Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Working Paper Series

16/17

DROPOUT AND TIME TO DEGREE IN ITALIAN UNIVERSITIES AROUND THE ECONOMIC CRISIS

DALIT CONTINI, GUIDO SALZA and ANDREA SCAGNI

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Dropout and time to degree in Italian universities around the economic crisis

Dalit Contini

Guido Salza

Andrea Scagni

University of Torino

Abstract

In this contribution, we address the following research questions: (i) Do parental education, parental occupation and economic conditions influence the students' university dropout probability, given students' previous educational attainment and performance? What is the relative role of these alternative indicators of family background? (ii) Do we observe any changes when comparing different cohorts of entrants – before and after the beginning of the economic crisis – in degree attainment and dropout hazard rates? Can we ascribe these changes to different behavior or to the different composition of the university student body? Is there any evidence that the observed behavioral changes can be ascribed to the economic crisis? To this aim, we use administrative data of the University of Torino – recently integrated with information on parental background – containing detailed information on the academic careers of the cohorts of first enrolled in years 2004-2014. These data include information on transfers to other degree programs within the University of Torino and to other universities, allowing to distinguish between withdrawal from the university system and changes of institution. On research question (i), we find that economic constraints negatively affect the dropout risk, while parental education and occupation seem to exert little further influence. Our findings on research question (ii) indicate that substantial improvements in both dropout probabilities and time to degree have occurred, and that these improvements are partly due to changes in the composition of the student body and partly due to changes in individual behavior.

1. Introduction[1]

The share of young individuals with a university degree in Italy has increased substantially over the past decades, but it is still around 24% (Eurostat, 2015), far from the goal set up by Horizon 2020 of 40% and much lower than the majority of EU countries. This critical outcome is not due to low participation rates in the university system, which are not far from the majority of EU countries (Eurydice 2012). Indeed, although somewhat declining (from 56% in the mid-2000s to 53% in

[1] Graphs and tables in sections 7.2, 7.3 and in Appendix 2 were drawn from the MA dissertation of Guido Salza "Unequal academic trajectories. Dropout phenomena at the University of Torino" (supervisor Dalit Contini), graduation day September 16th 2016.

2014), transition rates from upper secondary to tertiary education are not particularly low in Italy. Instead, a major reason for the low share of university graduates may be traced in the very high non-completion probabilities experienced by Italian students (ANVUR, 2016). Another negative distinctive feature of the Italian tertiary education system is the long time to degree attainment, far above the institutional length of the degrees (Almalaurea, 2016).

An extensive literature from the sociology and economics of education has analyzed university enrolment and dropout, highlighting large differentials between students from different upper secondary educational programs and of different family backgrounds. This literature is largely based on retrospective survey data, in particular on the *Survey of Upper Secondary School Graduates*, held at regular intervals from 1995 to 2011. This survey has many advantages, the most remarkable being that it allows to study enrolment decisions and that it provides information on parental background.

However, the survey also has severe limitations, related to the fact that it is based on self-reported data and that, being held 3-4 years after the attainment of the upper secondary degree, it does not allow to investigate the tertiary degree attainment process and the behavior of the non-negligible subpopulation of students enrolling in university with substantial delay.[2]

In-depth analyses of dropout behavior and the timing of degree attainment have suffered from the difficulty to obtain good quality and well organized micro-level longitudinal data on students' university careers. Administrative data have been analyzed in few contexts at the level of single degree courses; notably, the main shortcoming of this approach is that degree course changes and transfers to other institutions cannot be distinguished from dropout.

Fortunately, the Italian Ministry of Education has been constructing a harmonized administrative archive on university careers at the national level (*ANS*), and each institution may now obtain data releases regarding its own students. Although lacking information on the family of origin, these administrative data include detailed and reliable information on the entire university careers within one institution (thus, including changes of degree programs) from different and recent cohorts of newly matriculated students.[3]

In this contribution, we exploit the administrative data released by ANS relative to the University of Torino with the general aim to analyze the determinants of university dropout and time to degree

[2] For example, dropout rates are substantially lower than the figures deriving from administrative data (Contini et al, 2016).

[3] Dropout and time to degree have been analyzed at the national level by the National Agency for the Evaluation of University and Research (ANVUR) with the data collected by all the national institutions. Their reports include descriptive evidence in aggregate form (see for example, ANVUR 2016).

from the perspective of individuals, i.e. in terms of their capacity to carry on and complete their study programs. Thus, important explanatory factors are the previous schooling career and age at first enrolment.

It is worth noticing that the original ANS data only comprise information on the academic careers accomplished at the University of Torino. If students move to other national institutions before degree attainment, they cannot be distinguished from dropouts. Yet, from the perspective of individuals, they are not dropouts. To overcome this problem, we have obtained from the Ministry of Education an additional release of data regarding the segments of careers started at the University of Torino and continued in other national institutions.[4] Ultimately, our data archive contains the careers of all students first matriculated at the University of Torino between 2004 and 2013, including the relevant outcomes occurred after the transfer to another institution.

Moreover, in 2014 and for the first time ever, we have managed to record information on parental education and occupation at matriculation. Thus, we now have the extraordinary opportunity to analyze the influence of family background and economic conditions (as measured by ISEE, the *Indicator of the Economic Equivalent Situation*) on the students' academic careers after enrolment with high quality administrative data.

So far, we have collected and organized data for ten matriculation cohorts, from 2004 to 2014. Thus, we are able to analyze the changes occurred over time in the outcomes of interest: dropout and time to degree. A second aim of this contribution is to describe these changes. As we will see, we find a substantial *improvement* of these outcomes. To verify if these improvements are due to changes in the composition of the student body or to changes in the individuals' behavior, we have analyzed selected observed raw differentials in dropout and graduation probabilities with Blinder-Oaxaca-like decomposition methods. We may anticipate that both composition and behavioral effects are responsible for the observed changes.

Unfortunately, we will not be able to undertake empirical analyses of the causes of these improvements. One reason is that several important macro-level transformations have occurred in the past decade.

The first and more important is the economic downturn started in 2009. This devastating event has been producing huge changes in the labor market – already exposed to important institutional transformations promoting flexibility – substantially reducing job opportunities. However, as

[4] In particular, we have asked the national authorities in charge of the data to search for all students not obtaining a degree at the University of Torino within the other national institution archives; if found elsewhere, the relevant information on dropout or degree attainment was recorded.

discussed in Section 8, the effects of the crisis on university outcomes are far from being theoretically clear-cut.

Another potentially relevant macro-level change regards the university system. Since 2001, the tertiary education system in Italy has undergone a substantial number of reforms, in line with the so-called Bologna Process, aimed at harmonizing and enhancing university education across Europe. The so-called "3+2" structure – 3 year bachelor (BA) degree and 2 year master (MA) degree was enforced in 2001, allowing only few degree course (e.g. the medical and low schools) to preserve the traditional 5- or 6- year programs. As a result, the number of enrollments grew markedly, although only temporarily. Subsequent reforms have enforced further changes in the structure of Italian university (for example, moving teaching management responsibilities to departments, previously devoted only to research activities).[5]

As we have said, we will not offer sound empirical analyses of the reasons underlying the observed improvements – not only due to changes in the composition of university students – in the academic outcomes dropout and time to degree. We observe changes in these academic outcomes while witnessing transformations in the labor market and transformation in the university system. There is no simple way to relate these changes and disentangle the effects of these concurring causes (or other potential sources of change). Still, we will attempt to discuss the potential effects of the economic downturn and reconcile our results with possible alternative explanations of the observed improvements in dropout and time to degree.

2. Background

Sociological theories of cultural reproduction (Bordieau, Passeron, 1990) and rational action (Boudon, 1973, Breen, Golthorpe, 1997) provide competing explanations for the persisting different educational choices of children of different social origin. Indeed, in all western world countries we observe that the children of advantaged backgrounds are better performing in school, and are more inclined to make prestigious educational choices given prior achievement (e.g. Jackson, 2013). Consequently, the educational careers of the children of high socioeconomic levels are consistently more successful than the careers of the children from less privileged households already from early schooling stages.

Economists also refer to rational choices, but put more emphasis on income: individuals take decisions by comparing direct and indirect costs of education, in particular, tuition fees and

[5] The 3+2 reform was enforced with the Ministerial Act DM509/99; the two main subsequent reforms with DM 270/2004 and DM 240/2010. Other minor but numerous normative changes have been enforced over these years with the aim to improve the general functioning of the system.

foregone earnings, with benefits in terms of future wages (e.g. Blundell *et al.*, 2001). Lower income individuals may make less prestigious choices because they are more risk averse (Checchi *et al.*, 2014) or because of credit constraints and financial hardship. Instead, Carneiro and Heckman, (2002) find that what matters is not current income, but the long-run factors associated with higher income families, enabling to ensure better quality education and environments that foster children's cognitive and non-cognitive skills.

There is widespread evidence of socioeconomic differentials also in retention and completion probabilities (eg. Ishitani 2006, Vignoles, Powdthavee 2009), related to better prior academic preparation (Robin and Naylor 2001; Arulampalam *et al.* 2004) and social integration (Tinto, 1975). Other potential explanations are related to information asymmetries on the higher education system (if higher background children make better-informed choices, they will face lower dropout risks), or to credit constraints (Stinebrickner, Stinebrickner 2008), tuition fees and financial aid (Dynarski 2003).

Evidence on Italy

Research on enrolment in Italy on the demand side, largely focuses on intergenerational transmission of educational attainment over time (among others, Checchi *et al.* 2013. The data mostly used is the *Survey of Upper Secondary School Graduates* (SUSG) collecting retrospective data on specific cohorts 3-4 years after graduation on a representative sample at the national level. The advantage of this source is that, contrary to administrative data on university students, it allows to study the enrolment process. Moreover, detailed information on family background (parental education and occupation) is available, allowing to analyse social inequalities. Instead, no data on household income is provided. The existing literature offers consistent empirical evidence of large social inequalities in university enrolment, even net of prior schooling characteristics. Some contributions analyse the effect of the effects of the Bologna process implemented in 2001 on enrolment (Cappellari, Lucifora, 2009; Bratti *et al.*, 2010), finding a substantial (but transitory) increase in enrolment rates, in particular among diplomats from the technical track.

Research on university dropout has suffered from the unavailability of longitudinal data at the national level, only recently overcame by the release of ANS administrative micro-data. Some studies rely on administrative data from specific departments or degree-programs in given institutions. However these archives do not allow distinguishing between change of degree-program and withdrawal from higher education altogether. Using administrative data from large universities, other studies analyse dropout and change of programs with competing risks (eg. Clerici *et al.*,

2014), reporting a large influence of prior schooling and different patterns by field of study and degree programs.

