

Disentangling the secondary school effects in the assessment of university inefficiencies

Cristian Usala, Isabella Sulis, Nicola Tedesco, Mariano Porcu, Barbara Barbieri

ABSTRACT

The inefficacy of the Italian university system has been discussed in the relevant scientific literature, mainly using data from single universities or survey data. This paper uses the data provided by the National Student Archive related to eight cohorts of first-year students who enrolled in Italian universities between 2010 and 2018 to study the risk of dropout and churn during the 1st year. Revisiting the model of Tinto, to include in the analysis the possible effect of university services and linking each student to the high school attended before enrolment, the work has two primary aims: (i) to investigate the determinants of university failure, focusing on how strategic factors - such as students' mobility choices, previous educational background, and university policies addressed to improve digital communication and services - affect the probability of experiencing churn or dropout; (ii) to disentangle the secondary school effect on these two outcomes by adopting a two steps procedure which relays on multilevel logistic models. Results allowed us to estimate adjusted indicators of inefficiency, which are suitable for comparisons across universities that offer degree programs in the same disciplinary field and provide evidence on the role played by the secondary school attended in determining substantial inequalities at entrance.

Keywords: *university dropout, university services, secondary school effect, multilevel models, university failure*

1. INTRODUCTION

According to the OECD statistics in 2020 (OECD, 2022) in Italy, the completion rate of full-time students who entered a bachelor (or equivalent level) program within three years from the enrolment was about 50% for men and 56% for women (these rates are respectively equal to 19% and 22% if we look only at those who graduate in time). These rates place the country in the last position on the international scene. The indicators commonly adopted to monitor the inefficiency of university systems (delay in graduation and dropout rate) depict an alarming picture of the Italian tertiary education national system, with values of the key indicators for monitoring the system well below the average of the main OECD countries (OECD 2010; 2022; Allegro & Giambalvo 2020; Aina et al. 2022).

Several studies have shown that the factors influencing this phenomenon are manifold and point at students' socioeconomic and educational backgrounds, as well as institutional support structures, which play a relevant role in shaping academic careers (e.g., Di Pietro & Cutillo 2008; Belloc et al. 2011; Burgalassi et al. 2016; Meggiolaro et al. 2017). At the enrolment, university guidance services play a crucial role to correct the mismatching between what has been learned at school and what is required at university. Through communication and digital services (university websites, official social profiles), universities may intervene in redirecting students to a more suitable course and avoid early abandonment (Viale Tuleda 2014).

However, the variability of the dropout rates between students from different high schools emphasises that the university guidance services are not the only ones responsible. Nevertheless, as highlighted in the literature, the meaning of educational experiences before entering university provides information on how students will respond to their educational environment and their ability to “fit in” at university. The first phase of Tinto's model (1993) considers the relative background built during the pre-entry phase, labelled by the author as “pre-college schooling”, as a critical element to consider. Therefore, we believe it is essential to consider the so-called “school effect” as an indirect indicator

of the school's effectiveness in providing students with the skills required by the university. The second phase also considers external factors as crucial in the choice to continue or quit university studies, in which universities play a fundamental role in supporting the motivation to remain and in compensating for any lack of cultural and human capital due to the educational and family background. Literature has already highlighted how formal guidance strategies and peer support help students. As far as we know, insufficient research explores how much communication and digital services the university offers can influence the choice to continue or drop out.

In light of the above, this work aims to study the risk of early dropout and churn in Italian universities, taking into account students' previous educational background and the field of study chosen at the university through revisiting the model of Tinto (1975; 1993) to include in the analysis the possible effect of university services and linking each student to the attended school. The main research questions addressed in this study are: (i) How much of the inefficiency in academic carrier disruption is ascribable to the university peculiarities, and how much to differences in school effectiveness?; (ii) Can university policies addressed to improve communication and digital services prevent early dropout or churn?; (iii) What are the main determinants of students dropping out and churning in the 1st year of their career?

To this end, we used the microdata provided by the Italian Ministry of University and Research in its National Student Archive (ANS – *Anagrafe Nazionale Studenti*). We considered different cohorts of students enrolled in Italian universities from 2010-2018 to classify students according to their careers after one year (MOBYSU.IT Database 2022). The following categories of students were defined: regular, dropout, and churn. We considered information on students' demographic characteristics, previous education experiences, mobility experience, and the field of studies at the university. Interaction between the specific secondary school attended and the field of study has been specifically considered to analyse divergences in the secondary school effect according to the field of study. To investigate the possible role of university support structures, CENSIS composite indicators related to

the quality of university facilities, communication and digital services, and grants for 2010-2018 have been considered (CENSIS 2022). Moreover, information on the labour market conditions of the area of origin and destination have been included as control variables to take into account that the observed divergences in the dropout rates between students coming from different schools may, at least partially, reflect the differences in the labour opportunities for technical and professional schools compared with academic schools (the so called *Licei*) (Di Pietro, 2006; Contini & Zotti, 2022).

To consider how the secondary school attended affects university dropout over and above the range of factors described, a two-step procedure based on two multilevel (ML) logit models (ML logit model in Step 1 and ML multinomial logit model in Step 2) was used to disentangle the secondary school effect from the influence of the university services in preventing dropout and churn by accounting also for individual characteristics and the chosen field of study. In Step 1 we consider divergences in the dropout rates across students who attended the same secondary school and enrolled in the same disciplinary fields (i.e. in different universities). In contrast, in Step 2, the focus is on assessing divergences in dropout and churn rates of students enrolled in different disciplinary fields of different universities. Thus, in the 1st step, a proxy indicator of the effect of the high school attended on the propensity to drop out in the different disciplinary fields is estimated and entered in the 2nd Step as a predictor.

The present study contributes to the debate on the determinants of university dropout. It uses a unique national dataset related to nine cohorts of students in Italy considered longitudinally. Our contribution to this literature is threefold and addressed to (i) disentangle the high schools' effect from the university effect in estimating university inefficiency outcomes. In pursuing this task, we (ii) control the role played by students' mobility choices and previous educational background. We also assess the role of university policies addressed to improve digital communication and services in preventing the disruption of students' careers. Finally, the analysis allowed us to (iii) provide a set of adjusted indicators of university efficiency suitable for comparisons across universities.

The paper is organised as follows. In Section 2, we discuss the theoretical background. The data and variables considered in the analysis are described in Section 3. Section 4 presents the methodological approach adopted to assess the determinant of students' career disruption and to disentangle the high school effect from the university effect in university inefficiency. Section 5 reports the primary empirical evidence of the analysis, while Section 6 discusses the results and provides concluding remarks.

2. THEORETICAL BACKGROUND

The phenomenon of university dropout represents a vital issue in the global agenda, as it calls into question the efficiency of the higher education system in many countries.

Numerous studies on this subject first reflect the phenomenon's complexity in its definition. Among the different existing classifications, undoubtedly, a fundamental distinction is identifiable between dropout in relative terms (i.e., the student leaves a course of study to redirect to another study program) and dropout in absolute terms (the student quits university) (Hadjar et al. 2022). Additional classifications look at the time and space dimensions distinguishing between premature, early, and late dropout and internal, institutional, and total dropout in the space dimension (Viale Tudela 2014). The phenomenon's complexity also emerges from the evolution of the theoretical frames that frequently have as a conceptual starting point the work of Spady (1970) and Tinto (1975; 1993), who were the first to focus on students' academic and social integration. Tinto's model assumes that individual characteristics such as achievement in secondary school, family background, and individual attributes are connected to students' commitments and output at the beginning of their university career. Institutional commitments are associated with integrating the student into the academic system, which makes the interaction with student peers and faculty members (Hadjar et al. 2022). The importance of this last aspect justifies that scholars often integrate into Tinto's model institution-related factors to account for conditions such as institution expenditure, study programs,

activities of support, and so on (Pascarella et al. 2004; Georg 2009; Hadjar et al. 2022). Accordingly, some studies focus on the importance of favouring the integration of students from an academic and social point of view as an element to increase student commitment (Aina, Baici, Casalone, & Pastore, 2022). Moreover, several studies investigate the possible role of the dropout probability, the level of students' engagement, and their personal best goals and motivations (e.g., Reschly & Christenson 2022; Bernardo et al. 2022; Bardach et al. 2019). The literature highlights how the expectations about future labour market conditions, together with the capability of the universities to support students in their process of integration into the university environment, are both aspects that shed light on the process that brings students to make the decisions to dropout (Aina et al. 2022). Aina and colleagues (2022) identify five macro-categories of factors: i. student demographic characteristics and abilities; ii. parental background; iii. academic/social integration and institutional commitment; iv. features of the tertiary education system; v. labour market performance, which interact with university withdrawal, describing the process of career disruption as a complex phenomenon that needs to be analysed from multivariate perspectives.

Based on the available data, dropout evolution might be considered to develop through several stages (e.g. Leiviskä & Siponen 2013). Although several studies have contributed to explaining the phenomenon of dropout, most analyses used data from the local archives of single universities or survey data with the main aim of identifying students' profiles at risk of failure (Meggiolaro et al. 2017; Belloc et al. 2010; Di Pietro & Cutillo 2008), detecting the effect of inequalities in family background on dropout rates (Contini et al. 2018) and profile degree programs with dispersion rate higher and lower than the average (Clerici et al. 2015).

As regards the Italian context, few studies have been carried out using national data to shape the long-term trends of dropout and to assess its determinants (Tedesco & Salaris 2020), or to identify policies and practices that can prevent inefficiencies (Mealli & Rampichini 2012) and to assess the role played by secondary schools and universities in determining university inefficiencies (Perchinunno et al. 2019; Ballarino 2011; Triventi & Trivellato 2015).

In light of the above, the study contributes to advancing studies on this phenomenon, establishing a nuanced understanding of university dropout and churn within the Italian university system. In particular, the study proposes an advancement over Tinto's model. While Tinto's original model focuses on integrating the individual into the university environment as the main factor influencing student retention, this study explicitly extends the model to incorporate the impact of secondary education. Linking students' college outcomes to their high school background highlights the critical role of pre-college education in shaping college success. This aspect has received limited attention in previous applications of Tinto's framework.

From a methodological perspective, previous studies have often used single-level analyses that may not adequately account for the hierarchical nature of educational data. This article advances the field by employing multilevel logistic models, allowing for a more sophisticated analysis that disentangles the effects of individual, secondary school, and university characteristics. This approach provides a more accurate estimate of the factors contributing to dropout and churn, effectively addressing the nested structure of educational data.

