From South to North ... and Beyond: Educational Selectivity and Migration Trajectories in Italy

Roberto Basile
Sapienza University of Rome
roberto.basile@uniroma1.it

Francesca Centofanti
University of Rome Tor Vergata
francesca.centofanti@students.uniroma2.eu

Francesca Licari
Italian National Institute of Statistics
francesca.licari@istat.it

Abstract

This paper analyzes the migration trajectories of young individuals born in Southern Italy who moved to the North between 2011 and 2014. Using longitudinal microdata and discrete-time competing risks models, we examine whether these internal migrants are more likely to return to the South or emigrate abroad. Results reveal a strong educational gradient: highly educated individuals are significantly more likely to use the North as a springboard for international migration, while less-educated individuals tend to return home. These findings shed light on the dynamic interplay between internal and international mobility, and the enduring challenges of brain drain in the Mezzogiorno.

Keywords: *Migration trajectories, Mezzogiorno of Italy, Competing Risk Models*

1 Introduction

The literature has long treated internal and international migration as separate domains. Early theoretical frameworks emphasized internal mobility [e.g., Ravenstein, 1889, Todaro, 1969], while more recent work focused on cross-border flows despite the fact that internal moves remain larger in scale [Castles et al., 2005]. This separation has hindered our ability to capture the interdependencies between mobility forms and their cumulative effects on regional development.

Growing evidence indicates that internal and international migration are linked and often unfold sequentially along the life course [Skeldon, 2017, Bernard and Vidal, 2023, Bernard and Perales, 2021]. A common pattern is "stepwise migration". For instance, a young graduate may first relocate from a disadvantaged or rural region to a more prosperous city within the same country, and subsequently leverage the acquired skills, networks, and knowledge to migrate abroad. This supports the notion that internal migration can serve as a stepping stone to international migration. Once robust social connections have been established between origin and destination countries by earlier migrant groups, future migrants may bypass internal migration altogether and move directly abroad [Lindstrom and Lauster, 2001].

Southern Italy (the Mezzogiorno) provides a salient case. The area faces a *double* brain drain: it loses young graduates to international destinations and—uniquely—also to the Centre-North within Italy [Basile et al., 2025]. Understanding whether the initial South $\to$ North move culminates in long-term settlement, triggers onward emigration, or results in return to the South is thus crucial for assessing the human capital dynamics of Italian regions.

Using a unique longitudinal dataset tracking official changes of residence, we reconstruct migration trajectories of Southern-born individuals, differentiated by educational attainment. Moreover, we employ discrete-time competing risks models to explore how educational attainment and other individual and contextual factors influence the probability that a young Southerner who migrated to the North will either emigrate abroad or return to the South, relative to the probability of remaining in the North. The findings provide robust evidence of strong skill selectivity in migration choices: highly educated individuals (bachelor's degree or higher) are markedly more likely to use the North as a springboard for international migration, while less-educated individuals are more likely to return home. Internal migration from the South to the North does not necessarily mitigate the brain drain; instead, for the highly educated it may *amplify* it by facilitating subsequent emigration abroad. Policies aiming to retain and attract talent to Southern regions should therefore complement measures improving job quality, research infrastructure, and service accessibility with instruments that encourage high-skill return.

In this study, we make three main contributions to the literature. First, we develop a typology of migration trajectories that is explicitly data-driven, allowing us to capture the wide variety of mobility paths observed in the longitudinal records rather than relying solely on predefined categories. Second, we propose an integrated analytical framework that models post-internal moves within a discrete-time competing-risks setting, formally grounded in random-utility theory. This approach enables us to account for the sequential and mutually exclusive nature of onward and return migration decisions. Third, we provide novel empirical evidence on the role of education in shaping these trajectories, highlighting how differences in educational attainment strongly influence whether individuals use the North as a springboard for international migration or, alternatively, return to their region of origin.

The dataset and its structure are described in Section 2. Section 3 describes the theoretical and empirical framework, and the factors that influence the migration choice. Section 4 presents and discusses the empirical findings, while Section 5 concludes.

2 Data and trajectories

This study draws on a longitudinal dataset — provided by the Italian National Institute of Statistics (Istat) — that enables the reconstruction of migration trajectories for all young individuals (aged 25–34) from Southern Italy who migrated at least once — either to Northern Italy or abroad — between 2011 and 2024.[1] For each individual, the dataset provides information on year and place of birth, gender, year of first change of residence (i.e., first migration), origin (i.e., municipality in the South), destination (either the municipality in the North for internal migration or the country for international migration), and educational attainment.[2] Subsequent relocations — including year, origin, destination, and educational attainment — are also recorded up to 2024.

We focus on long-distance internal moves from the Mezzogiorno to the Centre-North, while international moves are defined as any relocation outside Italy.[3] For descriptive clarity, we consider four cohorts by year of first migration (2011–2014) and track individuals for ten years after the first South → North (S→N) move, or after a direct South → Abroad (S→A) move when applicable.

2.1 Typology of trajectories

We identify eleven primary trajectories capturing simple, return, and circular patterns (Table 1). Trajectory 1 is a one-time internal migration from the South to the North (S→N), with no subsequent move; Trajectory 2 is direct international migration from the South (S→A). Trajectory 3 represents an initial internal move followed by an international one ($S \to N \to A$), while Trajectory 6 reflects an international move followed by a relocation to Northern Italy ($S \to A \to N$). In essence, Trajectories 3 and 6 exemplify cases in which one form of migration — internal or international — acts as a precursor to the other. Return migrations from the North and from abroad are captured by Trajectories 4 and 5, respectively, while Trajectories 7 through 11 encompass various forms of circular or repeated migration.

[1] The South of Italy, or Mezzogiorno, includes eight Italian NUTS-2 regions: Abruzzo, Molise, Campania, Basilicata, Puglia, Calabria, Sicilia, and Sardegna. In this study, we define the North as encompassing all regions outside the Mezzogiorno, i.e., what is typically referred to as the Centre-North.

[2] Registrations capture inflows into a municipality; cancellations capture outflows. The municipality of cancellation is treated as *origin* and the new municipality of registration as *destination*. We follow United Nation recommendations and only consider migration events involving a change of residence lasting at least 12 months, thereby excluding movements of shorter duration. *De facto* moves without formal registration are not observed.

[3] While short-distance mobility (e.g., rural-to-urban transitions within Southern regions) may be relevant in other contexts, it is irrelevant for our analysis aimed at capturing the micro-foundations of the brain drain in the South. Similarly, intra-national mobility occurring within foreign countries is not captured by our dataset and is excluded from our typology.

Insert Table 1 here

Between 2011 and 2024, a total of 645 thousand young individuals (aged 25–34) migrated from Southern Italy. Approximately 90% were born in the Mezzogiorno, while the remaining 10% were almost evenly split between individuals born in Northern Italy and abroad. Male migrants account for approximately 54% of the sample.

In the 2011 cohort, approximately 31% of migrants held a high level of education (i.e., at least a bachelor's degree or higher qualifications such as a PhD), 35% had medium-level education (i.e., an upper secondary school diploma), and 34% had a low level of education (i.e., up to a lower secondary school leaving certificate) (Table 2). Over time, the human capital composition of migrants from the South has improved: by the 2014 cohort, the share of highly educated individuals had risen to 38%, while the proportion of low-educated migrants had declined to 26%.

Insert Table 2 here

The trajectory analysis focuses only on individuals born in the South. In this case, cohort sizes range from 33,329 to 40,369 individuals (Table 3). Across cohorts, S→N is the modal trajectory but declines from about three quarters of cases in 2011 to roughly two thirds in 2014. Direct S→A rises notably over the same period. Among complex paths, S→N→S (return) is most frequent (about one in eight). These patterns suggest that international destinations became progressively more attractive to young Southerners over the early 2010s, while return from the North remains a sizeable component of mobility.

Insert Table 3 here

Apparently, direct internal migration ($S \rightarrow N$) and direct international migration ($S \rightarrow A$) can be viewed as competing strategies to maximize opportunities. However, when comparing the two migration trajectories, it emerges that, during the first four cohorts, individuals with lower levels of education tend to migrate abroad, while those with higher levels of education are more likely to relocate to the North (see Table 4). This pattern is likely influenced by labor market dynamics: during the early 2010s, the Northern labor market tends to absorb more highly skilled workers from the South, whereas low-skilled individuals, facing limited employment opportunities within the domestic market, are more inclined to seek employment abroad, where demand for low-skilled labor may be higher.[4] However, it must be observed that over time the share of highly educated individuals among international migrants has strongly increased (see Figure 4 in *Appendix A*).[5]

[4] During that period, labor demand in Northern and Central Italy was relatively *skill-biased*, with firms reporting persistent recruitment difficulties for university graduates, especially in STEM and professional occupations [Unioncamere, 2019]. Empirical studies confirm that South-to-North migration flows were strongly selective with respect to education, leading to an *internal "brain drain"* from the South [Capuano, 2012, Fratesi and Percoco, 2014]. By contrast, foreign labor markets such as the United Kingdom primarily absorbed Southern Italian migrants into low-skilled jobs in sectors like hospitality, retail, and construction [Metcalf, 2013, D'angelo and Kofman, 2016].

[5] This last evidence is consistent with the literature highlighting that, in the most recent period, the international demand for Italian migrants has become increasingly *skill-selective*. In particular, post-Brexit immigration policies

Insert Table 4 here

In the following analysis, we focus on a subset of migration trajectories that are particularly relevant for understanding the relationship between internal and international migration in the context of brain drain. These include the simple path $S \to N$ (primarily involving highly skilled migrants) and its immediate extensions ($S \to N \to A$, and $S \to N \to S$), which together account for approximately 85% of all movements.

The main *research questions* are therefore: Do young people from the South who go to the North stay there (*permanent migration*), or do they use this first step as a springboard to expatriate or return to the South (*temporary migration*)? Is the choice of the trajectory affected by the level of education or the characteristics of the origin?

For each selected trajectory, we examine the timing and duration of migration events. A discrete-time survival analysis is applied to assess how individual and territorial characteristics influence migration decisions.