Data based on SUSG have been widely used to study system-level dropout. The existing literature highlights the importance of prior schooling and the role played by socioeconomic background over and beyond prior schooling (Di Pietro 2004; Cingano, Cipollone 2007; Di Pietro, Cutillo, 2008; Argentin and Triventi, 2011; Contini et al 2015). However, due to the absence of data on income, these studies cannot analyse the effects of household economic conditions. Moreover, the national character of the survey – obviously an advantage from many standpoints – has some limitations. Prior schooling characteristics (school track and final examination mark), also regarded as indicators of acquired skills and ability, may not be fully comparable across the country (for example, it is well known that marks are generally higher in the South, even if national and international standardized assessments highlight that Southern children display substantially lower achievement levels). In addition, the within-track heterogeneity in school-quality at the country level is likely to be larger than in specific areas. In this perspective, analyses at smaller scale are likely to ensure that students with the same prior schooling characteristics have uniform skill levels. To the extent that students from more advantaged backgrounds tend to attend higher quality institutions (even within the same track), this would allow to better disentangling the effects of prior schooling from the effects of family background operating after university enrolment.

Cappellari, Lucifora, (2009) and D'Hombres, (2007) study the impact of the Bologna process and show that the reform has contributed to a small reduction of the dropout probability. Mealli, Rampichini (2012) analyse the effect of grants with regression discontinuity design, and show that at the threshold, grants contribute preventing dropout, providing empirical evidence that family income matters.

The effects of the economic crisis on university attendance have been explored by Ghignoni (2016) with SUSG data from different waves of the survey. By applying probit selection models and decomposition techniques, she finds that changes in students' background and students' characteristics play a major role in the recent reduction of the aggregate dropout rate.

The literature on time to degree is scant. Aina *et al.* (2011) highlight the role of individual and family factors, and find that weak labour market prospects contribute to lengthening time to degree, while Garibaldi *et al.* (2012) observe a negative relation between tuition costs and timely completion probability.

3. The Italian educational system

In Italy, children enter formal schooling at 6 and are required to attend school up to age 16. Compulsory education is comprehensive for the first eight years, with five years of primary school followed by three years of lower secondary school, ending with a final exam, partially standardized at the national level. Upper secondary school offers many educational programs, differing significantly in content and approach; they are usually grouped into academic, technical and vocational tracks. The academic track includes various types of lyceums, and is considered a specific preparatory curriculum for tertiary education. Technical schools usually provide academic education together with job-oriented instruction. Vocational institutes are academically less demanding, and put more emphasis on the training for low qualification technical jobs. All secondary schools lasting 5 years end with *maturità*, a national non-standardized examination. All students attaining the degree, are allowed to enrol into all university programs. However, given the different academic content and the widely recognized prestige, transition rates to university vary markedly across tracks.

Schooling institutions in Italy are mostly public; private schools exists but are attended by a rather small minority of students. Most often, private schooling has remedial goals, allowing low-proficiency students who can afford it to overcome previous failures in public institutes (Bertola *et al.* 2008).

Entrance to the various secondary school tracks is unrestricted, regardless of the children ability in compulsory school. The same holds for university enrolment, whatever was the student's educational program and proficiency in upper secondary school. There are only a few exceptions regarding tertiary education programs where the number of available positions is restricted (*numero chiuso*): in this case, students are selected through *ad hoc* ability based entrance tests.

In Italy, tertiary education is basically equivalent to university: there are some “polytechnic” schools, but they are run and administered as universities, without any formal or substantial distinction from them. Tertiary education institutions are mostly public, although at this level most of the few existing private universities are perceived as high-quality institutions.

University programs have been reformed since the start of the “Bologna process” in 2001, and now are generally divided into 3-year bachelor courses, which may be followed by a 2-year master program. A few exceptions exist however, as the most prestigious programs in the medical and law schools are 5- or 6-years programs. Most students choose a single broad field of study (e.g. literature, law or chemistry) and have relatively limited freedom in choosing courses and exams. There are generally no time constraints to get a degree; accordingly, each exam can be taken

repeatedly in case of failure or unsatisfactory results. This allows many students to exceed the official time-to-degree set for each program (average time to completion for 3-years programs was 4.6 years in 2014, Almalaurea 2014).

Public universities have rather similar tuition fees – generally low, as in various other European countries – although some geographical differences exist.[6] Tuition rates depend on household income and wealth (as measured by the ISEE indicator, see in section 6.4) and are very low for disadvantaged students; thus, individuals coming from low-income families should not be prevented to enrol by the direct costs of education. Some scholarships and grants are assigned to disadvantaged (but proficient) students, but in general, institutional financial support for tertiary studies is limited.

4. Research questions

In this contribution, we address the following research questions:

RQ1) *As demonstrated in the literature, in Italy family background exerts an enormous influence on schooling choices and outcomes up to tertiary education enrolment. Does family background keep having an influence on educational careers thereon? What is the relative role of parental education, occupation and economic conditions in university completion?*

Family background is likely to play a minor direct role at this stage, as students are now grown-ups and the involvement of parents in tutoring low performing offspring is quite uncommon in university. However, parents might keep influencing students' motivations and aspirations: thus, despite the previous strong social selection, individuals from high social backgrounds could be somewhat more determined to complete tertiary education degrees as compared to their less advantaged peers. Instead, we expect household economic conditions to have a substantial influence on individuals' behavior, as students in vulnerable households are more likely to be in financial hardship and be forced to drop college to sustain themselves or contribute to income provision to the family of origin.

RQ2) *What is the role of prior schooling characteristics in the probability of withdrawal from university and time to degree?*

Since no entry barriers related to previous schooling or ability exist in the Italian educational system, neither at the transition to upper secondary education, nor at the transition to tertiary education, all the students holding a high school diploma (*maturità*) are entitled to attend university.

[6] Tuition fees in Northern public universities are on average 17% higher than in the South (Federconsumatori, 2015).

Thus, we expect prior schooling characteristics to be extremely relevant in determining students' university outcomes after enrolment.

RQ3) *Do we observe any changes in dropout hazard rates and in the timing of degree attainment between cohorts? Are these changes due to differences in the composition of the student body or to changes in individuals' behavior?*

As already mentioned, we have witnessed a reduction in aggregate dropout rates and in time to degree over recent cohorts. The student composition has also changed over the observed time span, possibly due to reduced perceived returns to tertiary education and/or to the effects of the economic crisis. In particular, the share of students holding non-strictly academic high school degrees or enrolling with delay has reduced markedly. Thus, we expect compositional changes to have played a role in the improvement of overall dropout rates and time to degree. Instead, it is difficult to make predictions on behavior. The first reason is that, as we will discuss below, the effects of the economic crisis are ambiguous in themselves. Secondly, the ongoing institutional changes in the tertiary education system undergone since 2004 might also have played a role, entailing positive effects on both completion rates and timing.

Finally, we will attempt to address the following issue:

Can we relate our findings to the recent economic crisis?

Two conditions must hold to be able to ascribe the observed changes to the economic crisis. Firstly, there must be evidence of a change in individuals' behavior, and we will see that this is indeed the case. Secondly, it must be possible to disentangle the effects due to all potential reasons behind these changes; however, as anticipated above and discussed more in depth in Section 8, this is a difficult task to accomplish, so only speculative conclusions can be drawn.

5. Data

Administrative data on tertiary education careers in Italy were recorded as standalone archives by single universities for many years. The National Statistics Institute (ISTAT) collected only limited information at an aggregate level, based on periodic transmission by each institutions. This situation changed in 2004 with the birth of ANS (Anagrafe Nazionale Studenti), setup by MIUR (Ministry of Education, University and Research), which started to collect micro-data from all universities with a common format. This innovation brought along new opportunities for research on education in Italy, but the database was initially plagued by problems of poor coverage and differing interpretations concerning the data to be loaded. It was only in 2013 that MIUR finally recognized ANS as the official data source on tertiary education in Italy. Data quality of ANS relies on the effort of each

institution to provide correct information: fortunately, the careers of the students at the University of Torino have now been recorded with reasonable accuracy for about a decade.[7]

We have constructed two databases from the original ANS release.

Archive 1 regards the cohort of students first enrolled in 2014 (approximately 13,000 students), for whom we observe whether they have re-enrolled in 2015 or withdrawn from university. In addition to the relevant information on previous educational history and degree program characteristics, we have linked data on the standardized economic indicator of the household (ISEE), used to determine tuition fees, and – for the first time ever – we have accomplished to collect information on maternal and paternal education and occupation.

Archive 2 includes data on the entire careers – degree programs, exams and grades, number of credits – of the cohorts of students first enrolled at the University of Torino in years 2004-2013 (approximately 108,000 individuals). These data include information on students who transferred to other universities, allowing us to distinguish between dropout from the university system and change of institution. Hence, we have information on all the students' career progress and on the timing of degree attainment or withdrawal. In addition, we have data on individual characteristics: gender, age at first enrollment, type of high school and graduation marks.

6. Methods

We use Archive 1 to address research questions 1 and Archive 2 to address research questions 2 and 3. More specifically, in the former we analyze 2014 cohort of newly enrolled, whereas in the latter we analyze and compare 2004, 2008 and 2011 cohorts: the first one regards the pre-crisis period; the second considers individuals entering tertiary education at the onset of the economic crisis; the third considers individuals in the midst of the downturn.

6.1 Research question 1

Focusing on students first matriculated in 2014, we analyze the role of the different family background indicators – parental education, parental occupation and economic conditions (ISEE) – with logit models for the probability of dropout versus re-enrollment after the first year. We also include control variables on previous educational history, gender and age at matriculation, as well as

[7] Since our overall research goals stretch from tertiary education to labor market entry and subsequent working careers, two other data sources are currently being linked with ANS: the Almalaurea archive on university graduates, as well as SILP and CCIA administrative databases on employment contracts and entrepreneurship in Piedmont. Almalaurea is a national Consortium that records the working condition of university graduates at three points in time (1, 3 and 5 years after graduation). SILP (Servizio Informazioni Lavoro Piemonte) is the official register for all private sector labor contracts in the region. CCIA is the chamber of Commerce, Agriculture and Industry, which records information on company creation.

characteristics of the degree program (field of study and institutional length).

Starting in 2014, we asked the University of Torino administration to include in the online enrollment procedure some additional items regarding the educational level and most recent job of both of the students' parents. Response is given on a voluntary basis, but a carefully structured enrollment form has allowed to obtain very high response rates (>98%).[8]

Additional information about the family background of new students was collected through ISEE (*Indicatore della Situazione Economica Equivalente*), an economic index (computed directly in Euros) that measures the total wealth of the family, including both income sources and other endowments.[9] All student grants and tuition fees exemptions are allocated based on ISEE, favouring students with lower income: those not trying to access such benefits are not required to disclose it. Thus, the figure is expected to be generally available for economically disadvantaged individuals, but not for higher ISEE students. Accordingly, this variable is available only for a part of the students population (63%).

