Too few university graduates. Inclusiveness and effectiveness of the Italian higher education system

Published version:

DOI:10.1016/j.seps.2020.100803

(Article begins on next page)

Too few university graduates. Inclusiveness and effectiveness of the Italian higher education system

Dalit Contini
Guido Salza

University of Torino

Abstract

Despite the absence of formal entry barriers, Italy is lagging behind the majority of the developed countries in the share of young individuals with tertiary education. Exploiting the administrative data of a large public Italian institution, we analyze student academic careers across recent matriculation cohorts. We propose a flexible discrete-time competing risks estimation that allows overcoming some major limitations of conventional competing risks models. We find that student pathways vary tremendously across prior schooling profiles: for example, the within 4-year dropout probability ranges between 10% and 77%. We observe improvements over time in retention and time-to-degree, and by decomposing changes into components related to the composition of the enrolled population, the choice of the field of study and 'individual behavior' after enrolment, we find that the latter plays a major role. However, the improvement is limited in size and does not interest students from the vocational track. Since this progress is not accompanied by an increase in the share of students making the transition from high school to university, altogether our results call for great concern over the inclusiveness and effectiveness of the Italian university system.

Keywords: educational economics, school choice, student academic careers, dropout, time to degree, longitudinal modelling

JEL codes: I2, I23, I24

1. Introduction

A large empirical body of research has shown that human capital is a key factor in individual wellbeing and national economic growth (Blundell et al. 1999). Acknowledging this evidence, supranational institutions emphasize the role of education in their policy promotion agendas. Higher education is receiving a special attention. The OECD highlights the growing importance of higher education in modern societies and strictly monitors its expansion across member countries (OECD 2017a); in the perspective of promoting economies based on knowledge and innovation, the EU2020 strategy defined in 2010 set the goal of 40% individuals of age 25-34 with tertiary education by 2020. Although it is now widely recognized that the cognitive skills rather than mere school attainment are drivers of economic growth (Barro 2001, Hanushek and Woessmann 2008), it is also clear that formal education plays a fundamental role in developing skills (OECD 2017a).

To attain a large share of young people with tertiary education, higher education systems need to be both *inclusive* and *effective*. Inclusiveness can be defined in terms of the capability of providing wide opportunities along race, ethnicity, gender, income and social class lines to attain higher education degrees[1], and also responds to the aim of promoting equity and social cohesion. Yet, although related to equity (entailing balanced participation), the concept of inclusiveness also implies the broad participation of all social groups. Inclusiveness can be defined in terms of *access*, if the enrolment probability is high for all social groups, or *completion*, if the (timely) completion probability given enrolment is high for all social groups.[2] We define a system effective if the majority of the students who enroll eventually attain the degree and within a reasonable amount of time. In this sense, inclusiveness in completion can be conceived as 'effectiveness for all'.

These are not independent dimensions. In a synchronic perspective, inclusiveness at access and effectiveness may be negatively related: dealing with a less selected student body, more inclusive systems are likely to display poorer outcomes in terms of retention and time to degree. Instead, inclusiveness at access and in completion may be positively intertwined. Incentives to university enrolment differ across social groups because students from more advantaged backgrounds have more economic resources and higher aspirations, and being on average better performing, they have a higher probability of succeeding in degree completion (Breen and Goldthorpe 1997). If individuals' educational decisions are influenced by the experience of the older cohorts and in particular of those considered as 'peers', the presence of de-facto or perceived barriers to the successful progression could discourage prospective students of the less disadvantaged groups from entering higher education. Instead, a system that appears to work well for all is likely to foster the development of positive attitudes, favoring enrolment and active engagement in the university life.

Despite the educational expansion that has interested all advanced economies in the past decades and that has translated in the large growth of the share of young individuals with tertiary education, variability across countries is wide. At the OECD level, the average share for year 2018 is 44% (OECD 2018). At the top of the

[1] European Commission/EACEA/Eurydice (2018).

[2] Inclusiveness at access can be evaluated in terms of norms, according to whether there are entry barriers based on ability, prior school performance or type of upper secondary education. However, actual barriers to participation may exist even if no formal restrictions to enrolment are in place, as shown by the low participation rates of students from disadvantaged backgrounds in many countries with formally open systems.

ranking there is Korea (70%) followed by Canada (61%) and Japan (60%), and among European countries, Ireland (53%), Luxembourg (51%) and Switzerland (50%). Italy (27%) is placed at the bottom of EU and OECD's rankings, and displays a much lower proportion even when compared to Mediterranean countries with similar educational systems like France (44%), Spain (43%) and Portugal (34%). This critical situation results not only from a low share of young individuals entering higher education (De Santis et al. 2019), but also to high non-completion rates (Schnepf 2014, ANVUR, 2016). Dropout is a severe problem in Italian higher education; moreover, time to degree attainment is on average far above the institutional length of the degree courses (Almalaurea, 2016). This raises an issue of system effectiveness.

There is also an issue of equity. The Italian upper secondary education system channels students into school-types with different learning targets; however, all individuals with a 5-year diploma have access to university and no ability-based restrictions are in place. Despite this formal openness, transition rates to university are much higher for individuals coming from the academic track than for students from more vocationally oriented high-schools. Given the different educational content of the various school-types, this evidence is hardly surprising and could be considered an inevitable and perhaps even intended consequence of the student tracking policy. However, as demonstrated by an extensive literature, high-school choices in Italy are persistently characterized by an extremely strong social stratification, even in comparison to other tracked systems (Jackson, 2013). In this scenario, large differences in the participation and completion of higher education across tracks cannot be interpreted only in terms of different preparedness, but must be considered also in the perspective of inequality of opportunity in educational attainment.

Against this background, in this work we analyze student academic careers in Italy, a country that despite having an open educational system in terms of access barriers to higher education, is still lagging behind most other developed countries in the share of the population with higher education, even in the most recent cohorts.

When looking at national-level aggregate data on the higher education incoming student population, we may unambiguously conclude that the Italian university system is *not* becoming more inclusive at access. The overall transition rates from upper secondary to tertiary education were rather stable around 56% at the national level between 2004 and 2010, then fell to 53% in 2015 and only slightly increased since then.[3] The decline has interested all school types, but to a larger extent the technical and vocational tracks. As a result, the composition of the newly matriculated population has changed markedly over time, being increasingly represented by students coming from the academic track (see Figure 2).

While there are no signs of increasing inclusiveness at access, there is evidence of increasing effectiveness: in the past decade, aggregate dropout rates have been steadily decreasing and time to graduation seems to be reducing (ANVUR 2016). These improvements might be entailed by organizational changes, better teaching practices, widened tutoring, or simply by lowered educational targets. In principle, the observed changes in educational outcomes could also be due to changes in student attitudes, related to cultural factors or perception of the labor market needs. However, there is also a competing mechanism, related to the (self-) selection process into university. Since students have become more positively selected, at the aggregate level we expect

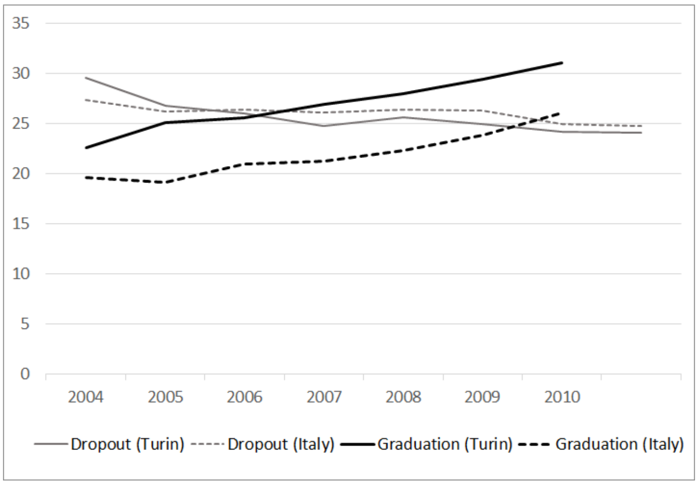
[3] Official reports of the National University Evaluation System (ANVUR 2016 and 2018).

more favorable outcomes in terms of retention and time to degree, even in the absence of behavioral changes on part of any of the actors involved, students or university institutions. To understand how academic careers are shaped across students' background and to analyze recent changes over time, in-depth analyzes of student trajectories in higher education are needed.

Our empirical analysis is based on the case-study of the University of Torino, a large public institution delivering degree courses in most fields of study in the North West of Italy. This is the larger of the two universities in Torino[4] – a large city in the North-West of Italy with a longstanding industrial tradition, partially reconverted to the tertiary sector from the 1980s – offering degree courses at the BA, MA and PhD level in most disciplines. This institution occupies a middle-high position in national university rankings,[5] has aggregate level dropout probabilities similar to the average national level and slightly higher timely completion rates (Figure 1).

We exploit the administrative longitudinal data released by the national university student registry. Our data archive contains the entire careers of all students first enrolled at the University of Torino between years 2004 and 2015, recording degree attainments, system dropout, degree changes and transfers to other institutions. We analyze the careers of students enrolled in bachelor (BA) programs up to either BA completion or dropout, with flexible discrete-time competing risks models, in order to evaluate the determinants of successful vs. less successful trajectories and assess the extent to which trajectories differ across previous schooling backgrounds. We then analyze trends over time. Consistently with national-level aggregate evidence, we find a moderate increase in both retention and timely degree probabilities. To shed some light on the drivers of this improvement, we decompose the observed changes into three distinct components, related to the composition of the student body, the choice of the field of study and a residual component that we may interpret in terms of 'individual behavior' after enrolment.

Figure 1. Share of students dropping out and graduating within 3 academic years, by cohort



Source: Turin, own calculations from ANS data. Italy, ANVUR (2016)

[4] The other university is the Polytechnic, delivering degrees in Engineering and Architecture. There are no major private higher education institutions in the region.
[5] 37° percentile according to the ranking of the "Il Sole 24 Ore", 2016.