2.2 Return heterogeneity: genuine vs. non-genuine returns

We extend the trajectory analysis by disaggregating the $S \to N \to S$ pattern into two distinct types of return migration (see *Appendix B*): *a*) *genuine returns*, where the destination municipality upon return matches the municipality of origin and *b*) *relocations within the South*, where the return is to a different southern municipality. The distinction is informative: genuine returns align with temporary migration (with the aim of acquiring skills or capital to be reinvested upon returning to the place of origin when they have strong family ties) or failed matches hypotheses (where the intended economic or professional outcomes were not achieved) [see, e.g., Borjas and Bratsberg, 1996], whereas non-genuine returns may be driven by destination-side opportunities within the South (such as local labor market conditions and service availability).

In the pooled sample, among the roughly 7,200 cases of return to the South, about 5,100 individuals — corresponding to nearly 70% — can be classified as genuine returns, while the remaining 2,100 involve settlement in a different municipality. The distribution of the two types is similar across the four cohorts considered.

3 A Discrete-Time Competing-Risks Model

We analyze the post-internal trajectories of young individuals whose first move is S→N at $t = 0$. In each subsequent discrete period $t = 1, \ldots, T$, an individual may: *i*) remain in the North (reference category, right-censored), *ii*) emigrate abroad (event 1), or *iii*) return to the South (event 2). These outcomes are mutually exclusive and constitute *competing risks*, as only one event can occur in a period, and the occurrence of one precludes the others.

in the UK introduced skill-based visa regimes that discourage low-skilled inflows and favor highly educated workers [Cuibus, 2023]. At the same time, Northern Italy has been attracting a growing number of foreign graduates through both international student inflows and skilled immigration, which partly reduces its reliance on university graduates from Southern Italy, as reported by Istat (Istat, 2024). Recent evidence shows that the share of Italian emigrants with a university degree has increased significantly in the last decade, consolidating the pattern of a high-skilled brain drain abroad (Istat, 2025).

3.1 Theoretical foundations

Let $P = \{i = 1, \ldots, N\}$ be the set of individuals who moved from the South to the North at $t = 0$. Each individual i is characterized by educational level $E_i \in \{L, M, H\}$, where L denotes low-skilled (lower secondary or less), M medium-skilled (upper secondary), and H high-skilled (tertiary education or more). We denote by $D_{it} \in \{0, 1, 2\}$ the migration status at time t: $D_{it} = 0$ if the individual remains in the North, $D_{it} = 1$ for international emigration, and $D_{it} = 2$ for return to the South. Once an event occurs, the individual exits the risk set.

Following the random-utility tradition [Beine et al., 2016, Bertoli et al., 2020], migration decisions are modeled as the outcome of a utility-maximizing process. At each time t, individual i compares the utility of remaining in the North (baseline) with the utilities of transitioning to the absorbing destinations $j \in \{1, 2\}$, corresponding to emigration abroad and return migration, respectively. The latent utility of choosing destination j is given by:

$$U_{ijt} = \alpha_j(t) + w_j(E_i) + \phi_j(E_i) + \eta_{ijt} \tag{1}$$

with $\alpha_j(t)$ a destination-specific baseline hazard, $w_j(E_i)$ the expected income conditional on education E_i, $\phi_j(E_i)$ non-pecuniary returns (such as the quality of life), and η_{ijt} i.i.d. type-I extreme value shocks. Education shapes both pecuniary and non-pecuniary components and thus the sequencing of moves. High-skill individuals face lower information and financing frictions for international mobility [Docquier and Rapoport, 2012, Ciriaci, 2014, Beine et al., 2014] and may use northern cities as "escalators" before moving abroad [Fielding, 1992, Recchi, 2015, King, 2002]. Low-skill individuals often face more constrained opportunities and may pursue targeted temporary migration to accumulate work experience or savings before returning to the South [Dustmann, 2003, Piracha and Vadean, 2010]. Medium-skilled individuals are expected to display more ambivalent strategies due to relatively balanced trade-offs between destinations.

Under the extreme-value assumption, these utility functions translate into *discrete-time cause-specific hazard functions*, which express the probability of choosing event j in period t, conditional on being at risk at the start of t:

$$h_{ij}(t) = \Pr(D_{it} = j \mid D_{i1} = 0, \ldots, D_{i,t-1} = 0) = \frac{\exp(U_{ijt})}{1 + \sum_{k=1}^{2} \exp(U_{ikt})} \tag{2}$$

The denominator includes the utility of remaining in the North, which is normalized to zero. In this framework, the probability of each competing migration event depends on the latent utilities associated with the alternatives and is explicitly conditioned on survival in the risk set.

3.2 From theory to empirics

In discrete-time statistical survival models, the target events are treated as categorical responses. Therefore, the cause-specific hazard function can be estimated using a *multinomial logit model* with $J + 1$ response categories [Tutz and Schmid, 2016, Schmid and Berger, 2021]. In our context, $J = 2$, and the reference category is "remaining in the North". Let $\mathbf{X}_i$ be a vector of covariates including individual characteristics (e.g., education, age, gender) and time-invariant contextual factors (origin and destination features). The hazard for event type j is:

$$h_j(t|\mathbf{X}_i) = \frac{\exp(\beta_{0jt} + \mathbf{X}_i'\boldsymbol{\beta}_j)}{1 + \sum_{j=1}^{2} \exp(\beta_{0jt} + \mathbf{X}_i'\boldsymbol{\beta}_j)} \tag{3}$$

where β_{0jt} are the cause-specific baseline hazards (i.e., the empirical counterpart of $\alpha_j(t)$ in the theoretical framework) and $\boldsymbol{\beta}_j$ are the event-specific covariate effects. As in the single-event case, the baseline hazard can be simplified by assuming a smooth functional form. Estimation is typically performed via maximum likelihood, treating each time period as a separate observation (i.e., a person-period data format).

Although $J+1$ response categories are considered, it suffices to specify the conditional probabilities of the target events $1, \ldots, J$. The probability of the reference category (i.e., remaining at risk at time t) is given by:

$$h_0(t|\mathbf{X}_i) = 1 - \sum_{j=1}^{2} h_j(t|\mathbf{X}_i) = \frac{1}{1 + \sum_{j=1}^{2} \exp(\beta_{0jt} + \mathbf{X}_i'\boldsymbol{\beta}_j)} \tag{4}$$

The ratio of cause-specific to survival hazard yields the odds of event j versus staying in the North:

$$\frac{h_j(t|\mathbf{X}_i)}{h_0(t|\mathbf{X}_i)} = \exp(\beta_{0jt}) \cdot \prod_{k=1}^{p} \exp(\beta_{jk})^{x_{ki}} \tag{5}$$

The coefficient β_{jk} thus reflects the change in log-odds of experiencing event j relative to conditional survival associated with a one-unit increase in covariate x_k:

$$\log\left(\frac{h_j(t|\mathbf{X}_i)}{h_0(t|\mathbf{X}_i)}\right) = \beta_{0jt} + \beta_{j1}x_{1i} + \cdots + \beta_{jp}x_{pi} \tag{6}$$

Finally, the survival function, i.e., the probability of remaining in the North up to time t, is:

$$S(t|\mathbf{X}_i) = \prod_{s=1}^{t-1} (1 - h_1(s|\mathbf{X}_i) - h_2(s|\mathbf{X}_i)) \tag{7}$$

and the cumulative incidence function (CIF) for event j, i.e., the probability that an individual experiences that particular event by a given time t, in the presence of other competing events, is:

$$F_j(t|\mathbf{X}_i) = \sum_{s=1}^{t} h_j(s|\mathbf{X}_i) \cdot S(s|\mathbf{X}_i)) \tag{8}$$

This integrated framework provides a theoretically grounded and statistically tractable model of migration trajectories, linking education-based utility maximization to observed transition probabilities via competing risks hazard functions.

3.3 Covariates selection

While our analysis focuses on the role of education in shaping migration trajectories, the hypotheses are grounded more broadly in established theoretical frameworks on migration decision-making. Specifically, we draw on two interrelated strands of literature: (*i*) the Neoclassical Economic Theory of Migration [Harris and Todaro, 1970, Borjas, 1989] and (*ii*) the New Economics of Labor Migration [Stark, 1991]. Both approaches view migration not as a single, one-time event but as a dynamic process shaped by evolving individual circumstances and structural conditions. They highlight how both individual-level factors — such as education, gender, and age — and contextual-level factors — such as the socioeconomic environment of sending and receiving areas — influence migration decisions over time. Life-course perspectives further reinforce this view, noting that the same factor (e.g., income or quality of life) can act as either a push or pull force depending on the timing and sequence of migration events [Bernard and Vidal, 2023].

Beyond education, we control for **gender** and **age**. Prior studies suggest that women are generally less mobile than men due to gendered labor market dynamics and family responsibilities [Bonifazi and Heins, 2000], while younger individuals are more mobile and more willing to undertake international moves [Zaiceva and Zimmermann, 2008]. This aligns with the neoclassical view of migration as an investment in human capital, where the benefits accrue over time and are greater if migration occurs early in the career [Sjaastad, 1962].

The choice of contextual variables is crucial. Unlike the gravity approach, where dyadic data simplify the classification of push and pull factors [Piras, 2021], a longitudinal framework complicates this selection. In our setting: (*i*) the characteristics of the origin municipality (South) at the first move represent push conditions; (*ii*) the characteristics of the destination municipality (North) initially serve as pull factors but may later become push factors if further migration occurs; (*iii*) the attributes of the second destination (abroad or return South) represent new pull conditions; and (*iv*) in genuine return migration, push and pull factors overlap, offering insights into cyclical mobility. For simplicity, we consider only the push conditions at the origin and the pull conditions at the first destination.

According to the New Economics of Labor Migration, individuals from **rural** or economically deprived areas are less likely to return, especially if their acquired skills cannot be used at home [Hugo, 2017]. Rural dwellers, often with lower social and economic capital, are also more likely to follow stepwise migration patterns — first internally, then internationally [Paul, 2015]. We therefore control for the urban/rural classification of origin and destination municipalities.[6] Specifically, we construct two dummy variables: one for rural-to-urban ($S \rightarrow N$) moves and one for urban-to-rural ($S \rightarrow N$) moves; the reference is urban-to-urban. Pure rural-to-rural cases are excluded due to their limited number and estimation issues.