Linking this additional data to the main archive on students' educational career and their re-enrolment choice at the beginning of their second year (2015/16), we were able to study the influence of parental background on the decision to withdraw after one year since matriculation through logit models[10]. Since some students drop out in later years, a more complete analysis will be possible only in the years to come. However, the first year dropout is indubitably an outcome of interest in itself since (as we will see in the results of research question 2) the dropout incidence is highest in the first year.

6.2 Research question 2

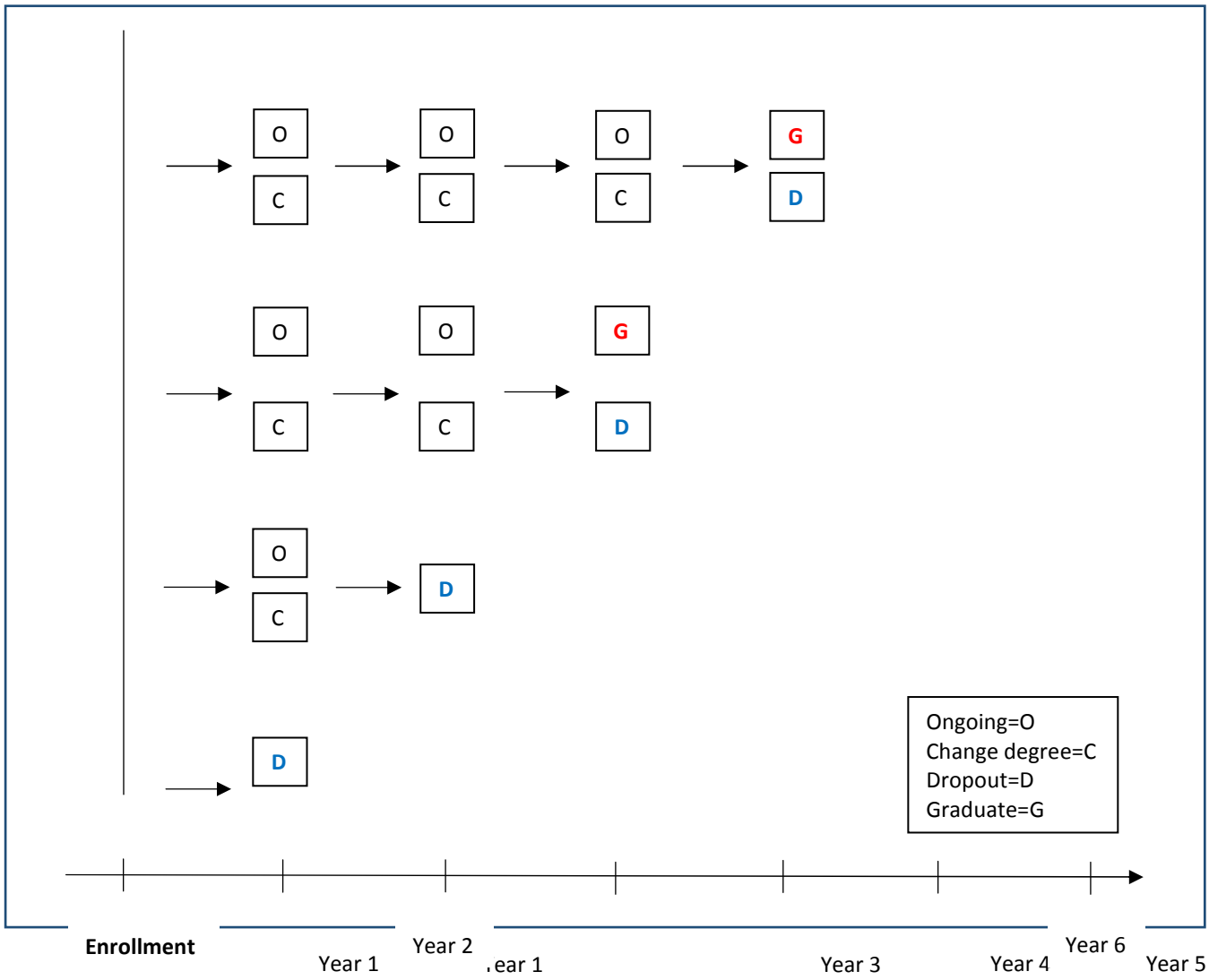
We investigate the educational outcomes of the different cohorts of entrants in dropout and graduation probabilities using a step-by-step longitudinal approach that mimics the estimation of discrete-time hazard functions in a competing risk framework, where the coefficients are allowed to vary freely at each step. (For details on the strategy and comparison with a conventional competing risks estimation, see Appendix 1).

Let us follow Figure 1, where we depict possible careers of students first enrolled in 3-year degree programs. At each year, we use multinomial logit models to estimate the probabilities of all possible outcomes. At year 1, we analyze the probability of ongoing (i.e. renovating enrollment) versus

[8] To ensure comparability, we used the same response items included in the Almalaurea Survey on university graduates.
[9] This information is not recorded in ANS, but was obtained directly from the University archives.
[10] Differently from the subsequent analyses, due to the lack of data, we focus here on the dichotomy between re-enrolling vs. withdrawing, without distinguishing between re-enrolment in the original degree program and degree program changes.

changing degree course versus dropout. In case the student drops out, she falls out of the risk set. At year 2, considering only students still in the risk set, we model again the probability of ongoing versus changing degree course versus dropout. At year 3, the student can also graduate. If this event occurs, she falls out of the risk set. We follow the process up to year 6.

Figure 1. University careers for students enrolled into 3-year degree programs.



Actually, our model is somewhat more complex than that depicted in Figure 1, because students changing degree course may transfer to another 3-year degree or move up to 5- or 6-year degrees. Since the characteristics of students making upward changes are substantially different from those making horizontal changes, we keep the two options distinct.

To acknowledge that the processes might vary between students first enrolled in short and long programs, we carry out separate analyses for individuals who started with a 3- year degree programs from those who started a 5- or 6- year program.[11]

More specifically, at each step we estimate multinomial logit models with the following time-invariant explanatory variables: gender, age at first matriculation, type of high school (lyceums, other general schools, technical, vocational) and high-school final examination mark. We also consider a limited number of time-varying covariates: field of study and an indicator of whether the student has changed degree program in the past (distinguishing between upward and horizontal moves).

***

The most informative output in competing risks models are *Cumulative Incidence Functions (CIF)*, describing the estimated probability that a given destination occurs within a given time span. For each matriculation cohort we estimate the probability that a given individual with a specific profile of explanatory variables will exit the system by dropping out or graduating within 1-6 years:

$$CIF_D(j) = P(T \leq j, Dest = D)$$

$$CIF_G(j) = P(T \leq j, Dest = G)$$

The complementary function:

$$P(T > j) = 1 - P(T \leq j, Dest = D) - P(T \leq j, Dest = G)$$

is the probability of being still enrolled after j years.

The estimation can be accomplished by multiplying the step-by-step probabilities of the single outcomes giving rise to a specific sequence compatible with the career of interest, and then adding up the probabilities of all compatible sequences.[12] For instance, the probability of withdrawal within year two is:

$$P(T \leq 2, Dest = D) = P(T = 1, Dest = D) + P(T = 2, Dest = D|T > 1)P(T > 1).$$

[11] In this case, we treat do not distinguish between horizontal degree changes (nearly inexistent) and downward changes.

[12] A simpler and straightforward way to evaluate the joint probabilities of interest consists in simulating a large number of educational careers using the estimated probabilities at each step and then simply count the share of cases experiencing the desired outcomes. In practice, we have used this approach, that turns out to be equivalent to the analytic computation of probabilities for large enough N. We have computed these probabilities for specific profiles of explanatory variables and for the entire observed populations, with their actual population composition in terms of all the relevant individual characteristics.

Using the notation of Figure 1, this probability is estimated as: $P(D_1) + P(D_2|O_1)P(O_1) + P(D_2|C_1)P(C_1)$. Once computed, we compare the *CIF*s of individuals with different profiles, also by disciplinary field.

6.3 Research question 3

Starting from the estimates of the multinomial logit models derived to address RQ2 for matriculation cohorts 2004, 2008 and 2011, we apply a Blinder-Oaxaca-like decomposition into explained and unexplained components. The aim is to evaluate the extent to which the differences in selected outcomes for two cohorts are related to the different composition of the populations of students enrolled (i.e. are *explained* by the different composition) or instead to different "behavior" (i.e. to changing model parameters).[13] The outcomes of main interest are the *CIF*s: we will analyze the probability of dropping out within 1-6 years and the probability of graduating within 3-6 years.[14]

In essence, the decomposition is carried out by comparing the observed probabilities for a given cohort to the so-called counterfactuals, i.e. the probabilities obtained by applying the estimated parameters of a given cohort to the population of the other cohort (see Table 1).

Table 1. Elements of the decomposition. Composition and behavior effects

Parameters (behavior) X-variables (composition)	2004	2008
2004	"observed" 2004	counterfactual 1
2008	counterfactual 2	"observed" 2008

When applying the parameter estimates of cohort 2004 to individuals of cohort 2004 we obtain "observed" probabilities (the double quotes refer to the fact that these are not the actual observed raw probabilities but model estimates). Instead, when applying the parameter estimates of cohort 2004 to individuals of cohort 2008 we obtain counterfactual probabilities (counterfactual 2), representing the probabilities that the students first enrolled in 2008 would experiment – *ceteris*

[13] This method was originally proposed by Blinder (1973) and Oaxaca (1973) has been widely applied in particular in the labor economics literature to a variety of issues (e.g. wage differentials).

[14] More specifically, when we compare cohorts 2004 and 2008 we consider what happens up to year 6, when we compare cohorts 2004 and 2011 we consider what happens up to year 3.

paribus – if they behaved like the students first enrolled in 2004[15]. Similarly, counterfactual 1 represents the probabilities experimented by cohort 2004 if they behaved like individual in cohort 2008. Comparisons of outcome probabilities in the same column inform on composition effects, whereas comparisons of outcome probabilities in the same row inform on behavioral effects.

6.4 Variables

We start with a brief description of the variables used throughout the paper. From ANS archives, we obtained all individual information regarding demographics, previous schooling and university career.

Gender.

A dummy variable taking value 1 when female and 0 when male.

Age at enrolment (age_19 - age_20_21 - age_22_25 - age_over_25).

The age at enrolment was classified in four groups. In the models, the lowest age interval is the reference category.

Type of diploma (lyceum - other_academic - technical – vocational).

The type of diploma obtained in secondary school was summarised into 4 groups, academic lyceums (classical, scientific and linguistic, explicitly conceived as programs preparing for tertiary education), other lyceums (socio-pedagogic, artistic), technical schools and vocational schools, mainly preparing for low-qualification jobs (e.g. mechanic, waiter, hair stylist, etc.). In the models, academic lyceums act as reference category.

