



**Electronic Journal of Applied Statistical Analysis  
EJASA, Electron. J. App. Stat. Anal.**

<http://siba-ese.unisalento.it/index.php/ejasa/index>

e-ISSN: 2070-5948

DOI: 10.1285/i20705948v12n4p826

**Exploring the determinants and trends of STEM students' internal mobility. Some evidence from Italy**

By D'Agostino, Ghellini, Longobardi

Published: 15 December 2019

This work is copyrighted by Università del Salento, and is licensed under a Creative Commons Attribution - Non commerciale - Non opere derivate 3.0 Italia License.

For more information see:

<http://creativecommons.org/licenses/by-nc-nd/3.0/it/>

# Exploring the determinants and trends of STEM students' internal mobility. Some evidence from Italy

A. D'Agostino<sup>\*a</sup>, G. Ghellini<sup>b</sup>, and S. Longobardi<sup>a</sup>

<sup>a</sup>*University Napoli "Parthenope" Dept. of Management and Quantitative Studies, Via Generale  
Parisi 13, 80133 Napoli*

<sup>b</sup>*Siena University, Dept. Economics and Statistics, P.zza San francesco 8, 53100 Siena*

Published: 15 December 2019

In recent years, there has been a widespread consensus that science, technology, engineering and math (STEM) education is crucial for the long-term productivity and growth of a country. In this light, the present paper aims to explore the phenomenon of the mobility of Italian STEM students, namely, the flows of graduated students from the Southern regions who enrol in the universities in the Northern/Central area and choose a STEM degree course. We exploit microdata from the Italian Ministry of Education, University and Research (MIUR) on 8 cohorts (from a.y. 2008/2009 to a.y. 2015/16) of enrolled students in STEM fields who obtained high school diplomas in Southern Italy. The main results of our analysis show that the flow of STEM movers increases from year to year. This flow particularly affects the top-performing students and, therefore, is a threat to the socio-economic growth prospects of the Southern regions whose gap with respect to the Central and Northern regions is expected to grow year by year.

**keywords:** STEM education, Students' mobility, Regional development, Multilevel modelling

## 1 Introduction

STEM is an acronym commonly used to describe education or professional practice in the areas of science, technology, engineering, and mathematics. Considering that it is

---

\*antonella.dagostino@uniparthenope.it

recognized worldwide that STEM workers could play a key role in sustaining economic growth and stability, during the past decade, the broad educational fields of STEM have received growing attention by both EU and Member State institutions. Indeed, STEM education creates critical thinkers, increases science literacy, and enables the next generation of innovators. The Office for the Chief Scientist in Australia stated in its 2014 report that the country's long-term prosperity compared to that of other countries is strictly related to STEM education building new jobs, creating growth and driving innovation. Atkinson and Mayo (2010) stressed the same concept in the U.S., namely, that the presence of STEM graduates stimulates innovation and therefore economic growth. There is also evidence that STEM graduates create positive externalities by improving the earning conditions of less educated individuals (Winters, 2018). From this perspective, given the importance of STEM education on local development, both the participation rate in STEM disciplines and the out-migration of STEM students, especially from less developed areas, are issues of great relevance in Europe and in other countries. Unfortunately, a well-established feature of the Italian academic system is the large incidence of movers from southern regions to universities located in the Centre and North of the country<sup>1</sup> (De Santis et al., 2019; De Angelis et al., 2017). In this critical framework, the specific analysis of the dynamics of the inter-regional mobility of STEM students also becomes more interesting due to its implications for workforce development in Southern Italy. Indeed, a descriptive picture of the trend of movers from southern regions showed a higher incidence of movers in STEM courses than in non-STEM courses (see Figure 2 of the next section). Consequently, there is a double penalty for the Southern Italian regions. First, the flow of movers from these areas of the country has been growing in recent years. Second, the decision to move towards the universities of the North is mostly undertaken by STEM students who can be considered "catalysts" of the socio-economic development of a region. Furthermore, student mobility from southern regions could be a result not only of individual characteristics but also of some sub-regional-level variables. Indeed, students attending high school in the same province tend to share certain common territorial characteristics (see, for instance, D'Agostino et al., 2019) that can reduce or increase the propensity to move. In this light, it is imperative to better understand the main individual characteristics that influence STEM choice for enhancing the enrolment in STEM areas and the individual and contextual factors that influence STEM students' mobility in order to anticipate and fight potential regional impoverishment.

In Italy, there is a gap in the empirical quantitative research on both of these two issues. The gap is due mainly to the lack of data that allow a specific analysis of STEM students. As such, this paper seeks to examine and integrate findings from this body of research with special attention given to the second of these two issues.

To the best of our knowledge, no study has explicitly addressed the critical problem of the internal mobility of Italian STEM students. In this paper, we aim to fill this gap.

---

<sup>1</sup>Indeed, international student mobility, e.g., students who enrol in degree programmes abroad, is becoming a real issue in Italy, at least in the North of the country, but unfortunately, it is quite difficult to measure this phenomenon using the available data.

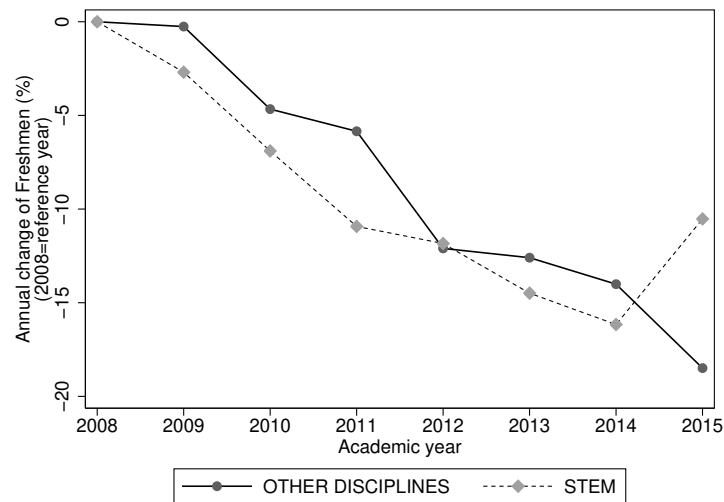
By exploiting a recent high-quality administrative dataset on the population of Italian graduates, our study contributes to understanding the relationship between STEM education and student mobility.