Beyond urbanization, migration is shaped by accessibility to essential services [Diamond, 2016], which contribute to determining the quality of life. Instead of composite indicators of the availability of local services, we rely on the 2014 classification of Italian municipalities by the Agency for Territorial Cohesion, which categorizes municipalities as "Poles" (providers

[6]We use the DEGURBA classification in three categories — "Cities," "Towns and suburbs," and "Rural areas." The first two are combined into a single "Urban" category.

of health, education, and mobility services) and others according to proximity to Poles: "Belts" (close) and "Inner areas" (remote, often mountainous or rural). Inner areas are targeted by policy interventions to counter depopulation.

This classification also incorporates the degree of urbanization and is therefore employed as an alternative to the DEGURBA classification rather than in conjunction with it. Following the approach outlined above, we construct dyadic dummies for the first south-to-north move: (*i*) Pole–Belt, (*ii*) Pole–Inner, (*iii*) Belt–Inner, (*iv*) Belt–Belt, (*v*) Belt–Pole, (*vi*) Inner–Pole, and (*vii*) Inner–Belt. The reference category is Pole–Pole, while the rare Inner–Inner is excluded.

We hypothesize that individuals from non-Pole areas are less likely to migrate abroad than those from Poles, as Poles (mainly urban) facilitate networking with foreign labor markets and access to information about international opportunities. Conversely, individuals from Inner areas are expected to show a lower propensity to return south.

To capture classic economic push/pull factors consistent with neoclassical theory [Harris and Todaro, 1970, Borjas, 1989], we include: (*a*) the log ratio of per capita income at destination versus origin[7] and (*b*) two dummies equal to one if the origin or destination municipality of the first move experienced above-average employment growth in the subsequent decade (as proxies of job opportunities or probability to find a job).

Finally, to further account for unobserved heterogeneity, we include two dummies for whether the first destination municipality is in the Northeast or Northwest (NUTS1); the reference category is the Center.

4 Econometric results

The Competing Risk Model (CRM) is employed to examine how educational attainment, together with other individual factors and contextual economic conditions, influences the likelihood that a young individual from the South who migrated to the North between 2011 and 2014 will either (*a*) emigrate abroad or (*b*) return to the South, relative to remaining in the North. Thus, three migration trajectories are considered over the ten-year observation period (up to 2024). Individuals who did not move again before 2024 are classified as right-censored ($S \to N$). A second group emigrated abroad from the North ($S \to N \to A$), the *first competing event*, while a third returned to the South ($S \to N \to S$), the *second competing event*. This framework provides valuable insights into internal brain drain, selective international migration, and return migration in the context of persistent regional disparities within Italy and between Italy and foreign destinations.

We estimate two alternative model specifications. The first (Table 5) uses contextual variables based on the European DEGURBA classification of urbanization, which facilitates international comparison and tests hypotheses from the New Economics of Labor Migration. The second (Table 6) replaces DEGURBA with the SNAI classification developed by the Italian Agency for Territorial Cohesion, which also accounts for access to local services (see Section 3.3).

[7] Municipal-level per capita income is based on IRPEF tax returns collected by the Ministry of Economy and Finance (MEF).

For an overview of the results, Figures 1–3 present cumulative incidence functions (CIFs), survival functions, and baseline hazards, estimated for the pooled cohorts using the DEGURBA specification. Patterns from the SNAI model and single cohorts are similar and thus omitted. Overall, the plots indicate that most competing events occur within the first few years after the initial move, after which individuals tend to settle in the North. This pattern aligns with classic migration theories, which suggest that onward or return migration risks are highest early in the settlement process and decline as integration increases.

Figure 1 shows CIFs for emigration abroad (blue) and return to the South (red). The probability of returning rises gradually, reaching about 13% by 2024, whereas the probability of emigrating abroad remains below 1.5%. Conditional on staying in the North, individuals are therefore more likely to return South than to emigrate abroad. This lower incidence of international migration may reflect stronger ties to national labor markets or higher barriers to emigration.

Insert Figure 1 here

Figure 2 shows the estimated survival function for the Southern-born individuals residing in the North of Italy. The survival function represents the probability of not experiencing any migration event (either returning to the South or emigrating abroad) up to each time point, as estimated from a discrete-time CRM. The curve exhibits a gradual decline over time, reflecting the cumulative effect of migration events. The decline is steepest in the first years after the initial move, then stabilizes, suggesting that those who remain are increasingly likely to settle permanently in the North.

Insert Figure 2 here

Baseline hazards, estimated via cubic smoothing splines, following the default implementation in the R package VGAM, are shown in Figure 3. The risk of return migration is higher during the first three years, while the hazard of emigration abroad initially rises before declining and converging toward that of return. Beyond three years, both hazards decrease monotonically, indicating growing stability of residence in the North.

Insert Figure 3 here

Coefficient estimates with 95% confidence intervals for the model specification using the DEGURBA classification are reported in Table 5. As expected, educational attainment plays a central role in shaping post-internal migration trajectories. Highly educated individuals are more than twice as likely to emigrate abroad, while medium education levels show no significant effect. Conversely, both higher and medium education reduce the likelihood of return migration by roughly 40% and 20%, respectively. This supports the view that internal migration can act as a stepping stone for international mobility among the highly skilled, while southern regions struggle to attract educated returnees, reinforcing the persistent brain drain.

Insert Table 5 here

Other individual characteristics also matter. Age consistently reduces mobility: each additional year lowers the odds of emigration abroad by 20–25% and of return by 15–20%. Women are significantly less likely to emigrate abroad (-35%) but show no gender difference in return migration.

Contextual factors have limited effects on international emigration but strongly affect return migration. Specifically, individuals from rural origins moving to urban destinations are less likely to return to the South, while those from urban origins moving to rural destinations are more likely to return (reference category: urban–urban). This pattern aligns with the New Economics of Labor Migration [Massey et al., 1993], emphasizing temporary migration from areas with greater economic opportunities, and partially contrasts with neoclassical predictions. Nevertheless, consistent with neoclassical theory, a higher income ratio at the destination reduces the probability of return, while stronger employment opportunities at the origin (destination) encourage (discourage) it. Finally, migration to the Northeast reduces, and to the Northwest increases, the likelihood of emigration abroad (reference category: Center).

Results from the SNAI specification (Table 6) broadly confirm these findings. In line with expectations, individuals settling in Northern areas other than poles are less likely to emigrate abroad. Interestingly, those from southern poles moving to northern belts are also less likely to expatriate. Moreover, individuals from southern inner areas are less likely to return South, especially when settling in poles, due to the higher accessibility to services with respect to belts, whereas those from non-inner areas show a higher likelihood of return, particularly when the northern destination is an inner area.

Insert Table 6 here

Finally, we replicated the estimates by distinguishing between *genuine* and *non-genuine* return migration, as described in Section 2.2. The results reported in Tables 7 and 8 in Appendix B are fully consistent with those discussed so far. The only noteworthy difference that appears to emerge concerns the coefficients of the dummies "Urban–Rural" in Table 7 and "Pole–Belt" and "Pole–Inner" in Table 8: while in the previous estimates originating from an urban (pole) municipality in the South appeared to encourage return migration, the new results show that, if an individual originates from an urban (pole) municipality in the South and has settled in a rural (belt or inner) municipality in the North, the probability of returning to the municipality of origin increases, whereas the probability of returning to another southern municipality does not (the coefficient of the dummy is not statistically significant in the case of non-genuine return).

5 Conclusions

This study provides new evidence on the migration trajectories of young individuals born in Southern Italy, with particular attention to the role of education in shaping post-migration decisions. Using longitudinal microdata and discrete-time competing risks models, we examined whether internal migrants to the North are more likely to emigrate abroad or return to the South.

The results confirm that education is a decisive factor in migration behavior and direction. Highly educated individuals are significantly more likely to use internal migration as a stepping stone to international mobility, reinforcing the role of the North as a transitional rather than final destination. This dynamic reflects growing global competition for talent and suggests that internal migration may intensify, rather than mitigate, brain drain by facilitating subsequent emigration abroad. At the same time, the low probability of return among high-skilled individuals underscores persistent structural weaknesses in Southern Italy's ability to retain or attract talent, despite temporary returns among less-educated groups.

These findings highlight the need for a life-course perspective and an integrated approach to internal and international migration. Policies aimed at regional development and demographic resilience should not only address the drivers of out-migration but also create conditions that encourage return or circular mobility. Investments in high-quality employment, research capacity, and innovation ecosystems in Southern regions are essential to reverse current trends.

More broadly, the analysis confirms that migration trajectories are neither linear nor unidirectional. Instead, they consist of interdependent decisions shaped by personal, educational, and contextual factors. Recognizing this complexity is crucial for designing effective policy responses that move beyond static views of migration as a one-time event and instead capture the full arc of individuals' geographic and professional mobility.

Future research should enrich this framework with more detailed information on labor market integration, job quality, and individual motivations. The growing availability of administrative microdata offers an opportunity to better understand how mobility and skills interact across space and time, and how territorial inequalities both drive and result from human capital flows.

Finally, evidence shows that the share of highly educated individuals among international migrants has been rising. This supports the idea that, once social networks link origin and destination countries, new migrants may bypass internal migration altogether and move directly abroad. Extending the Competing Risk Model to consider the $S \to A$ trajectory as the reference category, with $S \to A \to N$ and $S \to S$ as competing events, would shed light on these dynamics. At present, however, this analysis remains premature: since the trend of high-skilled dominance in the $S \to A$ trajectory has emerged only after 2020, sufficient time has not yet passed to observe subsequent outcomes such as staying abroad, returning South, or resettling in the North.

Funding details

This work was supported by the European Union – Next Generation EU" through the PRIN 2022 project titled "Reloading city: a new systemic approach to urban and territorial regeneration" (CUP: E53D23010120006).

References

R. Basile, F. Centofanti, and F. Licari. The brain drain of italians: are the southern regions doomed? *Unpublished article*, 2025.

M. Beine, R. Noël, and L. Ragot. Determinants of the international mobility of students. *Economics of Education review*, 41:40–54, 2014.