The novel contribution of this paper to the existing literature is twofold. From a substantive perspective, it aims at contributing to a deeper understanding of the ongoing processes underlying student academic careers in a European country that despite a recent substantial increase, still displays a very low share of young individuals with higher education. The case of Italy is of particular interest because while in general countries with low entry barriers to tertiary education tend to have higher shares of graduates (OECD 2018), Italy combines a situation with low formal barriers and few graduates. From a methodological point of view, we relate to a literature analyzing academic trajectories with administrative data within the framework of survival modelling. With the aim of providing an accurate depiction of university trajectories of students with different characteristics, we propose a flexible discrete time competing risks approach that overcomes some of the limitations of conventional competing risks models (Scott and Kennedy 2005) employed in the existing literature. Firstly, following Desjardins et al. (2002), we allow for time-specific model parameters. Secondly we allow different choice options across academic years, and distinguish between events of interest determining an exit from the state of interest (dropout and graduation) and events of interest that we wish to model but not consider as exit states (change of degree program). This flexibility comes at the price of a large number of parameters; however, this is not a major issue with large sample size. To obtain meaningful interpretation of the results, we then simulate educational careers for different individual types and estimate Cumulative Incidence Functions and graduation profiles given first year performance.

The remainder of the paper is organized as follows. In section 2 we describe the Italian educational system and review the existing literature on higher education outcomes. In section 3 we describe data and methods. Results are illustrated in section 4. In section 5 we discuss contextual factors potentially driving the observed time trends: labor market and university institutional changes. Section 6 is devoted to conclusions.

2. The Italian context

2.1 Educational system

In Italy, children enter formal schooling at 6 and must attend school up to age 16. Compulsory education is comprehensive up to grade eight, with two cycles: a five-years primary school and a three-years lower secondary school. Upper secondary school offers a variety of educational programs, differing in content and learning targets, and broadly grouped into lyceums (academic track), technical schools and vocational schools. Lyceums may be further divided into *traditional lyceums* (the classical lyceum, with a focus on humanities, and the scientific lyceum, with more emphasis on math and sciences) and *other lyceums* considered less prestigious (linguistic, socio-pedagogical, artistic). *Technical schools* provide academic education together with some job-oriented instruction. *Vocational schools* are academically less demanding, and put more emphasis on the training for low qualification technical jobs. There are no ability-based access restrictions to the different school types. At the end of high school, students take a high-stakes national (non-standardized) examination that determines their exit mark.

Despite the strongly stratified character of upper secondary education, higher education is formally open: all students with a 5-year high school diploma have access to university (although for few majors, enrolment is limited and regulated by admission tests), regardless of prior performance. However, transition rates vary markedly across tracks. Excluding the law and medical schools and few other degree-programs still lasting 5-6 years, since the implementation of the EU *Bologna Process* in 2001 the system has been organized into a 3-year bachelor program followed by a 2-year master program. Students choose a degree program focusing on a field of study (e.g. physics, philosophy or economics) and have limited leeway in choosing courses and exams. There are generally no limits to time-to-degree. Subject to some restrictions in their sequence, students take exams after the end of courses or whenever they decide to do so; if they fail an exam (or get a poor grade), they are allowed to repeat it until they pass (or obtain a result that they consider satisfactory). Consequently, many students largely exceed the institutional length of the degrees.[6]

In the Italian higher educational system there is no formal divide between universities and polytechnic institutions. Higher education is mainly composed by public institutions but there are also few very prestigious private ones. University qualifications have all the same "legal value". Hence, although different studies show that the reputation of a university is relevant in attracting students (Agasisti 2009, Cattaneo *et al.* 2016), the role played by reputation is less important in Italy than in other countries with more differentiated systems.

Tuition costs in public universities are far lower than in most Anglo-Saxon countries, but are above the EU average (OECD 2013). University fees depend heavily on household income, and very low-income students are totally exempted[7]. However, financial aid in the form of grants is limited.[8] Some scholarships are provided by regional authorities to low-income students fulfilling mild conditions on credit progression and grade point average, but since supply is subject to budget constraints, full coverage of those meeting the requirements has not always been assured. Overall, although the direct costs of education are generally low for disadvantaged students, living expenses (and foregone earnings) could prevent some young individuals to engage in tertiary education.

2.2 Evidence on schooling choices and student academic careers

There is a huge sociological and economic literature on educational choices focusing on the role of social class (Boudon 1974, Collins 1979, Bourdieu and Passeron, 1990; Breen and Goldthorpe 1997, Barone et al. 2018), family income and credit constraints (Carneiro and Heckman 2002, Dynarsky 2003), risk aversion (Checchi et al. 2014). These studies provide theoretical explanations of the overwhelming evidence that more prestigious educational choices including university enrolment are strongly influenced by family background.

The literature on student academic careers upon enrolment mainly focuses on the study of the determinants of university dropout and timely degree completion. A systematic review of the personal and institutional factors affecting dropout is in Larsen et al (2103). Tinto (1975) focuses on the role of academic and social integration. Smith and Naylor (2001) and Arulampalam *et al.* (2005) highlight the role of individual's prior

[6] Average time to completion for 3-years programs was 4.6 years in 2014 (Almalaurea 2016)
[7] There is some variability across institutions. Substantial differences exist along the North-South divide: average annual fees per student are approximately 30% higher in the North (ANVUR 2016). Since internal student mobility is largely from South to North (Enea 2018) there is no evidence of self-selection driven by lower tuition fees.
[8] In Italy 19% of the students benefitted of public financial aid in 2010-11 (OECD, 2013).

performance as an indicator of preparedness. Due to their better academic preparation, students from advantaged backgrounds experience higher retention and completion probabilities, although they often display better outcomes also net of prior schooling (Ishitani 2006, Vignoles and Powdthavee 2009, Bowen et al. 2009). Manski (1989) explained university dropout as a rational behavior: due to incomplete knowledge of learning targets and their actual chances to attain the degree, individuals may first enroll and then reevaluate their choices once they have acquired more information. In this perspective, Stinebrickner and Stinebrickner (2013), show that students update their beliefs about their own ability over time and that this process plays a decisive role in dropout decisions. Other push factors may be related to intervening financial constraints (Stinebrickner and Stinebrickner 2008). In this perspective, financial aid plays a role in favoring study progression (Dynarski 2003, Glocker 2011). In favorable labor market conditions pull factors may also operate, as students might be induced to accept good job offers and leave university. Some scholars focus on system selection. For example, analyzing data for Belgium Flanders, a country where higher education has low admission standards and strong ex-post selection, Declerq and Verboven (2018) argue that moderate admission standards may reduce unsuccessful careers without decreasing degree completion. Instead, Denning et al. (2019) argue that the recent increase in college completion rates in the US is due to reducing standards to degree receipt. Time to degree is the object of another line of research; studies on the United States show that time to degree has increased markedly over time outside the most selective universities and that this is due to a rising share of students in employment (Bound et al. 2012).

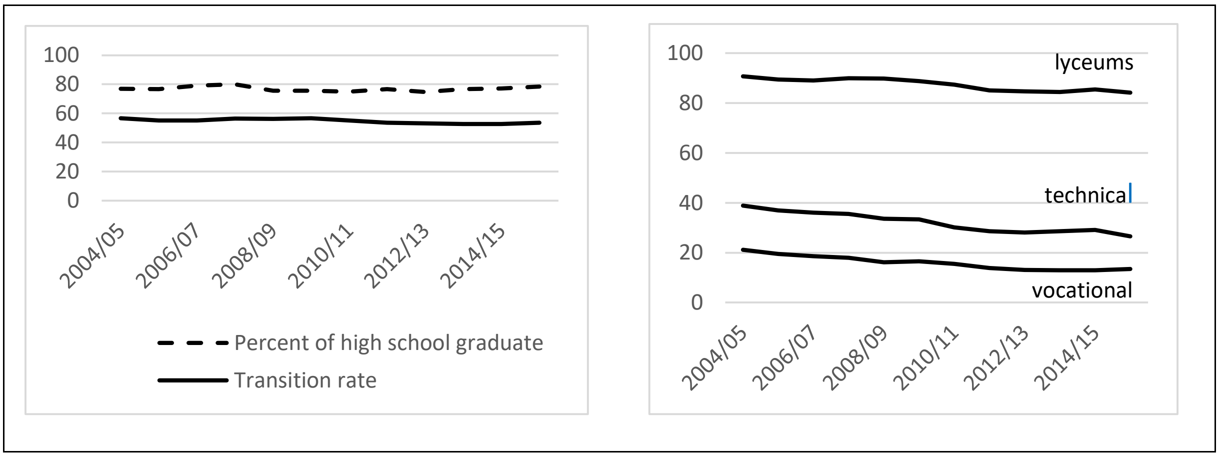
Within this scenario, Italy stands as a country with particularly large inequalities in upper secondary school choices and access to tertiary education (Jackson, 2013). Despite the absence on formal barriers to track choice and access to university, the Italian educational system is flawed by strong socioeconomic inequalities (Cobalti and Schizzerotto 1993; Contini and Scagni 2013). Social background critically influences students' high school choices (Gambetta 1987; Schizzerotto and Barone 2006). Even if inequalities in access to upper secondary education have witnessed a consistent reduction and a moderate increase in the share of students enrolling to the academic track[9], inequalities in track choices have not changed much over time (Panichella and Triventi 2014). Horizontal segregation in high-school has strong consequences on inequalities on university enrolment, as the transition rate to tertiary education varies largely across tracks (see Figure 2). Overall, there is evidence of increasing participation to higher education and slightly decreasing inequalities up to the 2000s (Argentin and Triventi 2011, Guetto and Vergolini 2017), but in the most recent decade, probably due to the economic crisis, transition rates have been declining and differences across high school tracks have increased, determining a change in the composition of the enrolled population.