As mentioned above, STEM student mobility can be determined by different levels of factors, ranging from the individual level to the sub-regional level. A problem that concerns the relationships between the variables measured at a number of different hierarchical levels is the multilevel problem (Hox, 2010). Therefore, a random intercept model was used to take into account the hierarchical structure of the data (students nested in the provinces where they received their high school diplomas) and address the following research question: RH1) Which individual-level and contextual-level variables affect STEM student mobility? Moreover, we estimate a model for each cohort to observe whether the individual and contextual-level variables have the same effects over time. The remainder of the paper is organized as follows. In the next section, the details of the main features of the recent literature concerning the issue of STEM education and internal student mobility are introduced. A general description of the data and methodological details is provided in Section 3. The main results are discussed in Section 4. Some final remarks are given at the end the paper.

## 2 Background

There is an international common agreement that innovation leads to new products and processes that sustain the economy at the local and national levels. This innovation depends on a solid knowledge base in the STEM fields. Indeed, the STEM skills associated with advanced technical skills can be considered to be strong drivers for technology and productivity gains not only in high-tech sectors but also in several other economic sectors. Most of the background literature concerning STEM education is concentrated on this issue. The important role of STEM innovations in generating economic productivity and growth has been recognized, at least since Robert Solow (1957) seminal work. Over the years, several other studies have empirically explored the role of STEM education as a driver of economic performance. Recently, for instance, (Ray, 2015) stressed that in the U.S., the presence of STEM graduates stimulates innovation and therefore economic growth. (Peri et al., 2014) found that STEM workers increased the total factor productivity in U.S. cities. Therefore, it seems imperative to have an adequate number of graduates in these disciplines to improve local development. It is also clear that most jobs in the future will require a basic understanding of math and science. In Europe, STEM professional and related professional occupations are expected to grow by 13% and 7%, respectively, over the 2015-2025 period, whereas all occupations are expected to grow by 3% over the same period (ICF et al., 2015). In the U.S., the projected growth of STEM occupations is even more rapid (Fayer and Watson, 2015). Moreover, (Goos et al., 2013) outlined that STEM high school education can be considered to be an important predictor of labour market success and that STEM occupations are recognized to pay more than all other disciplines in the job market. Despite these compelling facts, the general consensus among Italian policy-makers is that the current supply of STEM

skills in Italy is insufficient and presents a potentially significant constraint on future economic development. In particular, the enrolment of students from Southern regions in STEM education follows a significantly decreasing trend, as reported in Figure 1. When compared to all other disciplines, this trend is also steeper. There is only a slight recovery in the last year that needs to be carefully monitored in the future. Therefore, in the literature, we can find empirical evidence that supports monitoring and seeking to understand the reasons behind this issue with the aim to promote student interests in and attitudes towards careers in STEM (see, for instance, Winckler and Fieder, 2006; ICF et al., 2015).



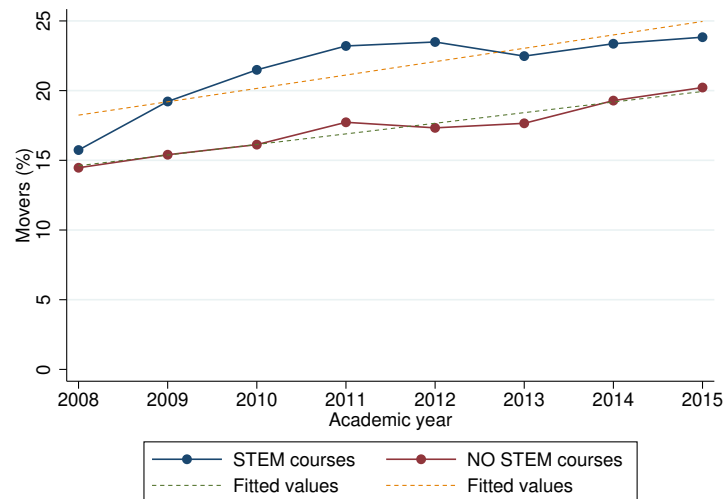
Note: Our elaboration on the MIUR data.

Figure 1: Trend of the university enrolment of southern students over the years

However, the crucial role of STEM education in local development also encourages reflecting on the effects that the internal migration of STEM students can have in areas that suffer from the constant out-migration of these students. The U.S. literature on this topic concentrates its attention on explaining the reasons behind international students' decisions to stay in or leave the United States after graduating with a STEM degree (Gesing and Glass, 2019). In Italy, it is also very interesting to focus on the internal migration of students from the south to the north of the country based on university enrolment. This phenomenon is detrimental to the southern regions because this migratory flow is essentially a "one-way trip". In 2009, the Association for the Industrial Development of Southern Italy (Svimez, 2009) warned that "only a third of Southern migrant students returns to Southern regions after graduation", which means that graduates who return to Southern regions do not balance the number of outgoing "local" graduates. Therefore, while international students' mobility is surely a great opportunity from the students' internal perspective, from a different point of view, it is creating a new form of inequality between these two geographical areas. Namely, there