M. Beine, S. Bertoli, and J. Fernández-Huertas Moraga. A practitioners' guide to gravity models of international migration. *The World Economy*, 39(4):496–512, 2016.

A. Bernard and F. Perales. Is migration a learned behavior? understanding the impact of past migration on future migration. *Population and Development Review*, 47(2):449–474, 2021.

A. Bernard and S. Vidal. Linking internal and international migration over the life course: A sequence analysis of individual migration trajectories in europe. *Population studies*, 77(3): 515–537, 2023.

S. Bertoli, J. F.-H. Moraga, and L. Guichard. Rational inattention and migration decisions. *Journal of International Economics*, 126:103364, 2020.

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

G. J. Borjas. Economic theory and international migration. *International migration review*, 23 (3):457–485, 1989.

G. J. Borjas and B. Bratsberg. Who leaves? the outmigration of the foreign-born. *Review of Economics and Statistics*, (78):165–176, 1996.

S. Capuano. The south–north mobility of italian college graduates. an empirical analysis. *European Sociological Review*, 28(4):538–549, 2012.

S. Castles, M. J. Miller, and G. Ammendola. The age of migration: international population movements in the modern world: New york: the guilford press,(2003), 338 pages. 2005.

D. Ciriaci. Does university quality influence the interregional mobility of students and graduates? the case of italy. *Regional Studies*, 48(10):1592–1608, 2014.

M. Cuibus. Eu migration to and from the uk. *Migration Observatory briefing, COMPAS, University of Oxford*, 2023.

R. Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *American economic review*, 106(3):479–524, 2016.

F. Docquier and H. Rapoport. Globalization, brain drain, and development. *Journal of Economic Literature*, 50(3):681–730, 2012.

C. Dustmann. Return migration, wage differentials, and the optimal migration duration. *European economic review*, 47(2):353–369, 2003.

A. D'angelo and E. Kofman. Uk: Large-scale european migration and the challenge to eu free movement. In *South-north migration of EU citizens in times of crisis*, pages 175–192. Springer International Publishing Cham, 2016.

T. Fielding. Migration and social mobility: South east england as an escalator region. *Regional Studies*, 26(1):1–15, 1992.

U. Fratesi and M. Percoco. Selective migration, regional growth and convergence: evidence from italy. *Regional Studies*, 48(10):1650–1668, 2014.

J. R. Harris and M. P. Todaro. Migration, unemployment and development: a two-sector analysis. *The American economic review*, 60(1):126–142, 1970.

G. Hugo. *New forms of urbanization: beyond the urban-rural dichotomy*. Routledge, 2017.

R. King. Towards a new map of european migration. In *European migration: What do we know?*, pages 3–26. Routledge, 2002.

D. P. Lindstrom and N. Lauster. Local economic opportunity and the competing risks of internal and us migration in zacatecas, mexico 1. *International migration review*, 35(4):1232–1256, 2001.

D. S. Massey, J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor. Theories of international migration: A review and appraisal. *Population and development review*, pages 431–466, 1993.

D. Metcalf. Immigration and the british labour market: the role of the migration advisory committee. *Migration Advisory Committee. London School of Economics*, 2013.

A. M. Paul. Capital and mobility in the stepwise international migrations of filipino migrant domestic workers. *Migration Studies*, 3(3):438–459, 2015.

M. Piracha and F. Vadean. Return migration and occupational choice: Evidence from albania. *World Development*, 38(8):1141–1155, 2010.

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

E. G. Ravenstein. The laws of migration. *Journal of the royal statistical society*, 52(2):241–305, 1889.

E. Recchi. *Mobile Europe: The theory and practice of free movement in the EU*. Springer, 2015.

M. Schmid and M. Berger. Competing risks analysis for discrete time-to-event data. *Wiley Interdisciplinary Reviews: Computational Statistics*, 13(5):e1529, 2021.

L. A. Sjaastad. The costs and returns of human migration. *Journal of political Economy*, 70(5, Part 2):80–93, 1962.

R. Skeldon. International migration, internal migration, mobility and urbanization: Towards more integrated approaches. In *United Nations expert group meeting on sustainable cities, human mobility and international migration population division department of economic and social affairs United Nations Secretariat New York*, pages 7–8, 2017.

O. Stark. *The migration of labor*. Basil Blackwell, 1991.

M. P. Todaro. A model of labor migration and urban unemployment in less developed countries. *The American economic review*, 59(1):138–148, 1969.

G. Tutz and M. Schmid. *Modeling Discrete Time-to-Event Data*. Springer, 2016.

A. Unioncamere. La domanda di professioni e di formazione delle imprese italiane nel 2018. monitoraggio dei flussi e delle competenze per favorire l’occupabilità. *Sistema Informativo Excelsior, Roma, Unioncamere< https://bit. ly/3FBYlJb*, 2019.

A. Zaiceva and K. F. Zimmermann. Scale, diversity, and determinants of labour migration in europe. *Oxford Review of Economic Policy*, 24(3):427–451, 2008.

	Trajectory	Type
1	S → N	Single internal migration
2	S → A	Single international migration
3	S → N → A	Internal migration followed by international migration
4	S → N → S	Return to the South
5	S → A → S	Return to the South
6	S → A → N	International migration followed by internal migration
7	S → A → S → A	Circular migration
8	S → A → S → N	Circular migration
9	S → N → A → N	Circular migration
10	S → N → S → A	Circular migration
11	S → N → S → N	Circular migration

TABLE 1
Migration Trajectories from the South

Education	Cohort 2011	Cohort 2012	Cohort 2013	Cohort 2014	Overall
Low	33.87	26.96	22.24	25.91	18.63
Medium	35.09	39.71	38.74	35.89	37.77
High	31.04	33.33	39.02	38.20	43.60
Number	37,565	45,295	39,601	39,185	644,998

TABLE 2
Percentage distribution of Italian migrants aged 25–34 from the South by education. Source: Istat

Trajectories	Cohort 2011	Cohort 2012	Cohort 2013	Cohort 2014	Pooling
S_N	77.26	74.6	69.49	67.77	72.29
S_A	5.9	8.2	12.12	14.75	10.23
S_N_A	0.98	1.07	1.04	0.98	1.02
S_N_S	12.3	12.1	12.52	11.02	11.98
S_A_S	0.44	0.71	1.28	1.99	1.1
S_A_N	0.12	0.14	0.27	0.39	0.23
S_A_S_A	0.08	0.12	0.19	0.23	0.15
S_A_S_N	0.04	0.04	0.13	0.16	0.09
S_N_A_S	0.1	0.12	0.13	0.11	0.12
S_N_S_A	0.26	0.3	0.24	0.15	0.24
S_N_S_N	2.04	2.11	1.95	1.8	1.98
Number	33,329	40,369	35,681	35,233	144,612

TABLE 3

Migration trajectories from the South (born in the South selected). Percentage values. Source: Istat

	Cohort 2011	Cohort 2012	Cohort 2013	Cohort 2014	Pooling
			$S \rightarrow N$ trajectory		
Low	28.90	24.40	18.70	20.70	23.30
Medium	36.30	39.30	37.80	36.10	37.50
High	34.80	36.30	43.50	43.20	39.20
			$S \rightarrow A$ trajectory		
Low	67.60	38.20	35.30	41.50	42.40
Medium	19.40	36.50	38.40	32.90	33.50
High	13.00	25.40	26.40	25.60	24.10

TABLE 4

Internal and international migration by education level (born in the South selected). Percentage values. Source: Istat

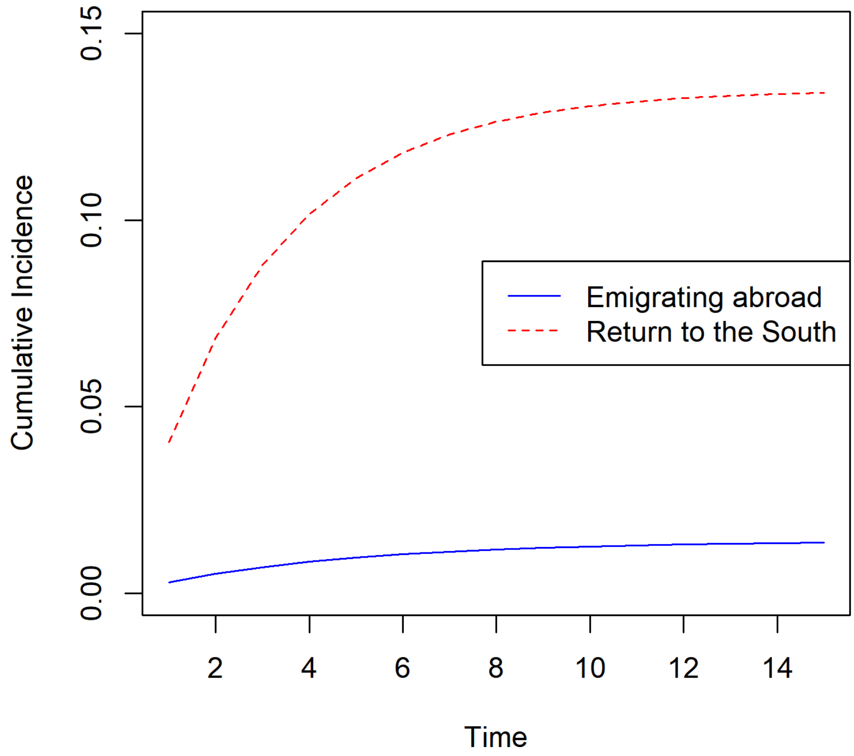


FIGURE 1
Cumulative incidence function. Estimated using the pooling specification based on the DEGURBA classification.