Research on student academic careers has been limited by the lack of appropriate longitudinal data at the national level. For this reason, the existing literature on university dropout has been largely based on survey data, and in particular on a retrospective survey data on high school graduates (e.g. Di Pietro and Cutillo 2008; Cingano and Cipollone 2007, Contini *et al.* 2018)[10]. This literature reports substantial differentials related to

[9] Up to the beginning of the 2000s (ANVUR 2016)
[10] Survey of Upper Secondary School Graduates (ISTAT, 2015).

family background and prior schooling. Other studies (Cappellari and Lucifora 2009) focus on the '3+2' reform in 2001, and show that it led to a small reduction of the dropout probability. Some recent studies are based on universities administrative data[11]. Focusing on different case-studies, Clerici *et al.* (2014) highlight that determinants of academic careers differ across the fields of study and Carrieri et al (2015) find that a stronger selection at entrance considerably reduces dropout risks. Indirect evidence of the role of family income is provided by a small literature analyzing the impact of various forms of scholarships and financial aid for low-income students, showing that income support favors study progression and degree completion (Mealli and Rampichini 2012, Azzolini et al 2018, Vergolini and Zanini 2015). Only few contributions focus on time to degree. Aina *et al.* (2011) highlight the role of individual and family factors and find that weak labor market conditions tend to delay degree completion. Garibaldi *et al.* (2012) find that time to degree is negatively related to tuition costs. Contini et al. (2018) show that timely completion is a more stringent outcome than retention and depends heavily on prior scholastic performance.

Figure 2. Transition rates from upper secondary to tertiary education



Source: Own elaboration from data reported by ANVUR (2016)

3. Data and methods

3.1 Data

We exploit administrative data provided by the Ministry of Education 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 a bachelor's program (BA) in the academic years 2004/05-2013/14 (approximately 90,000 individuals) up to 2015. As mentioned above, each institution has access only to its own micro-data. Hence, if the aim is to study student academic careers from a system-level perspective one major limitation is that it is generally not possible to distinguish between change of institution and withdrawal from higher education altogether. To overcome this problem and identify dropout correctly, we obtained a special data release including the career

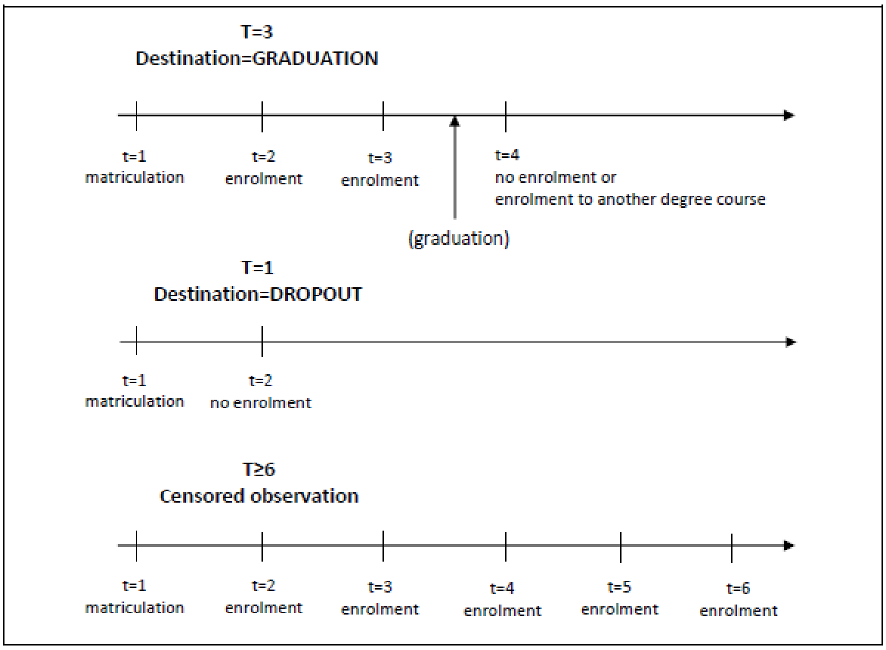
[11] Although a national archive collecting and harmonizing the data of all universities exists, in general institutions have access only to their own data.

segments occurred after transfers to any other national higher education institution. The administrative data provides full information on the students' progression (including exam transcripts and credits earned), degree changes, change of institution, timing of degree attainment or withdrawal, as well as demographic information on individuals (gender, age at first enrolment, place of birth) and information on previous schooling (type of high school and final examination marks).

Academic trajectories

We say that a student remains in the system for $T=t$ years if she enrolls for t subsequent academic years and does not re-enroll (at the University of Torino or at any other national higher education institution) in year t+1. If she attains the degree between t and t+1, we say that she graduates at year t (she attains the degree after being enrolled t years). If she withdraws before completion, we say that she drops out a t. Examples of student trajectories, time in the university system T and destination states are shown in Figure 3. Enrolment spells following a first dropout are not analyzed in this paper. However, the share of dropout students re-enrolling within a few years is low (around 10%) and only a minority of them eventually attain the degree.

Figure 3. Time in the system and exit states



Explanatory variables

We include individual variables capturing demographic characteristics and past schooling history: gender, age at matriculation, type of high school, final grade in the high school exam ('maturità' in Italian), as well as the field of study of the university degree course. Age at matriculation enters the models in a categorical form: 18-19, 20-21, 22 and more. 18-19 is the modal category and includes regular students, i.e. those who enroll in university immediately after the end of high school. The 20-21 age group includes students enrolling with 1-2 years of delay, possibly due to late high school enrolment, prior grade repetitions, or gap years. Students in the

22+ category experienced major delays (e.g. returns after significant spells in the labor market). We classify high schools into four broad types: traditional lyceums, other lyceums, technical and vocational schools. As a proxy of academic ability, we consider the *high school final grade*, a quantitative variable ranging between 60 and 100. The field of study – time varying, as students can change degree at the beginning of each academic year – are categorized into five broad groups: Health, (Hard) Sciences, Economics, Humanities and Social Sciences. In Table 1 we report descriptive statistics summarizing the explanatory variables and some outcomes, by matriculation cohort. The share of regular students and students coming from lyceums have increased over time (conversely, the share of students from technical and vocational high schools has decreased sharply)[12]. In parallel, both the retention and the graduation probabilities have increased.

Table 1. Explanatory variables at matriculation and selected outcomes, by matriculation cohort

	2004	2005	2006	2007	2008	2009	2010	2011	2012
Explanatory variables									
Age 18-19	67.5%	69.0%	70.8%	71.5%	71.5%	71.0%	71.0%	69.7%	70.1%
Age 20-21	18.0%	17.6%	17.0%	16.9%	17.9%	19.1%	19.7%	21.8%	22.5%
Age ≥22	14.5%	13.4%	12.3%	11.5%	10.6%	9.8%	9.3%	8.5%	7.4%
Female	60.9%	62.4%	62.8%	60.9%	63.1%	62.1%	61.7%	61.4%	62.2%
Trad. Lyceum	47.1%	48.5%	51.7%	52.5%	55.5%	56.7%	57.3%	58.8%	59.3%
Other Lyceum	8.8%	9.6%	9.3%	11.0%	12.8%	12.5%	13.0%	13.3%	14.9%
Technical	23.3%	20.3%	19.6%	19.9%	18.1%	17.7%	16.3%	16.1%	15.1%
Vocational	17.2%	17.5%	15.7%	12.1%	10.3%	10.0%	9.8%	8.3%	7.6%
Average HS mark	79.0	78.8	78.9	77.7	77.7	76.6	77.0	77.2	77.1
Health	9.1%	10.5%	10.8%	10.8%	10.5%	11.5%	10.9%	11.1%	10.5%
Science	20.0%	19.4%	19.6%	21.9%	21.2%	23.3%	24.0%	25.2%	24.6%
Economics	18.3%	14.8%	14.7%	13.5%	13.3%	14.7%	14.8%	14.4%	13.5%
Humanities	22.7%	25.1%	24.9%	24.3%	26.7%	23.3%	23.9%	24.7%	27.1%
Social Science	18.5%	18.6%	18.7%	18.5%	18.2%	17.3%	16.6%	13.4%	14.5%
Selected outcomes									
% dropout Y1	19.0%	15.9%	15.7%	15.2%	15.3%	14.8%	14.6%	14.9%	12.9%
% dropout Y1-2	25.4%	22.6%	22.4%	21.3%	22.0%	20.9%	20.4%	20.4%	18.1%
% dropout Y1-3	29.6%	26.8%	25.9%	24.7%	25.6%	24.9%	24.2%	24.0%	-
% dropout Y1-4	32.9%	29.8%	29.1%	27.8%	28.5%	27.6%	26.8%	-	-
% dropout Y1-5	35.3%	32.5%	31.4%	30.4%	31.0%	30.2%	-	-	-
% dropout Y1-6	37.6%	34.4%	34.0%	32.6%	33.3%	-	-	-	-
% change Y1	7.0%	8.8%	8.6%	9.4%	12.0%	11.2%	11.6%	11.4%	11.3%
% change Y1-2	8.9%	10.8%	11.0%	12.4%	14.6%	13.9%	14.0%	13.6%	13.4%
% change Y1-3	9.8%	11.7%	12.3%	13.8%	15.7%	15.2%	15.5%	14.7%	-
% graduation Y3	22.6%	25.0%	25.6%	26.9%	28.0%	29.4%	31.1%	30.3%	-
% graduation Y3-4	37.2%	40.0%	41.3%	43.2%	43.9%	45.3%	47.0%	-	-
% graduation Y3-5	45.1%	48.2%	48.8%	51.3%	51.3%	52.7%	-	-	-
% graduation Y3-6	49.3%	52.2%	53.0%	55.0%	54.5%	-	-	-	-
N	9242	8650	8509	8361	8471	8820	9002	8447	8667

[12] Note that the share of students in lyceums (both traditional and others) has increased in the past decades, but has only changed marginally in the period under study (from 43% in 2004 to 45% in 2013).