is an important and growing flow of graduated students from the regions of southern Italy who migrate towards the universities in the Northern areas, which are the richest in the country ( D'Agostino et al., 2019; De Santis et al., 2019; Giambona et al., 2017). The southern regions that suffer from this constant out-migration of more outstanding and/or richer students have to cope with a substantial decrease in their human capital. Consequently, in the southern regions, the human capital depletion process is discussed because the effects of the decreased university enrolment are combined with those of the migration towards the central and northern areas. According to Res and Viesti (2016), an emptying of southern universities is taking place, which constitutes a real threat to the socio-economic development of these regions. The negative effects of student migration between the Italian regions are further increased if the analysis focuses on STEM degrees due to their crucial role in long-term productivity and growth (De Philippis, 2017). Figure 2 compares the observed trend of movers between STEM vs non-STEM education across the years in our working sample. These results reveal a rather higher prevalence of movers for STEM education than that of the others.

Therefore, the goal of the current article is to dynamically investigate which individual and contextual determinants affect the migration of STEM students from southern to northern/central regions. This analysis indeed integrates two research directions, namely, STEM education research and migration studies, with the aim to illuminate the role of STEM education in the era of global migration in Italy. Specifically, we will focus on STEM enrolment.



Note: Our elaboration on the MIUR data.

Figure 2: Trend of movers from the South in STEM and non-STEM education

## 3 Data and Methods

### 3.1 Sample selection and description

The STEM definition/acronym was originally used by the education-related programmes of the National Science Foundation (NSF), but it was not explicitly defined by the NSF. In Italy, indeed, no common and detailed definition of which fields of study constitute the STEM core disciplines exists. In this paper, we updated the grouping of the STEM macro-categories provided by the EU 2015 report. The equivalent STEM fields of study can be classified into the following 3 categories: (i) natural sciences, mathematics and statistics; (ii) information and communication technologies; and (iii) engineering, manufacturing and construction. Similar to Chise et al. (2019), we also included architecture because it has many interactions with civil engineering studies. This study relies on micro data from the Italian Ministry of Education, University and Research (MIUR)<sup>2</sup>. To analyse the mobility of STEM student enrolment, 8 cohorts of students enrolled in STEM fields Southern Italy from 2008/2009 to 2015/2016 were used. These cohorts comprised freshmen who chose STEM education in and beyond the regions of southern Italy. A further selection criterion was that students had not been enrolled in online universities. Furthermore, we focus on Italian high school graduates and exclude from our sample foreign students who completed their high school abroad.

We retain students who chose a 3-year cycle degree (i.e., “Laurea”) and students who chose a 5-year cycle degree (i.e., “Laurea magistrale a ciclo unico”) with no missing covariate values in the empirical analysis (dropped missing values account for 2% on average of the initial sample size). We end up with a final annual sample that varies from a maximum of 30,651 to a minimum of 25,874 records in 2008 and 2014, of which approximately 43% are female students. We use both MIUR and ISTAT data. The MIUR data include information on high school degree, gender, high school grades, and the timing of their university enrolment following high school graduation. The ISTAT data include information on the youth unemployment rates across provinces/years and provide information on the distance matrix in kilometres between municipalities in Italy. The province (NUTS3 level) where students attended high school is also included in the database; accordingly, the freshmen are grouped into the 38 provinces located in southern Italy. The out-migration mobility phenomenon was monitored using a dummy variable that takes a value equal to 1 if a student who chose a university STEM degree migrated from a southern province to a university in a central or northern region. On average, migration accounts for approximately 22% of the total number of students under investigation ( $n=223,552$ ).

### 3.2 Econometric model

For the purposes of this present study, multilevel logit models were used (Rabe-Hesketh and Skrondal, 2012). This was done to take into account the hierarchical structure of

---

<sup>2</sup>Database MOBYSU. IT [Mobilità degli Studi Universitari in Italia], research protocol MIUR - Universities of Cagliari, Palermo, Siena, Torino, Sassari, Firenze and Napoli Federico II, scientific reference Prof. Massimo Attanasio (UNIPA), Data Source ANS-MIUR/CINECA.

the data, in which freshmen are grouped according to the province where they attended high school and to explicitly consider the effects of contextual factors on STEM student mobility. This approach has been applied by D'Agostino et al. (2019) in the same framework and by other empirical analyses concerning inter-regional mobility (see, for instance, Ferrante et al., 2019; Luo and Kwok, 2012). In this statistical framework, the probability of being a mover  $\Pi_{ij}$  for the  $i$ -th ( $i=1\dots n_j$ ) student that attended high school in the  $j$ -th ( $j=1\dots J$ ) province was therefore modelled as a function of individual and contextual characteristics. The response variable was coded as a dummy variable that indicates whether a student is a mover or a stayer (stayer – baseline –, or mover).

The individual characteristics include the covariate GENDER (man – baseline –, or woman); high school grades (GRADE), which was a continuous variable centred around the grand mean; and an indicator of whether the student enrolled immediately after high school graduation (LATE), which was coded as a dummy variable (enrolled immediately after the high school graduation – baseline– or enrolled after 1 year or more). The type of school was classified into three categories: scientific lyceum (baseline), which provides mostly theoretical training and prepares students for STEM career; classic lyceum (CLASSIC Lyceum) and technical and vocational (OTHER SCHOOL). Finally, the distance between the municipality where the student attended high school and the capital city of the region in which the province belongs (DISTANCE) was also computed from the ISTAT information on the distance matrix between municipalities and included in the model as a crude control for living in more outlying areas.