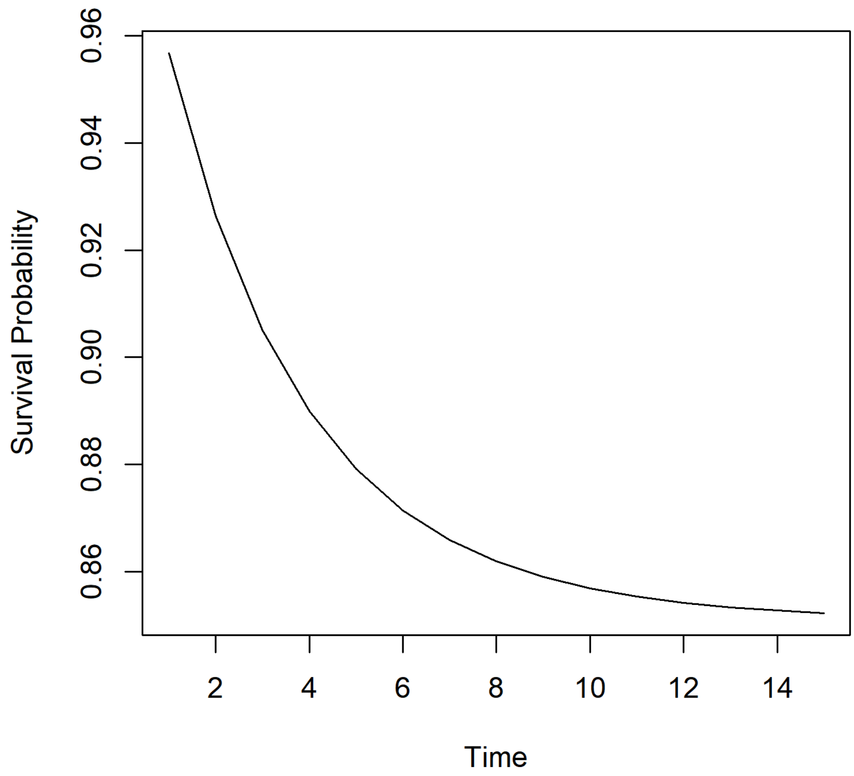


FIGURE 2
Survival function. Estimated using the pooling specification based on the DEGURBA classification.

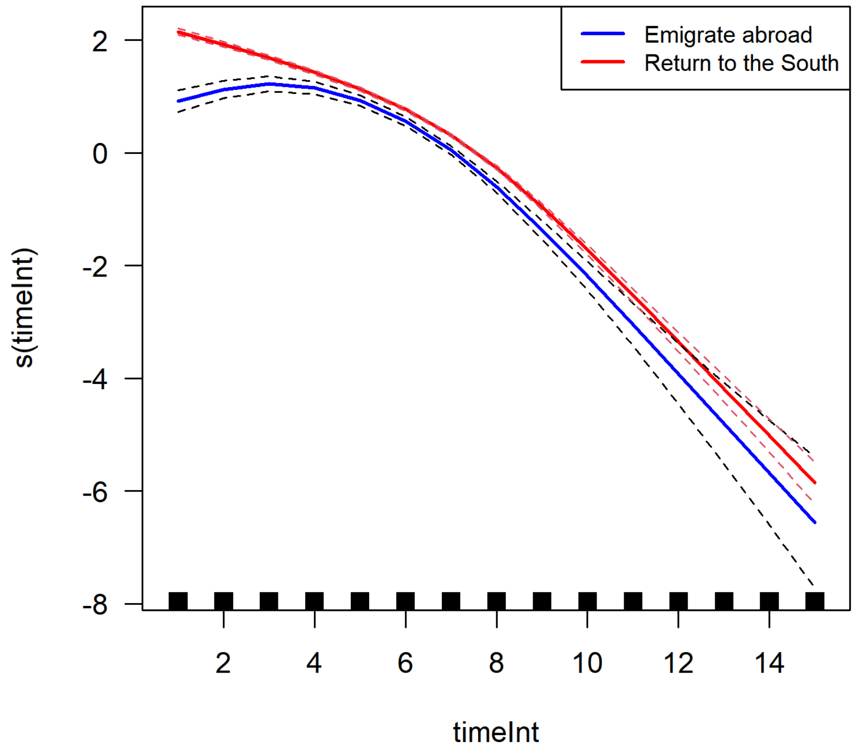


FIGURE 3
Baseline hazards. Estimated using the pooling specification based on the DEGURBA classification.

TABLE 5

Discrete-time competing risks model estimates. Specification with DEGURBA classification.

	2011			2012			2013			2014			Pooling		
Variables	$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI	
	Emigrate abroad														
Intercept	0.010	[0.006	0.016]	0.006	[0.004	0.010]	0.004	[0.002	0.007]	0.011	[0.007	0.019]	0.007	[0.005	0.009]
High education	2.619	[1.949	3.519]	2.212	[1.676	2.919]	2.746	[1.895	3.981]	1.528	[1.125	2.076]	2.188	[1.877	2.549]
Medium education	0.819	[0.582	1.152]	0.967	[0.716	1.307]	1.223	[0.820	1.823]	1.021	[0.740	1.407]	0.995	[0.843	1.176]
Female	0.587	[0.464	0.741]	0.679	[0.554	0.833]	0.661	[0.530	0.825]	0.639	[0.510	0.801]	0.643	[0.576	0.718]
Age	0.758	[0.723	0.794]	0.779	[0.748	0.812]	0.782	[0.748	0.818]	0.782	[0.748	0.818]	0.776	[0.759	0.793]
Rural-Urban	0.888	[0.635	1.241]	0.709	[0.517	0.972]	0.555	[0.379	0.812]	0.875	[0.633	1.209]	0.748	[0.632	0.884]
Urban-Rural	0.472	[0.232	0.961]	0.788	[0.459	1.354]	0.752	[0.397	1.427]	0.882	[0.511	1.521]	0.724	[0.537	0.976]
Income differential	0.838	[0.600	1.170]	1.217	[0.908	1.631]	1.183	[0.861	1.624]	0.955	[0.693	1.317]	1.057	[0.903	1.237]
High empl. growth origin	0.894	[0.713	1.120]	0.694	[0.567	0.848]	0.819	[0.659	1.017]	0.965	[0.775	1.201]	0.826	[0.742	0.920]
High empl. growth destination	0.935	[0.707	1.237]	1.281	[0.972	1.687]	1.444	[1.038	2.007]	1.027	[0.770	1.371]	1.142	[0.988	1.320]
Northeast	0.624	[0.443	0.879]	0.778	[0.575	1.052]	0.694	[0.482	0.998]	0.683	[0.497	0.938]	0.689	[0.585	0.812]
Northwest	1.198	[0.892	1.609]	1.457	[1.129	1.881]	1.736	[1.297	2.324]	1.134	[0.866	1.484]	1.354	[1.180	1.553]
	Return to the South														
Intercept	0.199	[0.172	0.230]	0.185	[0.162	0.211]	0.225	[0.195	0.260]	0.206	[0.177	0.240]	0.186	[0.172	0.201]
High education	0.560	[0.511	0.613]	0.652	[0.601	0.708]	0.700	[0.640	0.765]	0.652	[0.596	0.713]	0.641	[0.614	0.669]
Medium education	0.791	[0.732	0.854]	0.892	[0.830	0.958]	0.909	[0.836	0.988]	0.814	[0.748	0.887]	0.849	[0.816	0.883]
Female	1.031	[0.963	1.105]	0.951	[0.894	1.011]	0.946	[0.887	1.010]	0.981	[0.916	1.050]	0.974	[0.943	1.007]
Age	0.821	[0.810	0.832]	0.829	[0.820	0.839]	0.809	[0.799	0.820]	0.813	[0.802	0.824]	0.818	[0.813	0.824]
Rural-Urban	0.815	[0.733	0.906]	0.723	[0.654	0.801]	0.740	[0.666	0.821]	0.739	[0.663	0.824]	0.751	[0.713	0.792]
Urban-Rural	1.146	[1.009	1.302]	1.346	[1.203	1.507]	1.201	[1.057	1.364]	1.164	[1.014	1.336]	1.220	[1.146	1.299]
Income differential	0.970	[0.877	1.074]	0.915	[0.836	1.001]	0.803	[0.730	0.883]	0.815	[0.736	0.901]	0.872	[0.831	0.915]
High empl. growth origin	1.023	[0.955	1.096]	1.128	[1.061	1.199]	1.018	[0.954	1.087]	1.087	[1.015	1.164]	1.066	[1.031	1.101]
High empl. growth destination	0.782	[0.722	0.847]	0.804	[0.748	0.864]	0.827	[0.764	0.895]	0.854	[0.785	0.929]	0.814	[0.783	0.847]
Northeast	0.961	[0.879	1.051]	1.028	[0.950	1.112]	0.940	[0.862	1.024]	1.021	[0.933	1.118]	0.993	[0.951	1.037]
Northwest	0.910	[0.832	0.995]	1.014	[0.940	1.095]	0.991	[0.914	1.075]	1.052	[0.965	1.146]	0.994	[0.954	1.036]

Notes: Base outcome = remain in the North.

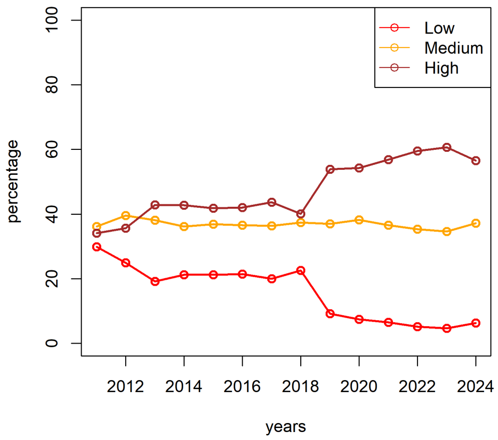
TABLE 6

Discrete-time competing risks model estimates. Specification with SNAI classification.