Finally, while past research on university dropouts in Italy has often relied on data from individual universities or specific regions, this study uses the National Student Archive, which includes information on the university student population in Italy. This comprehensive use of data allows for a more generalized and robust analysis of school dropout and churn trends.

3. DATA AND DESCRIPTIVE STATISTICS

The analysis relies on the administrative database MOBYSU.IT (MOBYSU.IT 2022), which includes information on the careers of the population of students enrolled in a bachelor's degree program at an Italian university between 2010 and 2018. Since we are interested in the elements that affect students' career disruption, our analysis focuses on students enrolled between 2010 and 2018 and the 2018 cohort is the last one for which we observe the information on 2nd year students. Using this data,

we classify students according to a three-category classification: “*regular*” (i.e., students who enrolled in 2nd year in the same university), “*dropout*” (i.e., students who have quit), and “*churn*” (i.e., students who changed university between the first and second year). The analysis considered the following covariates related to students’ characteristics: age, gender, foreigner status, macro area of residence, year of enrolment, field of study according ISCED classification, and mobility status. Students’ mobility status is defined according to the supply of degree programs in their local area and the distance between the city of residence and the chosen universities. MOBYSU.IT contains information on car travel distances between students’ cities of residence and all Italian universities (computed using data from the National Statistical Institute - ISTAT). This information allows us to define, for each student, her/his local area as the area between her/his city of residence and the city that hosts the nearest university. Moreover, to account for the fact that many universities may be close to each other, we consider a broader definition of students’ local area by adding 15 minutes to the minimum distance. For example, suppose a student resides in a city within 30 minutes of travel from the nearest university. In that case, we consider all the universities located within 45 minutes of her/his city as local. Accordingly, students are defined as movers if they enrol in a university located outside their local area. We also distinguish between forced and free movers to account for the different supply of degree programs in students’ local areas, following a similar logic to that used in the study by Usala and colleagues (2023). Forced movers have chosen a university outside their local area but within 15 minutes from the nearest university that provides a degree program in their chosen field. Free movers are the residual category. This definition allows us to account for the differences between students with different mobility backgrounds without relying on administrative borders and considering the variability of universities’ and degree programs’ supply in the country. Moreover, it includes information about their educational background, associating data about high school grades, the type of high school, and the institute they attended. Indicators of the labour market conditions in the area where the university is located have been integrated into the database and considered as control variables to account for the influence of labour market dynamics on the decision to drop out.

The three indicators considered are the municipal income per capita, the unemployment rate at the province level, and the regional gross domestic product (GDP) measured one year before students' enrolment. Data on the provincial unemployment rate and regional GDP has been obtained from ISTAT, while the municipal income per capita has been calculated based on the information on the tax returns available from the open data portal of the Italian Ministry of Economy and Finance. However, it is worth highlighting that it is out of the scope of this paper to develop an econometric analysis of the effect of labour market supply on university dropout as we do not consider specific information about the occupation rates and salaries of students attending different types of high schools in the years and for the cohorts considered. Labour market indicators at province, regional, and municipality levels are considered to adjust the "secondary school effect" on dropout rates for the possible influence of general economic conditions of the local labour market.

Additionally, the literature recognizes the importance of considering integration aspects to prevent withdrawal which depend mainly on the range of activities offered within the campus. The information available at a national level for comparing university institutions in the time span considered is mainly related to the institution's characteristics in terms of facilities. Accordingly, CENSIS composite indicators on grants, communication, service, and structures were used to integrate information on university contributions, communication, service, and facilities in the dataset (CENSIS 2022). CENSIS indicators are defined by aggregating several pieces of information in a single standardized score ranging between 66 and 110. Specifically, the grants indicator considers universities' expenditures and policies on scholarships, awards, financial contributions for students' transportation, and other subsidies provided to students. The communication indicator is based on 66 indices that measure students' satisfaction regarding universities' websites and social networks. The service indicator considers the information on the number of meals provided in universities' canteens, the number of places in students' dormitories, and the financial contributions for students' accommodation. Finally, the structure's indicator is based on the availability of services such as study

rooms, libraries, laboratories and student satisfaction with these services. Table 2 provides the descriptive statistics of the variables used in the empirical analysis. Unfortunately, no reliable indicators at the university level are available to monitor universities' policies and activities to foster the integration of students from an academic and social point of view.

The analysis considers nine years (2010-2018) with an average of 200k students enrolled each year for a total of 1.8 million students (Figure 1).

[FIGURE 1 – NEAR HERE]

Table 1 reports their distribution according to ISCED field of study and career outcome after the 1st year. On average, 80.9% of students are regular and continue their studies in the same degree program and university, 14.9% dropped out, and 4.2% experienced a churn. Generally, we observe gender differences in dropout propensity, with males recording a higher percentage (17.5%) than females (12.7%). The dropout share has decreased in time, from 17.6% in 2010 to 14.2% in 2018 (-3.4%), while the proportion of churning students does not show significant variations.

[TABLE 1 – NEAR HERE]

Differences were also detected concerning the type of high school attended. Students from the academic schools (classical and the scientific *Licei*) recorded lower percentages of dropout (respectively 7.9% and 8.6%), and higher churn. By contrast, students from technical (23.7%) and vocational institutes (29.6%) registered a higher proportion of dropouts. Finally, regarding the ISCED field (Figure 2), Information Communication Technologies (ICT) and Agriculture and veterinary have the highest dropout percentages. It is interesting to note that there is a significant share of churn for Natural Sciences, Mathematics and Statistics.

[FIGURE 2 – NEAR HERE]

4. THE MODEL

We adopt a two steps model fitting based on multi-level logistic models to reach the threefold aim of (i) disentangling the school effect from the university effect; (ii) assessing key determinants related to students' socio-demographic characteristics and educational background on the probability to churn and dropout; (ii) sizing the impact of university policies.

Step 1 estimates the high school effect on the probability of dropout and churn (Skronidal & Rabe-Hasketh 2004; Goldestein 2011). For this sake, two variance component logistic models have been estimated by considering a hierarchical two-level structure where students (Level-1 units) are clustered in 116,300 groups (Level-2 units) given by the high-school (20,233) field of study (10) combination. This approach allows us to understand how much of the observed variability across students lies at each level of the multilevel structure. As a result of this step, we obtain the set of Posterior Predictions that measure the effect of high schools on students' dropout and churn probabilities by accounting for the differences among fields of study (Goldstein & Healy 1995; Goldstein & Spiegelhalter 1996; Goldstein 2008). These estimates are then used in Step 2 where the university effect on the probability of observing one of the two outcomes (dropout and churn) has been estimated using a Multilevel Multinomial Logit Model (Goldstein 2011). In this case, we consider a two-level structure where students (Level-1 units) are clustered in 4,151 Level-2 groups provided by the University $\times$ Field of study $\times$ Academic Year combinations. Accordingly, Level-2 groups consider in the same cluster students who, in an academic year, attended a degree program in the same disciplinary field in the same university. The approach allows us to consider the between universities' variability and the heterogeneity in students' composition concerning demographic characteristics, educational background, and university choices.

Namely, in Step 1 the probability of student i ($i=1, \dots, n$) belonging to school k ($k=1, \dots, K$) and who selects a degree in the field of study m to dropout (1) or churn (2) is estimated using two separate logistic functions:

$$\text{logit}(\pi_{ikm}^{1,2}) = \alpha^{1,2} + v_{km}^{1,2}$$

where $\pi_{ikm}^{1,2}$ indicates the probability of students to dropout (π_{ikm}^1) or to churn (π_{ikm}^2) with respect to the baseline, given by all students that are not classified in the considered category. For example, to model the probability of dropout, the baseline is given by all students that are classified as regulars or churners (i.e., non-dropouts). The probabilities $\pi_{ikm}^{1,2}$ are estimated as functions of a fixed constant $\alpha^{1,2}$ and the set of the Level-2 random intercepts $v_{km}^{1,2} \sim N(0, \sigma_{vm}^{1,2})$ that account for differences between high schools in the risk of dropout (1) and churn (2) according to the field of study selected at the university. Posterior estimates of both random effects are obtained as Empirical Bayes Estimates and used in Step 2 to provide information on the risk associated to students attending a specific high school to dropout/churn according to the disciplinary field selected at the university. Namely, these estimates are considered indirect indicators of the effect of secondary school attendance on the probability of dropout and churn. Indeed, the expected posterior predictions of the group random effects are considered in the literature as adjusted indicators suitable for pair comparisons among units under observation and concerning the average standard (Goldstein & Spiegelhalter, 1996; Goldstein & Haley, 1995). Models' parameters are obtained by using the STATA routine *melogit*, which estimates the parameters using a maximum likelihood technique based on the mean-variance adaptive Gauss-Hermite quadrature.

In Step 2, the probability of student i ($i=1, \dots, n$) to dropout (π_{i1kg}) or to churn (π_{i2kg}) is estimated using a multilevel multinomial logistic model and the following logistic function:

$$\log \left(\frac{\pi_{ijk g}}{\pi_{i0kg}} \right) = \alpha_j + X' \beta_j + \gamma_1 \hat{v}_{kmj}^1 + \gamma_2 \hat{v}_{kmj}^2 + u_{jg}$$

where $j=0, 1, 2$ represent the three observed events (regular=0, dropout=1 and churn=2), g ($g=1, \dots, G$) indicates the cluster to which student belong to given by the combination of the academic year, field of study and the university, and k ($k=1, \dots, K$) indicates the combination between the high school and

the field of study. This approach allows us to estimate two equations or contrasts based on the differences between the considered group (dropouts or churners) and the baseline given by the regular students. Each equation informs on the elements that characterize the considered category and that affect the probability of being classified as a dropout or churner. As shown in the logistic function, these probabilities are estimated as functions of a constant α_j , a set of fixed effect covariates $\mathbf{X}$, the posterior estimates of $\hat{v}_k^{(1=drop)}$ e $\hat{v}_k^{(2=churn)}$ obtained in Step 1, and the set of random effects at the cluster g level u_{jg} . $\mathbf{X}$ includes the CENSIS indicators on communication, grants, services and structures in each university, and the information on students' socio-demographic characteristics (Gender, Age, Foreigner status), macro area of residence (North, Centre, South and Islands), high school background (diploma grade, curricula), students' mobility status (Forced Mover, Free Mover), and labour market condition indicators (Regional GDP, Provincial Unemployment rate, Municipal Income). Table 2 provides the definition and descriptive statistics for the variables included in the estimation. To account for the differences in scales among the predictors, all the continuous variables have been standardized in z-scores.