Score at the secondary final exam (maturità_mark)

This is the result for the final secondary school exam, ranging from 60 to 100.

Field of study and institutional length (Social_3 - Social-5_6 - Human-3 - Human-5_6 - Healthcare-3 - Healthcare-5_6 - Scientific-3 - Scientific-5_6).

Fields of study at university were broadly grouped in four areas: healthcare, scientific, social (including economics and law) and humanistic (including both literature and pedagogical studies). Moreover, programs were distinguished according to their institutional length and the degree level attained: 3-year programs (corresponding to a BA) and 5-6 year programs (at the end of which students attain a Master's level degree). In some models we consider all programs, including all the

[15] Using the approach described in footnote 2, we simulate individual educational careers by applying the estimates of a given cohort to the population of the other cohort, and then look at the desired outcomes for the entire populations or for subpopulations of interest.

interactions between field of study and institutional length; in others we focus on three-year programs and thus only refer to the disciplinary area. Three-year programs in the *Social* area are the reference category in models.

***

The analysis for research question 1 involved the following additional covariates, obtained through direct enquiries into local archives (ISEE student declarations) or newly included in the enrolment form questionnaire (parental education and occupation[16]).

Parental education (par_ed_compuls - par_ed_secondary – par_ed_tertiary).

Parental education was originally classified into four groups, i.e. primary school, full compulsory schooling, upper secondary school, tertiary education, but for all analyses the first two groups are merged together. As detailed in note 19, a number of alternative covariates were defined, based of the original information. In the models presented in the paper, parental education is specified as the highest qualification between mother and father. Families where at least one parent has a university degree make up the reference category in the models.

Parental occupation (par_job_qual_high - par_job_qual_medium - par_job_qual_low)

Parental occupation also was recorded in a quite detailed fashion, distinguishing among house-worker, blue collar, low level white collar, high level white collar, teacher, medical doctor, liberal profession, entrepreneur, manager, commercial or artisanal autonomous worker. For the models we focused on a graded scale that classified jobs into three levels of professional qualification, low, medium and high[17] (with the latter as reference category).

ISEE (<10000/10000-20000/20000-30000/30000-45000/45000-70000/>70000/not_disclosed).

ISEE is a wealth index at the family level, expressed in monetary values, that combines the sources of income with other endowments, real estate properties, etc. through a specific algorithmthat takes into account the number of individuals in the household. When available (disclosure is optional), it was classified into 6 levels, plus a "not available" category. ISEE was designed at the national level

[16]Students were asked to qualify the most recent job held by their mother and father. Thus, this variable does *not* record occupational status, because it does not distinguish *retired* or *unemployed* individuals form those belonging to the active workforce.

[17]See Appendix 3 for details.

to represent a comparable measure of a family wealth to be employed in many contexts. Indeed, its definition is obviously arbitrary, and may be subject to measurement error in cases of fiscal fraud[18].

7. Results

7.1 Research question 1

To investigate the factors affecting the first year dropout probability and in particular on the various indicators of family background, we estimated a set of logit models. Results are generally in line with the expectations. In the model without economic conditions, parental education has a non-significant effect, no matter how it is operationalised.[19] Instead, a small and weakly significant effect is found in the model with ISEE. This finding suggests that parental education exerts its full influence on previous schooling decisions and outcomes, including the choice to enrol in tertiary education, but has little further effects after this stage. Parental occupation is weakly significant in Model 1 (lower class students experience a dropout probability on average 2 percentage points above that of higher class students), while it is no longer significant when including the ISEE indicator (Model 2).

Table 2. Effects of family background and households' economic conditions.

Pr(dropout)	**Model 1**					
	Coef.	**Std. Err.**	**A.M.E.**	**pv_AME**	**[95% AME C.I.]**	
par_ed_compuls.	-0.0040	0.1069	-0.0005	0.970	-0.0286	0.0275
par_ed_secondary	0.0010	0.0813	0.0001	0.990	-0.0212	0.0215
par_job_qual_low	0.1684	0.0912	0.0225	0.066*	-0.0015	0.0465
par_job_qual_medium	0.1067	0.0795	0.0140	0.179	-0.0064	0.0344
_cons	0.0371	0.2593				
	Model 2					
	Coef.	**Std.Err.**	**A.M.E.**	**pv_AME**	**[95% AME C.I.]**	
par_ed_compuls.	0.1527	0.1104	0.0198	0.166	-0.0082	0.0478
par_ed_secondary	0.1512	0.0839	0.0196	0.071*	-0.0017	0.0409
par_job_qual_low	0.1187	0.0961	0.0154	0.218	-0.0091	0.0400
par_job_qual_medium	0.0521	0.0817	0.0067	0.523	-0.0138	0.0271

[18]Notice that the algorithm to compute ISEE was revised in 2015. As a result, some students ended up declaring the old ISEE and some others the new one. In this version of the work, we have disregarded this issue. According to preliminary analyses, differences between the two measures are small.

[19] As robustness checks, we estimated a set of models including alternative classifications. We started including dummy variables for each parent (distinguishing among tertiary, upper secondary and compulsory education), also allowing for interaction effects between maternal and paternal education. In addition, we considered simple dichotomies like at least one parent having a degree vs. all others combinations. Only negligible differences were found. Notice that the non-significance holds even if parental education is the only covariate representing family background in the model. A clear effect of parental education emerges only when omitting the covariates related to previous student educational career (type of diploma and final mark in secondary school).

ISEE_10000-20000	-0.1048	0.1244	-0.0122	0.404	-0.0408	0.0164
ISEE_20000-30000	-0.2657	0.1363	-0.0294	0.054*	-0.0592	0.0005
ISEE_30000-45000	-0.3825	0.1503	-0.0408	0.011**	-0.0721	-0.0094
ISEE_45000-70000	-0.2480	0.1810	-0.0276	0.160	-0.0660	0.0109
ISEE_>70000	-0.2497	0.3414	-0.0278	0.435	-0.0974	0.0419
ISEE_Not_disclosed	0.7399	0.1123	0.1095	0.000***	0.0803	0.1386
_cons	-0.3629	0.2815				

NOTES. Control variables included as indicated in section 6: gender, age at matriculation, upper secondary school track, mark at the final examination in upper secondary school, field of study and institutional length of the degree.
* Statistical significance at the 0.10 level; ** significance at the 0.05 level; *** significance at the 0.01 level.
Model 1: N=7352; pseudo-R^2=0.0701. Model 2: N=7352; pseudo-R^2=0.0992.

Instead, the household's economic conditions – grouped in classes to allow for non-linear effects – are relevant. The dropout probability is highest at low-income levels, decreases up to income 30000-45000, and then increases again. (Notice that when considered as a quantitative variable, ISEE has a negative and significant coefficient). The relatively higher dropout propensities of the students coming from affluent families may be related to their lesser feeling of obligation towards their families to succeed in their studies, because university costs are negligible relative to their budgets, and they have no urgency to make their own living or contribute to family income.

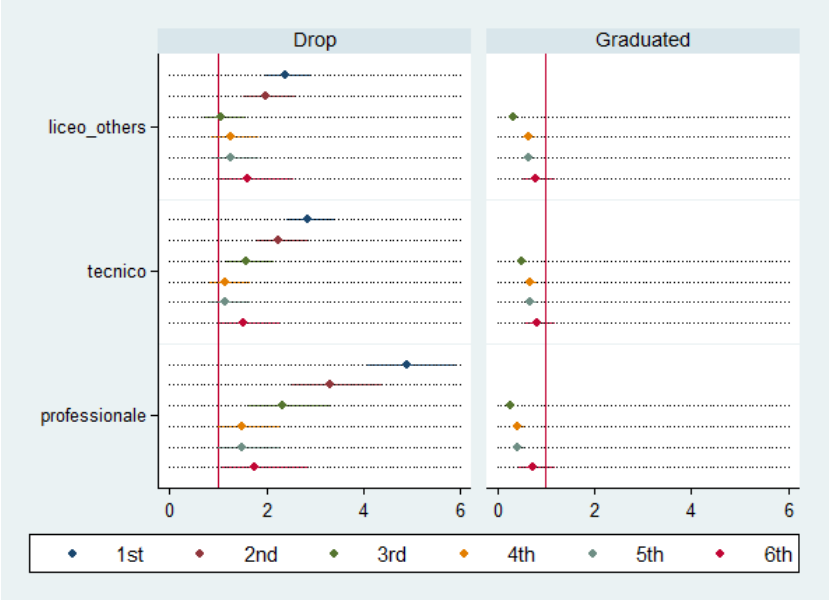
Non-disclosure of ISEE was treated as an additional possible response. Findings relative to this category are rather surprising: those who did not disclose ISEE have a much higher probability of withdrawal even if compared to the students declaring the lowest ISEE levels. This result is difficult to interpret and – despite the speculations made above on the behaviour of high-income students – it makes us question our initial idea that the students not disclosing ISEE are the most affluent. A possible explanation could be that different students compose this group. On the one side, those having no incentives to declare their income because it is too high to get any benefit; on the other side, poorly motivated students (probably not very-low-income), who might live the university experience with little determination and may even "forget" to go through the necessary steps to reduce participation costs.

We may summarize our findings on the role of family conditions as follows: once we control for economic conditions, parental education and occupation seem to matter little after university enrolment. Instead, and consistently with the expectations, economic conditions do play a role, although in a non-linear way. Even if they are highly positively selected, low-income students are exposed to substantially larger dropout probabilities than middle-income students (by 3-4 percentage points), supporting the hypothesis that the lack of economic resources can constitute a barrier to the completion of university studies.

7.2 Research question 2

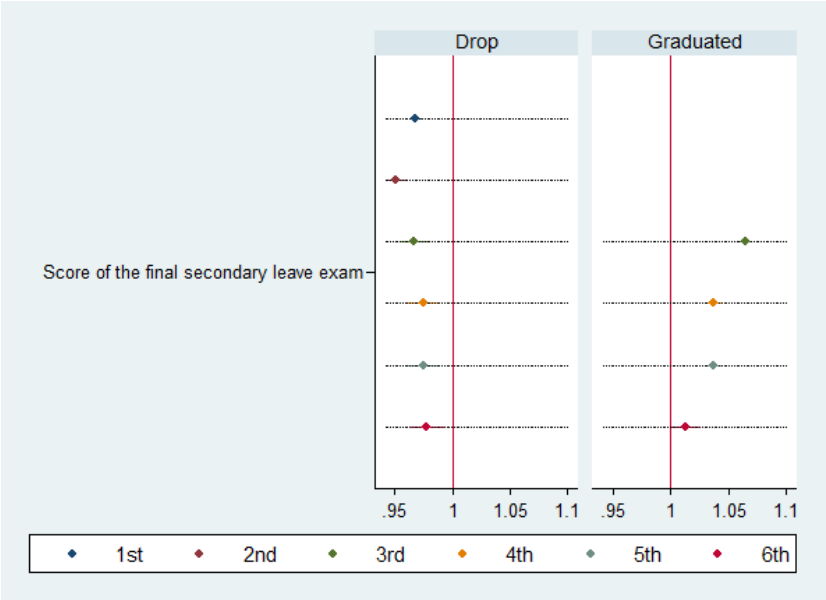
To date, we have carried out analyses only for the students matriculated in 3-year degree programs. The results for research question 2 are organized as follows. Firstly, we discuss a selection of the results of the multinomial regressions for each step of the academic career (according to Figure 1) as explained in section 6.2. For illustrative purposes, we will focus on matriculation cohort 2008. Secondly, we will present the Cumulative Incidence Functions (CIFs) for chosen profiles of students, for all the cohorts analyzed.