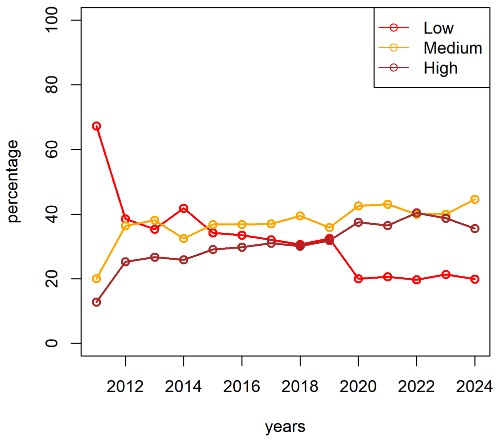
	2011			2012			2013			2014			Pooling		
Variables	$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI	
	Emigrate abroad														
Intercept	0.014	[0.008	0.024]	0.008	[0.005	0.013]	0.005	[0.003	0.009]	0.014	[0.009	0.025]	0.009	[0.007	0.012]
High education	2.359	[1.750	3.179]	2.033	[1.534	2.694]	2.567	[1.769	3.725]	1.447	[1.063	1.969]	2.023	[1.734	2.360]
Medium education	0.808	[0.574	1.138]	0.937	[0.693	1.268]	1.177	[0.789	1.755]	1.001	[0.726	1.381]	0.968	[0.819	1.143]
Female	0.606	[0.479	0.765]	0.682	[0.556	0.836]	0.668	[0.535	0.834]	0.649	[0.518	0.813]	0.651	[0.583	0.727]
Age	0.758	[0.723	0.795]	0.779	[0.748	0.812]	0.781	[0.747	0.817]	0.780	[0.746	0.816]	0.775	[0.758	0.792]
Pole-Belt	0.671	[0.436	1.032]	0.510	[0.322	0.810]	0.520	[0.312	0.866]	0.599	[0.375	0.957]	0.578	[0.459	0.729]
Pole-Inner	0.326	[0.119	0.897]	0.821	[0.413	1.633]	1.154	[0.573	2.326]	0.871	[0.449	1.691]	0.759	[0.526	1.094]
Belt-Belt	0.435	[0.253	0.747]	0.718	[0.474	1.088]	0.508	[0.291	0.886]	0.522	[0.309	0.881]	0.550	[0.429	0.705]
Belt-Pole	0.862	[0.622	1.193]	0.866	[0.650	1.154]	1.172	[0.870	1.578]	0.916	[0.672	1.248]	0.948	[0.815	1.103]
Belt-Inner	0.377	[0.138	1.027]	0.536	[0.219	1.314]	0.709	[0.288	1.749]	0.379	[0.140	1.031]	0.480	[0.299	0.770]
Inner-Pole	0.859	[0.621	1.187]	0.762	[0.571	1.018]	0.922	[0.673	1.264]	0.843	[0.612	1.160]	0.840	[0.719	0.982]
Inner-Belt	0.321	[0.176	0.584]	0.477	[0.303	0.751]	0.451	[0.263	0.775]	0.516	[0.313	0.850]	0.440	[0.340	0.569]
Income differential	0.731	[0.491	1.088]	1.100	[0.778	1.556]	0.886	[0.606	1.294]	0.837	[0.572	1.226]	0.895	[0.742	1.079]
High empl. growth origin	0.888	[0.706	1.118]	0.688	[0.560	0.844]	0.826	[0.662	1.031]	0.967	[0.775	1.208]	0.827	[0.741	0.922]
High empl. growth destination	0.960	[0.724	1.274]	1.295	[0.982	1.708]	1.529	[1.097	2.130]	1.033	[0.772	1.382]	1.167	[1.008	1.350]
Northeast	0.624	[0.442	0.880]	0.818	[0.604	1.108]	0.746	[0.518	1.076]	0.719	[0.522	0.991]	0.722	[0.612	0.851]
Northwest	1.231	[0.912	1.663]	1.561	[1.201	2.029]	2.000	[1.481	2.702]	1.219	[0.924	1.610]	1.464	[1.271	1.685]
	Return to the South														
Intercept	0.195	[0.166	0.228]	0.177	[0.154	0.204]	0.206	[0.177	0.240]	0.182	[0.155	0.214]	0.173	[0.159	0.187]
High education	0.562	[0.513	0.616]	0.658	[0.605	0.714]	0.706	[0.645	0.772]	0.666	[0.608	0.728]	0.649	[0.621	0.678]
Medium education	0.790	[0.731	0.854]	0.890	[0.828	0.956]	0.908	[0.835	0.987]	0.820	[0.753	0.893]	0.850	[0.817	0.884]
Female	1.030	[0.961	1.103]	0.946	[0.890	1.006]	0.944	[0.884	1.008]	0.980	[0.915	1.050]	0.972	[0.940	1.004]
Age	0.821	[0.810	0.832]	0.829	[0.819	0.839]	0.809	[0.799	0.820]	0.813	[0.802	0.824]	0.818	[0.813	0.824]
Pole-Belt	1.063	[0.939	1.204]	1.118	[1.002	1.247]	1.126	[0.999	1.270]	1.228	[1.082	1.392]	1.132	[1.067	1.202]
Pole-Inner	1.095	[0.904	1.325]	1.212	[1.023	1.435]	1.306	[1.089	1.565]	1.206	[0.990	1.470]	1.209	[1.103	1.325]
Belt-Belt	1.166	[1.033	1.316]	1.113	[0.997	1.243]	1.241	[1.103	1.396]	1.268	[1.116	1.440]	1.192	[1.123	1.265]
Belt-Pole	1.171	[1.053	1.302]	1.021	[0.930	1.121]	1.074	[0.973	1.186]	1.168	[1.054	1.294]	1.101	[1.047	1.157]
Belt-Inner	1.050	[0.862	1.279]	1.364	[1.155	1.611]	1.310	[1.085	1.582]	1.416	[1.171	1.712]	1.287	[1.174	1.411]
Inner-Pole	0.938	[0.838	1.050]	0.822	[0.744	0.909]	0.872	[0.785	0.968]	0.849	[0.759	0.950]	0.865	[0.820	0.913]
Inner-Belt	0.900	[0.787	1.028]	0.899	[0.797	1.014]	0.932	[0.819	1.060]	1.035	[0.906	1.183]	0.936	[0.878	0.999]
Income differential	0.941	[0.837	1.059]	0.970	[0.873	1.077]	0.845	[0.756	0.945]	0.865	[0.768	0.973]	0.905	[0.856	0.957]
High empl. growth origin	1.006	[0.938	1.079]	1.107	[1.040	1.178]	1.001	[0.937	1.069]	1.060	[0.989	1.136]	1.045	[1.011	1.081]
High empl. growth destination	0.780	[0.720	0.845]	0.788	[0.734	0.847]	0.819	[0.756	0.886]	0.850	[0.782	0.925]	0.806	[0.775	0.838]
Northeast	0.968	[0.884	1.059]	1.042	[0.962	1.128]	0.950	[0.871	1.037]	1.010	[0.922	1.107]	1.000	[0.958	1.044]
Northwest	0.914	[0.835	1.002]	1.016	[0.940	1.099]	0.999	[0.919	1.086]	1.031	[0.944	1.126]	0.994	[0.953	1.037]

Notes: Base outcome = remain in the North.

Appendix A



(a) $S \rightarrow N$ trajectory



(b) $S \rightarrow A$ trajectory

FIGURE 4
Internal and international migration by education level (born in the South selected). Percentage values. Years: 2011-2024

Appendix B

Extending the analysis of trajectories: distinguishing between genuine and non-genuine return to the South

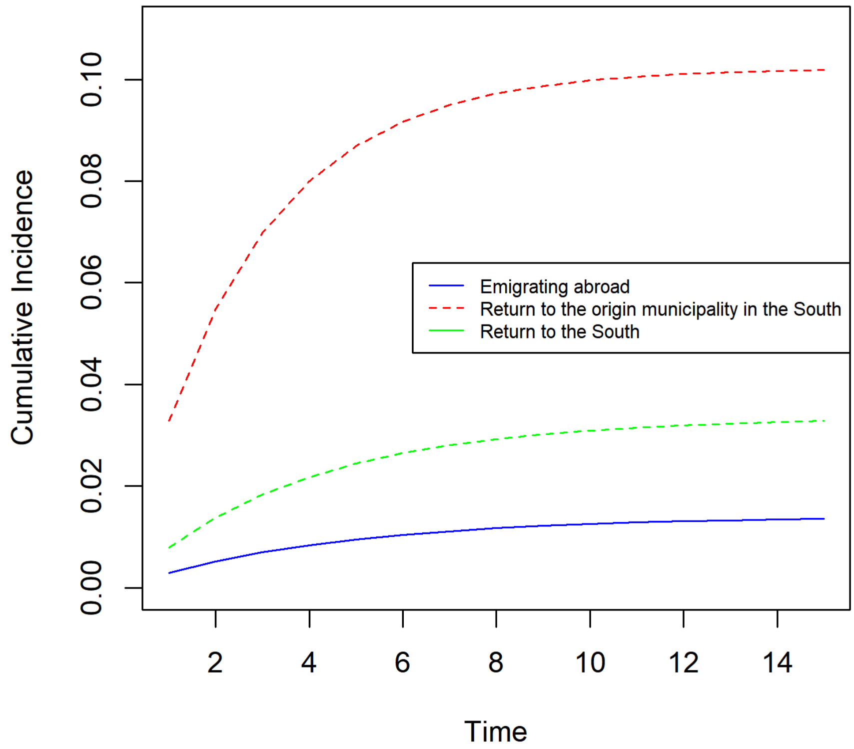


FIGURE 5
Cumulative incidence function. Estimated using the pooling specification based on the DEGURBA classification.

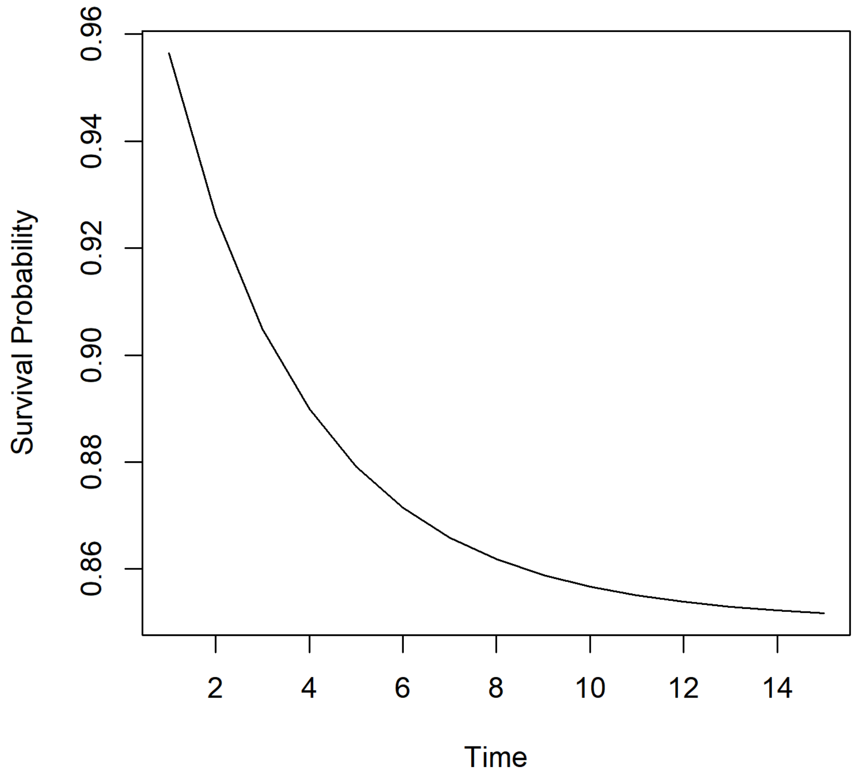


FIGURE 6
Survival function. Estimated using the pooling specification based on the DEGURBA classification.