In addition to student-level characteristics, the key explanatory variables at the province levels are as follows. The macro region (ISLAND) indicates where the province is located (mainland – baseline – or one of the two major islands), and the supply of STEM courses (SUPPLY) and the youth unemployment rate (UNEMPL) at the NUTS3 level are also included. These variables represent salient aspects of the local conditions that are particularly relevant for students. The descriptive statistics of individual variables across years are reported in Table 1 and Table 2.



Table 1: Descriptive Statistics - a.y. 2008-2011

	2008		2009		2010		2011	
	Movers	Stayers	Movers	Stayers	Movers	Stayers	Movers	Stayers
	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd
LATE	0.05	0.06	0.03	0.08	0.04	0.07	0.04	0.08
GRADE	84.82	82.35	83.46	81.05	83.89	81.31	83.76	81.74
GENDER	12.72	12.74	12.21	12.24	12.18	12.20	12.20	12.13
DISTANCE	0.41	0.44	0.42	0.44	0.43	0.43	0.42	0.44
	116.39	79.98	118.02	76.31	114.87	74.54	115.23	76.24
	68.32	75.28	68.14	74.24	66.99	73.02	68.66	73.80
CLASSIC (Lyceum)	0.15	0.11	0.16	0.11	0.16	0.12	0.15	0.13
SCIENTIFIC (Lyceum)	0.59	0.55	0.62	0.55	0.62	0.57	0.64	0.58
OTHER (SCHOOL)	0.25	0.35	0.23	0.34	0.23	0.31	0.21	0.29
<i>N</i>	4857	25783	5752	24104	6192	22800	6350	21280

Table 2: Descriptive Statistics - a.y. 2012-2015

	2012		2013		2014		2015	
	Movers	Stayers	Movers	Stayers	Movers	Stayers	Movers	Stayers
	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd
LATE	0.05	0.07	0.05	0.09	0.06	0.09	0.05	0.10
GRADE	83.46	81.79	84.98	82.60	85.02	82.41	85.40	82.80
	12.03	11.78	11.95	11.85	11.88	11.86	11.72	11.88
GENDER	0.44	0.44	0.41	0.42	0.41	0.42	0.42	0.41
DISTANCE	117.01	74.72	116.93	75.45	117.22	74.24	115.65	76.81
	68.49	73.56	70.26	73.62	71.68	73.54	68.75	75.44
CLASSIC (Lyceum)	0.16	0.13	0.15	0.13	0.15	0.12	0.13	0.11
SCIENTIFIC (Lyceum)	0.64	0.59	0.63	0.58	0.64	0.59	0.65	0.60
OTHER (SCHOOL)	0.20	0.28	0.22	0.29	0.22	0.29	0.21	0.29
<i>N</i>	6358	20603	5972	20212	6067	19757	6611	20701

The considered statistics are the mean and the standard deviation (SD), which are only for the two continuous variables. Each statistic was computed using the sample of students excluding the missing values. On the whole, as expected, movers appear to have better school career characteristics than stayers, with a certain amount of variability across years. Indeed, the proportion of stayers who enrol late at a university increases in the 2008-2015 period. However, this proportion is not only constant over time between the group of movers but is also lower by more than 3% on average compared to that of the stayers, highlighting that migrating students have more regular school careers. The average high school grades of movers are always higher than those of stayers. Moreover, while the value remains quite stable across the years for stayers, it increases for movers in the last three years.

Notable differences are found with respect to the variable that expresses the distance between the municipality where students attended high school and the capital city of the region in which the municipality belongs. It appears that a greater distance from the regional capital is a characteristic of the group of movers. On average, the movers reside at a distance of over 115 km from the capital, while the average distance of the stayers is equal to 76 km. Even the type of school seems to characterize the two groups. The group of movers is composed to a greater extent by students who have attended a classic or scientific lyceum, while the group of stayers is composed of a higher proportion of students attending other schools (30%).

The econometric model can be written as follows:

$$\text{logit}(\Pi_{ij}) = \beta_0 + \beta_1 \text{LATE}_{ij} + \beta_2 \text{GRADE}_{ij} + \beta_3 \text{GRADE}_{ij}^2 + \beta_4 \text{GENDER}_{ij} + \beta_5 \text{CLASSIC}_{ij} + \beta_6 \text{OTHER}_{ij} + \beta_7 \text{DISTANCE}_{ij} + \beta_8 \text{UNEMPL}_j + \beta_9 \text{SUPPLY}_j + \beta_{10} \text{ISLANDS}_j + u_j \quad (1)$$

where the  $\beta_s$  ( $s=1..10$ ) are the regression coefficients to be estimated and the random effect  $u_j$  is specific for province  $j$ . The random effect is assumed to follow a normal distribution with a variance of  $\sigma_u^2$ .

Separate models were estimated for each cohort. The variance partition coefficient (VCP) was applied to quantify the proportion of observed variation in the outcome that is attributable to the effect of clustering<sup>3</sup>. Finally, the predicted subject-specific probabilities were computed. The predictions refer to hypothetical students with specific values of  $x$  ( $x=x^*$ ) and  $z$  ( $z=z^*$ ) and with a random effect  $u_j$  set to percentiles of its estimated distribution. A hypothetical mean (median) province is defined when the group-level residual  $u_j$  is assumed to be equal to 0 (i.e., it has an average level of movers).

<sup>3</sup>There are a variety of procedures for calculating the VPC for binary responses (Goldstein et al., 2002); among these different procedures, we used the VPC based on the latent response formulation of the model, i.e.,  $\text{VPC} = \sigma_u^2 / (\sigma_u^2 + \pi^2 / 3)$ .