[TABLE 2 – NEAR HERE]

The posterior estimates of the level two random terms $\hat{v}_k^{(1=drop)}$ e $\hat{v}_k^{(2=churn)}$ obtained in Step 1 provide, for each secondary school, a measure of its effect on the individual risk of dropping out or churning in each of the fields of study considered in the analysis (Goldstein & Spiegelhalter, 1996). These two indicators, which allow us to consider that students coming from different schools who enrolled in different fields of study are at different risks of dropping out and churning out, are considered as fixed covariates in Step 2, with the primary aim to disentangle the high school effects from other risk factors in modelling the probability to dropout and churn (Leckie, 2009). The interaction between high school and field of study allows us to take into account the efficacy of each

high school in preventing dropout and churning with respect to the possible fields of study where students may enrol at university. The posterior prediction of the random term $\mathbf{u}_{jg}^{1,2} \sim N(\mathbf{0}, \sigma_u^{1,2})$ accounts for differences in the probability to experience an adverse event (namely, dropout or churn) comparing fields of study offered by the considered universities, across the academic years. Thus, they measure the university effect on the probability of dropout and churning over and above the effect of the secondary school attended. The assumption of normality of the prior distribution of the random effects is the standard approach in the multilevel and latent variable model literature (for instance, Goldstein, 2011; Skrondal & Rabe-Hesketh, 2004). As the school effect on the risk of dropping out is a measure of the latent high school inefficacy in ensuring a successful completion of 1st year at the university, it is reasonable to assume a prior normal distribution in the absence of information that may suggest other distributions. Moreover, the expected posterior predictions are empirical Bayes' estimates obtained as the expected value of the posterior distribution of the latent trait that is calculated considering the observed loglikelihood and the prior distribution. Posterior estimates of $\hat{\mathbf{u}}_{jg}^{(1,2)}$ can be considered indicators of university inefficiency because they allow us to compare degree programs in the same field of study offered by the universities based on the risk of their students to drop out or to churn and to assess trends of university efficiency in terms of their capability in preventing these events over time. Step 2 model's parameters are obtained by using the *runmlwin* routine in STATA. This routine allows us to exploit the computational power of the software MLwiN, which is designed to estimate multilevel models (Leckie & Charlton, 2013; Rasbash et al., 2009). We estimated the model's parameters by using a Markov Chain Monte Carlo (MCMC) method using, as starting values, the set of coefficients estimated using an Iterative Generalized Least Squares method. MCMC method is a Bayesian estimation technique that can be used to estimate complex multilevel multinomial logit models (for more details see Browne, 2012).

5. EMPIRICAL FINDINGS

In Step 1, two logit models have been estimated to separately model the probability of dropout and churn as a function of the secondary school attended and the field of study where student enrolled at the university. As a result, we obtained two indicators for each high school $\times$ field of study combination, which capture the differences between secondary schools on the risk that their students will experience the event of interest (dropout or churn) during their first year at the university. In Step 2, this information has been incorporated among the predictors of the multinomial logit model, which considers as Level-1 units the students and as Level-2 units the combination of the academic year, the attended university, and the chosen field of study.

The results of the multinomial multilevel logit model have been exploited to study factors mainly associated with the probability of dropping out or churning out during the first year at the university and to quantify the risk associated with the secondary school attended over and above students' demographic characteristics and the peculiarities related to the specific attended degree programs. The baseline event models the probability of enrolling in the second year at the same university, namely, the probability of being classified as regular at the end of the first year.

We considered several factors associated with students' regular careers to shed light on internal and external factors influencing outcomes (churn and dropout) and their trends over time. We include three sets of predictors that allow us to account for divergences in socio-demographic factors (gender, age, area of residence, international student status), educational background (high school and high school final mark), university choices (in terms of local and non-local university and field of study) and university policies in supporting students at the entrance (CENSIS indicators on grants, communication, services, and structures). Moreover, as explained above, we consider in the Step 2 model the posterior predictions of the high school $\times$ field of study random effects estimated in Step 1 to consider that the risk of dropping out or churning out are related to the secondary attended school

and to capture for the differences in the effect that attending a specific high school can have according to the field of study chosen at the university.

Nonetheless, we introduce three labour market indicators to remove the effect of labour market dynamics on students' higher education career disruptions (Becker, 1964; Duncan, 1965; Di Pietro, 2006; Contini et al. 2022). Namely, the following indicators have been considered for describing the labour market in the area of origin and destination: the provincial unemployment rate (Provincial Unemployment), the municipal income per capita provided by the Ministry of Economic and Finance (Municipal Income), and the regional per-capita gross domestic product (Regional GDP). The two indicators related to the area of origin and destination coincides for stayer students. Thus, after assessing the effect of both sets on model estimation, we retained the model that considers the labour market conditions of the destination area to account for the role played by the labour market conditions in the areas where students are located during their university careers. Table 2 provides the descriptive statistics and the definitions of all the covariates included in the estimation.

The results of the multinomial logit model reported in Table 3 show that, in terms of socio-demographic characteristics, male students have a higher risk of dropping out than females, even adjusting for the different propensity of males and females to enrol in a specific field of studies. Once we control for other characteristics, foreign students dropout less than others, highlighting a lower risk compared to Italian students. According to the area of origin, students from the Main Islands (Sicily and Sardinia) have a higher probability of dropout, which increases with age. If we look at the educational background, students from vocational and technical schools have the highest risk of dropping out, and this event is negatively affected by the high school grade; nonetheless, students from classical and scientific high schools register the lowest dropout rates. The sign of the labour market indicators suggests that, if on the one hand, better economic conditions of the destination area in terms of regional GDP and municipal income per capita are associated with lower dropout rates, on the other hand, an increase in the rate of unemployment is associated with a greater propensity to withdraw, highlighting how positive economic performances can impact students' future

expectations. The signs of the coefficients of the labour market indicators confirm that the higher university dropout rate could result from a rational decision after comparing the actual costs (known with certainty only once enrolled in the chosen study course) and the expected benefits from the continuation of studies. It is worth highlighting that when considering the indicators of labour market conditions at origin (rather than at destination), the sign and magnitude of the effect of all covariates, except for the unemployment rate, remain almost constant. A higher unemployment rate in the province of residence appears to be a restraint to withdrawal. The invariance of the results related to the other covariates highlights the robustness of the findings to set of predictors considered in terms of origin or destination area.

This analysis's most interesting and innovative results are related to the secondary school effect parameter estimates, among which the specific attended school is the main predictor of the probability of dropping out and churning, even controlling for differences in high school curricula. The size and the sign of the secondary school effect show that this is the factor with the strongest influence on the probability of dropping out. This finding is also supported by the results obtained estimating two other specifications. In the first, reported in Table A1 in the Appendix, we have estimated the multilevel model in Table 3 without accounting for the high school effect. In the second, presented in Table A2 in the Appendix, we modified the multilevel structure by considering a three-levels cross-classified model where students are the Level 1 units, the combination of high schools and field of studies are the Level 2 units and the combination between universities, academic years, and fields of study are considered as Level 3 units. The results show that the variability of dropout and churn rates ascribable to differences between field of study and university groups (Level-2 in Table 3) is lower once the posterior estimates of the schools' effect are introduced among the predictors. This result suggests that, controlling for the fields of study selected at the university, inequalities at the entrance are mainly determined by divergences in the competencies provided by the secondary school attended, even accounting for the type of secondary school. Namely, model comparisons show that

the variance of the random terms at the university and field of study level, which captures the heterogeneity in the level of inefficiencies across disciplinary fields and universities, decreases from 0.23 in Table A1 and 0.25 in Table A2 to 0.10 if the high school effect is included among the predictors (Table 3). This result suggests that not accounting for the high school effect in the model would lead to an overestimation of the importance related to the heterogeneity among universities and fields of study.

Looking at the university policies, that can prevent dropout, it is interesting to highlight that the main findings suggest that the quality of communication has an average positive, even if weak, effect in preventing churning and dropout events. This result supports our thesis that easy access to information and policies addressed to improve the clearness of websites and communication can help students and prevent career disruption during the first year. Moreover, lower dropout rates are also associated with the quality of the structures, showing that a good environment in terms of structures helps students prevent inefficiency in the 1st year of their university career.

The churn rate shows a different behaviour concerning the high school of origin, with the highest churn rates associated with students from the academic school Liceo-Classic (the baseline) or Others High Schools and who selected non-local universities. The negative sign of the coefficients of the CENSIS indicators shows that the quality of communication and services has an average positive, even if weak, effect in preventing churn and dropout events.

[TABLE 3 – NEAR HERE]

The churn rate decreases with age, pointing out that older students are more likely to dropout or to continue their studies in the chosen field of study rather than to change their choices in terms of university during the first year. At the same time, students who selected a non-local university have a higher churn rate than local students, and this effect is stronger for free mover students, pointing at a different propensity to avoid a disruption of the university career between the two categories of students which are in mobility at the university. The divergences in churn rates across universities

and disciplinary fields (Level-2 units) is 35% higher than the divergences in their dropout rates, indicating that the churning event is linked to the choice of specific disciplinary fields that experience high churn rates during their 1st year.

The posterior predictions of the Level-2 random terms can be considered as adjusted indicators of field of study efficiency for dropout and churn rates and have been used to make adjusted comparisons across universities that supply homologous degree programs (Goldstein & Healy 1995; Leckie 2009) and adjusting for the influence of the labour market conditions on withdraw decisions.