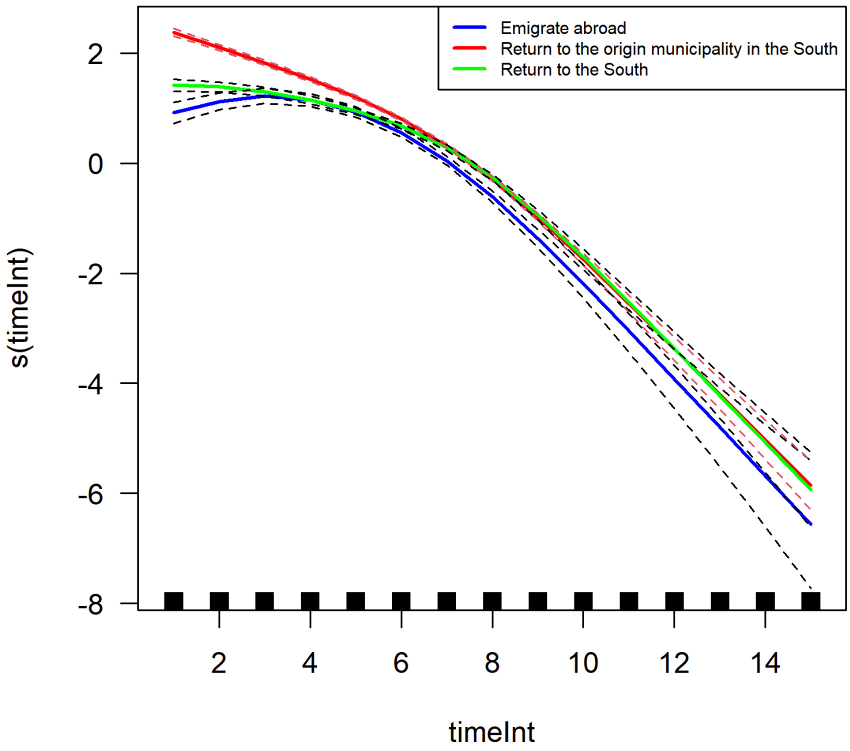


FIGURE 7
Baseline hazards. Estimated using the pooling specification based on the DEGURBA classification.

TABLE 7

Discrete-time competing risks model estimates. Specification with DEGURBA classification.

	2011			2012			2013			2014			Pooling		
Variables	$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI	
	Emigrate abroad														
Intercept	0.010	[0.006	0.016]	0.006	[0.004	0.010]	0.004	[0.002	0.007]	0.011	[0.007	0.019]	0.007	[0.005	0.009]
High education	2.619	[1.949	3.519]	2.212	[1.676	2.919]	2.746	[1.895	3.981]	1.528	[1.125	2.076]	2.188	[1.878	2.549]
Medium education	0.819	[0.582	1.152]	0.967	[0.716	1.307]	1.223	[0.820	1.822]	1.021	[0.740	1.407]	0.995	[0.843	1.176]
Female	0.587	[0.464	0.741]	0.679	[0.554	0.833]	0.661	[0.530	0.825]	0.639	[0.510	0.801]	0.643	[0.576	0.718]
Age	0.758	[0.723	0.794]	0.779	[0.748	0.812]	0.782	[0.748	0.818]	0.782	[0.748	0.818]	0.776	[0.759	0.793]
Rural-Urban	0.888	[0.635	1.241]	0.709	[0.517	0.972]	0.555	[0.379	0.811]	0.875	[0.633	1.209]	0.748	[0.632	0.884]
Urban-Rural	0.472	[0.232	0.961]	0.788	[0.459	1.354]	0.752	[0.397	1.427]	0.882	[0.511	1.521]	0.724	[0.537	0.976]
Income differential	0.838	[0.600	1.170]	1.217	[0.908	1.631]	1.183	[0.861	1.624]	0.955	[0.693	1.317]	1.057	[0.903	1.237]
High empl. growth origin	0.894	[0.713	1.120]	0.694	[0.567	0.848]	0.819	[0.659	1.017]	0.965	[0.775	1.201]	0.826	[0.742	0.920]
High empl. growth destination	0.935	[0.707	1.237]	1.281	[0.972	1.687]	1.444	[1.038	2.007]	1.027	[0.770	1.371]	1.142	[0.988	1.320]
Northeast	0.624	[0.443	0.879]	0.778	[0.575	1.052]	0.694	[0.482	0.998]	0.683	[0.497	0.938]	0.689	[0.585	0.812]
Northwest	1.198	[0.892	1.609]	1.457	[1.129	1.881]	1.736	[1.297	2.324]	1.134	[0.867	1.485]	1.354	[1.180	1.553]
	Return to the origin municipality in the South														
Intercept	0.152	[0.128	0.179]	0.157	[0.135	0.182]	0.175	[0.149	0.207]	0.170	[0.144	0.202]	0.148	[0.136	0.161]
High education	0.563	[0.507	0.624]	0.651	[0.593	0.714]	0.689	[0.623	0.762]	0.647	[0.585	0.716]	0.638	[0.607	0.670]
Medium education	0.784	[0.717	0.856]	0.868	[0.800	0.941]	0.892	[0.811	0.981]	0.809	[0.735	0.891]	0.836	[0.800	0.874]
Female	1.042	[0.964	1.128]	0.943	[0.879	1.011]	0.966	[0.897	1.041]	0.969	[0.896	1.047]	0.976	[0.940	1.014]
Age	0.825	[0.812	0.837]	0.838	[0.827	0.850]	0.816	[0.804	0.828]	0.818	[0.805	0.830]	0.825	[0.819	0.831]
Rural-Urban	0.824	[0.731	0.930]	0.703	[0.626	0.791]	0.711	[0.629	0.803]	0.705	[0.622	0.799]	0.732	[0.689	0.778]
Urban-Rural	1.116	[0.963	1.295]	1.398	[1.234	1.585]	1.301	[1.130	1.498]	1.172	[1.003	1.368]	1.254	[1.168	1.345]
Income differential	1.013	[0.902	1.138]	0.890	[0.804	0.986]	0.820	[0.735	0.914]	0.815	[0.727	0.914]	0.879	[0.832	0.928]
High empl. growth origin	1.007	[0.931	1.089]	1.112	[1.037	1.191]	1.039	[0.964	1.119]	1.088	[1.007	1.176]	1.063	[1.024	1.104]
High empl. growth destination	0.799	[0.729	0.876]	0.794	[0.732	0.862]	0.831	[0.759	0.910]	0.859	[0.781	0.945]	0.818	[0.782	0.855]
Northeast	0.987	[0.891	1.094]	1.072	[0.980	1.172]	1.001	[0.907	1.106]	1.032	[0.932	1.144]	1.031	[0.982	1.083]
Northwest	0.910	[0.822	1.009]	1.015	[0.931	1.107]	1.030	[0.938	1.132]	1.066	[0.967	1.175]	1.010	[0.963	1.058]
	Return to the South														
Intercept	0.048	[0.036	0.064]	0.031	[0.024	0.041]	0.051	[0.038	0.068]	0.038	[0.028	0.052]	0.039	[0.033	0.045]
High education	0.549	[0.456	0.661]	0.660	[0.556	0.783]	0.737	[0.615	0.882]	0.670	[0.557	0.807]	0.653	[0.597	0.714]
Medium education	0.812	[0.695	0.948]	0.978	[0.843	1.133]	0.965	[0.814	1.143]	0.833	[0.698	0.993]	0.892	[0.823	0.967]
Female	0.997	[0.868	1.146]	0.979	[0.863	1.110]	0.887	[0.778	1.011]	1.023	[0.888	1.179]	0.968	[0.905	1.035]
Age	0.809	[0.787	0.831]	0.800	[0.780	0.820]	0.788	[0.767	0.809]	0.796	[0.773	0.819]	0.798	[0.787	0.809]
Rural-Urban	0.783	[0.628	0.977]	0.789	[0.646	0.965]	0.830	[0.680	1.014]	0.857	[0.693	1.061]	0.814	[0.733	0.903]
Urban-Rural	1.235	[0.967	1.578]	1.167	[0.912	1.495]	0.894	[0.670	1.193]	1.137	[0.851	1.517]	1.112	[0.974	1.269]
Income differential	0.847	[0.690	1.039]	1.001	[0.832	1.204]	0.751	[0.620	0.909]	0.813	[0.659	1.002]	0.852	[0.772	0.940]
High empl. growth origin	1.077	[0.937	1.238]	1.183	[1.043	1.342]	0.957	[0.840	1.090]	1.081	[0.938	1.245]	1.074	[1.005	1.148]
High empl. growth destination	0.732	[0.626	0.857]	0.835	[0.719	0.970]	0.813	[0.694	0.953]	0.838	[0.704	0.997]	0.804	[0.742	0.870]
Northeast	0.884	[0.739	1.058]	0.889	[0.754	1.047]	0.771	[0.648	0.917]	0.983	[0.816	1.185]	0.877	[0.804	0.958]
Northwest	0.908	[0.760	1.084]	1.010	[0.866	1.178]	0.883	[0.753	1.035]	1.005	[0.842	1.201]	0.948	[0.873	1.030]

Notes: Base outcome = remain in the North.