## 4 Results

We fit eight respective Random Effect (RE) logistic regression models using the respective data for each a.y<sup>4</sup>. Table 3 displays the estimated regression coefficients of the model specified in equation (1) as well as the estimates of the variances of the distributions of the random effects and the VPCs. For each model, the likelihood ratio *chisq*<sup>2</sup> indicates that there is a significant difference between the standard logistic estimate and the multilevel logistic estimate<sup>5</sup>. Thus, the unmeasured province-specific factors affecting the propensity of moving have been partially captured by including province-specific random effects. The residual variation in outcomes that remains after accounting for the variables in the model does not vary much over time. Systematic differences between provinces account for 15% of the total variation in 2008 and 19% in 2015. Most of the individual variables are statistically significant at the 1% level over time and have the expected signs.

Table 4 presents the odds ratios obtained by exponentiating only the estimated regression coefficients related to the student level variables.<sup>6</sup> These odds ratios are cluster-specific measures of association. This means that they are province-adjusted associations between student characteristics and mobility.

Accordingly, the estimated odds ratios suggest that when comparing two students within the same province who have different genders but who share identical remaining covariates, the odds of being a mover for a women is 26% and 13% less likely than the odds of being a mover for a man in 2008 and 2015, respectively. This finding is consistent with other studies related to the mobility of Italian students (Contini et al., 2015) and, more generally, to gender inequalities in Italy, where there are large gender-related differences in terms of female labour force participation, unemployment rates and earnings (Belloc and Tilli, 2013). In addition, high school grades have a quadratic effect on the probability of moving at least in the last five years. In particular, the coefficient associated with the squared grade is significant and positive. This U-shaped relationship with the mobility likelihood indicates that the probability of being a mover strongly increases for the best-performing students.

By contrast, enrolling later decreases the probability of moving. The odds ratios vary from 0.6470 in 2011 to 0.8488 in 2015. Moreover, distance plays a significant role, as students residing in more isolated municipalities (with greater distances from the regional capital) will be more willing to enrol in a university in the northern/central regions than to move (or change residence) within southern regions. When examining the odds ratios in Table 4, one could interpret the odds ratio for distance in 2008 (1.8497 per 50-km increase in distance) as suggesting that the odds of being a mover is approximately 2

---

<sup>4</sup>It is worth noting that the choice to estimate a different model for each a.y. was motivated by preliminary analyses based on a pooled multilevel model with time dummies and interaction effects between them and some relevant covariate. Most of the interaction effects resulted statistical significant, so we decide to leave all coefficients to vary over a.y. and we estimated different models.

<sup>5</sup>We do not report the *chisq*<sup>2</sup> statistics in Table 3 for the sake of simplicity.

<sup>6</sup>We do not report the odds ratios of the provincial characteristic variables because their interpretations present some difficulties, as reported in Neuhaus et al. (1991) and Neuhaus and McCulloch (2006).

times higher for a student who lives in a municipality that is 50 km from the capital city of its region compared to the odds of being a mover for a student who lives closer to the capital city. These odds ratios are quite stable over the years.

Finally, the positive and negative impacts of the two dummies for school imply that students attending classic lyceum are more likely to move than those attending scientific lyceum, whereas students attending other schools are less likely to move than those attending scientific lyceum.

The classic lyceum can be interpreted as a proxy of students from more affluent families (Schizzerotto, 2006) who have more economic resources to guarantee subsistence for studying abroad and are more eager to invest in their children's human capital. In this light, these results are broadly consistent with several studies that highlight the role of familiar socio-economic background on the probability of choosing a university in a different area (Lupi and Ordine, 2009; Ishitani, 2011).

Regarding the second-level characteristics, some interesting results emerge in Table 3. First, the probability of moving decreases as the number of STEM degrees increases in the province of origin; therefore, as shown by Lupi and Ordine (2009), in provinces with a greater supply of tertiary education, there will be a lower probability of moving; second, students who attended high school in Sardinia or Sicily are less likely to leave their regions of origin for college education, as was also highlighted by Contini et al. (2015). Third, the youth unemployment rate appears to not have any influence on the probability of moving.