3.2 Modelling trajectories

A large body of work in the existing literature on university dropout uses logit or probit modelling to analyze whether students withdraw the system within a given amount of time after enrolment (Naylor and Smith 2001, Vignoles et al 2009, Chies et al 2014), sometimes in a bivariate framework to analyze enrolment and dropout (Ghignoni 2017, Di Pietro and Cutillo 2008), or credits earned and dropout (Belloc et al. 2011). Event history is a longitudinal analytic technique that is particularly well suited to study the temporal nature of student academic careers (Scott and Kennedy 2005, DesJardins 2006). In this perspective, many contributions analyze dropout with hazard models (e.g. Arumpalam et al. 2004, Aina 2013). Other contributions analyze the entire trajectories in university, including dropout, stopout, degree change and completion, transition to second level degrees using discrete time competing risks models (DesJardins et al 2002, Ishitani 2006, Lassibille et al. 2009, Gloecker 2011, Ortiz and Dehon 2013, Clerici et al. 2014, Enea 2018).[13] In parallel, a different analytic tradition employs discrete choice models to analyze individuals' educational choices within a behavioral utility maximization framework (eg. Cameron, Heckman 2001, Declercq and Verboven 2018).[14] Following the strand of the literature using event history analysis, we investigate the educational outcomes of the different cohorts of BA degree entrants using a flexible 'step-by-step' longitudinal approach that mimics the estimation of discrete-time hazard functions within a competing risk (CR) framework.

Event history analysis builds on survival analysis. Survival analysis aims at modelling time elapsed in a given state of a discrete-state process. When the event "exit" can occur only at isolated time points (say, at times T =1, 2, 3…), we refer to discrete time models. In this context, the functions of main interest are the survival function $S(j) = P(T > j)$, the probability that exit will not occur up to time $T = j$, and the hazard function $h(j) = P(T = j|T > j - 1)$, the probability of exiting the state at time j given survival up to time $j - 1$. Common specifications including explanatory variables are the discrete-time version of the proportional hazard model or, within the logit modelling framework, the proportional odds model (Jenkins, 2005).

Competing risks (CR) models are extensions of survival models when 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 the exit event. Our case study fits this situation. Students remain in the university system until they either drop out (one possible destination) or attain the degree (alternative possible destination). Individuals are censored if they are still enrolled at the end of the observation window. More specifically, CR modelling is centred on the *cause-specific hazards* (Collett 2015)[15], defined as:

[13] There is also a recent literature aimed at early prediction of student dropout with machine learning techniques like neural networks (Delen 2010, Berens et al. 2018), but this is out of the scope of the present work.

[14] Despite the different framing, there are similarities between discrete time CR models and discrete choice models for panel data: in fact, they both often end up using different versions of multinomial logit models. A very general dynamic discrete choice model that seems to encompass all (although survival/CR models are never mentioned) has been proposed by Cameron, Heckman (2001, pg. 456-69). Our own model can be seen as a special case of theirs, with time varying coefficients, a particular set of choice options and no unobserved heterogeneity.

[15] The competing risks literature often refers to (1) as hazard rates of the latent variables $T^{(v)}$ (failure time until event v), under the assumption that latent variables are independent (Cox and Oakes, 1990; Collett 2015). This interpretation is not of interest in the present context as it is not needed for identifiability of (1).

$$h_v(j) = P(T = j, D = v | T > j - 1) \qquad + \qquad (1)$$

the probability of exiting at time j towards destination $D = v$ given survival up to $j - 1$. The unconditional probability that a given individual exits the system at time j towards destination v is:

$$P(T = j, D = v) = h_v(j) \prod_{k=1}^{j-1} (1 - h_v(k)) \qquad (2)$$

i.e. 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.

The literature on CR models has been largely developed within the framework of continuous time models. Discrete time models can be specified with appropriate multinomial logit modelling, applied to person-time data (Jenkins 2005, Scott and Kennedy 2005, Ambrogi et al. 2009). As described in Figure 2, time is measured as the number of subsequent enrolments (students are required to enroll at the beginning of each year) before the occurrence of one of the two possible events defining the exit from the system: dropout or graduation. Considering matriculation cohorts separately, at each academic year, we estimate multinomial logit models to derive the probability of all possible options (Table 2). At year 1, we estimate the probability of continuing in the chosen degree program, versus switching to another degree, versus dropping out. Transfers to other institutions (rare) are included in the degree change option, and we distinguish between changes to 3 year programs and to 5/6 year programs. If the student drops out, she falls out of the risk set. At year 2, we consider only students still in the risk set and model choices within the same set of options. From year 3 graduation is also possible. Graduation, like dropout, is an exit state.

Table 2. Options and exit states by academic year after enrolment

Academic year	Options	Exit states
1	CONTINUE, DROPOUT, CHANGE TO OTHER 3Y, CHANGE TO 5/6Y	DROPOUT
2	CONTINUE, DROPOUT, CHANGE TO OTHER 3Y, CHANGE TO 5/6Y	DROPOUT
3+	CONTINUE, DROPOUT, CHANGE TO OTHER 3Y, CHANGE TO 5/6Y GRADUATE	DROPOUT GRADUATE

NOTE. Transfers to other institutions (rare) are treated as degree changes.

At each step, and for each matriculation cohort, we include the following time-invariant explanatory variables: gender, age at first matriculation, high school track (traditional lyceums, other lyceums, technical schools, vocational schools) and high-school final examination mark, as well as the field of study at first enrolment.[16] We also include time-varying information on previous degree program switches.

Formally, the CR model, defined as a multinomial step-by-step model, may be written as:

[16] Fields of study are grouped into the following categories: health, economics, social sciences, scientific, humanities. Due to institutional changes occurred within the period under study, law is excluded from the analyzes.

$$P(T = j, D = v | T > j-1) = \frac{\exp(\beta_{vj} X_{j-1})}{1+\sum_v \exp(\beta_{vj} X_{j-1})} \qquad \forall v \in set\ of\ options \neq 'continue'$$

$$P(T > j | T > j-1) = \frac{1}{1+\sum_v \exp(\beta_{vj} X_{j-1})} \qquad (3)$$

where $T > j$ means that at time j the students continues in the original degree course and $T = j, D = v$ represents exit at time j towards one the other possible options v. If the destination belongs to the set of exit states, the student falls out of the risk set thereafter.

The advantages of step-by-step estimation over conventional competing-risk (CR) models are related to its greater flexibility. First, model parameters β_{vj} are time-specific: hence, we relax the usual assumption in CR models that explanatory variables have the same effect at all years after enrolment. This assumption is often not corroborated by empirical evidence: DesJardins et al. (2002) find that the independent variables have effects that are different over the course of a student's academic careers[17]. We find ourselves that some explanatory variables are empirically relevant in first year, while have no effect or a larger effect in subsequent ones (see Appendix). Second, we can easily distinguish between events of interest determining an exit from the state of interest (*final destinations*) and events of interest that we want to model, but not consider as exit states. In conventional CR modelling, once individuals are assigned a destination, they fall out of the risk set. Consider the event 'change of degree program'. This event is not rare: approximately 10% of the students make a change during the first two years after matriculation. We are interested in modelling degree changes, because they are informative on individuals' attitudes and help predicting future outcomes. Horizontal moves (change to another BA program) usually come along with dissatisfaction over the current degree program or insufficient academic skills to continue, while upward moves (degree changes from BA to 5- or 6- year programs) usually demonstrate high aspirations and/or high ability[18]. We wish to model their occurrence without defining them as exit states. This is not possible within the conventional CR framework. Instead, with step-by-step analysis we are able to make degree changes endogenous, by including them as possible transitory outcomes within the process under study, and as explanatory variables upon occurrence. Once individuals experience the change, we do not force them out of the risk set, so we are still able to analyze their trajectories up to dropout or graduation (or censoring). Third, although graduation is possible only after the institutional length of the degree has elapsed, conventional CR analyzes do not allow including an exit state only at a certain point. Imposing this condition is straightforward with our approach. A limitation of our procedure is the large number of parameters, that might yield to inefficiency of the estimates. However, our samples should be large enough (approximately 9,000 individuals per matriculation cohort) to ensure the delivery of reliable estimates of the probabilities of the outcomes of interest.

For each matriculation cohort we then estimate the *Cumulative Incidence Functions (CIF)*, a commonly employed representation in CR modelling, that allows grasping timing and probability of the destination

---

[17] Desjardins et al. (2002) propose a parametric specification that allows for time-varying coefficients.

[18] The majority of these upward moves are from 3-year degrees like biology or chemistry to 6-year medical school. This is because entrance to medical school is subject to a national level examination and regulated by strict *numerus clausus*. Students not admitted in year 1 often matriculate into a BA degree program and try again the next year.

events, and highlighting differences in the trajectories across individual profiles in a better way than when focusing on hazard functions' coefficients. In our contest, *CIF* are defined as the probability that a newly enrolled student with a specific profile of explanatory variables will exit the university system by dropping out or graduating within 1-6-years:

$$CIF_D(j) = P(T \leq j, D = d)$$
$$CIF_G(j) = P(T \leq j, D = g) \tag{4}$$

The destination state can be either D=dropout (d) or G= graduation (g). The complementary function:

$$P(T > j) = 1 - P(T \leq j, D = d) - P(T \leq j, D = g) \tag{5}$$

is the probability of being still enrolled after j years.[19]

Simulations

Ideally, the *CIF* could be estimated by multiplying the step-by-step predicted probabilities of the single outcomes giving rise to each specific sequence consistent with the outcome of interest, and then adding up the probabilities of all these sequences. However, this approach is cumbersome when destination states are fewer than the available options, because the number of possible trajectories may rapidly increase with the time points considered.[20] A straightforward alternative is performing Monte Carlo simulations of a large number of educational careers – using the predicted probabilities of occupying a specific state at each step – and then computing the share of simulated cases experiencing the desired outcome. For large N, this approach is equivalent to the analytic computation of probabilities. We compute the *CIF* 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. More specifically, when we analyze specific profiles we use *N*=20000, when we focus on the overall population we expand the observed samples by a multiplier factor *N*=1000.[21]

More generally, our strategy allows to easily estimating the probability of any specific trajectory or any complex outcome of interest. For example, we could estimate the probability that an individual with given characteristics will change degree after year 1, continue in the same degree program after year 2 and then drop out after year 3; or instead the probability of graduating after 5 years after having experienced a degree change at some point in time. Alternatively, as we will do in section 4.2, we can estimate the probability that a student who attained a given number of credits in year 1 will eventually graduate within k years (we do this by including the number of credits earned in the first year in the choice models after each year, and evaluate the joint probability accordingly). Moreover, although in this paper we focus on single spells, we can easily handle more complex event history data including multiple spells (for example by considering re-enrolment after a drop out or enrolment into master programs).