Figure 3 and Figure 4 display for degree programs of the Social sciences, journalism, and information field of study the expected posterior predictions of the random effects at the university level estimated in the second step of the analysis with their relative 95% confidence intervals, computed based on estimates' (pseudo) standard errors. These predictions inform us of the effect of being enrolled in a specific field of study of a university on the probability of dropout and churn at two point times: 2010 and 2018. Both Figures allowed us to capture trends in the efficiency indicators and to identify which universities improved or worsened their performance in the years considered. Indeed, the performance of each university in reducing inefficiency in each disciplinary field is assessed by checking if the two 95% confidence intervals related to the two cohorts overlap. Namely, an improvement in the efficiency indicators related to dropout and churn is registered if the 95% confidence interval for 2018 lies entirely below the lower bound of the related confidence interval for the same unit in 2010 (Goldstein & Healy 1995). It is interesting to note that 11 out of 51 universities improve their performance concerning the risk of dropout for the field of study. Figures 5 and 6 report information for all fields of study considered in this analysis. Moreover, the expected posterior predictions plotted in Figure 3 and Figure 4 can be considered adjusted indicators of university performance suitable to make comparative evaluation across institutions that offer degree programs in the same field of study (Goldstein & Spiegelhalter 1996), taking into account the heterogeneity in the composition of students' population at the enrolment. Nonetheless, these indicators are built up

by removing the heterogeneity in the labour market conditions of the geographical area where universities are located.

[FIGURE 3 – NEAR HERE]

[FIGURE 4 – NEAR HERE]

[FIGURE 5A – NEAR HERE]

[FIGURE 5B – NEAR HERE]

[FIGURE 6A – NEAR HERE]

[FIGURE 6B – NEAR HERE]

6. DISCUSSION AND CONCLUSIONS

This study analysed the phenomenon of university dropout and churn in degree courses, including four sets of predictors that allow us to account for divergences in socio-demographic factors (gender, age, area of residence, international student status), educational background (high school, and high school grade), university choices (in terms of local and non-local university and field of study) and university policies in supporting students at the entrance (grants, communication, services, and structures), and labour market indicators in the origin area of student and destination area where the university is located (municipal income per capita, the unemployment rate at province level and the regional GDP). The results indicated that a series of elements connected with the student's socio-demographic characteristics play an important role in shaping the outcome of the first year of studies. In line with previous empirical evidence, male students are likelier to dropout than females, while international students dropout less than Italian students.

Furthermore, the results show that early dropout is experimented above all among students from the Main Islands, and the phenomenon becomes more acute with increasing age. Vocational and technical school students have the highest risk of dropping out. High school grades negatively affect this event;

academic high school students (the *Liceo*) register the lowest dropout rates. This empirical result also confirms what has already been highlighted in the literature. As for the university choices (in terms of local and non-local universities and fields of study), findings suggest that dropout rates are lower for students in mobility, especially considering the free movers. In contrast, churn rates are higher for students in mobility. The first of these two results, widely documented in the literature, considering the one-way migratory flow from southern to northern Italy (Attanasio et al, 2020), can be read as a selection process that pushes the most prepared and motivated students to choose mobility. On the other hand, the second could be read as an opportunity for students to achieve the educational objectives set at the time of enrolment better and/or in less time, a sort of ongoing reorientation.

The analyses have been carried out considering the labour market conditions of the area of the origin and destination to adapt the estimates of the school influence to the influence that economic conjecture has on rational decisions to invest in education.

The most interesting findings are related to the “secondary school effect” from which the specific high school attended is the main predictor of the probability of dropping out and churn, even controlling for differences in high school curricula. The magnitude and the sign of the secondary school effect show that this is the factor with the strongest influence on the probability of dropping out. This finding is also supported by the reduction in the variability of dropping out or churning ascribable to differences between the field of studies and university groups, which is observed once the posterior estimates of the schools’ effect is inserted among the predictors. The school effect is also relevant considering the differences between universities and high school fields of study and curricula. Other studies have also shown that the type of high school chosen by students and their final grades statistically affect the probability of dropping out (Boero et al. 2005; Bussu, Detotto & Serra 2020), but, to the best of our knowledge, this is the first time that the heterogeneity in the competencies at the entrance between students coming from schools with similar curricula is proved with empirical data, over and above the final mark achieved.

In essence, the results allow us to propose some reflections on the higher university dropout rate of students with a vocational and technical secondary school degree compared to students with an academic secondary school diploma. The choice to invest in students' cultural and human capital is not only linked to family background but probably also depends on the different investments of educational institutions. If it is true that students from an academic high school have tertiary education as a natural outlet, as they are less professionalised and therefore have less favourable conditions to enter the job market directly, in contrast, students with vocational and technical secondary school degrees have, hypothetically, a greater probability of entering the job market immediately. Instead, the worsening of the conditions of the Italian labour market has decreased the employment opportunities of high school graduates compared to university graduates. (ISTAT 2021).

The higher university dropout rate of these students could be due to the acquisition, already in the first year of university, of the awareness of possessing inadequate educational background to undertake a university career. Therefore, rather than a rational decision, enrolling in university would seem to be a "leap into the void" that quickly leads to dropping out of the tertiary education system. Furthermore, if we consider the unemployment levels of graduates in Italy (ISTAT, 2021), it could also be said that the choice to drop out is the consequence of a rational decision after comparing the actual costs (known with certainty only once enrolled in the course chosen study) and the benefits expected from continuing education.

Thus, our findings suggest the crucial importance of more significant investment by schools in building a more effective bridge in the school-university transition, especially concerning adequate entry skills compared to those expected from specific degree courses. Moreover, during the first year at the university, these shortcomings add to students' difficulties in managing their studies. Adapting to the new academic context seems to be affected by the support actions put in place by academic institutions. Indeed, looking at the university policies that can prevent dropout, our findings suggest that the quality of communication has a positive effect, even if weak, in preventing churn and dropout events. This result supports our thesis that easy access to information and policies addressed to

improve the website's clarity and communication can help students and prevent career disruption during the first year. Considering dropout as a possible consequence of less than rational choices in which there is little awareness of having an educational background that is inadequate to fit into the university context, it appears clear how important it becomes to strengthen not only guidance, counselling and welcome services during entry and the first year of university but also all those information and communication services that facilitate students in the adaptation process.

Furthermore, lower dropout rates are also associated with the quality of the structures, showing that the availability of a suitable environment in terms of structures helps students to prevent disruptions in the 1st year of their university career.

Finally, there are substantial differences in efficiency between the universities, even considering the heterogeneity of the origin of the secondary school, the educational pathway, and the field of study. This finding points out that few preventive measures are implemented to counter the phenomenon of university dropout.

This research's first general policy indication is the strategic function that university guidance actions carry out to counter dropout. The co-factors that can feed the phenomenon seem to have a common breeding ground in the lack of guidance towards the choice of university path and the resulting cultural and instrumental weakness with which some students enter the tertiary education system. In summary, universities should implement more services to mitigate the “school effect”, which strongly predicts dropout and churn. This paper has several limitations. Moreover, this strategy allowed us to use all the available data on students' careers in MOBYSU.IT. Indeed, using more than two years of students' careers would have required selecting fewer cohorts, limiting the scope of the paper and the precision of our estimates. Nonetheless, because of the absence of more updated data on students' careers, our work cannot contribute to understanding the effects of the COVID-19 pandemic on students' career choices. A further limitation is the absence of information to monitor the possible impact of policies addressed to promote peer interaction within universities. This aspect needs to be

explored in further research. It will also be interesting to assess how much of the variability in the school effect, over and above the type of school, is related to family socioeconomic conditions. This information is provided by the MOBYSU.it database only for subsets of cohorts, and thus it could not be considered in this study. Finally, further research development will be addressed to disentangle in Step 1 the school effect from the family effect by exploiting data from different sources, such as those provided in Italy by the National Institute for the Evaluation of the Educational System.

7. PRACTICAL AND POLICY IMPLICATION

The study delves into the complexity of the inefficiencies of the Italian university system, focusing mainly on the phenomena of student dropout and churn during the first year of university. By revisiting and integrating Tinto's model, initially focused on academic and social integration, this research adds a new dimension by considering the influence of university services and the central role of secondary education background. The implications of this study are multiple and shed light on potential strategic interventions and policy adjustments aimed at mitigating inefficiencies and promoting a more equitable and effective higher education landscape.

The investigation into the effects of secondary school on university success, for example, highlights substantial disparities depending on the type of secondary school attended. This reveals a critical area for policy intervention to ensure a level playing field for students from different educational backgrounds. Initiatives could include developing standardized preparatory courses for students from lower-performing schools, providing them with the skills needed for university studies. Closing this educational gap is essential to reducing dropout rates and ensuring that all students have a fair chance of succeeding in higher education.

The results of the study also highlight the variability in dropout rates between different socio-demographic groups and educational contexts. This variability requires a nuanced approach to policy and support mechanisms. Tailored programs that meet the specific needs of these different groups, such as older students or those from professional backgrounds, could include flexible scheduling

options and additional academic support services. Such targeted interventions are critical to addressing the unique challenges faced by these students and increasing their chances of success in higher education.

Finally, there is an impact, albeit moderate, that university policies on communication and digital services have on reducing dropouts and churn rates. This aspect highlights the urgent need for universities to prioritize investments in digital infrastructure. Improving the clarity of websites, streamlining online support services, and ensuring digital communication channels are robust and student-friendly can make a substantial difference. Such improvements would help provide better guidance and support students, particularly during the transition from the first year to the second year, which is often the most challenging.

Acknowledgments

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU– Project Title From high school to university: Assessing peers' influence in educational inequalities and performances, PI Isabella Sulis, CUP F53D23006150006 – Grant Assignment Decree No. 1060 adopted on 07/17/2023 by the Italian Ministry of Ministry of University and Research (MUR).

The data used in this study have been processed under the RESEARCH PROTOCOL for the study “From high school to the job placement: analysis of university careers and university mobility from Southern to Northern Italy” among the Ministry of University and Research, the Ministry of Education and Merit, the University of Palermo as the lead institution, and the INVALSI Institute. The reference researcher is Mariano Porcu.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Additional information

This work was supported by Ministero dell’Istruzione, dell’Università e della Ricerca: [grant no 2017HBTK5P - CUP B78D1900018000].