Table 3: Results of the RE logistic regression over the years

	2008	2009	2010	2011	2012	2013	2014	2015
LATE	0.1300 (0.0834)	-0.3226** (0.0853)	*0.0342 (0.0795)	-0.4354** (0.0757)	*0.2843** (0.0747)	*0.4240** (0.0735)	*0.1954** (0.0704)	-0.1639* (0.0690)
GRADE	0.0154** (0.0014)	*0.0175** (0.0014)	*0.0127** (0.0014)	*0.0107** (0.0014)	*0.0180** (0.0015)	*0.0204** (0.0015)	*0.0205** (0.0015)	*0.0205** (0.0015)
GRADE(SQUARED)	0.0003* (0.0001)	-0.0001 (0.0001)	0.0003* (0.0001)	0.0003* (0.0001)	0.0006** (0.0001)	*0.0005** (0.0001)	*0.0004** (0.0001)	0.0004** (0.0001)
GENDER	-0.2989** (0.0366)	*0.2473** (0.0354)	*0.1319** (0.0344)	*0.2136** (0.0342)	*0.1510** (0.0340)	*0.2330** (0.0354)	*0.2037** (0.0356)	*0.1333** (0.0345)
DISTANCE	0.0123** (0.0006)	*0.0132** (0.0006)	*0.0134** (0.0006)	*0.0118** (0.0006)	*0.0132** (0.0006)	*0.0115** (0.0006)	*0.0123** (0.0006)	*0.0124** (0.0006)
CLASSIC Lyceum	0.3421** (0.0527)	*0.3753** (0.0506)	*0.3501** (0.0490)	*0.2200** (0.0489)	*0.2079** (0.0477)	*0.1272* (0.0506)	0.3018** (0.0512)	0.1633** (0.0518)
OTHER SCHOOL	-0.5321** (0.0413)	*0.6248** (0.0407)	*0.5381** (0.0401)	*0.5847** (0.0404)	*0.5629** (0.0409)	*0.5065** (0.0411)	*0.5953** (0.0418)	*0.6286** (0.0404)
UNEMPLOYMENT RATE	0.0030 (0.0150)	0.0081 (0.0172)	-0.0162 (0.0181)	-0.0225 (0.0176)	0.0129 (0.0145)	0.0027 (0.0129)	-0.0325* (0.0154)	-0.0098 (0.0159)
SUPPLY	-0.0603** (0.0126)	*0.0707** (0.0154)	*0.0631** (0.0163)	*0.0653** (0.0146)	*0.0721** (0.0160)	*0.0719** (0.0150)	*0.0650** (0.0153)	*0.0664** (0.0172)
ISLANDS	-1.6160** (0.2866)	*1.6576** (0.3069)	*1.4004** (0.3119)	*1.1295** (0.2828)	*1.3337** (0.2954)	*1.0828** (0.2848)	*1.0455** (0.2722)	*1.3449** (0.3068)
Constant	-2.5473** (0.1996)	*2.2972** (0.2218)	*2.3483** (0.2281)	*2.1460** (0.2025)	*2.2286** (0.2073)	*2.1438** (0.2022)	*2.2096** (0.2044)	*2.0454** (0.2268)
$\sigma_u^2$	0.57	0.72	0.75	0.59	0.63	0.59	0.60	0.76
VPC	0.15	0.18	0.19	0.15	0.17	0.15	0.15	0.19
N	30651	29866	29003	27640	26966	26223	25874	27329

Notes: \* p&lt;0.10; \*\*p&lt;0.05; \*\*\*p&lt;0.001.

Table 4: Estimated odds ratios for the RE logistic regression over the years

	2008	2009	2010	2011	2012	2013	2014	2015
LATE	1.1388 (0.0950)	0.7243** (0.0617)	0.9664 (0.0768)	0.6470** (0.0490)	0.7525** (0.0562)	0.6544** (0.0481)	0.8225** (0.0579)	0.8488* (0.0586)
GRADE	1.0156** (0.0014)	1.0163** (0.0015)	1.0176** (0.0014)	1.0127** (0.0014)	1.0108** (0.0014)	1.0182** (0.0015)	1.0206** (0.0015)	1.0207** (0.0015)
GRAND_GRADE2	1.0003* (0.0001)	0.9999 (0.0001)	1.0001 (0.0001)	1.0003* (0.0001)	1.0006** (0.0001)	1.0005** (0.0001)	1.0004** (0.0001)	1.0004** (0.0001)
FEMALE	0.7416** (0.0271)	0.7809** (0.0277)	0.8764** (0.0302)	0.8077** (0.0276)	0.8599** (0.0292)	0.7922** (0.0281)	0.8157** (0.0291)	0.8752** (0.0302)
DISTANCE (per 50-km increase)	1.8497** (0.0006)	1.9348** (0.0006)	1.9542** (0.0006)	1.8040** (0.0006)	1.9348** (0.0006)	1.7771** (0.0006)	1.8497** (0.0006)	1.8589** (0.0006)
LICEO CLASSICO	1.4078** (0.0742)	1.4554** (0.0736)	1.4192** (0.0696)	1.2461** (0.0610)	1.2311** (0.0587)	1.1356* (0.0575)	1.3522** (0.0693)	1.1773** (0.0610)
OTHER SCHOOLS	0.5874** (0.0242)	0.5354** (0.0218)	0.5838** (0.0234)	0.5572** (0.0225)	0.5695** (0.0233)	0.6026** (0.0248)	0.5514** (0.0231)	0.5333** (0.0215)

Notes: \* p<0.10; \*\*p<0.05; \*\*\*p<0.001.

The estimated regression coefficients were then used in the prediction to better highlight, from a descriptive point of view, the effects of the covariates over time. The predicted subject-specific probabilities comparing different hypothetical individual profiles are presented in Figures 3 and 4. The baseline profile was defined as follows: a male student living on the mainland who attended a high school in the capital city and graduated from a scientific lyceum in the same year when he enrolled in a university with average grades. In addition, he obtained his high school diploma in a province with an average level of movers (the random effect is fixed to zero). In Figure 3, this base profile is compared with three other profiles that differ for just one specific characteristic.

Figure 3 shows the trend in the predicted probabilities recorded from 2008 to 2015 according to gender, better high school grades and an irregular university enrolment time. According to our findings, the estimated probabilities follow an upward trend for the entire period, apart from a more irregular pattern for students who enrol in college late, especially for the first four years. The predicted probability of moving for the baseline student grew by a total of approximately 2% from 2008 to 2015. This figure, however, hides large differences between the best performing students, where the predicted probability grew by a total of approximately 6%. In addition to the more rapid growth in the probability of moving, the estimated probability of moving for the best performing students remains substantially greater than the baseline. The differences in estimated probabilities over the period between men and women indeed remain constant and women always have lower probabilities of moving than men. Next, Figure 4 compares different individual profiles with the baseline. A good/poor provincial context is characterized, according to our findings, by the availability of many/few university courses, by the youth unemployment rate being statistically insignificant and by the random effect  $\mu_j$  being equal to  $-1.96\sigma_u + 1.96\sigma_u$ . The predicted probabilities show a constant advantage for good provinces in terms of retaining STEM students over the entire period of time.