[19] The estimation of the CIFs in parametric continuous time CR models is not straightforward and has been the object of debate (Austin and Fine 2017, Allison 2018). The debate does not seem to involve discrete time modelling. In this paper we directly estimate CIFs via simulation methods.
[20] For example, with our set of options, there are 9 possible sequences compatible with the outcome ($T = 3, D = d$) and 243 sequences with the outcome ($T > 6$).
[21] These values were determined after different trials, in order to reach substantial stability of the estimates.

Finally, we can easily decompose observed differences across subgroup or over time into changes due to different internal composition (explanatory variables) or different 'behavior' (parameters), by simulating trajectories of individuals with the characteristics of one group and the parameters of another group and comparing the 'counterfactual' outcomes. As explained below, we will analyze time trends with this strategy.

3.3 Analyzing time trends

To analyze time trends we focus on 4-year outcomes. Starting from the estimates of the step-by-step multinomial logit models for, we apply a Blinder-Oaxaca-like decomposition of the observed changes in the population-level *CIF* estimates at time 4, i.e. on the probability of dropping out within 4 years and the probability of graduating within 4 years. We choose this threshold, because 4 years is a reasonable amount of time for graduating in a 3-year program, given that the share of students graduating within the institutional time is actually quite low. 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), possibly entailed by organizational changes or different teaching practices or by changes in student attitudes.[22]

In particular, we assume that the decision process of the individual who attained a high-school diploma first consists in choosing whether to enrol in university and then the field of study: this defines the composition of the enrolled population in each field of study. What happens next – university careers – is expression of the behavior of the actors involved: students in first place, and the university institutions. Since data on all students eligible for university enrolment are not available, we cannot study the enrolment process, but only account for the distribution of the characteristics of the higher education student population at matriculation $f(X|enrolled)$. We may then analyze the probability of choosing a specific field of study given enrolment $P(F|X, enrolled)$, and the probability of different academic careers' outcomes $P(Y|F, X, enrolled)$ given the field of study[23].

In the end, the outcome's Y distribution for each matriculation cohort may be expressed as follows:

$$P_C(Y) = \sum_F \sum_X \sum_{track} P_C(Y|F, X, track) P_C(F|X, track) P_C(X, track) \quad (6a)$$

$$P_C(Y|track) = \sum_F \sum_X P_C(Y|F, X, track) P_C(F|X, track) P_C(X|track) \quad (6b)$$

where *X* are time invariant explanatory variables, *F* the field of study at matriculation and *c* the cohort of matriculation. Expression (6a) refers to the entire student population, expression (6b) to each track separately. In this framework, we may study how university careers change over time across cohorts by decomposing the observed change into three parts: a first one related to changes in the composition of the student body, a second related to changes in the choice of the field of study and a third one related to changes in "behavior". Since

[22] A version of the Blinder-Oaxaca decomposition adapted for a binary outcome has been proposed by Fairlie (1999). Cameron and Heckman (2001) apply a similar decomposition to examine the black/white gap. Our decomposition follows the same logic, but it is applied to a 3-stages process.
[23] We consider here the choice of the field of study at matriculation. As described above, students may change degree course; the occurrence of this event is modeled within the CR modeling of students' academic careers.

this component is defined residually, rigorously speaking it can be interpreted as truly behavioral only conditional on all the relevant determinants being accounted for.

We apply this decomposition to the change between the outcomes of the two extreme cohorts for which we observe 4 years after matriculation, i.e. cohorts 2004 and 2010:

$$P_{222} - P_{111} = (P_{222} - P_{122}) + (P_{122} - P_{112}) + (P_{112} - P_{111}) \quad (7)$$

Subscript 2 represents cohort 2010 and subscript 1 represents cohort 2004, the first subscript refers to the composition of the enrolled, the second to the choice of the field and the third to behavior. In practice, this decomposition is performed by means of simulations. Each "counterfactual" probability is estimated by combining probabilities relative to different cohorts. For example, we estimate P_{122} by applying to the population of the first cohort (2004) the estimated parameters of the model for the choice of the field of study of the second cohort (2010), and then the parameters of the academic careers' models of the second cohort. Since what varies is only the student body composition, the first term in parenthesis refers to compositional changes; similarly, the second term refers to changes in the pattern of choice of the field and the third to changes in student trajectories given individual characteristics and field of study.[24]

4. Results

The empirical analyzes are organized as follows. In section 4.1, we examine the academic careers of students matriculated in 2008, the latest cohort for which we observe up to 6 academic years after enrolment. After estimating the step-by-step competing risks model, we derive the Cumulative Incidence Functions for different students with different profiles. Then, to gain understanding of the process underlying the differences across profiles, in section 4.2 we examine the number of credits earned in the first year after matriculation and analyze subsequent choices – continuation, dropout from university, change of degree program – as well as the probabilities to eventually graduate within six years. Finally, in section 4.3 we examine time trends in the dropout risk and graduation probabilities and analyze the extent to which these changes can be ascribed to behavioral changes, or instead if they depend on variations in the composition of the enrolled student population and in the different fields.

4.1 Students' academic careers

The estimates of the regression coefficients are shown in the Appendix (Figure A1). At each step, we find the effects usually reported in the literature: less dropout and higher graduation probabilities for students enrolled at younger age, for students coming from traditional lyceums (followed by other lyceums, technical schools and vocational schools) and with good prior scholastic performance. Gender differences are substantial (girls do better than boys), but reduce and often disappear once we control for prior schooling and fields of study. Different outcomes are observed also by field of study. Moreover, coefficients vary considerably over time,

[24] In principle, 6 decompositions are possible. We report only results relative to (7), as it seems to be the most meaningful one. The others are less salient because do not follow the hypothesized time ordering of the decision process.

being often larger in the first years. Due to strong previous selection, university participation and choice of institution are endogenous, thus our results are descriptive rather than necessarily causal (as argued in Carneiro and Heckman 2002, these estimates are likely to be conservative estimates of the true causal effects).

We observe that the risk of experiencing a degree change towards another 3-year program relative to continuing in the original one is larger for students from non-academic tracks and with low high school marks. Instead, the probability of changing towards a 5/6-year program (medical and law schools) is higher for students of more advantaged schooling backgrounds. Older students have a lower probability to make a degree change of any type. Students who have experienced a horizontal degree change are more likely to drop out in subsequent years and also to make another change and, not surprisingly, take more time to attain the degree.

The estimates of the Cumulative Incidence Functions for students of selected profiles defined by age at matriculation, school track and final high-school examination grade of the students matriculated in 2008 and enrolled in 3-year degree courses (BA) are depicted in Figure 4a-b[25]. Females and males are merged together (females do better than males, but when adding controls, the differences largely fade away). Figure 4a refers to individuals matriculated at age 18-19 (the regular age of end of high-school), Figure 4b to individuals matriculated at age 22-25. The solid line refers to students with a median high school mark, the two dashed lines to the 75th percentile ("high mark") or the 25th percentile ("low mark").

Both dropout probabilities and graduation probabilities vary markedly along these dimensions. Consider younger students from traditional lyceums as an example. Students with good marks in high school perform quite well: their probability of attaining the degree within 3 years is 54% and within 6 years is 82%. Their withdrawal probability within 6 years is 10%. The remaining 8% is the quota of students still enrolled after 6 years: in general, this is an indicator of system ineffectiveness.[26] As we move towards lower high school marks, the picture deteriorates sharply, as well as for other school types. Students from other lyceums and technical schools display similar patterns, and high mark students from these tracks behave similarly to medium mark students from traditional lyceums. Students from the vocational track are those with the poorest performance; if we exclude those with very high marks, within 6 years from enrolment the share of dropouts is larger than the share of graduates. The large majority of those delaying entrance to 22-25 years old leaves the system before degree completion, whatever the previous schooling. An intermediate pattern (not shown) applies to students enrolling at age 20-21.

Altogether, it seems that only the outcomes of very well performers from lyceums, enrolling with no delay after high school, may be considered "good" outcomes. Our estimates show that the issue of effectiveness involves students from all schooling backgrounds, including those from academic high schools with average prior performance levels.

[25] The reported CIFs refer to the average student with a specific age and schooling profile (where "average" is with respect to the field of study). In practice, this is obtained by taking a population of N=20000 individuals with a given profile, simulating the choice of the field of study and then simulating the trajectories given the fields of study. Observed and predicted probabilities at the population level are very similar, and reported in the Appendix, Table A2.

[26] However, this share also includes students who switched degree course from a 3- to a 5/6-year program, a group largely composed by well performing high school students from lyceums.

