(13.479)	(20.317)	(11.806)	(17.863)	(28.237)
Above 300 km	41.605***	35.990**	77.595***	40.834***	36.376*	77.210**
	(9.000)	(14.209)	(21.121)	(13.551)	(21.318)	(33.174)
Above 400 km	44.297***	37.713**	82.010***	43.307***	38.266*	81.574**
	(9.278)	(14.681)	(21.618)	(13.432)	(21.843)	(33.528)
Above 500 km	45.814***	41.537**	87.350***	56.347***	52.792*	109.139**
	(9.992)	(16.708)	(24.625)	(17.083)	(29.640)	(44.410)

Notes: Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels, respectively.

Since we are considering only migrants in the working-age class, these results seem to confirm the hypothesis that longer-distance movements are mainly related to economic and labor market conditions. Indeed, as expected, the higher the distance threshold, the higher the impact of migration on growth. The results are robust to controlling for endogeneity.

4.3. Further robustness checks

We repeated all the estimates including the log of Gross Fixed Investments (GFI) on value-added ratio as covariate in the model. Capital accumulation is recognised to play a role in stimulating productivity growth, technological innovation, and overall economic development, as elucidated by both neoclassical and endogenous growth theories. Consequently, migration patterns are closely linked to capital dynamics, as workers often move in search of better economic prospects in regions with higher levels of growth opportunities. Unfortunately, the National Institute of Statistics (ISTAT) provides data on Gross Fixed Investments (GFI) only at NUTS2 level and not at NUTS3. The information at NUTS3 level was estimated using microdata (from AIDA database). With these balance sheet data, yearly weights for each province on total regional fixed investments were first computed. Then, these weights were used to disaggregate at NUTS3 level the NUTS2 GFI data provided by ISTAT. The results obtained including the log of the Investment on Value-added ratio (based on this estimate) are reported in the new *Appendix B* and are strongly consistent to those reported in the main text. Since the NUTS3 GFI level is a reconstructed measure, we preferred to report these results only in Appendix as robustness check.

We also repeat the estimates using the gross migration rates of Italian and foreign citizens as migration variables, to look separately at in-migration and out-migration and size the growth effect of migration. The estimation results of the dynamic spatial model 1, obtained using QMLE without control for endogeneity, are reported in *Appendix C*. As before, we first look at the model with the unweighted migration variables. Differentiating between immigration and emigration variables reveals a symmetric impact: the effects of the in-migration rate of Italian citizens on regional growth are positive and significant, while the out-migration rate has a significant negative impact. Again, the gross-internal migration rate of foreign citizens does not enter significantly. This reinforces our findings that migration tends to increase regional disparities, in line with the existing literature showing that “emigration” economies are penalized in favor of the “immigration” ones (Huber and Tondl, 2012). Indeed, emigration provinces are often also provinces with a low GDP per capita, while immigration provinces have a higher GDP per capita. Let us look at the model with HK-weighted gross migration. The results are overall confirmed and, as for the net migration rates, the coefficient of the HK-weighted gross migration is higher than that of the unweighted variables, reinforcing the evidence that the growth impact of migration is greater when it refers to highly skilled workers. These estimates support the argument that migration lowers growth in the source economy when the highly skilled workers emigrate, and the opposite happens in the destination economy (Drinkwater, Levine, Lotti, Pearlman, et al., 2003).

A significant source of bias could be the existence of regional heterogeneous responses to common shocks (Basile, Girardi, Mantuano, and Russo, 2017). In fact, different provinces may react to business cycles or other time-varying (common) shocks in different ways, and this heterogeneity affects migration and growth. We re-estimate Model 1 including interactions between the dummy South and the year dummies, to capture the North-South heterogeneous responses to common business cycle effects. The estimation results, reported in *Appendix D*, confirm the robustness to common-factors effects. In particular, the estimate of the migration parameter of HK-weighted net migration suggests that unobserved time-invariant and Southspecific factors do not confound the negative impact of skill-selective migration.