## 5 Some final remarks

The economic and workforce development of southern regions in Italy requires a greater understanding of STEM students' intentions to stay in the south for their higher education studies or leave the south after their completion of high school. This is mainly because STEM students can be considered to be the "catalysts" of the socio-economic development of a region. Our study highlighted that the decision to move towards the universities in the North is mostly undertaken by STEM students, and this phenomenon has grown in recent years.

Therefore, a negative spiral is at work with respect to the socio-economic development of the southern regions facing the progressive reduction of the most specialized human resources with the highest potential (such as STEM students). This increasing reduction, in turn, could generate a constant decrease and worsening of the socio-economic conditions of these areas and, therefore, a greater push to move elsewhere, which results in a loop with very negative effects for these territories.



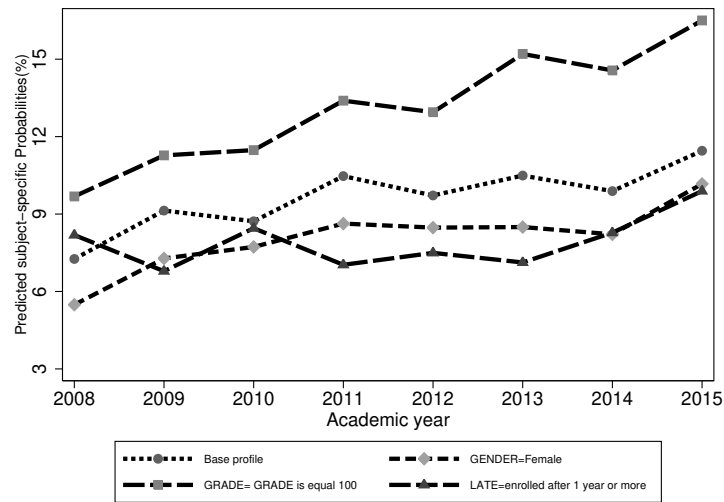


Figure 3: Estimated probabilities of different individual profiles over the years

Accordingly, this study explores the individual and contextual determinants of STEM mobility using a multilevel logistic regression model that accounts for the hierarchical structure of the data. From a policy point of view, the identification of the key determinants of this issue could be central to designing efficient policies aimed at reducing the number of students moving from the southern regions.

Our findings suggest that gender disparities persist over the period. In addition, students with a better (higher high school grades), more regular (they enrol in a university without delay) school career who obtained a classic lyceum and obtained their high school diploma in municipalities with greater distances from the regional capital migrate more than the others.

These results are basically in line with those of other studies, and therefore, they may also contribute to the literature on social equity in higher education by highlighting that gender and social class (indirectly measured by the type of school) inequalities are additional issues that can be addressed in this particular framework.

Beyond the individual characteristics mainly related to family background, the supply of tertiary education also appears to play a remarkable role in determining the probability of student migration, and its effect is always negative, as expected. This means that a rich and varying supply of educational courses should discourage the loss of students, at least at enrolment. We find that the local youth unemployment rate does not have an effect.

In terms of temporal dynamics, interesting conclusions can be made at least from a descriptive point of view. Indeed, the signs of the effects of the individual and contextual variables do not change over time, but the estimated probability for different individual profiles generally has a positive trend. Namely, over the years, the effects of the covariates tend to become more pronounced. The phenomenon is, however, remarkable for top-

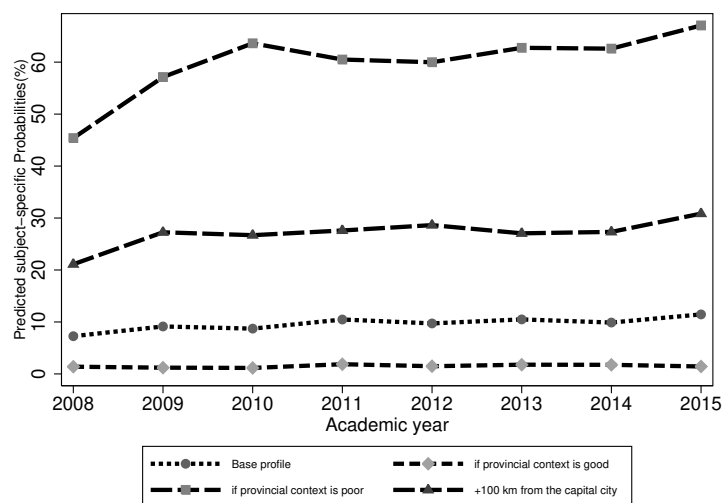


Figure 4: Estimated probabilities of different individual profiles over the years

performing students whose probability of moving grows constantly from year to year.

These results constitute a wake-up call for political decision-makers for at least a few reasons.

First, these results highlight the need for policies that are able to stop these human capital losses that generate a negative spiral affecting both the survival of the universities in the southern regions and the economic growth prospects of these areas. In fact, in southern universities, we are witnessing a progressive deprivation mechanism that, starting with the constant decrease in student enrolment, leads to a lower availability of financial resources for the universities. This situation, in turn, leads to a lower variety and supply of educational courses and fewer services, which in turn negatively influence university enrolment.

Second, our results may suggest the need of extraordinary efforts towards the southern regions, both with educational policies (i.e. investments for increasing the number of STEM courses for reducing the mobility and brain drain), and some economic reform (i.e. economic support to innovation, digitalization process, industry 4.0, etc.).

The recent European Community report (Commission, 2019) indeed shows that the divergence between the North and South with regard to investments is particularly marked in the sphere of intangible assets and innovation. The number of individuals employed in the high-tech sector is almost double that in the North (3.7% vs. 2% in the South), and R&D expenditures are 1.5 times greater (1.4% of regional GDP vs. 0.9% in the South). The number of patents per capita is 10 times greater in the North than in the South (106.8 vs. 10.1 per million inhabitants, respectively).