TABLE 8

Discrete-time competing risks model estimates. Specification with SNAI classification. (continue ...)

	2011			2012			2013			2014			Pooling		
Variables	$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI	
	Emigrate abroad														
Intercept	0.014	[0.008	0.024]	0.008	[0.005	0.013]	0.005	[0.003	0.009]	0.014	[0.009	0.025]	0.009	[0.007	0.012]
High education	2.359	[1.750	3.179]	2.033	[1.534	2.695]	2.567	[1.769	3.725]	1.447	[1.063	1.969]	2.023	[1.734	2.360]
Medium education	0.808	[0.574	1.138]	0.937	[0.693	1.268]	1.177	[0.789	1.755]	1.001	[0.726	1.381]	0.968	[0.819	1.143]
Female	0.606	[0.479	0.765]	0.682	[0.556	0.836]	0.668	[0.535	0.834]	0.649	[0.518	0.813]	0.651	[0.583	0.727]
Age	0.758	[0.723	0.795]	0.779	[0.748	0.812]	0.781	[0.747	0.817]	0.780	[0.746	0.816]	0.775	[0.758	0.792]
Pole-Belt	0.671	[0.436	1.032]	0.510	[0.322	0.810]	0.520	[0.312	0.866]	0.599	[0.375	0.957]	0.578	[0.459	0.729]
Pole-Inner	0.326	[0.119	0.897]	0.821	[0.413	1.633]	1.154	[0.573	2.326]	0.871	[0.449	1.691]	0.759	[0.526	1.094]
Belt-Belt	0.435	[0.253	0.747]	0.718	[0.474	1.088]	0.508	[0.291	0.886]	0.522	[0.309	0.881]	0.550	[0.429	0.705]
Belt-Pole	0.862	[0.622	1.193]	0.866	[0.650	1.154]	1.172	[0.870	1.578]	0.916	[0.672	1.248]	0.948	[0.815	1.103]
Belt-Inner	0.377	[0.138	1.027]	0.536	[0.219	1.314]	0.709	[0.288	1.749]	0.379	[0.140	1.031]	0.480	[0.299	0.770]
Inner-Pole	0.858	[0.621	1.187]	0.762	[0.571	1.018]	0.922	[0.673	1.264]	0.843	[0.612	1.160]	0.840	[0.719	0.982]
Inner-Belt	0.321	[0.176	0.584]	0.477	[0.303	0.751]	0.451	[0.263	0.775]	0.516	[0.313	0.850]	0.440	[0.340	0.569]
Income differential	0.731	[0.491	1.088]	1.100	[0.778	1.556]	0.886	[0.606	1.294]	0.837	[0.572	1.226]	0.895	[0.742	1.079]
High empl. growth origin	0.888	[0.706	1.118]	0.688	[0.560	0.844]	0.826	[0.662	1.031]	0.967	[0.775	1.208]	0.827	[0.741	0.922]
High empl. growth destination	0.960	[0.724	1.274]	1.295	[0.982	1.708]	1.529	[1.097	2.130]	1.033	[0.772	1.382]	1.167	[1.008	1.350]
Northeast	0.624	[0.442	0.880]	0.818	[0.604	1.108]	0.746	[0.518	1.076]	0.719	[0.522	0.991]	0.722	[0.612	0.851]
Northwest	1.231	[0.912	1.663]	1.561	[1.201	2.029]	2.000	[1.481	2.702]	1.219	[0.924	1.610]	1.464	[1.271	1.685]
	Return to the origin municipality in the South														
Intercept	0.148	[0.123	0.177]	0.153	[0.131	0.180]	0.158	[0.132	0.188]	0.155	[0.130	0.186]	0.139	[0.126	0.152]
High education	0.564	[0.508	0.625]	0.654	[0.595	0.718]	0.695	[0.628	0.769]	0.658	[0.594	0.728]	0.644	[0.613	0.677]
Medium education	0.783	[0.717	0.856]	0.866	[0.799	0.940]	0.892	[0.811	0.982]	0.814	[0.740	0.896]	0.837	[0.800	0.875]
Female	1.040	[0.962	1.125]	0.938	[0.875	1.006]	0.962	[0.893	1.037]	0.967	[0.895	1.046]	0.973	[0.937	1.010]
Age	0.825	[0.812	0.838]	0.838	[0.827	0.849]	0.816	[0.804	0.828]	0.818	[0.805	0.830]	0.825	[0.819	0.831]
Pole-Belt	1.057	[0.916	1.220]	1.141	[1.010	1.288]	1.164	[1.016	1.334]	1.186	[1.029	1.367]	1.138	[1.064	1.217]
Pole-Inner	1.130	[0.909	1.404]	1.162	[0.958	1.409]	1.473	[1.207	1.798]	1.131	[0.902	1.418]	1.223	[1.103	1.357]
Belt-Belt	1.130	[0.983	1.300]	1.018	[0.896	1.156]	1.204	[1.050	1.379]	1.159	[1.001	1.342]	1.120	[1.046	1.199]
Belt-Pole	1.140	[1.010	1.287]	0.967	[0.870	1.076]	1.020	[0.909	1.144]	1.091	[0.971	1.225]	1.046	[0.988	1.107]
Belt-Inner	1.011	[0.803	1.272]	1.203	[0.989	1.464]	1.310	[1.055	1.627]	1.243	[0.994	1.554]	1.192	[1.071	1.328]
Inner-Pole	0.982	[0.864	1.115]	0.778	[0.694	0.872]	0.880	[0.781	0.991]	0.840	[0.741	0.953]	0.860	[0.809	0.914]
Inner-Belt	0.901	[0.774	1.050]	0.929	[0.812	1.061]	0.975	[0.843	1.128]	1.036	[0.893	1.203]	0.956	[0.890	1.028]
Income differential	0.980	[0.857	1.121]	0.965	[0.856	1.086]	0.883	[0.777	1.003]	0.860	[0.752	0.983]	0.921	[0.864	0.982]
High empl. growth origin	0.998	[0.921	1.081]	1.099	[1.024	1.180]	1.031	[0.955	1.112]	1.077	[0.995	1.166]	1.053	[1.013	1.094]
High empl. growth destination	0.797	[0.727	0.874]	0.775	[0.715	0.841]	0.818	[0.748	0.896]	0.854	[0.776	0.939]	0.807	[0.772	0.844]
Northeast	0.995	[0.898	1.103]	1.080	[0.986	1.182]	1.015	[0.918	1.122]	1.023	[0.922	1.135]	1.037	[0.987	1.090]
Northwest	0.918	[0.827	1.020]	1.005	[0.919	1.099]	1.037	[0.942	1.142]	1.044	[0.944	1.154]	1.005	[0.957	1.054]

(... continue) Discrete-time competing risks model estimates. Specification with SNAI classification.

	2011			2012			2013			2014			Pooling		
Variables	$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI		$exp(\beta)$	95% CI	
	Return to the South														
Intercept	0.047	[0.035	0.065]	0.028	[0.021	0.038]	0.049	[0.036	0.066]	0.030	[0.021	0.042]	0.035	[0.030	0.041]
High education	0.557	[0.462	0.671]	0.672	[0.565	0.799]	0.743	[0.620	0.891]	0.692	[0.575	0.834]	0.666	[0.609	0.729]
Medium education	0.812	[0.696	0.949]	0.973	[0.839	1.128]	0.961	[0.811	1.138]	0.841	[0.705	1.004]	0.894	[0.825	0.969]
Female	0.997	[0.867	1.145]	0.974	[0.859	1.105]	0.888	[0.779	1.013]	1.023	[0.887	1.179]	0.967	[0.904	1.034]
Age	0.809	[0.787	0.831]	0.799	[0.779	0.820]	0.789	[0.768	0.810]	0.798	[0.775	0.821]	0.798	[0.788	0.809]
Pole-Belt	1.080	[0.844	1.382]	1.029	[0.808	1.311]	1.010	[0.789	1.293]	1.392	[1.067	1.817]	1.112	[0.982	1.260]
Pole-Inner	0.988	[0.668	1.464]	1.400	[0.994	1.972]	0.826	[0.539	1.266]	1.510	[1.016	2.246]	1.162	[0.958	1.409]
Belt-Belt	1.278	[1.010	1.618]	1.472	[1.186	1.828]	1.359	[1.081	1.708]	1.706	[1.323	2.199]	1.445	[1.286	1.623]
Belt-Pole	1.272	[1.030	1.570]	1.222	[1.009	1.479]	1.252	[1.032	1.518]	1.475	[1.193	1.825]	1.295	[1.171	1.432]
Belt-Inner	1.170	[0.803	1.705]	1.976	[1.453	2.688]	1.311	[0.904	1.899]	2.106	[1.480	2.996]	1.613	[1.356	1.919]
Inner-Pole	0.791	[0.621	1.008]	0.991	[0.808	1.215]	0.847	[0.685	1.048]	0.885	[0.694	1.129]	0.882	[0.789	0.986]
Inner-Belt	0.893	[0.682	1.171]	0.785	[0.597	1.032]	0.798	[0.607	1.050]	1.032	[0.769	1.385]	0.867	[0.755	0.996]
Income differential	0.828	[0.654	1.049]	0.986	[0.795	1.222]	0.734	[0.586	0.919]	0.881	[0.690	1.125]	0.855	[0.762	0.958]
High empl. growth origin	1.032	[0.896	1.190]	1.134	[0.998	1.290]	0.913	[0.799	1.042]	1.004	[0.869	1.160]	1.020	[0.953	1.092]
High empl. growth destination	0.729	[0.623	0.853]	0.833	[0.718	0.967]	0.821	[0.701	0.962]	0.838	[0.705	0.996]	0.804	[0.743	0.871]
Northeast	0.885	[0.738	1.060]	0.920	[0.779	1.086]	0.775	[0.650	0.924]	0.969	[0.801	1.171]	0.887	[0.812	0.969]
Northwest	0.903	[0.754	1.083]	1.056	[0.901	1.237]	0.895	[0.760	1.054]	0.988	[0.822	1.186]	0.961	[0.883	1.047]

Notes: Base outcome = remain in the North.