Figure 4a. CIF estimates for individuals enrolled in 3-year degree programs, by high school type and final examination mark, age at enrolment 19 (matriculation cohort 2008)

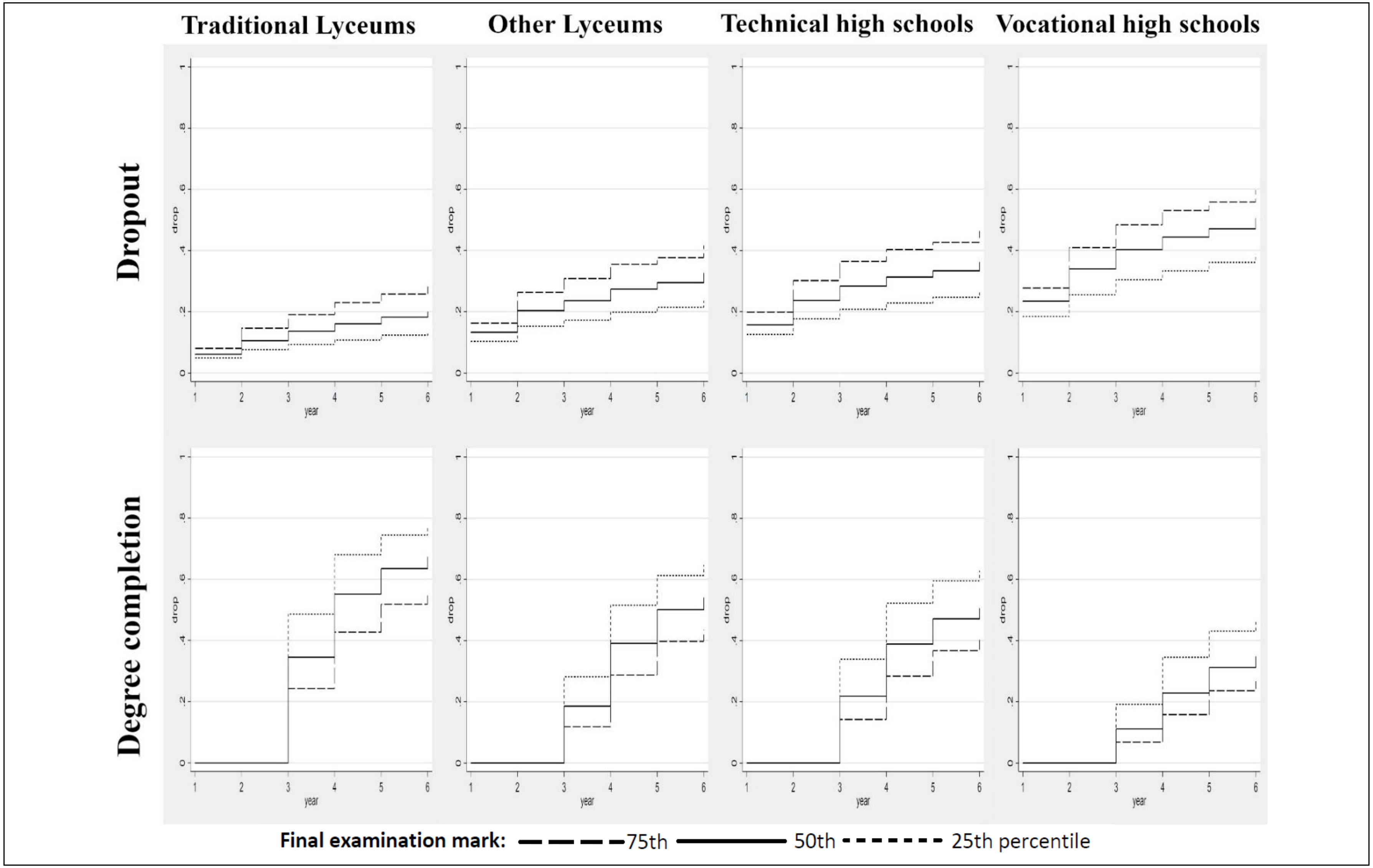
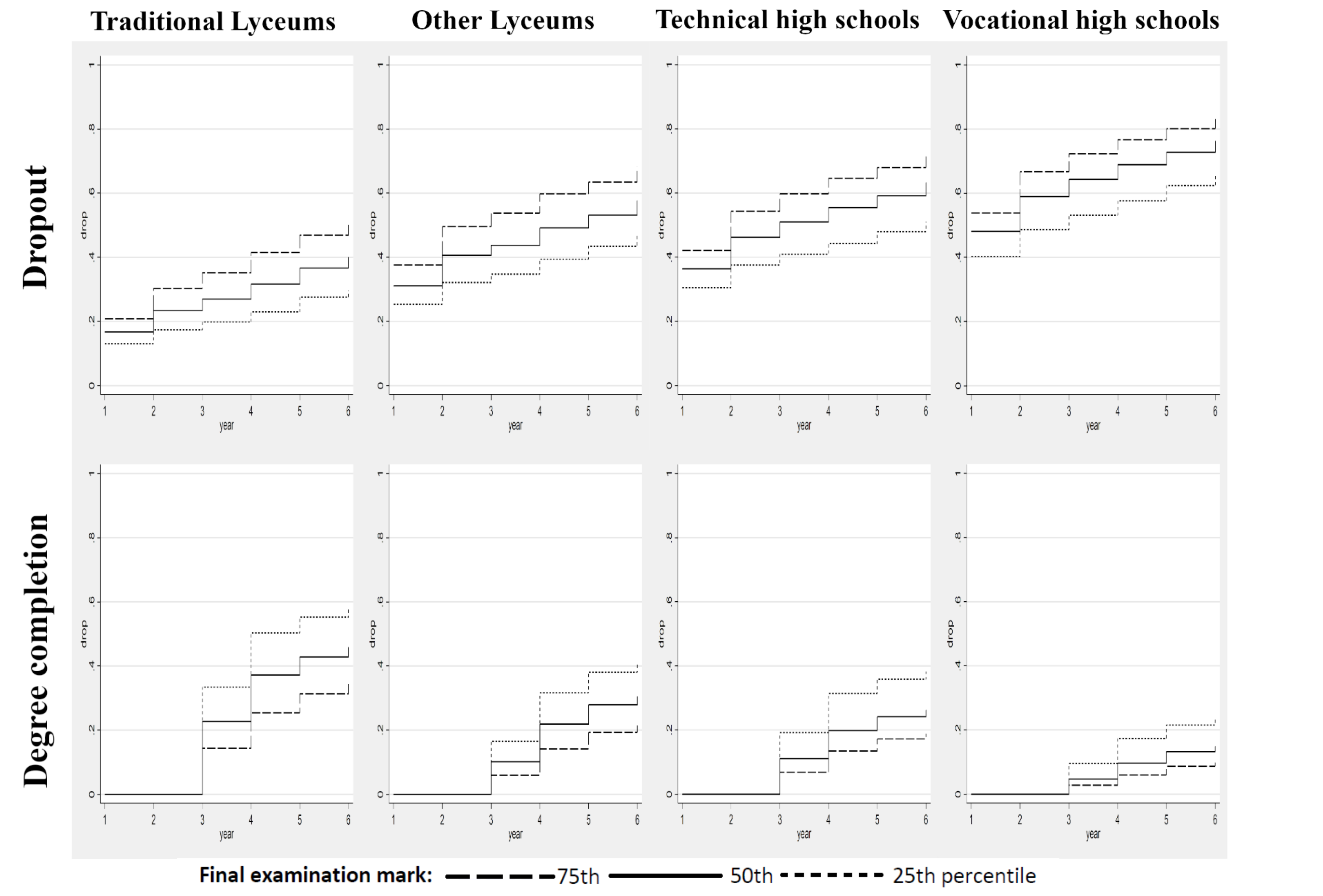


Figure 4b. CIF estimates for individuals enrolled in 3-year degree programs, by high school type and final examination mark, age at enrolment 22-25 (matriculation cohort 2008)



4.2 First year performance and future outcomes

We now analyze credits earned in the first academic year and how the subsequent pathways relate to this initial outcome. The scheduled number of credits per academic year is 60; we analyze the probability of attaining ≤20 credits (at most one third) and ≥ 40 credits (at least two thirds) of the total workload.

i) Credit attainment varies greatly across high school types (Table 3). We find traditional lyceums ranking at the top, other lyceums and technical schools in the middle (with similar results), and the vocational track at the bottom. The probability of earning <=20 credits (poor result) varies between 25% for traditional lyceums and 56% for vocational schools. At the opposite, more than 53% of the students from traditional lyceums earn at least 40 credits (good result), against 26-39% of those from the other tracks.[27]

Table 3. Credit attainment in the first academic year, by high school track (percentage points)

	Raw probabilities		Raw difference from Traditional lyceums		AME ref Traditional lyceums	
	≤ 20 credits	≥ 40 credits	≤ 20 credits	≥ 40 credits	≤ 20 credits	≥ 40 credits
Traditional Lyceum	25.46	53.14	-	-	-	-
Other lyceum	38.44	39.42	12.98	-13.72	12.86	-15.14
Technical	42.12	35.67	16.66	-17.47	12.21	-14.52
Vocational	56.35	26.47	30.89	-26.67	25.35	-22.23

NOTES. AME from logit model with gender, age at enrolment, final high-school mark and field of study as controls.

ii) How do students behave after the first academic year, given the number of credits earned? According to our estimates shown in Table 4, approximately 95% of the students earning at least 40 credits in year 1 continue in the same degree course in year 2, regardless of the high school background. Focusing on poor performing students (≤20 credits), we find that almost half (44%) do not re-enroll the following year and nearly 22% change degree. Students from traditional lyceums are much more likely to change degree program rather than dropping out. The differential across high-school tracks is large even after controlling for individual characteristics and the actual number of earned credits (see AME in the lower panel of Table 4). Qualitative research (Romito *et al.* 2019) provides an explanation to this evidence. Students from more privileged backgrounds tend to choose traditional lyceums as the natural route to university entry that is often taken for granted since childhood: when confronted with a failure, these students will adjust their choices in order to achieve this goal. Instead, students with less favorable backgrounds, whose educational plans take shape later on, are more inclined to give up. Interestingly, other things being equal, the higher the final high-school examination mark, the higher the probabilities to both drop out and (in particular) change degree program rather than continuing in the same program[28]. This suggests that students who are used to performing well are more likely to perceive a small number of earned credits as a failure, and react to this failure with an action: the most resilient ones change degree course, the more fragile withdraw.

---

[27] The final examination high-school mark is also relevant, showing that the number of credits earned in the first year depends heavily on previous preparedness.

[28] Full results of model estimation underlying Tables 3 and 4 are available upon request.