5. Conclusions

This work has investigated the role of internal migration in affecting the regional growth and the convergence process in Italy over the years 2002-2019. The Italian case is interesting because of Italy’s distinctive pattern of disparities and migrations (Fratesi and Percoco, 2014). The net population movements are mainly oriented from the poorer provinces in the South to prosperous provinces with higher real income prospects in the North. Moreover, the composition of the internal migration flows of the working-age population during the sample period appears substantially heterogeneous in nationality and labor skills. The present paper has added to the previous literature by explicitly considering this heterogeneity. It has further contributed to the literature by considering spatial spillover effects to explore the role played by migration in spatial-economic development. In the analysis, various specifications of a dynamic spatial model have been estimated with QML estimators and we control for the endogeneity of migration variables through a control function approach.

It has been shown that migration of Italian citizens has a positive direct and indirect impact on GDP per capita growth, slowing down the convergence process. On the contrary, there is no evidence of a significant effect of migration of foreign citizens. A positive sign of a net migration coefficient is consistent with the perspective of the new endogenous growth theories and the new economic geography (which emphasize the strengthening benefits of agglomeration) rather than with the neoclassical model with homogenous labour. The nature of the mechanisms through which net migration increases growth still has to be further explored. The impact of migration on capital accumulation and technological change would be central issues in this context. The composition of the migration flows in terms of skills play an essential role too. Indeed, using an HK-weighted measure of net migration rate, internal movements of Italian citizens have an even higher positive impact on growth (a negative effect on convergence), thus supporting the skill-selective view (Haque and Kim, 1995). The results are robust after controlling for possible endogeneity problems. It is also interesting to remark that the spatial lag model controls for the endogeneity of net migration. The spatial lag term indeed captures the effects of unobserved space and time-varying variables, which may generate the endogeneity bias.

References

Badinger, H., Müller, W., and Tondl, G. (2004). Regional convergence in the European Union, 1985-1999: A spatial dynamic panel analysis. *Regional Studies*, *38*(3), 241–253.

Baldwin, R. E. (1999). Agglomeration and endogenous capital. *European Economic Review*, *43*(2), 253–280.

Barro, R., and Sala-i Martin, X. (2004). *Economic growth*, second edition.

Barro, R. J. (2015). Convergence and modernisation. *The Economic Journal*, *125*(585), 911–942.

Barro, R. J., and Lee, J. W. (2010). A New Data Set of Educational Attainment in the World, 1950-2010. NBER Working Paper No. 15902. *National Bureau of Economic Research*.

Bartel, A. P. (1989). Where do the new US immigrants live? *Journal of Labor Economics*, *7*(4), 371–391.

Basile, R., Girardi, A., Mantuano, M., and Russo, G. (2017). Interregional migration, human capital externalities and unemployment dynamics: evidence from Italian provinces. *ZBW-Deutsche Zentralbibliothek für Wirtschaftswissenschaften*.

Baum, C. F. (2006). *An introduction to modern econometrics using Stata*. Stata press.

Beine, M., Docquier, F., and Rapoport, H. (2008). Brain drain and human capital formation in developing countries: winners and losers. *The Economic Journal*, *118*(528), 631–652.

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.

Bonifazi, C., and Crisci, M. (2020). Correction to: Demography, Migration and the Economy at a Regional Level: Recent Evidence from Catalonia. *Applied Spatial Analysis and Policy*, *13*, 291–291.

Bonifazi, C., and Heins, F. (2000). Long-term trends of internal migration in Italy. *International Journal of Population Geography*, *6*(2), 111–131.

Bonifazi, C., Heins, F., Strozza, S., and Vitiello, M. (2009). The Italian transition from emigration to immigration country. *IRPPS working Papers, n. 24*.

Borjas, G. J. (2019). Immigration and economic growth. *National Bureau of Economic Research*.

Bouayad-Agha, S., and Vedrine, L. (2010). Estimation strategies for a spatial dynamic panel using GMM. A new approach to the convergence issue of European regions. *Spatial Economic Analysis*, *5*(2), 205–227.