Figure 2. Effects of school track on dropout and graduation probabilities in years 1-6 from enrolment (OR). Matriculation cohort 2008.



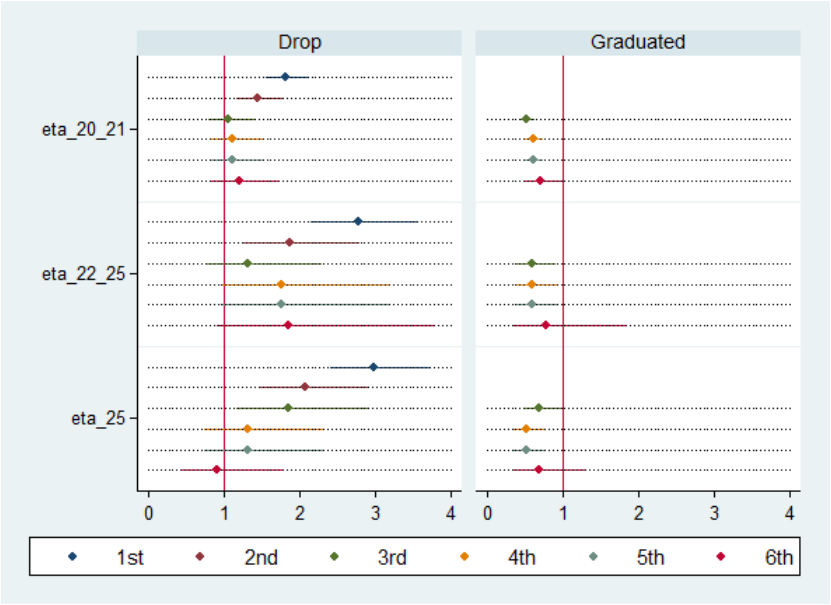
School type exerts strong effects on university outcomes. The dropout probability is much higher and the graduation probability is much lower for all school types other than liceo (the poorest outcomes are obtained by the students from the vocational track, OR against liceo nearly 5 at year 1). The effect on the dropout probability is strongest at year 1 after enrolment, diminishes up to year 3 and then slightly increases again (although differences after year 3 are non-significant). The effect on the graduation probability is stronger at year 3 and then reduces progressively.

Figure 3. Effects of the final upper secondary examination mark on dropout and graduation probabilities at years 1-6 from enrolment (OR). Matriculation cohort 2008.



Strong effects of grade at the upper secondary final examination (given school track): the dropout probability decreases and the graduation probability increases with grades. The effect on the dropout probability is larger at years 1-3; while the effect on the probability to attain the degree progressively reduces from year 3 on.

Figure 4. Effects of age at enrolment on dropout and graduation probabilities in years 1-6 from enrolment (OR). Matriculation cohort 2008.



The age at first enrolment is highly relevant. The dropout probability is much lower and the graduation probability is much higher for students enrolling at the regular age (19). The effect of age at matriculation on the dropout probability is strongest at year 1 after enrolment and then

declines up to year 3. Due to the small sample size of the oldest age groups, differences are often not statistically significant (most the confidence intervals overlap). The effect on the graduation probability is rather stable over time.

In addition, the dropout probability is consistently lower for girls at all years, although it is never statistically significant (consider that if we run a single model for all years, the gender coefficient becomes significant). Instead, gender has no effect on the graduation probability.

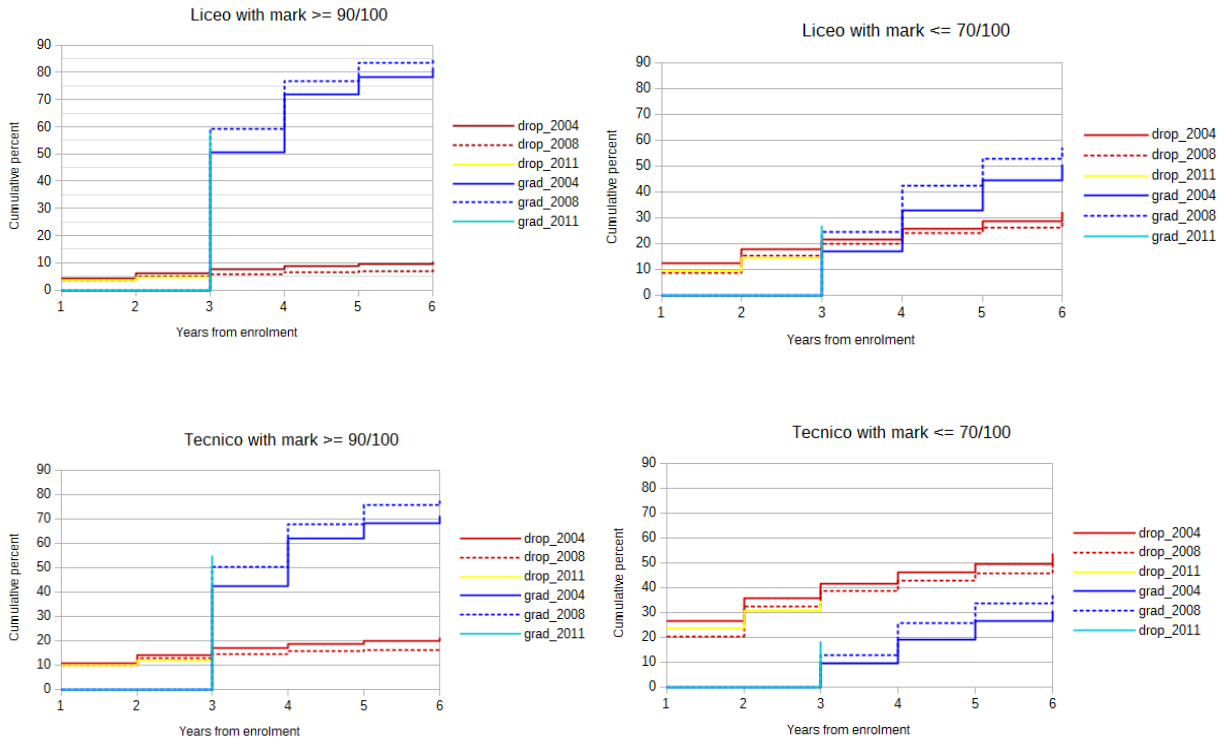
Differences across fields of study are also marked. In particular, the health field stands as the one with the lowest dropout and the highest graduation probabilities at all years (with the exception of year 6 for graduation). The better overall performances of the students in this field is presumably because they are subject to a severe ability based selection at enrollment.

Two additional remarks. First, the above results refer to matriculation cohort 2008, but these patterns hold approximately also for the other matriculation cohorts. Second, as we have seen, all the coefficients (with the exception of gender) change markedly as time from enrolment goes by. To allow for this variability in a standard survival/competing risk framework we would need to include non-linear interactions between all the explanatory variables with elapsed duration in the state. This is obviously possible, although many interaction effects would be involved seriously complicating interpretation.

***

In Figure 5 we depict the CIFs for selected student profiles. We consider individuals enrolled at the regular age of 19, comparing those with a lyceum diploma and those from the technical track, with a high mark (≥90) or a poor mark (≤70). Clearly, these are exemplificative cases, as many other different profiles are in the sample. In this section, we comment the findings in Figure 5 focusing on the differences across the profiles for cohort 2008, while we leave the description of changes across cohorts to the section devoted to research question 3.

Figure 5. CIFs for individuals enrolled at 19 in 3-year degree programs, coming from liceo vs. tecnichal schools, marks ≥90 vs ≤70. Matriculation cohorts 2004, 2008, 2011.



We estimate that very well performing students from liceo have 8% probability of withdrawing within 6 years from enrolment (half of which in year 1) and 84% probability to graduate within 6 years (yet, less than 60% attain the degree within the institutional time of 3 years).

Students with the same profile but that obtained a poor score in the final high school examination exhibit much poorer outcomes: 27% probability of withdrawing within 6 years from enrolment (a third of which in year 1), 56% probability of graduation within 6 years (24% within 3 years).

The role of performance in high school is even stronger when focusing on students from the technical track. While the well performing ones exhibit relatively good outcomes (17% and 76% dropout and graduation probabilities within 6 years respectively), the low performing end up with very poor ones (46% and 36% dropout and graduation within 6 years respectively).

Students from other_liceo types perform similarly to those from the technical track, and students from vocational tracks perform even more poorly (although those with high marks are doing better than those from traditional liceos with low marks). As we see in Appendix 2, however, students from the vocational track are strongly under-represented in university, due to the previous selection process.

The shares of students still in the system after 3 and 6 years (i.e., they have not withdrawn but have not graduated) are reported in Table 3. Interestingly, these figures seem to depend more on the high

school final mark than on the high school track. The shares after 6 years vary between 5-11% for those with high marks and 13-18% for those with low marks.

Table 3. Residual percentage of students remaining in the University system after 3 and 6 years from enrolment.

Profile:	Lyceum, High mark		Lyceum, Low mark		Other academic, High mark		Other academic, Low mark	
Year:	3th	6th	3th	6th	3th	6th	3th	6th
2004	41,6	7,7	61,8	17,3	51,3	10,3	56,0	18,4
2008	34,8	7,1	55,8	14,2	47,1	6,6	57,8	16,3
2011	35,4	-	55,3	-	38,4	-	54,4	-

Profile:	Technical, High mark		Technical, Low mark		Vocational, High mark		Vocational, Low mark	
Year:	3th	6th	3th	6th	3th	6th	3th	6th
2004	40,7	7,6	48,8	15,6	51,7	11,5	50,8	17,7
2008	35,3	5,6	48,4	13,2	46,9	10,7	41,8	13,7
2011	31,1	-	46,9	-	43,0	-	47,1	-

7.3 Research question 3

Let us go back to Figure 5 and focus on the changes of CIFs over time. For all the depicted profiles, the outcomes tend to improve substantially between cohorts 2004 and 2008. The graduation probability increases markedly and the dropout probability declines by a few percentage points. (Notice that these results hold also hold for the corresponding profiles in the other_liceo school types, whereas the opposite trends hold for the small share of students from the vocational track). The probability of degree attainment increases also between cohorts 2008 and 2011; instead, the dropout probability tends to remain stable. Finally, the share of students still in the system after 6 years decline consistently between cohorts 2004 and 2008 by a couple of percentage points.