Table 4. Choices after the first academic year, by n° credits earned and high school type (in %)

		≤20 CREDITS				≥40 CREDITS			
		Drop	Change 3Y	Change 5/6Y	Cont	Drop	Change 3Y	Change 5/6Y	Cont
RAW PROB.	Trad Lyceum	28.89	28.11	3.81	39.19	0.05	3.22	0.94	95.79
	Other lyceum	48.62	18.89	0.69	31.80	0.00	4.35	0.00	95.66
	Technical	53.93	13.96	0.27	31.84	0.62	3.26	0.67	95.45
	Vocational	59.40	12.23	0.18	28.19	0.00	4.08	0.19	95.73
	TOT	44.19	20.02	1.73	34.06	0.11	3.37	0.74	95.78
AME	Trad Lyceum	-	-	-	-	-	-	-	-
	Other lyceum	11.98***	-4.93**	-3.11***	-3.94**	-0.00	0.01	-0.01***	-0.00
	Technical	15.52***	-7.59***	-2.54***	-5.40***	0.01	0.00	-0.00	-0.00
	Vocational	18.76***	-8.68***	-3.32***	-6.76***	-0.00	0.01	-0.01**	-0.00

NOTES: *sig al 5% **sig al 1% ***sign al 0.1%
AME= average marginal effects from multinomial logit models for students who earned ≤20 or ≥40 credits
Controls: gender, age at matriculation, high school mark, field of study, n° credits earned the first year.

iii) We now apply step-by-step modelling with the inclusion of the number of credit progression in the first year as an explanatory variable, to estimate the probability of eventually attaining the BA degree within 3-6 years after enrolment. In the upper panel of Table 5 we focus on those for those who earned <=20 credits and did not drop out immediately afterwards[29]. We estimate that only less than 25% complete the degree course within 6 years, with some differences across high school tracks.[30] In the lower panel of the same table we show graduation probabilities for the students who earned ≥40 credits in the first year. Overall, these students have a 56% probability to attain the degree on time and 86% attain the degree within 6 years from enrolment. Differences across tracks are relevant, in particular in the timing of degree completion.

Table 5. Graduation probabilities within 3-6 years, by credits in year 1 and high-school track (%)

STUDENTS WHO EARNED <=20 CREDITS (AND DID NOT DROPOUT AFTER YEAR 1)				
	3 academic years	4 academic years	5 academic years	6 academic years
Traditional Lyceum	1.4	13.2	22.5	28.5
Other Lyceum	0.6	8.7	17.3	22.5
Technical	0.8	8.7	16.6	20.9
Vocational	0.5	5.8	12.4	16.5
TOTAL	1.1	10.8	19.4	24.7

STUDENTS WHO EARNED >=40 CREDITS				
	3 academic years	4 academic years	5 academic years	6 academic years
Traditional Lyceum	62.0	81.8	87.8	89.9
Other Lyceum	44.2	68.2	77.9	81.1
Technical	51.9	72.5	80.1	82.7
Vocational	33.2	54.6	65.1	68.9
TOTAL	56.0	76.6	83.7	86.1

NOTE. Step-by-step models estimated on all students. Models for all academic years include all previous covariates, the number of earned credits in the first year and indicators of previous degree change

[29] While reverse causation might affect result in Table 3 (withdrawal could be the cause and not the consequence of few credit earning) this issue does not apply to results in Table 4, because we are considering only students who, despite their poor performance in the first year, decided not to leave university.
[30] Students eventually re-enrolling after a dropout are not included. In related work we estimated a 10% probability of re-enrolment after withdrawal, and among them, a 30% probability of attaining the degree.

In sum, the amount of credits earned in the first year heavily depends on the high school track, and subsequent student academic careers are strongly related to first year's credit attainment[31]. Although these results cannot be interpreted strictly in causal terms, they do provide evidence that first year performance is highly predictive of the students' chances of success in higher education completion, but in different manners across previous students' schooling backgrounds.

4.3 Trends over time. Compositional or behavioral changes?

As shown in Table 1, we observe declining dropout probabilities at all durations and steadily increasing graduation probabilities. However, these changes are not homogeneous across high school types. Consider the within 4-year dropout and graduation probabilities. In Figure 5 we report these probabilities for matriculation cohorts 2004-2010 (we do not observe 4 years after enrolment for later cohorts). Students from the vocational track do not benefit as the other subgroups. Their dropout rates – already very high in earlier cohorts (43%) – increase by more than 6 percentage points between matriculation cohorts 2004 and 2010, and the share of degree completion is stable around 25%. Instead, all other groups improve substantially in dropout and particularly in graduation probabilities. Recalling that they have become increasingly fewer in number (Figure 2), altogether these results suggest that students from the vocational track are being progressively marginalized from the higher education system.

Is this descriptive picture driven by behavioral changes, or is it due to changes in the composition of the high-school subgroups enrolling at university, or to the pattern of choice of the field of study?

The student composition has changed markedly over the observed time span (Table 1). The share of students coming from lyceums has increased from 57% to 73% between matriculation cohorts 2004 and 2010, and students first enrolled before age 21 increased from 87% to 93%. This tendency also holds by field of study, although the changes are marked for some fields (economics, scientific and health) and mild for others (humanities and socio-politics). Thus, we expect compositional changes to have contributed to the reduction of overall dropout rates and time to degree.

Figure 6 shows the decomposition of the total difference in the 4-year dropout and graduation probabilities between matriculation cohorts 2010 and 2004. Since the share of younger students and from lyceums has been increasing, on the whole the university student-body population has become more positively selected. The reduction of the dropout probability can be partly ascribed to these compositional changes and partly to changes in individual behavior. No role is played instead by changes in the pattern of choice of the field of study. Instead, the improvement in the graduation probability is almost entirely due to changes in behavior.

[31] As shown by Enea and Attanasio (2016), better performing students in terms of number of credits earned also tend to have higher grades.

Figure 5. Probability of dropping out or completing the degree within 4 years from enrolment, BA students, by high-school track and matriculation cohort

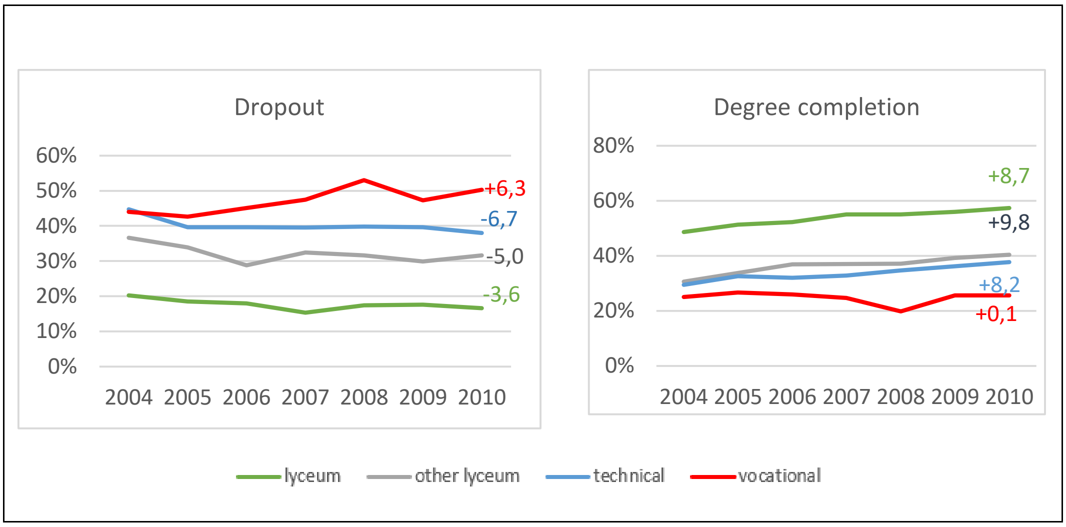
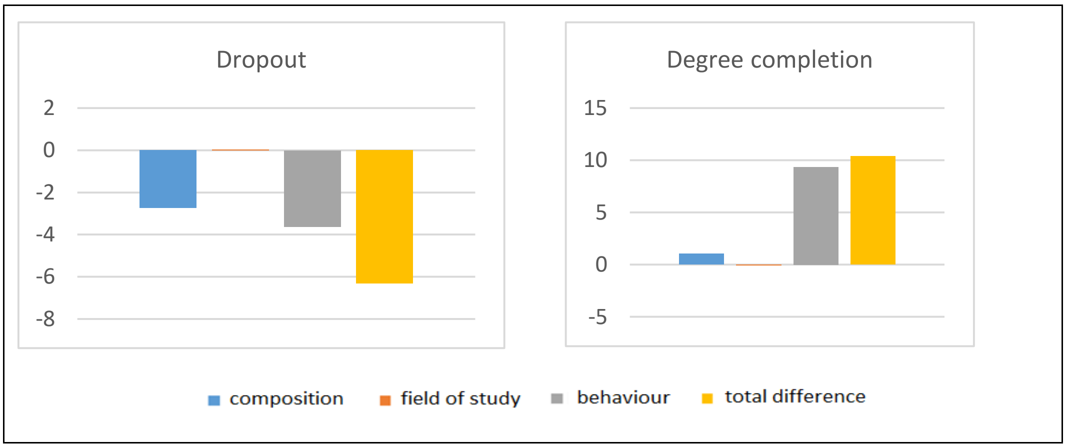
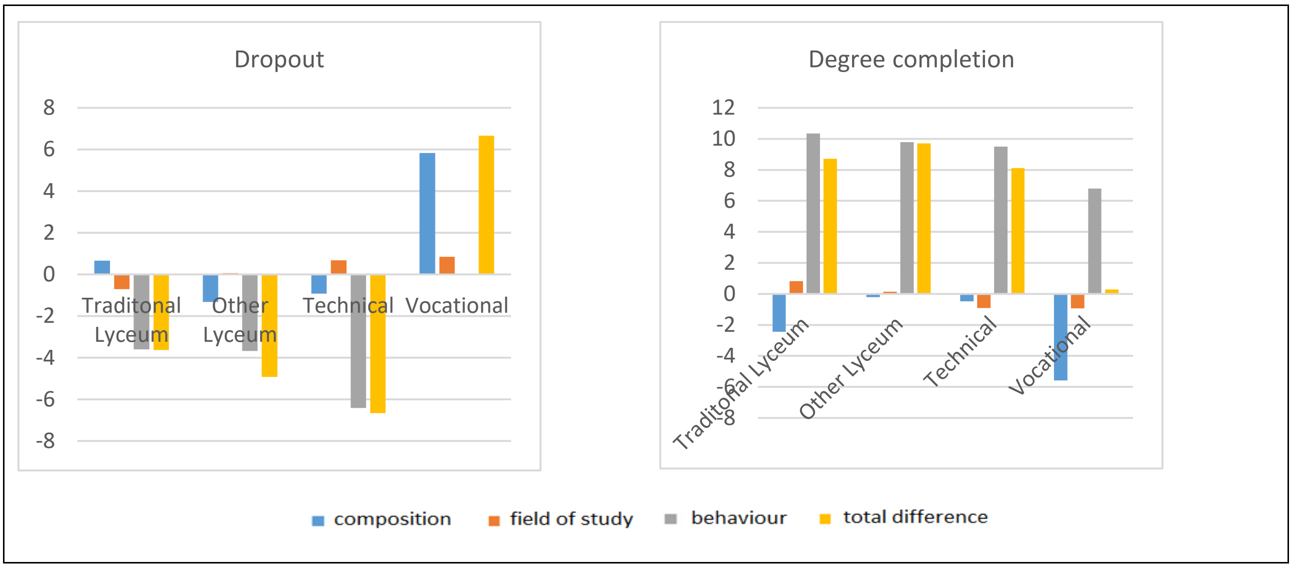