Brunow, S., Nijkamp, P., and Poot, J. (2015). The impact of international migration on economic growth in the global economy, in *Handbook of the economics of international migration*, vol. 1, pp. 1027–1075. Elsevier.

Campo, F., Gorte, G., and Portes, J. (2018). The impact of migration on productivity and native-born workers' training. *IZA Discussion Paper*.

Cantore, N., Clara, M., Lavopa, A., and Soare, C. (2017). Manufacturing as an engine of growth: Which is the best fuel? *Structural Change and Economic Dynamics*, *42*, 56–66.

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. (2009). Immigration and inequality. *American Economic Review*, *99*(2), 1–21.

Coulter, R., Ham, M. V., and Findlay, A. M. (2016). Re-thinking residential mobility: linking lives through time and space. *Progress in Human Geography*, *40*(3), 352–374.

Daniele, V., Malanima, P. et al. (2007). Il prodotto delle regioni e il divario Nord-Sud in Italia (1861-2004). *Rivista di politica economica*, *97*(2), 267–316.

Debarsy, N., and Ertur, C. (2010). Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics*, *40*(6), 453–470.

Debarsy, N., Ertur, C., and Lesage, J. P. (2012). Interpreting dynamic space–time panel data models. *Statistical Methodology*, *9*(1-2), 158–171.

Di Berardino, C., D'ingiullo, D., Pozzi, C., Quaglione, D., and Sarra, A. (2020). Are The Effects on Regional Growth of Intra-National and Inter-National Skill-Specific Immigration Flows Different? *L'industria*, *41*(4), 1675–697.

Drinkwater, S. J., Levine, P., Lotti, E., Pearlman, J. et al. (2003). *The economic impact of migration: A survey*. Citeseer.

Elhorst, J. P. (2001). Dynamic models in space and time. *Geographical Analysis*, *33*(2), 119–140.

Elhorst, J. P., et al. (2014). *Spatial econometrics: from cross-sectional data to spatial panels*, vol. 479. Springer.

Ertur, C., and Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, *22*(6), 1033–1062.

Etzo, I. (2008). Internal migration: a review of the literature. *Munich Personal RePEc Archive (MPRA)*.

Fratesi, U., and Percoco, M. (2014). Selective migration, regional growth and convergence: evidence from Italy. *Regional Studies*, *48*, 1650–1668.

Ganong, P., and Shoag, D. (2017). Why has regional income convergence in the US declined? *Journal of Urban Economics*, *102*, 76–90.

Griliches, Z. (2008). *Output measurement in the service sectors*, vol. 56. University of Chicago Press.

Haque, N. U., and Kim, S.-J. (1995). Human capital flight: Impact of migration on income and growth. *Staff Papers*, *42*(3), 577–607.

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and structural transformation. *Handbook of economic growth*, *2*, 855–941.

Hierro, M., and Maza, A. (2010). Per capita income convergence and internal migration in Spain: Are foreign-born migrants playing an important role? *Papers in Regional Science*, *89*(1), 89–107.

Huber, P., and Tondl, G. (2012). Migration and regional convergence in the European Union. *Empirica*, *39*, 439–460.

Impicciatore, R., and Strozza, S. (2016). Lasciare il Mezzogiorno. *il Mulino*, *65*(1), 125–132.

Incaltarau, C., Pascariu, G. C., Duarte, A., and Nijkamp, P. (2021). Migration, regional growth and convergence: a spatial econometric study on Romania. *The Annals of Regional Science*, *66*, 497–532.

Istat, D. (2021). Demografia in cifre. ISTAT.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and the impact of immigration. Discussion paper, *National Bureau of Economic Research*.

Johnston, B. F., and Mellor, J. W. (1961). The role of agriculture in economic development. *The American Economic Review*, *51*(4), 566–593.

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.

Larramona, G., and Sanso, M. (2014). Internal migration and spanish regional convergence (1972–1998). *International Migration*, *52*(6), 128–148.