We now turn to analyzing the decomposition results, starting with dropout. The first step is the computation of observed and "counterfactual" aggregate probabilities (see Table 4).

Table 4. Observed and "counterfactual" dropout probabilities within 1, 3, 6 years from enrolment (CIFs). All fields of study.

Year 1

Populations \ Coefficients	2004	2008	2011
2004	18,9	17,7	17,21
2008	16,92	15,33	15,25
2011	16,65	14,84	14,95

Year 3

Populations \ Coefficients	2004	2008	2011
2004	29,59	28,58	26,75
2008	27,06	25,63	24,28
2011	26,99	25,13	24,27

Year 6

Populations \ Coefficients	2004	2008
2004	37,78	35,92
2008	35,14	32,87
2011	35,37	32,65

Consider year 1 as an example. The "observed" dropout percentages in the samples are in the main diagonal: 18.90 for cohort 2004, 15.33 for cohort 2008 and 14.95 for cohort 2011. Hence, the overall trend is of sharp decrease between the first two cohorts (-3.57) and essential stability between the second and the third (-0.38). Off-diagonal elements represent the so-called counterfactuals. If we were to apply the coefficients estimated for 2008 to the population of cohort 2004, we would obtain an aggregate share of dropouts of 17.70. If we compare this figure to the observed share for 2008 (15.33), we obtain composition effects, because these two figures refer to different populations with the same behavior. Instead, if the population enrolled in 2004 behaved like that of 2008, we would observe a lower share (17.70) than that actually observed in 2004 (18.90). This comparison informs us on behavioral effects. Hence, we may decompose the total observed differential -3.57 into a component due to differences in composition (15.33-17.70=-2.37) and a component due to differences in behavior (17.70-18.90=-1.20). The component related to compositional effects accounts for 66.4% of the change, whereas the one related to behavioral differences accounts for the remaining 33.6%. As for the negligible gap between the dropout probability after year 1 between cohorts 2008 and 2011, we observe that it is largely due to compositional difference in the two populations.

Clearly, we could also make an alternative decomposition by comparing the observed shares to the symmetric counterfactual (the probability that we would obtain by applying the coefficients estimated for 2004 to the population of cohort 2008). For simplicity, we have chosen to focus on the first one.

By inspecting Table 5, we see that the largest change has occurred between cohorts 2008 and 2004 within the first year after enrolment (the gap only slightly increases after the first year).[20] Composition effects appear to prevail in all comparisons. Instead, only small changes have taken place between cohorts 2008 and 2011, mostly between years 1 and 3, and mainly due to behavioral effects.

Table 4. Decompositions into compositional and behavioral effects of dropout CIFs. Years 1, 3 and 6 after enrolment. All fields of study.

[20] The comparison between cohorts 2008 and 2011 cannot be made at year 6 because the latter is too recent and the available data goes up to year 2014.

I drop

Cohort compared	Total Difference	Difference due to composition	Difference due to behaviour	Percent of compositions effects	Percent of behavioural effect
2008-2004	-3,57	-2,37	-1,2	66,39	33,61
2011-2008	-0,38	-0,3	-0,08	78,95	21,05

III drop

Cohort compared	Total Difference	Difference due to composition	Difference due to behaviour	Percent of compositions effects	Percent of behavioural effect
2008-2004	-3,96	-2,95	-1,01	74,49	25,51
2011-2008	-1,36	-0,01	-1,35	0,74	99,26

VI drop

Cohort compared	Total Difference	Difference due to composition	Difference due to behaviour	Percent of compositions effects	Percent of behavioural effect
2008-2004	-4,91	-3,05	-1,86	62,12	37,88

We now examine the findings relative to graduation probabilities. In Table 6 we report the observed and counterfactual probabilities of obtaining the degree within 3 and 6 years.

Table 6. Observed and "counterfactual" graduation probabilities within 3 and 6 years from enrolment (CIFs). All fields of study.

Coefficients

Populations	2004	2008	2011
2004	22,43	26,87	30,82
2008	23,56	27,82	31,33
2011	22,98	27,08	30,25

Year 3

Coefficients

Populations	2004	2008
2004	49,17	52,76
2008	51,54	55,48
2011	51	55,05

Year 6

Table 7 shows the decomposition for graduation probabilities after year 3 and 6. The largest change has occurred between cohorts 2004 and 2008. Since the difference has changed little between year 3 and 6, we can conclude that the most relevant improvement regards graduation at year 3.

Here behavioral effects prevail. Notice the odd figures -44.44% and 144.44% result from the ratios -1,08/2.43 and 3.51/2.43 and tells us that if only composition effects were to apply, we would observe a *decline* in the share of graduates after year 3 between 2008 and 2011. The reason why we witness to a 2.43 percentage points increase is that behavioral effects are strong and counteract the negative effect due to compositional changes. Hence, in this case the improvement is *entirely* due to changes in behavior.

Table 7. Decompositions into compositional and behavioral effects of graduation CIFs. Years 3 and 6 after enrolment. All fields of study.

III grad

Cohort compared	Total Difference	Difference due to composition	Difference due to behaviour	Percent of compositions effects	Percent of behavioural effect
2008-2004	5,39	0,95	4,44	17,63	82,37
2011-2008	2,43	-1,08	3,51	-44,44	144,44

VI grad

Cohort compared	Total Difference	Difference due to composition	Difference due to behaviour	Percent of compositions effects	Percent of behavioural effect
2008-2004	6,31	2,72	3,59	43,11	56,89

We propose the following intuition behind the finding that composition effects mainly apply to the dropout probability and play only a minor role in the observed changes in the graduation probability. Since dropout probabilities decline in recent cohorts (even net of individual characteristics), more "weak" students remain in the university system. Thus, even if they students are on average better endowed at enrollment, they need not to be better endowed later on, after a further selection mechanism (dropout) has operated.

Decompositions by field of study.

In the decompositions presented above, the explanatory variables comprise individual characteristics and the disciplinary area. Therefore, in the composition effects we include also the allocation process of students to the available options. We now apply the above decompositions by analyzing one field at a time, so compositional effects refer to the changes in the population of students enrolling in a specific field.

Our results – that need further effort to be fully understood – can be summarized as follows. In the humanities field, improvements in the dropout probabilities are modest, whereas graduation probabilities improve markedly. Behavioral effects prevail for both outcomes. In the social-politics-law field, substantial improvements are observed for dropout and graduation probabilities. Relevant compositional and behavioral effects are found for dropout, but only behavioral effects for graduation. In the science field, improvements in the dropout and graduation probabilities occur only between 2004 and 2008, mainly due to compositional changes. In the healthcare field, for dropout improvements are observed only between cohorts 2008 and 2011, and they are mainly behavioral effects. For graduation, we have witnessed to a substantial pejorative process between cohorts 2004 and 2008, reducing the timely graduation probabilities (whereas the probabilities of

graduation within 6 years have been virtually unchanged). This has been entirely due to changes in behavior.

8. Effect of the economic crisis?

Theoretical predictions on the effects of the economic downturn on tertiary education outcomes are not clear-cut. Firstly, individuals may react differently to adverse conditions of the labour market, according to preferences, risk aversion and other psychological traits. Secondly, individuals are subject to different constraints, due to diverse economic and cultural endowments.

The effects of economic crisis on educational choices may operate at two levels, that we find useful to keep distinct. A downturn affects the conditions of the macroeconomic context and labour market *opportunities*, thereby changing the prospective students' evaluation of costs and benefits of tertiary education. However, a downturn also directly affects the conditions of individuals, *constraining* the educational options of prospective students experiencing financial hardship.

In this perspective, we now examine the potential mechanisms related to the economic crisis operating on university enrolment, dropout and time to degree.

Enrolment. According to economists' theoretical predictions, by reducing the opportunity costs of studying, poor labour market prospects should increase tertiary education participation. Yet, poor labour market prospects might also affect motivation and yield to discouragement, contributing to reduce participation. The balance between these opposite forces should depend on the (perceived) returns to tertiary education.

However, bad economic conditions might also have a direct negative influence on enrolment: individuals living in families experiencing severe financial hardship might have an urgent need of income and may not be able to attend university.

In this scenario, while the crisis increases the share of individuals in the population in poor economic conditions, the effects on the composition of university students is by no means clear-cut.

What does the empirical evidence tell us? It shows that participation to tertiary education has become more selective. The share of high school diplomats making the transition to university has decreased by 3-4 percentage points from 2009; correspondingly, the population of university students has become increasingly less composed by students from the technical and vocational tracks. Yet, whether these changes can be entirely ascribed to the crisis is unclear.[21]

[21] According to ANVUR (2016) the decline in enrolment rates is due to the rise in the incidence of foreign students among upper secondary diplomats, as these students have very low transition rates to tertiary studies.

Dropout. High unemployment rates may contribute to lower dropout rates because the opportunity costs of attending university decline. Moreover, a decline in employment opportunities should reduce the number of withdrawals related to labour market pull out factors. On the contrary, students lacking economic resources might be pushed out of the educational systems due to an immediate need to earn income. (This is particularly relevant in Italy, as the loan system to finance educational projects is very limited).

In this respect, our findings from RQ1 unambiguously tell us that income matters: households with lower economic resources experience higher dropout rates. (Interestingly, this occurs in spite of the compositional change in the university student population according to which students from non-strictly academic upper secondary tracks – where students from low socioeconomic backgrounds are overrepresented – are increasingly less inclined to enrol in college and hence, those among them who do enrol are more positively selected in terms of motivation and aspirations).

However, financial hardship in itself does not seem to be a major driver of the observed change in dropout behaviour. In fact, our findings from RQ3 show that dropout rates have declined substantially over matriculation cohorts, and that this decline is largely attributable to the "better" composition of students. This is possible because due to the selection process at enrolment the economic crisis does not necessarily lead to a larger share of university students in financial hardship.

We also observe some behavioural changes, and these reinforce the decline of dropout. This result supports the hypothesis that the lower opportunity costs of studying and reduced labour market pull out factors play a role in reducing university withdrawal.

Time to degree. On the one side, if the labour market prospects of university graduates are poor, opportunity costs decline and individuals should take longer time to complete their majors. However, the increasing competition over scarce job positions may also induce students to attain the degree in time, because this would be a good signal for prospective employers. Focusing on the role of current resources, students living in households with financial problems might feel urged to graduate, to cut university tuition expenses and start earning income, or, on the contrary they might need to work while studying, contributing to extend time to degree.

Results from RQ3 show that time to graduation has declined during the years of the economic crisis, and that the share of students attaining the degree within the institutional time (although still far below the wishes) has risen considerably. The improvement is due to both compositional and behavioural changes, but the latter definitely prevail. This evidence supports the hypothesis that the

crisis has led to increasing competition among students and increasing urge to complete studies in a short time span.