REFERENCES

- Aina, C., Baici, E., Casalone, G., & Pastore, F. (2022). The determinants of university dropout: A review of the socio-economic literature. *Socio-Economic Planning Sciences*, 79, 101102.
- Allegro, R., & Giambalvo, O. (2020). *University Performance Before and During Economic Crises*. UNIPA Springer Series.
- Attanasio M., Ragozini G., Porcu M. (2020) (Eds.) *Verso Nord. Le nuove e vecchie rotte delle migrazioni universitarie*, Franco Angeli, ISBN 978-88-35-10562-6
- Ballarino, G. (2011). Le politiche per l'università. In U. Ascoli (ed.), *Il welfare in Italia* (pp.197-224). Bologna: il Mulino.
- Bardach, L., Lüftenegger, M., Oczlon, S., Spiel, C., & Schober, B. (2019). Context-related problems and university students' dropout intentions—the buffering effect of personal best goals. *European Journal of Psychology of Education*, 35:477–493. <https://doi.org/10.1007/s10212-019-00433-9>
- Becker, G. S. (1991). *Human capital: a theoretical and empirical analysis*. Labour Economic & Labour Relations, Fifth Edition ed. USA, Prentice-Hall.
- Belloc, F., Maruotti, A., & Petrella, L. (2011). How individual characteristics affect university students dropout: a semiparametric mixed-effects model for an Italian case study, *Journal of Applied Statistics*, 38:10, 2225-2239.
- Bernardo, A. B., Galve-González, C., Núñez, J. C., & Almeida L. S. (2022). A Path Model of University Dropout Predictors: The Role of Satisfaction, the Use of Self-Regulation Learning Strategies and Students' Engagement. *Sustainability* 2022, 14, 1057. <https://doi.org/10.3390/su14031057>
- Boero, G., Laurenti, T., & Naylor, R. (2005). An econometric analysis of student withdrawal and progression in post-reform Italian universities. Working Papers 4, Centro Ricerche Economiche Nord Sud (CRENOS)
- Browne, W. J. (2012). *MCMC Estimation in MLwiN*, v2.26. Centre for Multilevel Modelling, University of Bristol.
- Burgalassi, M., Biasi, V., Capobianco, R., & Moretti, G. (2016). Il fenomeno dell'abbandono universitario precoce. Uno studio di caso sui corsi di laurea del Dipartimento di Scienze della Formazione dell'Università "Roma Tre". *Italian Journal of Educational Research*, (17), 105-126.
- Bussu, A., Detotto, C., & Serra, L. (2020). Indicators to prevent university drop-out and delayed graduation: An Italian case. *Journal of Applied Research in Higher Education*, 12(2), 230-249.
- CENSIS (2022) *Nota metodologica 2022. Le famiglie di valutazione e gli indicatori*. <https://www.censis.it/sites/default/files/downloads/Nota%20metodologica%20completa%202022.pdf>

- Clerici, R., Giraldo, A., & Meggiolaro, S. (2015). The determinants of academic outcomes in a competing risks approach: evidence from Italy, *Studies in Higher Education*, doi: 10.1080/03075079.2013.878835.
- Contini, D., Cugnata, F., & Scagni, A. (2018). Social selection in higher education. Enrolment, dropout and timely degree attainment in Italy. *High Educ* 75, 785–808.
- Contini, D., & Zotti, R. (2022). Do Financial Conditions Play a Role in University Dropout? New Evidence from Administrative Data. In: Checchi, D., Jappelli, T., Uricchio, A. (eds) *Teaching, Research and Academic Careers*. Springer, Cham. https://doi.org/10.1007/978-3-031-07438-7_3
- Di Pietro, G. (2006). Regional labour market conditions and university dropout rates: Evidence from Italy, *Regional Studies*, 40:6, 617-630
- Di Pietro, G. & Cutillo, A. (2008). Degree flexibility and university dropout: The Italian experience. *Economics of Education Review*, 27, 546-555.
- Di Pietro, G., Cutillo, A. (2008). Degree flexibility and university dropout: The Italian experience. *Economics of Education Review*, 27, 546-555.
- Duncan, B. (1965). Dropouts and the Unemployed. *Journal of Political Economy*, 73(2), 121-134.
- Georg, W. (2009). Individual and institutional factors in the tendency to dropout of higher education: a multilevel analysis using data from the Konstanz Student Survey, *Studies in Higher Education*, 34:6, 647-661. DOI: 10.1080/03075070802592730.
- Goldstein, H. (2011). *Multilevel Statistical Models*. Wiley Series in Probability and Statistics, 4 Edn. Wiley & Sons.
- Goldstein, H. (2008). School league tables: What can they really tell us. *Significance*, 5(2), 67–69.
- Goldstein, H., & Spiegelhalter, D. J. (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society: Series A*, 159,385–443.
- Goldstein, H., & Healy, M. J. R. (1995). The graphical presentation of a collection of means. *Journal of the Royal Statistical Society: Series A*, 158, 175–177.
- Hadjar, A., Haas, C., & Gewinner, C. (2022). Refining the Spady-Tinto approach: the role of individual characteristics and institutional support in students' higher education dropout intentions in Luxembourg. *European Journal of Higher Education*. DOI: 10.1080/21568235.2022.2056494
- ISTAT - Istituto Nazionale di Statistica. (2021). Labour Force Survey, <http://dati.istat.it/?lang=en&SubSessionId=13c1302b-d0a2-45d0-8ad1-0fd7930e9642>
- Leckie, G. (2009). The complexity of school and neighbourhood effects and movements of pupils on school differences in model educational achievement. *J. R. Stat. Soc. A* 172(3), 537–554.
- Leckie, G., & Charlton, C. (2013). runmlwin: A Program to Run the MLwiN Multilevel Modeling Software from within Stata. *Journal of Statistical Software*, 52(11), 1–40. <https://doi.org/10.18637/jss.v052.i11>.

- Leiviskä, K., & Siponen, M. (2013). Understanding why is students dropout: toward a process theory. ECIS 2013 Completed Research. 62. http://aisel.aisnet.org/ecis2013_cr/62
- Mealli, F., & Rampichini, C. (2012). Evaluating the effects of university grants by using regression discontinuity designs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(3), 775-798.
- Meggiolaro, S., Giraldo, A., & Clerici, R. (2017). A multilevel competing risks model for analysis of university students' careers in Italy. *Studies in higher education*, vol. 42, pp. 1259-1274.
- MOBISU.IT Database (2022) [Mobilità degli Studi Universitari in Italia], research protocol MUR - Universities of Cagliari, Palermo, Siena, Torino, Sassari, Firenze, Cattolica and Napoli Federico II, Scientific Coordinator Massimo Attanasio (UNIPA), Data Source ANS-MUR/CINECA.
- OECD (2010). How many students dropout of tertiary education? In: *Highlights from Education at a Glance 2010*, OECD Publishing, Paris. DOI: https://doi.org/10.1787/eag_highlights-2010-8-en
- OECD (2022), *Education at a Glance 2022: OECD Indicators*, OECD Publishing, Paris, <https://doi.org/10.1787/3197152b-en>.
- Pascarella, E. T., Pierson C. T., Wolniak G. C., & Terenzini P. T. (2004). First-Generation College Students, *The Journal of Higher Education*, 75:3, 249-284, DOI: 10.1080/00221546.2004.11772256
- Perchinunno, P., Bilancia, M., & Vitale, D. (2021). A statistical analysis of factors affecting higher education dropouts. *Social Indicators Research*, 156(2), 341-362.
- Rasbash, J., Charlton, C., Browne, W.J., Healy, M. and Cameron, B. 2009. *MLwiN Version 2.1*. Centre for Multilevel Modelling, University of Bristol.
- Reschly, A.L. & Christenson S.L. (Editors) (2022). *Handbook of Research on Student Engagement*. Springer, Switzerland. <https://doi.org/10.1007/978-3-031-07853-8>
- Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variables modeling*. Boca Raton: Chapman & Hall.
- Spady, W. G. (1970). Dropouts from higher education: an interdisciplinary review and synthesis. *Interchange*, 1:64-85.
- Tedesco, N., & Salaris, L. (2020). University dropout and mobility in Italy. First evidence on first level degrees, *SIS 2020 - Book of Short Papers*. PEARSON, pp.1601-1606.
- Tinto, V. (1975). Dropout from higher education: a theoretical synthesis of recent research. *Review of educational research*, 45(1):89-125.
- Tinto, V. (1993). *Leaving college: Rethinking the cause and cures of student attrition* (2nd ed.). Chicago: University of Chicago
- Trivellato, P., & Triventi, M. (2011). Differentiated trends in student access and performance during the Bologna Process. The case of universities in Milan. *Italian Journal of Sociology of Education*, 3(2).

- Usala, C., Porcu, M. & Sulis, I. (2023). The high school effect on students' mobility choices. *Stat Methods Appl* 32, 1259–1293. <https://doi.org/10.1007/s10260-023-00694-1>
- Viale Tuleda, H.E. (2014). A theoretial approach to the college student dropout. *Revista digital de investigation en docencia universitaria*, 8(1):59-73.

APPENDIX

[TABLE A1 – NEAR HERE]

[TABLE A2 – NEAR HERE]

TABLES

Table 1 - Distribution of first-year students by ISCED fields of study, gender, and second-year status

	All students		Regular	Dropout	Churn
	N/1000	%	%	%	%
Total	1,862	100	80.95	14.88	4.17
Females	1,007	54.08	83.05	12.69	4.26
Males	855	45.92	78.47	17.45	4.08
Education	84	4.53	80.52	16.97	2.52
Arts and humanities	334	17.93	81.27	15.70	3.02
Social sciences, journalism and information	333	17.90	80.90	15.97	3.13
Business, administration and law	268	14.40	81.48	15.55	2.96
Natural sciences, mathematics and statistics	238	12.77	76.06	14.13	9.80
Information and Communication Technologies (ICTs)	46	2.47	73.10	23.90	3.01
Engineering, manufacturing and construction	351	18.87	82.70	12.80	4.50
Agriculture, forestry, fisheries and veterinary	70	3.74	75.17	20.83	4.00
Health and welfare	162	8.69	87.63	9.37	3.00
Services	75	4.04	81.30	15.43	3.27

Table 2 – Definitions and descriptive statistics of variables used in estimation.