Lee, L.-F., and Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, *154*(2), 165–185.

Marchiori, L., Shen, I.-L., and Docquier, F. (2013). Brain drain in globalization: a general equilibrium analysis from the sending countries' perspective. *Economic Inquiry*, *51*(2), 1582–1602.

Maza, A. (2006). Migrations and regional convergence: The case of Spain. *Jahrbuch für Regionalwissenschaft*, *26*(2), 191–202.

Milio, S., Lattanzi, R., Casadio, F., Crosta, N., Raviglione, M., Ricci, P., and Scano, F. (2012). Brain drain, brain exchange and brain circulation. The case of Italy viewed from a global perspective. *National interest*.

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*(3), 537–561.

Piras, R. (2013). Can the Augmented Solow Model with Migration Explain the Italian Internal Brain Drain? *Labour*, *27*(2), 140–163.

Piras, R. (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.

Piras, R. (2021). Migration flows by educational attainment: Disentangling the heterogeneous role of push and pull factors. *Journal of Regional Science*, *61*(3), 515–542.

Ray, S., Haqiqi, I., Hill, A. E., Taylor, J. E. and Hertel, T. W. (2023). Labor markets: A critical link between global-local shocks and their impact on agriculture. *Environmental Research Letters*, *18*(3), 035007.

Reichlin, P., and Rustichini, A. (1998). Diverging patterns with endogenous labor migration. *Journal of Economic Dynamics and Control*, *22*(5), 703–728.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, *94*(5), 1002–1037.

Rosenthal, S. S., and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. in *Handbook of regional and urban economics*, vol. 4, (pp. 2119– 2171). Elsevier.

Shioji, E. (2001). Composition effect of migration and regional growth in Japan. *Journal of the Japanese and International Economies*, *15*(1), 29–49.

Spence, M. (1978). Job market signaling, in *Uncertainty in Economics*, (pp. 281–306). Elsevier.

Stark, O., and Lucas, R. E. (1988). Migration, remittances, and the family. *Economic development and cultural change*, *36*(3), 465–481.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, *49*(2), 326–365.

Thomas, M., Gillespie, B., and Lomax, N. (2019). Variations in migration motives over distance. *Demographic Research*, *40*, 1097–1110.

Vakulenko, E. (2016). Does migration lead to regional convergence in Russia? *International Journal of Economic Policy in Emerging Economies*, *9*(1), 1–25.

Wooldridge, J. M. (2014). Introduction to econometrics. *Cengage Learning: Andover*.

ORCID

Francesca Centofanti https://orcid.org/0000-0001-5671-6743

Roberto Basile https://orcid.org/0000-0002-4531-2820

Francesca Licari https://orcid.org/0000-0003-3380-4753

Annexes

Annex A. Migration distance according to migration motives

There is a growing body of research focused on the determinants of migration distance. The common assumption is that local moves are motivated by life-course transitions, such as family formation and dissolution (Kulu and Milewski, 2007), and associated shifts in housing consumption and neighbourhood preferences (Boyle, Kulu, Cooke, Gayle, and Mulder, 2008). On the contrary, long-distance migration has typically been assumed to be motivated by educational or employment-related factors, including job transfers or the promise of higher wages, better labour market prospects, more educational opportunities (Borjas, Bronars, and Trejo, 1992; Clark and Withers, 2007), or better climate (Niedomysl and Clark, 2014).

The subjectivity of this distinction makes the comparison between empirical studies problematic. A first approach to classify migration movements is defining moves within the geographical unit considered as residential mobility, while migration is any other move that crosses a geographical border. Yet, regardless of the scale used, this approach has the undesired effect of considering any short-distance moves crossing borders as longer-distance migrations. This misclassification bias is sometimes called "pseudo migration" (Thomas, Gillespie, and Lomax, 2019). The second approach uses distance threshold to avoid misclassification issues associated with boundary-based definitions of migration movements, yet inconsistencies in the distance thresholds used make the comparison of study results difficult. Lomax, Norman, and Darlington-Pollock (2021) showed that migration propensity varies across a number of distance thresholds, which differ in magnitude and direction depending on the migrant attributes being studied. Moreover, the relationships between migration motives and distances are likely to be context-specific. In the case of Italy, it is reasonable to believe that longer distance movements (from the South to the North) are mainly driven by economic and labor market variables, so they should have a grater impact on growth. The results we show in Section 4.1 of the main text seem to confirm this hypothesis.