University as a parking? The evidence from the University of Torino does not support the hypothesis that due to decreasing employment prospects high school leavers enter college because they lack alternatives and then remain in the university system without putting much effort in their studies. On the contrary, fewer youngsters enrol and, net of compositional effects, take less time to graduate.

A final remark. The considerations we have made on the potential effects of the economic crisis are largely speculative. An important concurring factor is that, parallel to the downturn, the Italian university has undergone to various institutional changes that followed the implementation of the Bologna process, aimed at ameliorating the system. These interventions could have contributed to the improvement of dropout and timely completion rates. Further research is needed to deepen our understanding on the reasons underlying the observed changes. Since institutional changes invested the entire Italian university system simultaneously, disentangling the effects of institutional changes from the effects of the crisis is a difficult task.

8. Conclusions

Our main findings can be sketched as follows:

(i) Family income matters: students in economically disadvantaged households experience higher dropout risks. This result holds despite the fact that, given observable characteristics, these students are likely to be more positively selected in terms of unobserved ability and motivations as compared to their more affluent peers. Hence, the estimated effect is likely to be an underestimate of the real "causal" effect of the lack of income.

(ii) Previous schooling matters: the upper secondary school track and the mark at the final high school examination have a huge effect on both dropout and time to degree. This is hardly surprising, since no entry barriers related to ability exist in the Italian university system. Yet, this mechanism ends up adding to the large previously established inequalities that make the children from advantaged family backgrounds much more likely to attend lyceums and to have good school performances, as disproportionate shares of individuals of low socio-economic status are pushed out the educational system after enrolment.

(iii) Dropout risks and time to degree have declined substantially across recent cohorts. These improvements can partly be ascribed to the changes underwent in the composition of the

student population – many more students now come from lyceums and have a regular previous schooling career – and partly to behavioral changes. More specifically, the decline in dropout probabilities seems to be largely due to composition effects, whereas the shortened time to degree is mostly due to changes in individual behavior.

(iv) There is no evidence that university is being increasingly considered as a "parking". The share of students still in the system a few years after enrolment has actually declined substantially over recent cohorts.

(v) How can we interpret these findings? In particular, do we have any evidence that the observed changes are related to the economic crisis? As argued in section 8, we can only make speculative considerations, because other macro-level factors may have positively influenced university outcomes; in particular, ongoing institutional changes in the university system. Disentangling the potential causes of change is extremely difficult because these changes have all occurred at the same time.

(vi) Keeping this in mind, our findings may support the hypotheses: i) that the lower university dropout risks are related to the reduced opportunity costs of studying and to the reduced labour market opportunities; ii) that the economic downturn has led to a stronger competition over limited job opportunities, so students tend to put more effort in their studies and reduce time to degree.

References

Aina C., E. Baici, G. Casalone (2011) Time to degree: students' abilities, university characteristics or something else? Evidence from Italy. *Education Economics*, 19, 3, 311-325

Almalaurea (2016), XVII indagine Profilo dei Laureati 2015, Consorzio AlmaLaurea, Bologna

Almalaurea (2015), XVII indagine Profilo dei Laureati 2014, Consorzio AlmaLaurea, Bologna

Almalaurea (2014), XVII indagine Profilo dei Laureati 2013, Consorzio AlmaLaurea, Bologna

ANVUR (2014) Rapporto sullo stato del sistema universitario e della ricerca 2013, Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca, Roma.

ANVUR (2016) Rapporto biennale sullo stato del sistema universitario e della ricerca 2016, Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca, Roma.

Arulampalam W., R. A. Naylor, J. P. Smith (2005) Effects of in-class variation and student rank on the probability of withdrawal: cross-section and time-series analysis for UK university students, *Economics of Education Review* 24, 251–262

Bertola G., D. Checchi, V. Oppedisano (2008) Private school quality in Italy, *Giornale degli Economisti e Annali di Economia*, 66-3, 375-400

Blundell R., L. Dearden, B. Sianesi (2001) Estimating the Returns to Education: Models, Methods and Results, Centre for the Economics of Education London School of Economics.

Bourdieu P., J.C. Passeron (1990) *Reproduction in education, society, and culture*, Vol 4, Sage

Boudon R. (1974) *Education*, *Opportunity, and Social Inequality: Changing Prospects in Western Society*, John Wiley & Sons

Bratti M., C. Broccolini, S. Staffolani (2010) Higher Education Reform, Student Time Allocation and
Academic Performance in Italy: Evidence from a Faculty of Economics, *Rivista Italiana degli Economisti*, 2, 275-304

Breen R. and J. Golthorpe, 1997 Explaining educational differentials: towards a formal rational action theory, *Rationality and Society,* 9, 275-305

Cappellari L., C. Lucifora (2009) The "Bologna Process" and college enrollment decisions, *Labour Economics* 16, 638–647

Carneiro, P., J. Heckman. (2002) The evidence on credit constraints in postsecondary schooling. *Economic Journal*. 112:705–734.

Checchi D., C. V. Fiorio, M. Leonardi (2014) Parent's risk aversion and children's educational attainment, *Labour Economics*

Cingano F., P. Cipollone (2007) University drop-out: The case of Italy, Temi di discussione Servizio Studi della Banca d'Italia, 626

Clerici R., A. Giraldo, S. Meggiolaro (2014) The determinants of academic outcomes in a competing risks approach: evidence from Italy, *Studies in Higher Education*

Contini Dalit, Cugnata Federica, Scagni Andrea (2015) Social selection in tertiary education. Enrolment, dropout and timely degree attainment in Italy, WP 9/16 Department of Economics and Statistics, University of Turin.

D'Hombres, (2007) The impact of university reforms on dropout rates and students' status: evidence from Italy, Joint Research Centre Scientific and Technical Reports, European Commission

Di Pietro G. (2004) The determinants of university dropout in Italy: a bivariate probability model with sample selection *Applied Economics Letters*, 11:3, 187-191

Di Pietro G., A. Cutillo (2008) Degree flexibility and university dropout: the Italian experience, *Economics of Education Review* 27, 546–555

Dynarski S. (2003) Does aid matter? Measuring the effect of student aid on college attendance and completion, *American Economic Review*, 93(1), 279-288.

Eurostat (2015) Education and training in the EU- facts and figures (http://ec.europa.eu/eurostat/statistics-explained/index.php/Education_and_training_in_the_EU_-_facts_and_figures)

Eurydice, 2012. Key data on education in Europe 2012, European Commission.

Federconsumatori (2015) 6° Rapporto Nazionale Federconsumatori sui costi degli atenei italiani, http://www.federconsumatori.it/news/foto/Rapporto%20Tasse%20universitarie%20Rapporto%202015-2016.pdf

Garibaldi P. F. Giavazzi, A. Ichino, E. Rettore (2012) College cost and time to complete a degree: evidence from tuition discontinuities, *The Review of Economics and Statistics,* 94(3): 699–711

Ghignoni, E., 2016. Family background and university dropouts during the crisis: the case of Italy. Higher Education

[Hillman N. W., E. Lee Orians (2013) Financial Aid's Role in Meeting State College Completion Goals. *Education Finance and Policy* 8:3, 349-363]

Ishitani T. T. (2006) Studying Attrition and Degree Completion Behavior among First-Generation College Students in the United States, *The Journal of Higher Education*, 77, 5, 861-885

Jackson M., J. O. Jonsson (2013) Why does inequality of educational opportunity vary across countries? Primary and secondary effects in comparative context, in Jackson M. eds. *Determined to Succeed? Performance, Choice and Education*, Stanford University Press

Mealli F., C. Rampichini (2012) Evaluating the effects of university grants by using regression discontinuity designs, *Journal of the Royal Statistical Society A*, 175, 3, 775–798

OECD (2013), *Education at a Glance* 2013: OECD Indicators, OECD Publishing

Stinebrickner R., T. Stinebrickner (2008) The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study, *American Economic Review* 98,5, 2163-84

Tinto V. (1975) Dropout from higher education: a theoretical synthesis of recent research, *Review of Educational Research*, 45, 1, 89-125

Vignoles A. F., N. Powdthavee (2009) The Socioeconomic Gap in University Dropouts, *The B.E. Journal of Economic Analysis & Policy*: Vol. 9: Iss. 1 (Topics), Article 19.

Appendix 1

Comparison between the conventional competing risk model and step-by-step estimation

Let us first recall the basics of survival and competing risks models.

Survival analysis aims at modeling time elapsed in a given state of a discrete state process. When the event “exit” can occur only at isolated time points (say, at time 1, 2, 3…), we speak of discrete

time models. In this context, the functions of main interest are the survival function $S(j) = P(T > j)$ and the hazard function $h(j) = P(T = j|T > j-1)$, representing the probability of exiting the state at time $T = j$ given survival up to time $j-1$.

Competing risks (CR) models are extensions of survival models applied when the exit from the state of interest may occur towards different destination states and we are interested in considering the destination along with the timing of exit event. Our case study fits this situation. Students remain in the system if they renovate enrolment while the exit event occurs if they either drop out of university (first destination) or attain the degree (second destination). Individuals are censored if they are still enrolled at the end of the observation window.

More specifically, CR models focus on the so-called transition intensities or cause-specific hazards, defined as $P(T = j, D = v|T > j-1)$, i.e. the probability of exiting at time $T = j$ towards destination $D = v$ given survival up to time $j-1$.

The probability that a given individual exits the system at time $T = j$ towards destination $D = v$ is:

$$P(T = j, D = v) = (1 - P(T = 1))(1 - P(T = 2|T > 1) \dots P(T = j, D = v|T > j-1)$$

which is equal to the probability of not exiting the system at time 1, times the probability of not exiting the system at time 2 given survival up to time 1, etc. …, times the probability of exiting the system towards a given destination (dropout or graduation) at time *j* given survival up to time *j*-1.

How do explanatory variables enter transition intensities? Usually, explanatory variables are assumed to have the same effect at all time points. Since in discrete-time modeling the probabilities are specified as binary or multinomial logit models, the most typical assumption in discrete-time modeling is that of proportional odds.

There are a number of advantages in adopting this step-by-step estimation approach over conventional discrete time competing-risk modelling in our case study.

A first shortcoming of conventional CR estimation is that the assumption that explanatory variables have the same effect at all time points is not corroborated by empirical evidence. Some explanatory variables are empirically relevant in first year choices while have no effect (or sometimes even a stronger effect) in subsequent years (see…). By analyzing each step separately, these constraints are not applied.

A second reason for adopting a step-by-step approach is that individuals are allowed to graduate only after the institutional time to degree has elapsed (3, 5, or 6 years after enrolment). This restriction cannot be applied in conventional CR analyses.