Variable	Definition		
Continuous variables:		Mean	Std dev
Grants	Indicator for university expenditures on grants (CENSIS)	83.648	12.095
Services	Composite indicator for service availability at university (CENSIS)	78.991	9.258
Structures	Composite indicator for the quality of university's structures (CENSIS)	82.919	9.454
Communication	Composite indicator for the quality of university's web site (CENSIS)	93.856	10.398
Age	Age of the student in years	19.865	3.520
Diploma Grade	Student's high school final mark	78.699	11.703
Municipal Income	Municipal taxable income per capita	23,798	3,917
Regional GDP	Regional Gross Domestic Product in millions of euros	135,166	94,297
Provincial Unemployment	Provincial unemployment rate	0.108	0.054
High School R.E.: Churn	Posterior estimates for high school random effect on churn	-0.010	0.548
High School R.E.: Dropout	Posterior estimates for high school random effect on dropout	-0.311	0.690
Categorical variables:		Proportion	
Female	1 if the student is female	0.541	
Foreigner	1 if the student is not Italian	0.028	
Islands	1 if the student resides in the Islands	0.105	
North	1 if the student resides in the North	0.422	
South	1 if the student resides in the South	0.263	
Centre	1 if the student resides in the Centre	0.210	
Diploma: Classic	1 if the student attended a classic lyceum	0.110	
Diploma: Vocational	1 if the student attended a vocational institute	0.074	
Diploma: Scientific	1 if the student attended a scientific lyceum	0.388	
Diploma: Technical	1 if the student attended a technical institute	0.229	
Diploma: Other	1 if the student attended another type of high school (residual)	0.199	
Forced Mover	1 if the student is a forced mover	0.165	
Free Mover	1 if the student is a free mover	0.312	

Table 3 - Results of the multinomial multilevel model: baseline category Regular, Level 2: University, field of study, and academic year combinations.

	Drop vs Regular		Churn vs Regular	
	Coeff.	C. I. 95%	Coeff.	C. I. 95%
α	-1.7370	[-1.763;-1.711]	-3.4966	[-3.534;-3.460]
Grants	-0.0614	[-0.075;-0.047]	-0.0674	[-0.086;-0.048]
Services	0.0274	[0.016;0.039]	-0.0467	[-0.063;-0.031]
Structures	-0.0182	[-0.030;-0.007]	-0.0017	[-0.017;0.014]
Communication	-0.0038	[-0.017;0.010]	-0.0479	[-0.065;-0.029]
Female	-0.1823	[-0.192;-0.172]	-0.0280	[-0.045;-0.011]
Age	0.1558	[0.152;0.159]	-0.2567	[-0.275;-0.238]
Foreigner	-0.1089	[-0.132;-0.085]	-0.1977	[-0.251;-0.140]
Islands	0.0346	[0.005;0.065]	0.0389	[0.001;0.076]
North	0.0163	[-0.006;0.037]	0.0272	[-0.005;0.059]
South	0.0251	[0.004;0.047]	0.0277	[-0.003;0.057]
Diploma Grade	-0.5258	[-0.531;-0.521]	-0.2013	[-0.209;-0.193]
Diploma: Technical	0.0881	[0.068;0.108]	-0.0371	[-0.068;-0.005]
Diploma: Vocational	0.3025	[0.279;0.325]	-0.0245	[-0.067;0.021]
Diploma: Scientific	-0.1139	[-0.132;-0.095]	-0.0655	[-0.088;-0.043]
Diploma: Other	0.2145	[0.195;0.235]	0.0469	[0.019;0.075]
Provincial Unemployment	0.0120	[-0.001;0.025]	0.0016	[-0.017;0.020]
Regional GDP	-0.0547	[-0.070;-0.039]	-0.0162	[-0.038;0.005]
Municipal Income	-0.0160	[-0.029;-0.002]	0.0613	[0.040;0.081]
Forced Mover	0.0406	[0.028;0.054]	0.1182	[0.097;0.140]
Free Mover	-0.0702	[-0.082;-0.059]	0.5049	[0.486;0.523]
High School R.E.: Churn	0.1721	[0.162;0.183]	1.5853	[1.569;1.602]
High School R.E.: Dropout	1.2717	[1.263;1.281]	0.3028	[0.287;0.318]
Level 2				
Variance	0.1018	[0.095;0.109]	0.1380	[0.128;0.149]
Cov(Drop, Churn)	0.0614	[0.055;0.068]		
Bayesian DIC	1829313			
Observations	1861720			

Table A1 - Results of the multinomial multilevel model without the High School effect: baseline category Regular, Level 2: University, field of study, and academic year combinations.

	Drop vs Regular		Churn vs Regular	
	Coeff.	C. I. 95%	Coeff.	C. I. 95%
α	-2.3032	[-2.333;-2.271]	-3.2073	[-3.249;-3.166]
Grants	-0.0789	[-0.096;-0.059]	-0.0776	[-0.105;-0.049]
Services	0.0013	[-0.016;0.019]	-0.0843	[-0.111;-0.058]
Structures	-0.0336	[-0.049;-0.016]	0.0319	[0.007;0.058]
Communication	-0.0200	[-0.042;-0.001]	-0.0869	[-0.126;-0.061]
Female	-0.2472	[-0.257;-0.238]	-0.0162	[-0.032;-0.000]
Age	0.1856	[0.182;0.189]	-0.2553	[-0.273;-0.237]
Foreigner	-0.0141	[-0.039;0.010]	-0.1652	[-0.223;-0.109]
Islands	0.0138	[-0.019;0.044]	0.0263	[-0.017;0.069]
North	-0.1544	[-0.179;-0.130]	-0.1142	[-0.150;-0.081]
South	-0.0287	[-0.052;-0.006]	0.2078	[0.177;0.239]
Diploma Grade	-0.5468	[-0.552;-0.542]	-0.1820	[-0.190;-0.174]
Diploma: Vocational	1.5124	[1.490;1.533]	-0.4707	[-0.513;-0.430]
Diploma: Other	0.8330	[0.814;0.851]	-0.2101	[-0.236;-0.183]
Diploma: Scientific	-0.0150	[-0.034;0.003]	-0.3232	[-0.346;-0.300]
Diploma: Technical	1.0992	[1.080;1.117]	-0.4736	[-0.502;-0.447]
Provincial Unemployment	0.0398	[0.021;0.058]	-0.0168	[-0.042;0.007]
Municipal Income	0.0063	[-0.006;0.019]	0.0702	[0.047;0.093]
Regional GDP	-0.0829	[-0.102;-0.063]	-0.0194	[-0.051;0.011]
Forced Mover	-0.0785	[-0.090;-0.066]	0.2346	[0.213;0.257]
Free Mover	-0.1750	[-0.187;-0.164]	0.6517	[0.634;0.670]
Level 2				
Variance	0.2283	[0.216;0.241]	0.4610	[0.436;0.488]
Cov(Drop, Churn)	0.2058	[0.191;0.221]		
Bayesian DIC	1950531			
Observations	1861720			

Table A2 - Results of the multinomial multilevel model: baseline category Regular; Level 2: High School and Field of study combinations; Level 3: University, field of study, and academic year combinations.

	Drop vs Regular		Churn vs Regular	
	Coeff.	C. I. 95%	Coeff.	C. I. 95%
α	-2.3699	[-2.400;-2.338]	-3.3414	[-3.393;-3.292]
Grants	-0.0855	[-0.109;-0.064]	-0.0844	[-0.116;-0.055]
Services	0.0024	[-0.016;0.020]	-0.0880	[-0.114;-0.062]
Structures	-0.0374	[-0.053;-0.022]	0.0291	[0.006;0.051]
Communication	-0.0227	[-0.044;-0.003]	-0.0870	[-0.117;-0.059]
Female	-0.2542	[-0.265;-0.244]	-0.0214	[-0.039;-0.005]
Age	0.1923	[0.189;0.196]	-0.2632	[-0.282;-0.245]
Foreigner	-0.0307	[-0.057;-0.006]	-0.1806	[-0.238;-0.120]
Islands	0.0075	[-0.023;0.038]	0.0102	[-0.032;0.054]
North	-0.1634	[-0.187;-0.139]	-0.1501	[-0.183;-0.118]
South	-0.0397	[-0.064;-0.018]	0.1956	[0.165;0.226]
Diploma Grade	-0.5678	[-0.573;-0.563]	-0.1936	[-0.202;-0.185]
Diploma: Vocational	1.5390	[1.515;1.562]	-0.4472	[-0.487;-0.408]
Diploma: Scientific	-0.0139	[-0.034;0.006]	-0.3111	[-0.335;-0.287]
Diploma: Technical	1.1159	[1.096;1.136]	-0.4642	[-0.493;-0.435]
Diploma: Other	0.8464	[0.826;0.867]	-0.1917	[-0.220;-0.165]
Regional GDP	-0.0894	[-0.107;-0.071]	-0.0240	[-0.054;0.006]
Provincial Unemployment	0.0359	[0.017;0.054]	-0.0207	[-0.047;0.009]
Municipal Income	0.0063	[-0.008;0.022]	0.0753	[0.051;0.102]
Forced Mover	-0.0874	[-0.101;-0.074]	0.2232	[0.200;0.247]
Free Mover	-0.1935	[-0.206;-0.182]	0.6529	[0.634;0.672]
Level 2				
Variance	0.3121	[0.306;0.319]	0.4488	[0.368;0.490]
Cov(Drop, Churn)	0.1893	[0.173;0.205]		
Level 3				
Variance	0.2541	[0.240;0.268]	0.4881	[0.461;0.517]
Cov(Drop, Churn)	0.2289	[0.213;0.246]		
Bayesian DIC	1935331			
Observations	1861720			

FIGURES

Figure 1 - Number of first-year students by year of enrolment (2010-2018)

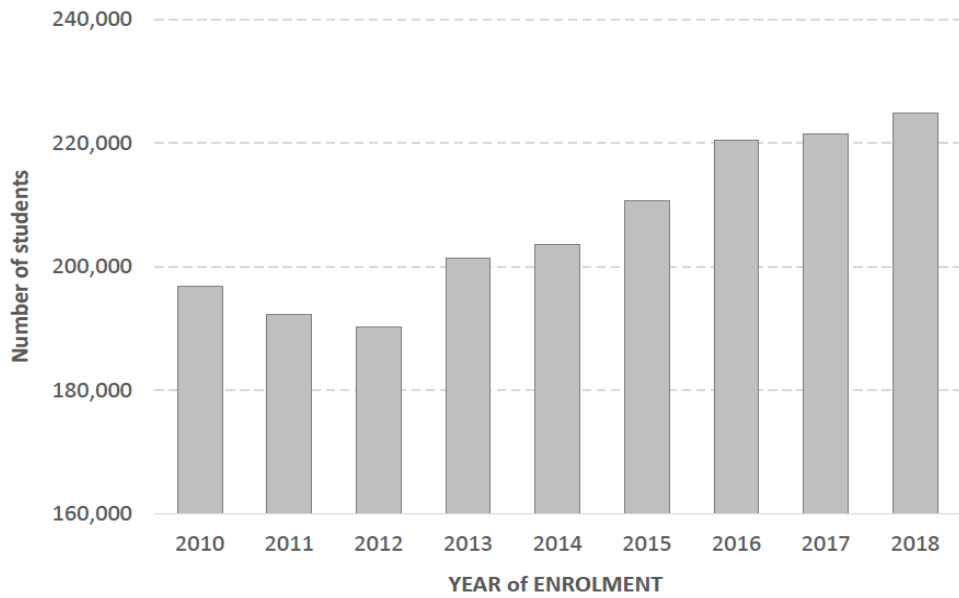


Figure 2 - First-year students’ dropout and churn by ISCED field (percentage)