Annex B. Estimates with the inclusion of capital investment

The role of capital in economic growth and migration dynamics is crucial and warrants careful consideration. The addition of a capital covariate enables us to explore how variations in capital investment levels influence the economic growth, knowing that such changes might also elucidate migration patterns.

Economic theory suggests that capital accumulation plays a significant role in shaping a country's productivity, income levels, and migration patterns. The neoclassical growth theory, pioneered by Solow (1956), emphasizes the importance of capital accumulation alongside labor inputs and technological progress in driving economic growth. According to this framework, an increase in capital stock leads to higher productivity levels, which, in turn, contribute to rising incomes and improved living standards. Consequently, migration decisions are influenced not only by labor market conditions but also by the availability of capital and investment opportunities in destination regions. Furthermore, the new economic geography theory, advanced by Krugman (1991), emphasizes the fact that regions with higher levels of capital often exhibit greater economic agglomeration and attract more businesses and skilled workers. Thus, migration patterns are influenced by the concentration of economic activity in capital-rich regions.

We repeat the estimates including the log of Gross Fixed Investments (GFI) as a covariate. This inclusion allows for a more comprehensive analysis, addressing the role of capital accumulation in stimulating economic growth. Overall, the results for each model specification, reported in tables 1 to 4 strengthen the validity and robustness of our analysis.

Table 1.
QMLE results - Net rates: no control for endogeneity of migration

Dependent variable:log of GDP per WP	**(1) total migration**	**(2) human capital**	**(3) human capitalno spatial**	**(4) without migration**	**(5) without migrationno spatial**
$Y_{(t-1)}$	0.532*** (0.045)	0.525*** (0.045)	0.612*** (0.066)	0.609*** (0.043)	0.726*** (0.086)
WY_t	0.233*** (0.035)	0.228*** (0.035)		0.244*** (0.035)	
WP growth rate	-0.017 **	-0.018**	-0.020**	-0.010	-0.013
	(0.007)	(0.007)	(0.008)	(0.006)	(0.009)
Agriculture employment share	-1.428***	-1.466**	-1.357***	-1.310***	-1.188***
	(0.266)	(0.266)	(0.250)	(0.259)	(0.243)
log of GFI	0.010*	0.011**	0.014**	0.011*	0.015**
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
Net-migration rate (*Italians*)	4.792***	14.750***	15.200***		
	(1.173)	(3.369)	(3.758)		
Net-migration rate (*Foreigners*)	0.098	0.170	0.115		
	(0.185)	(0.180)	(0.175)		
Implied convergence speed (%)	10.534	10.724	8.191	8.265	5.34
BIC	-2328.18	-2331.27	-1841.03	-2302.96	-1872.23

Notes: The spatial weights matrix adopted is an inverse distance matrix. Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels, respectively. All models include spatial and time fixed effects. Variables are defined in Section 3 of the main text.

TABLE 2.
QMLE results: Control function - Second stage. HK-weighted migration only

Dependent variable: log of GDP per WP	(6) No spatial		(7) SAR
$Y_{(t-1)}$	0.479**** (0.044))		0.527*** (0.045)
WY_t			0.223*** (0.034)
WP growth rate	-0.007		-0.017**
	(0.007)		(0.007)
Agriculture employment share	-1.444***		-1.452***
	(0.253)		(0.264)
log of GFI	0.016***		0.011*
	(0.006)		(0.006)
Net-Migration rate (*Italians*)	12.078**		12.114**
	(5.063)		(4.612)
Net-Migration rate (*Foreigners*)	-0.257		0.172
	(0.248)		(0.248)
First-stage residuals (*Italians*)	9.548		3.689
	(6.627)		(6.171)
F-test		339.75***	
		[0.000]	
First-stage residuals (*Foreigners*)	0.576*		-0.012
	(0.345)		(0.308)
F-test		550.50***	
		[0.000]	

Notes: The spatial weights matrix adopted is an inverse distance matrix. Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels, respectively. All models include spatial and time fixed effects. Variables are defined in Section 3 of the main text. For the first stages, the same model specification has been used.