Consequently, policies that aim to increase the supply of university courses must necessarily be framed in a broader context of structural reforms that aim to revive southern Italy and put it in a position to at least partially decrease the gap that separates

them from the most developed and advantaged regions of Italy and Europe.

Finally, our analysis has three main limitations. First, the data do not control for the economic standards of living of the families, which are recognized worldwide as an important explanatory factor for student migration. Nevertheless, controlling for high school diplomas could mitigate this lack. Second, the econometric model controls only for a limited number of contextual variables, namely, the supply of STEM courses and the youth unemployment rate. Nevertheless, obtaining reliable dynamic indicators at the NUTS3 level is not an easy task because of the comparability of data sources across time. Third, a more complex econometric specification could be implemented (i.e., allowing for a random slope model). However, the proposed model, due to its simplicity, offers some interesting points for consideration and relevant issues to be addressed in the near future.

## References

- Atkinson, R. D. and Mayo, M. J. (2010). Refueling the us innovation economy: Fresh approaches to science, technology, engineering and mathematics (stem) education. *The Information Technology & Innovation Foundation, Forthcoming*.
- Belloc, M. and Tilli, R. (2013). Unemployment by gender and gender catching-up: Empirical evidence from the italian regions. *Papers in Regional Science*, 92(3):481–494.
- Chise, D., Fort, M., and Monfardini, C. (2019). Scientifico! like dad: On the intergenerational transmission of stem education in italy.
- Commission, E. (2019). Country report italy 2019 including an in-depth review on the prevention and correction of macroeconomic imbalances. Technical report, European Commission, Brussels.
- Contini, D., Cugnata, F., and Scagni, A. (2015). From south to north: student internal migration in italy. should it be an issue?
- D’Agostino, A., Ghellini, G., and Longobardi, S. (2019). Out-migration of university enrolment: the mobility behaviour of italian students. *International Journal of Manpower*, 40(1):56–72.
- De Angelis, I., Mariani, V., and Torrini, R. (2017). New evidence on interregional mobility of students in tertiary education: the case of italy. *Politica economica*, 33(1):73–96.
- De Philippis, M. (2017). Stem graduates and secondary school curriculum: does early exposure to science matter? *Bank of Italy Temi di Discussione (Working Paper) No*, 1107.
- De Santis, G., Pirani, E., and Porcu, M. (2019). Rapporto sulla popolazione. l’istruzione in italia.
- Fayer, A. L. S. and Watson, A. (2015). Stem occupations: Past, present and future. technical report. Technical report, US Bureau of Labor Statistics.
- Ferrante, M., Sciuto, V., Pollina-Addario, S., and Parroco, A. M. (2019). Individual

- and contextual determinants of inter-regional mobility in cancer patients. *Electronic Journal of Applied Statistical Analysis*, 12(1):14–25.
- Gesing, P. and Glass, C. (2019). Stem student mobility intentions post-graduation and the role of reverse push-pull factors. *International Journal of Educational Development*, 65:227–236.
- Giambona, F., Porcu, M., and Sulis, I. (2017). Students mobility: Assessing the determinants of attractiveness across competing territorial areas. *Social indicators research*, 133(3):1105–1132.
- Goldstein, H., Browne, W., and Rasbash, J. (2002). Partitioning variation in multilevel models. *Understanding statistics: statistical issues in psychology, education, and the social sciences*, 1(4):223–231.
- Goos, M., Hathaway, I., Konings, J., and Vandeweyer, M. (2013). High-technology employment in the european union. *VIVES discussion paper 41*, pages 1–59.
- Hox, J. (2010). Multilevel analysis: Techniques and applications. routledge. *New York*.
- ICF, for the European Commission, C., et al. (2015). Eu skills panorama (2014) stem skills analytical highlight, 1-5.
- Ishitani, T. T. (2011). The determinants of out-migration among in-state college students in the united states. *Research in Higher Education*, 52(2):107–122.
- Luo, W. and Kwok, O.-m. (2012). The consequences of ignoring individuals' mobility in multilevel growth models: A monte carlo study. *Journal of Educational and Behavioral Statistics*, 37(1):31–56.
- Lupi, C. and Ordine, P. (2009). Family income and students' mobility. *Giornale degli Economisti e Annali di Economia*, pages 1–23.
- Neuhaus, J. M., Kalbfleisch, J. D., and Hauck, W. W. (1991). A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. *International Statistical Review/Revue Internationale de Statistique*, pages 25–35.
- Neuhaus, J. M. and McCulloch, C. E. (2006). Separating between-and within-cluster covariate effects by using conditional and partitioning methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(5):859–872.
- Peri, G., Shih, K. Y., and Sparber, C. (2014). Foreign stem workers and native wages and employment in us cities. Technical report, National Bureau of Economic Research.
- Rabe-Hesketh, S. and Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata*. STATA press.
- Ray, R. (2015). Stem education and economic performance in the american states.
- Res, F. and Viesti, G. (2016). *Università in declino: un'indagine sugli atenei da Nord a Sud*. Donzelli editore.
- Schizzerotto, A. (2006). Le disuguaglianze nelle opportunità di istruzione. *SCHIZZEROTTO A., BARONE C., Sociologia dell'istruzione, Bologna, Il Mulino*.
- Solow, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, pages 312–320.
- Svimez (2009). *Rapporto sull'economia del Mezzogiorno*. Il Mulino.

- Winckler, G. and Fieder, M. (2006). Declining demand among students for science and engineering. *Universities and business: Partnering for the knowledge society*, pages 233–241.
- Winters, J. V. (2018). Do higher levels of education and skills in an area benefit wider society? *IZA World of Labor*.