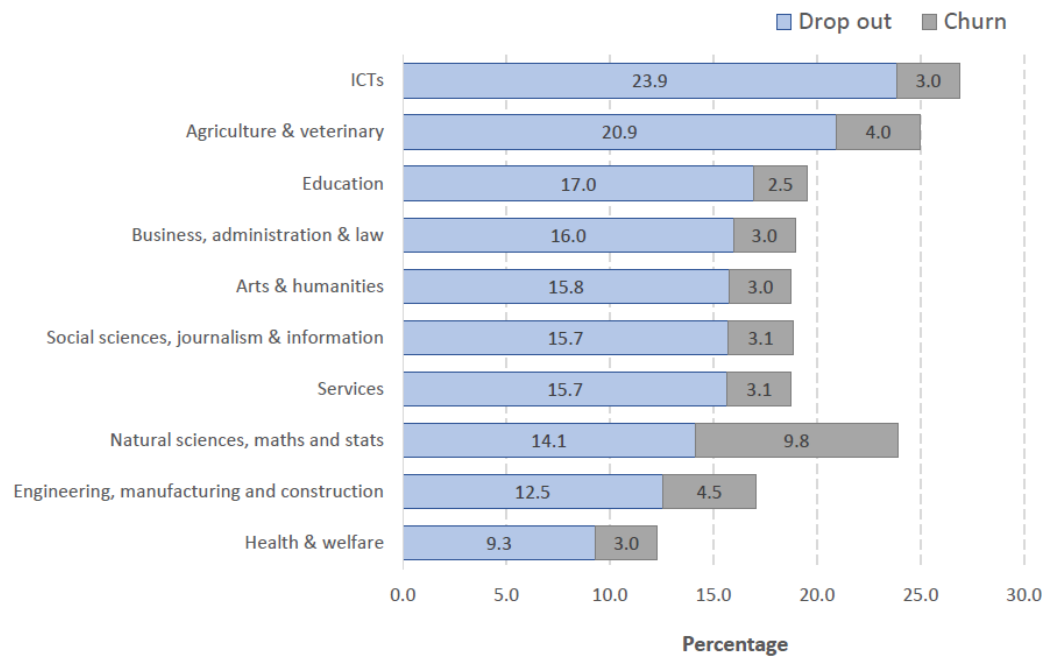


Figure 3 - The University Effect 2010-2018, dropout vs Regular

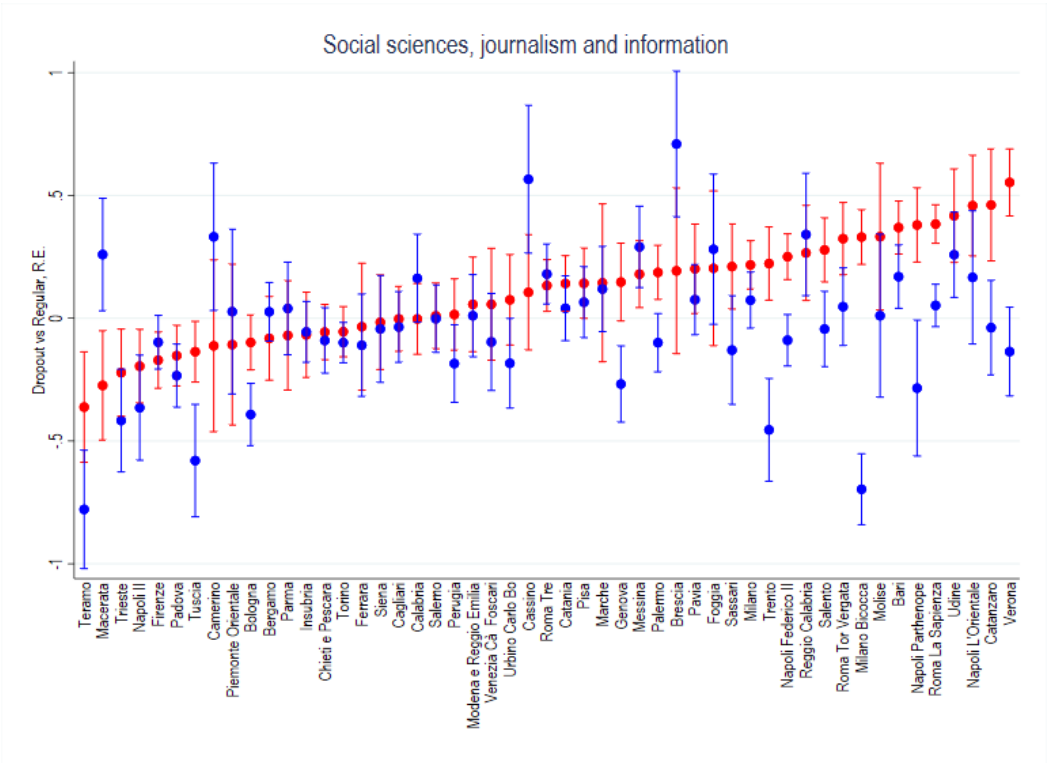


Figure 4 - The University Effect 2010-2018, Churn vs Regular

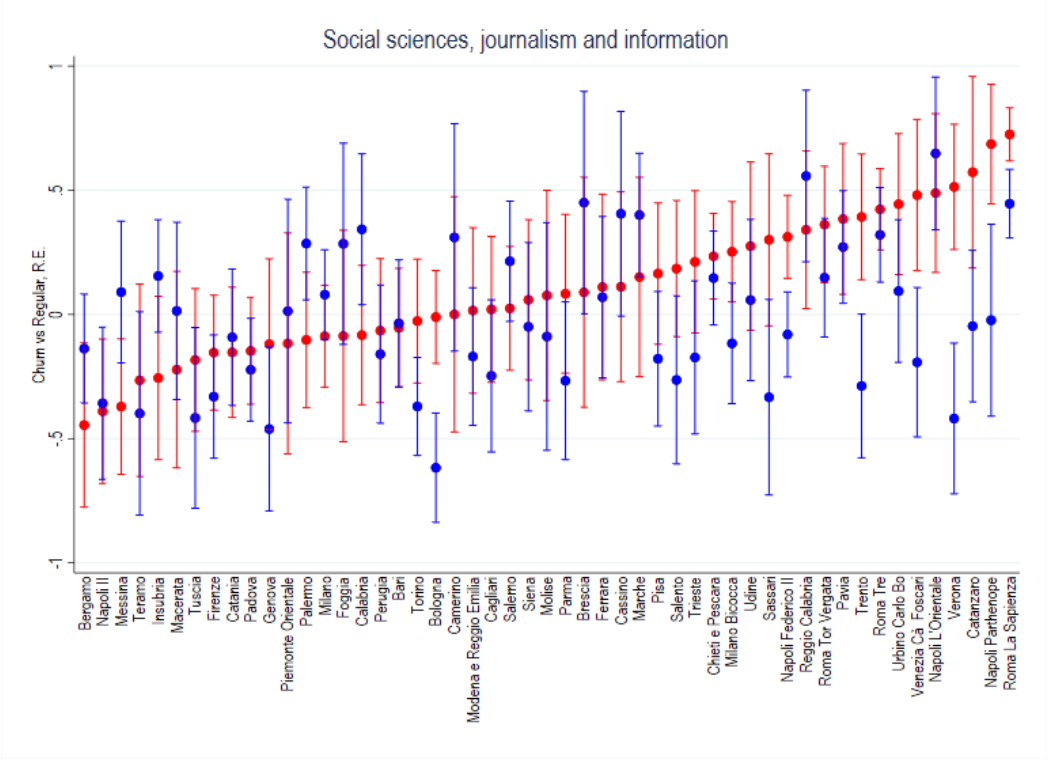


Figure 5A - The University Effect 2010-2018, dropout vs regular. All disciplinary fields

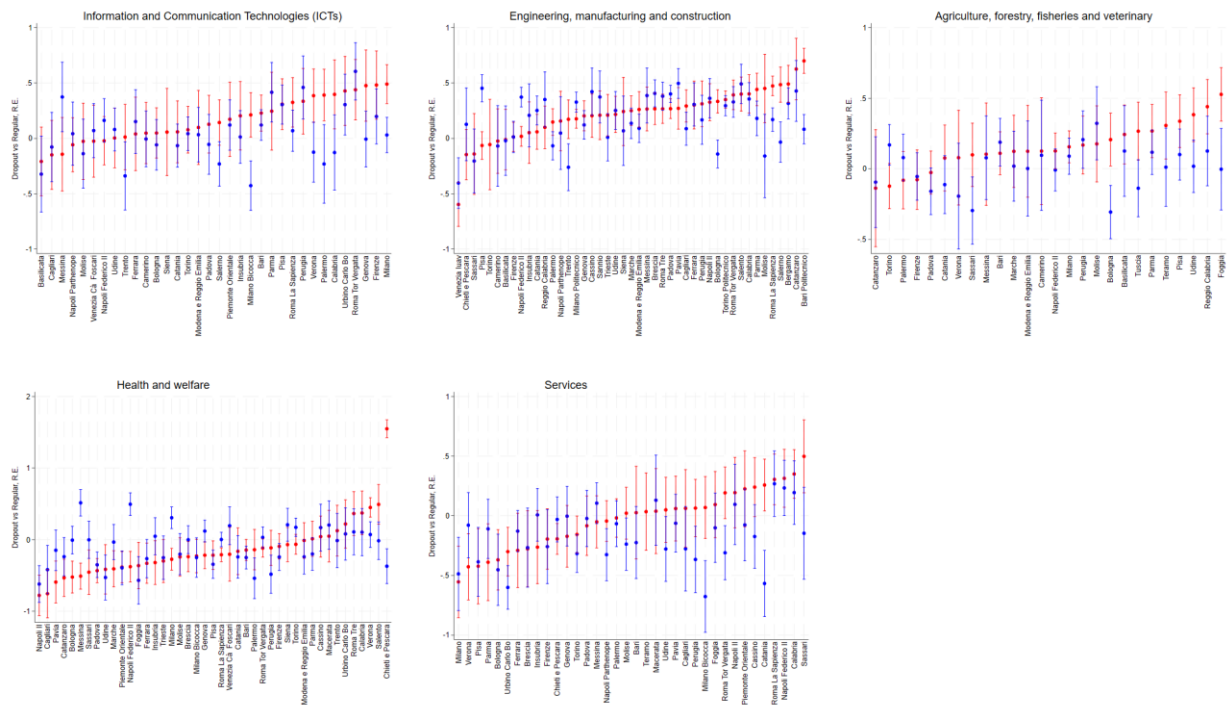


Figure 5B - The University Effect 2010-2018, dropout vs regular. All disciplinary fields