The third and most important reason for not using conventional CR models is that they do not allow distinguishing between different ways of remaining in the system, meaning that we cannot model the event of changing the degree program without defining it as a possible destination. However, this is not what we want to do, because once individuals are assigned a destination, they fall out of the risk set, and are no longer included in the analysis at following years. Yet, students changing degree are not exiting the system. Nonetheless, we are interested in modeling degree changes, because they are informative on individuals' attitudes and help predicting future outcomes. Horizontal moves usually come along with dissatisfaction over the current degree program or insufficient academic skills to continue, while upward moves (degree changes from 3- to 5- or 6-year programs) usually demonstrate high aspirations and/or high ability. Contrary to conventional CR models, with step-by-step analysis we make degree changes endogenous by including them as possible transitory outcomes within the process under study. Once individuals experience the change, we do not force them out of the risk set, so we are still able to analyze their dropout or graduation probabilities from thereon.

Clearly, the main limitation of the step-by-step estimation approach is that the number of parameters rises substantially, affecting the efficiency of the estimates.[22] However, given the large number of observations (approximately 10,000 individuals per matriculation cohort, followed for 1-6 years), we are confident that the estimates are reliable enough.

Appendix 2

Samples' descriptives.

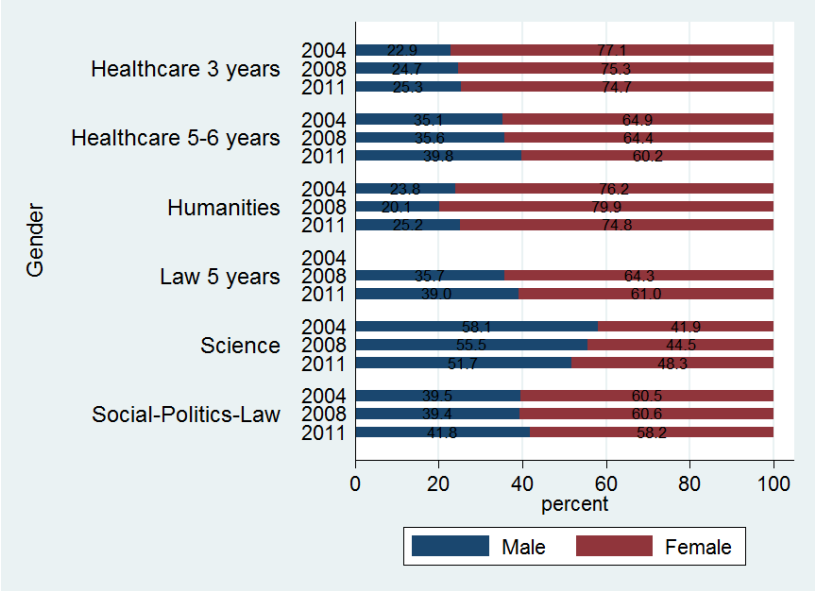
[22] Considering the set of explanatory variables summarized in Table 1, this means that to account for the effects of the explanatory variables instead of 2*k* parameters (*k* variables times 2 destinations), we specify 24*k* parameters (6 years by 4 options per year: O, C+, C=, D in years 1-2; O, C, D, G in subsequent years).

	2004		2008		2011		Total	
	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent
Gender								
Female	7021	60,88	6569	63,06	6506	61,39	20096	61,74
Male	4511	39,12	3848	36,94	4092	38,61	12451	38,26
Age at enrolment								
<= 19	7779	67,46	7449	71,51	7385	69,68	22613	69,48
20-21	2080	18,04	1863	17,88	2313	21,82	6256	19,22
22-25	671	5,82	476	4,57	471	4,44	1618	4,97
>25	1002	8,69	629	6,04	429	4,05	2060	6,33
Previous School								
Liceo	5437	47,15	5777	55,46	6230	58,78	17444	55,53
Other Liceo	1015	8,8	1333	12,8	1414	13,34	3762	11,56
Technical	2684	23,27	1885	18,1	1703	16,07	6272	19,27
Vocational	1981	17,18	1071	10,28	882	8,32	3934	12,09
Foreign	356	3,09	338	3,24	364	3,43	1058	3,25
Missing	59	0,51	13	0,12	5	0,05	77	0,24
Academic Field								
Healthcare	1417	12,29	1493	14,33	1518	14,32	4428	13,60
Scientific	2329	20,2	2047	19,65	2377	22,43	6753	20,75
Social-Politics-Law	5744	49,81	4781	45,9	4386	41,39	14911	45,81
Humanities	2042	17,71	2096	20,12	2317	21,86	6455	19,83
Total N	11532	35,43	10417	32,01	10598	32,56	32547	100

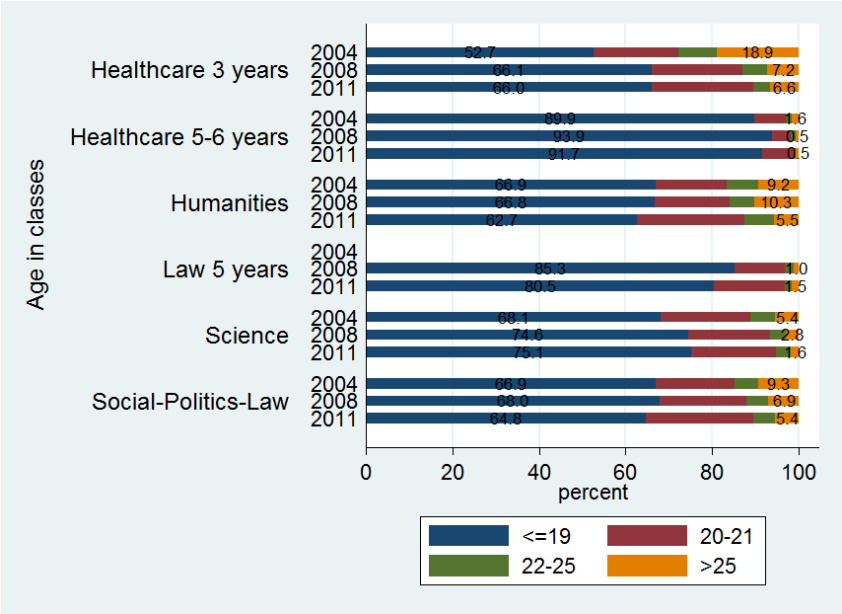
<u>Choice of the field of study over time</u>

We note a reduction of the gender gap in the science field (where girls are traditionally under-represented), in the healthcare and law fields (where girls are traditionally over-represented)….

Gender differences across fields of study over time

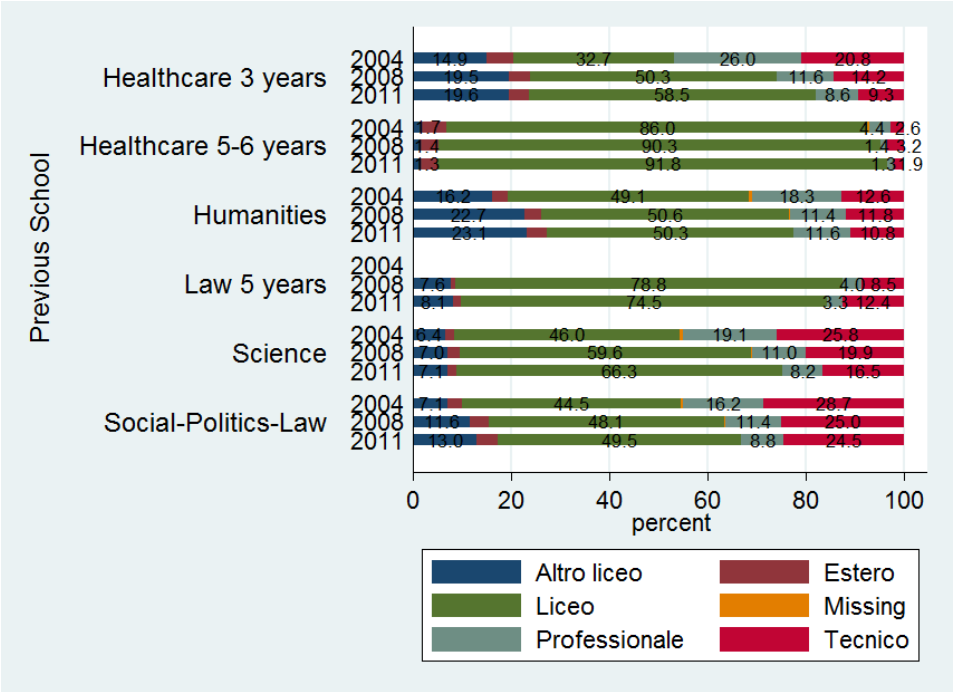


Differences in the age at enrolment across fields of study over time



Increase in the share of students aged 21 and below. Sharp reduction of over 25.

School track differences across fields of study over time



Sharp increase of the share of students from traditional liceo, in particular in the healthcare (and even more in 3-year programs) and science fields Sharp reduction of students from the vocational track in all fields. Students from the technical track are decreasingly represented in the healthcare 3-years and science, while they gain quotas in law 5-years.

Appendix 3

The classification scheme for parental occupation was defined by first merging professions into a 5-level scale per each parent, and then combining the categories for the two family members as in the following table.

Merging of parental job types into a three way qualification level classification

professional qualification level	***mother***				
father	*houseworker*	*blue collar, artisan, retailer*	*white collar, teacher*	*entrepreneur, professional*	*manager, medical doctor*
blue collar, artisan, retailer	**low**	**low**	*medium*	*medium*	high
white collar, teacher	**low**	*medium*	high	*medium*	high
entrepreneur, professional	**low**	*medium*	*medium*	*medium*	high
manager, medical doctor	high	high	high	high	high

Please notice that the term “entrepeneur” included in the third group very often refers to small or even very small (family based) firms; thus, some care should be taken in considering this as a fully “high qualification group. The apparently inconsistent final classification could be due to this feature.

Some descriptives on the size of groups and behaviour of student relative to the above classification are given below.

Group sizes according to job types, families of first-time students, 2014/15

n. of families	***mother***					
father	*house-worker*	*blue collar, artisan, retailer*	*white collar, teacher*	*entrepreneur, professional*	*manager, medical doctor*	Total
blue collar, artisan, retailer	560	1434	1100	106	63	3263
white collar, teacher	262	365	1634	122	103	2486
entrepreneur, professional	219	269	725	416	122	1751
manager, medical doctor	109	138	671	92	213	1223
Total	1150	2206	4130	736	501	8723

Dropout frequency by parental job qualification, first-time students, 2014/15

	continuing	*dropout*	*Total*
high qualification job	4023	765	4788
%	84.0	**16.0**	100.0

medium qualification job	3611	791	4402
%	82.0	**18.0**	100.0
low qualification job	2834	772	3606
%	78.6	**21.4**	100.0
Total	10468	2328	12796
	81.8	18.2	100.0