TABLE 3.
Marginal effects of HK-weighted Net-Migration rate of Italians

	No control for endogeneity			**Control function**		
Short run	DIRECT	INDIRECT	TOTAL	DIRECT	INDIRECT	TOTAL
	15.030***	4.149***	19.179***	12.114***	3.326**	15.441**
	(3.170)	(1.223)	(4.175)	(4.612)	(1.479)	(5.968)
Long run						
	34.002***	27.188***	61.190***	27.506***	21.979*	49.485**
	(7.296)	(10.425)	(16.386)	(10.575)	(11.675)	(21.216)

Notes: Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels, respectively.

TABLE 4.
Marginal effects of HK-weighted Net-Migration rate of Italians - Distance thresholds

	No control for endogeneity			Control function		
Short run	DIRECT	INDIRECT	TOTAL	DIRECT	INDIRECT	TOTAL
Above 100 km	15.007***	3.775***	18.782***	11.133**	2.893**	15.546**
	(3.227)	(1.151)	(4.145)	(4.377)	(1.393)	(5.656)
Above 200 km	16.297***	4.046***	20.342***	13.456***	3.503**	16.958***
	(3.417)	(1.210)	(4.357)	(5.043)	(1.674)	(6.578)
Above 300 km	16.333***	4.100***	20.434***	15.640***	4.115**	19.755***
	(3.409)	(1.240)	(4.379)	(5.705)	(1.927)	(7.471)
Above 400 km	17.485***	4.350***	21.835***	16.674***	4.353**	21.027***
	(3.539)	(1.287)	(4.515)	(5.585)	(1.930)	(7.334)
Above 500 km	17.757***	4.509***	22.266***	21.431***	5.699**	27.130***
	(3.726)	(1.348)	(4.758)	(6.904)	(2.422)	(9.090)
Long run						
Above 100 km	36.686***	29.430**	66.117***	26.785**	21.842*	48.627**
	(8.021)	(12.023)	(18.404)	(10.725)	(12.642)	(22.388)
Above 200 km	39.662***	31.130**	70.792***	32.241***	26.225*	58.466**
	(8.416)	(12.515)	(19.061)	(12.389)	(15.324)	(26.574)
Above 300 km	39.970***	32.084**	72.055***	37.796***	31.546*	69.342**
	(8.480)	(13.184)	(19.804)	(14.168)	(18.279)	(31.098)
Above 400 km	42.584***	33.605**	76.189***	40.268***	33.279*	73.547**
	(8.720)	(13.581)	(20.217)	(13.900)	(18.524)	(30.974)
Above 500 km	44.030***	36.629**	80.659***	52.642***	45.615*	98.257**
	(9.340)	(15.187)	(22.264)	(17.499)	(24.765)	(40.294)

Notes: Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels,respectively.

ANNEX C. ESTIMATES WITH GROSS MIGRATION RATES AS MIGRATION VARIABLES

As discussed in the main text, we also capture the possible asymmetrical impact of migration by differentiating between inflows and outflows in a province. This robustness check proved particularly important to size the growth effect of migration. The results are displayed in Table 5.