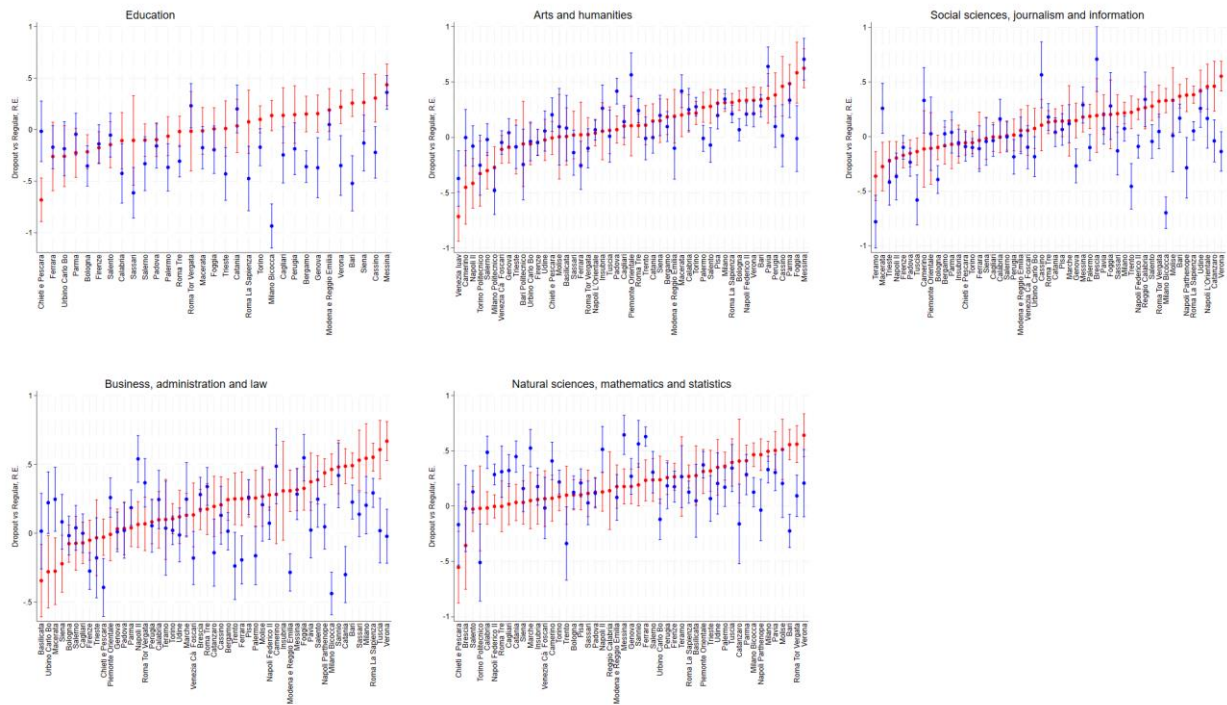


Figure 6A - The University Effect 2010-2018, Churn vs Regular. All disciplinary fields

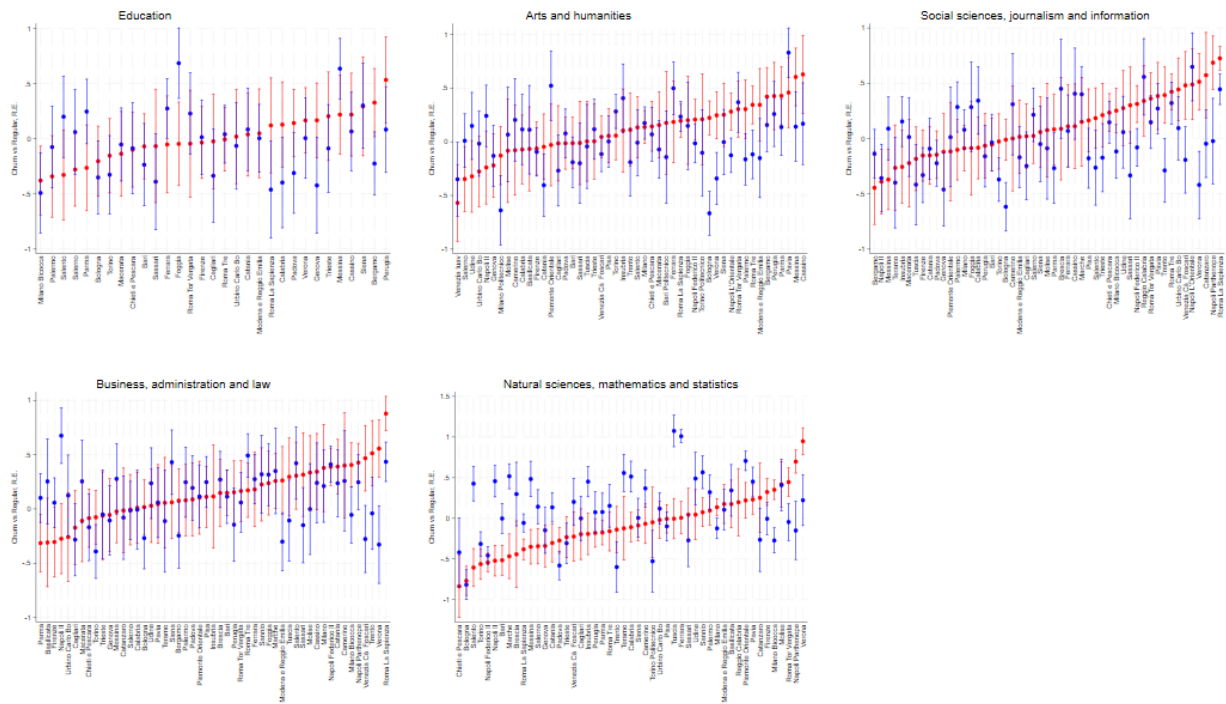


Figure 6B - The University Effect 2010-2018, Churn vs Regular. All disciplinary fields

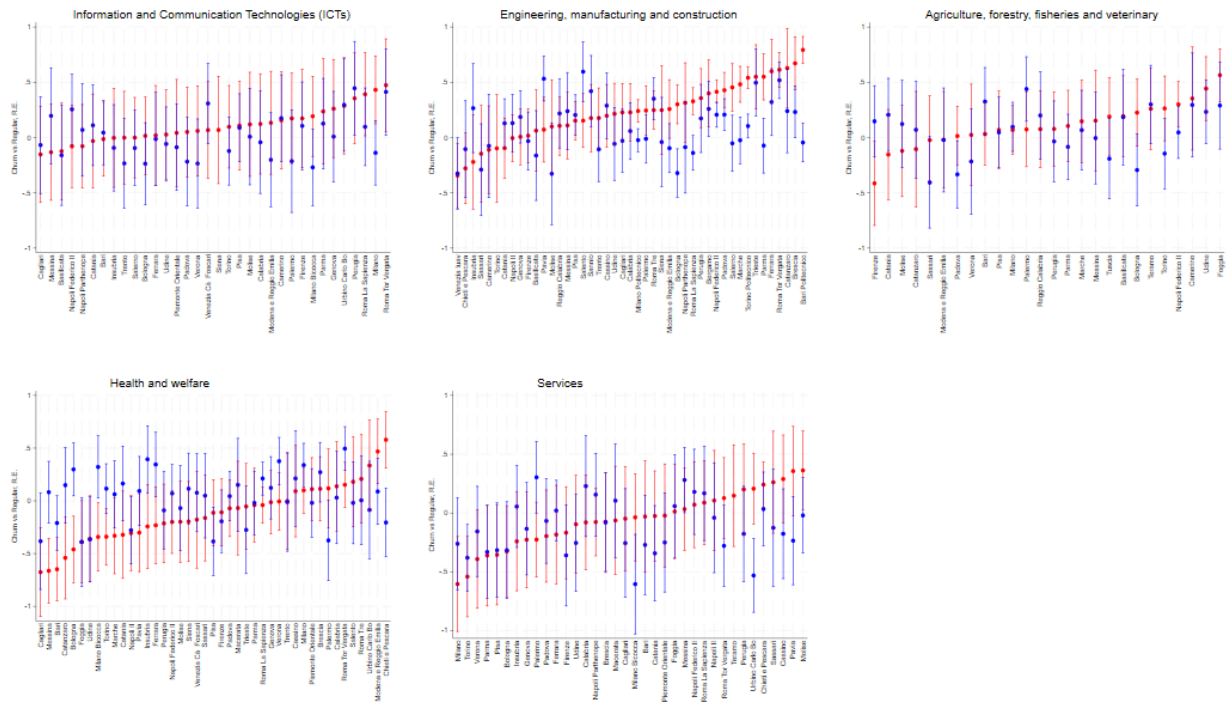


FIGURE CAPTIONS LIST

Figure 1 - Number of first-year students by year of enrolment (2010-2018)

Figure 2 - First-year students' dropout and churn by ISCED field (percentage)

Figure 3 - The University Effect 2010-2018, dropout vs Regular

Figure 4 - The University Effect 2010-2018, Churn vs Regular

Figure 5A - The University Effect 2010-2018, dropout vs regular. All disciplinary fields

Figure 5B - The University Effect 2010-2018, dropout vs regular. All disciplinary fields

Figure 6A - The University Effect 2010-2018, Churn vs Regular. All disciplinary fields

Figure 6B - The University Effect 2010-2018, Churn vs Regular. All disciplinary fields