TABLE 5.
QMLE results - Gross rates: no control for endogeneity of migration

Dependent variable: log of GDP per WP	(6) total migration	(7) human capital
$Y(t-1)$	0.524*** (0.047)	0.523*** (0.047)
WY_t	0.248*** (0.034)	0.241*** (0.034)
WP growth rate	-0.019**	-0.019**
	(0.007)	(0.007)

TABLE 5. CONT.
QMLE results - Gross rates: no control for endogeneity of migration

Dependent variable: log of GDP per WP	(6) total migration	(7) human capital
Agriculture employment share	-2.080***	-2.090***
	(0.629)	(0.627)
Immigration rate (*Italians*)	4.926***	15.144***
	(1.235)	(3.544)
Emigration rate (*Italians*)	-4.625***	-13.020***
	(1.615)	(4.999)
Immigration rate (*Foreigners*)	0.092	0.167
	(0.205)	(0.194)
Emigration rate (*Foreigners*)	-0.039	-0.082
	(0.218)	(0.212)
Implied convergence speed (%)	10.771	10.803
BIC	-2306.4	-2308.1

Notes: The spatial weights matrix adopted is an inverse distance matrix. Standard errors are in parentheses. *, **, and *** denote significance at the 1, 5 and 10 per cent levels, respectively. All models include spatial and time fixedeffects. Variables are defined in Section 3 of the main text.

ANNEX D. ESTIMATES WITH INTERACTIONS BETWEEN THE DUMMY SOUTH AND THE YEAR DUMMIES

To capture the North-South heterogeneous responses to common business cycle effects, we adapt our dynamic spatial model into the following specification:

$$y_t = \tau y_{t-1} + \delta W y_t + X_t\beta + \mu + \xi_t \iota_N + \pi_t \iota_N \times South + \varepsilon_t \quad (1)$$

where $\iota_N \times South$ is the interaction term between the dummy South, indicating whether the province belongs to the South area (Abruzzo, Molise, Campania, Calabria, Basilicata, Puglia, Sicily, and Sardinia) or not, and the yearly time dummies. The $K \times 1$ vector π_t includes the interaction term parameters, which capture the time-varing heterogeneity between the North-South repartitions. The control provided by this robustness check allows us to properly asses the role of unobserved time-invariant and South-specific factors in the convergence process. The results are displayed in Table 6.

TABLE 6.
QMLE results - South-time dummy interaction: no control for endogeneity of migration

Dependent variable: log of GDP per WP	(8) South-time dummies interaction
$Y(t-1)$	0.518*** (0.046)
WY_t	0.232*** (0.033)
WP growth rate	-0.018**
	(0.007)
Agriculture employment share	-2.185***

TABLE 6. CONT.
QMLE results - South-time dummy interaction: no control for endogeneity of migration

Dependent variable: log of GDP per WP	**(8)** **South-time dummies interaction**
	(0.626)
Net-migration rate (*Italians*)	13.945***
	(3.619)
Net-migration rate (*Foreigners*)	0.108
	(0.198)
Implied convergence speed (%)	10.963
BIC	-2297.3

Notes: The spatial weights matrix adopted is an inverse distance matrix. Standard errors are in parentheses. *, **,and *** denote significance at the 1, 5 and 10 per cent levels, respectively. The model includes spatial and time fixed effects. Variables are defined in Section 3 of the main text.

Annex References

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.

Boyle, P. J., Kulu, H., Cooke, T., Gayle, V., and Mulder, C. H. (2008). Moving and union dissolution. *Demography*, *45*(1), 209–222.

Clark, W. A., and Withers, S. D. (2007). Family migration and mobility sequences in the United States: Spatial mobility in the context of the life course. *Demographic Research*, *17*, 591–622.

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

Kulu, H., and Milewski, N. (2007). Family change and migration in the life course: An introduction. *Demographic Research*, *17*, 567–590.

Lomax, N., Norman, P., and Darlington-Pollock, F. (2021). Defining distance thresholds for migration research. *Population, Space and Place*, 27(4), e2440.

Niedomysl, T., and Clark, W. A. (2014). What matters for internal migration, jobs or amenities? Migration *Letters*, *11*(3), 377–386.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65–94.

Thomas, M., Gillespie, B., and Lomax, N. (2019). Variations in migration motives over distance. *Demographic Research*, *40*, 1097–1110.