Figure 6. Decomposition of the changes between matriculation cohorts 2010 and 2004 in the probability of dropping out and completing the degree within 4 years (population)



Inspection of the changes observed within high school tracks provides additional insights (Figure 7). The improvements observed for students from traditional lyceums, other lyceums and technical schools are largely due to behavioral effects. Compositional effects are small: they contribute to a dropout increase for the first group and to a small reduction for the other two, and to a graduation probability decrease for all. Instead, the role played by changes in the choice of the field of study is negligible. A completely different picture is observed for the students coming from the vocational track. As mentioned above, this is the only subgroup exhibiting an aggregate deterioration of dropout probabilities. This change is entirely due to compositional effects. Notice that the average age at enrolment increases substantially over time for this group (Table A1, Appendix). Instead, stability observed at the aggregate level in the graduation probability results from the sum of a negative compositional and a positive behavioral component. We may conclude that, other things being equal, students from the vocational track do not experience a reduction in their chances to dropout, while they share with the other students an improvement in (almost) timely graduation probability.

Figure 7. Decomposition of the changes between cohorts 2010 and 2004 in the probability of dropping out and completing the degree within 4 years, within high school tracks



5. Contextual factors

Our findings unambiguously point to an improvement in the system effectiveness, and that changes are not only nor mainly due to compositional effects. Although addressing the causes of the observed changes is out of the scope of the present paper, we briefly sketch out some possible channels related to contextual factors. Two major macro-level factors occurred in the past years might have triggered substantial changes in individual educational decisions and outcomes at the tertiary level: i) the institutional changes in the university system enforced by a season of reforms following the Bologna Process; ii) the changes in the labor market entailed by the economic crisis that hit Italy in 2008 and had a harsh impact on the Italian economy, followed by a modest recovery that has not brought back per-capita income to its pre-crisis levels (IMF 2019).

Institutional changes in the university system

In the past decade, the Italian university underwent various institutional rearrangements that started with the implementation of the Bologna process in 2001. This reform, also known as '3+2', aimed at homogenizing higher education systems at the European level, introduced a two-tier degree structure, organized into a 3-year first level degree (BA) and a 2-year second level degree (MA). Subsequently, two broad reforms where enacted in 2004 and in 2010. The first reorganized BA and MA degrees, leaving more leeway in the choice of the master's degrees after a given first-level degree and increasing autonomy in the degree program curricula. The second radically changed the institutional governance and internal organization of Italian state universities moving teaching management responsibilities to departments, previously devoted only to research. In the meanwhile, the system has become increasingly accountable; the allocation of financial resources and recruiting policies were linked to performance evaluations of scientific research in particular, but also of the number of students enrolled, teaching quality and students' academic careers. In 2006 the National Agency for Evaluation was established to govern the evaluation of the university system of teaching and research activities;

in 2013 the 'AVA' system was introduced, with the aim of improving teaching and research quality, increasing autonomy and responsibility of the institutions through the enforcement of a quality management model based on self-assessment procedures and external evaluation.

Overall, this long-lasting reorganization process could have contributed to the improvement of student outcomes. On the one hand, the increasing emphasis on quality and accountability could have entailed a positive change of the faculty's staff mind-set, expressed through organizational changes, better teaching practices, widened tutoring, at the benefit of didactics at large. On the other, the urge to meet the targets set by the national authorities (and the increased prominence of scientific research for both individual career progression and financial allocation to institutions) might have led to lowering learning standards, reducing the students' effort needed for degree attainment.

Notice that both explanations would be consistent with *all* students experiencing better outcomes. When relating these predictions to our empirical results, we observe that vocational track students have not experienced a reduction in dropout probabilities. Yet, this could be due to their stronger exposure to the effects of the economic crises, since these are the students with the highest risk of being in need of income (see below).

Labor market changes and 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 labor market, according to preferences, risk aversion and other psychological traits. Secondly, individuals are subject to different constraints, due to diverse economic and cultural endowments. Thus, the effects of economic crisis on educational choices may operate at two levels. A downturn affects the conditions of the macroeconomic context and labor market opportunities, influencing the evaluation of costs and benefits of tertiary education and/or the choice of the field of study. 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 briefly review the potential mechanisms related to the economic crisis operating on university enrolment (affecting the *composition* of the student body) and on student trajectories given specific profiles (affecting *behavior*).[32]

Enrolment: By reducing the opportunity costs of studying, poor labor market prospects should increase participation to higher education. Yet, poor labor market prospects might also affect motivation and yield to discouragement, contributing to reduce participation. The net effect 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, even if the crisis increases the share of individuals in poor economic conditions in the population, the effects on the composition of university students

[32] Another factor could be the increasing job insecurity and share of atypical and fixed-term jobs, ongoing since the mid-nineties. Consider that returns to higher education in Italy have been steadily lower than in most other countries (OECD, 2018). This might be responsible of low enrolment rates, while influence on completion is difficult to predict.

is not clear-cut. Overall, we have seen that the student population has become more selected in terms of previous schooling, although we have no direct information in terms of family income.

Dropout: Poor labor market prospects may contribute to lower dropout rates because the opportunity costs of attending university decline. Moreover, a decline in employment opportunities should reduce withdrawals related to labor market pull out factors. On the contrary, students lacking economic resources might be pushed out of the educational systems due to need to earn income. Our results suggest that although both mechanisms could be at work, the net effect is a reduction of dropout, supporting the lower opportunity costs explanation. Yet, this is not true for vocational track students, usually coming from disadvantaged family backgrounds: for them the lack of income could have played a major role.

Time to degree: If the labor market prospects of university graduates are poor, opportunity costs decline and individuals should take longer time to complete their degrees. However, the increasing competition over scarce job positions may also induce students to graduate in time, as a good signal for prospective employers. Yet, students living in low-income households might feel urged to graduate and start earning income; on the contrary, they might choose to work while studying, contributing to lengthen time to degree. Our evidence supports the hypothesis that the crisis has led to a greater urge to complete university studies. Either to become economically independent, or to face the competition on the graduates' labor market.

6. Conclusions

This study depicts the recent evolution of student academic careers in a country that has low formal entry barriers to tertiary education, but extremely large differences across social groups (Jackson 2013) and in the end, a very low share of young individuals with tertiary education. In particular, we analyze the university trajectories of BA students matriculated in 2004-2013 at the University of Torino. Our data include transfers to other institutions, so we are able to study system-level dropout. Given the diversity across institutions and territories, this case-study is not representative of the Italian higher education system at large. However, it can be considered illustrative of the situation in public institutions occupying middle range positions in national rankings and with student academic careers similar to the national average.

To model trajectories, we propose a flexible discrete-time competing risks model that allows studying academic careers without defining all states of the process as final destinations. This method can be easily applied to analyze academic careers with administrative data, and may be adapted to study other aspects of student careers, including re-enrolment after dropout, masters' degree enrolment and so on.

In our empirical analysis we have focused at first on matriculation cohort 2008, the most recent cohort for which we observe 6 full academic years after enrolment, with the aim of highlighting the determinants of successful vs. less successful academic careers and estimating relevant outcomes for different student profiles. Dropout probabilities – larger after the first academic year, but substantial also in later years – are high for most profiles, and the chances to attain the BA degree in a reasonable amount of time are generally low. At the aggregate level only 28% of the students attains the degree within the 3-year institutional length of time and 56% reach the goal within 6 years. The remaining 44% either drop out (two thirds) or are still enrolled after 6 years (one third). These figures point to a severe lack of effectiveness of the system, entailing a huge waste of public and individual resources.

Student trajectories largely differ across high school tracks and prior performance. Given the strong social stratification of high school choices in Italy, these differences must be interpreted not only in terms of different preparedness, but also in the perspective of social inequalities. Age at matriculation is also a powerful predictor of success: the large majority of the students delaying higher education entrance a few years after the regular age of high school completion, drops out at some point. This adds on to the gaps observed across high-school tracks, because delayed entrance is much more frequent among students from technical and vocational schools. Indeed, the direction of these results is not surprising, but the magnitude of the differentials across student profiles is impressive. Although defining what can be considered satisfactory is arbitrary, we believe we may speak of "good" outcomes only for the subgroup of students enrolling immediately after high school and coming from traditional lyceums with high marks. Degree completion may be too demanding for those with less successful prior schooling and in particular for more mature students. These findings are in explicit conflict with the goal of the EU agenda to increase the inclusiveness of higher education and to provide opportunities to individuals from more disadvantaged backgrounds to enter and complete higher education.

Examination of credit earning in the first academic year and subsequent outcomes allows shading some light on the processes at work. One third of the students earns at most 20 credits and only less than 50% reach 40 credits out of the planned 60. Indeed, differences across tracks are large. Moreover, when experiencing a poor result – a much more likely event for non-academic high school students – individuals react differently. Students from lyceums tend to remain in the university system and change degree course, whereas students from less advantaged backgrounds are more likely to drop out. Among those who do not leave the system, only a minority will attain the degree within 6 years from enrolment, with persisting differences across school types. Hence, in each step of the "chain" of events – credit earning, resilience to a "failure" and subsequent outcomes – inequalities across high school tracks cumulate, giving rise to the impressive large differentials observed in the Cumulative Incidence Functions represented in Figure 4.

When we examine changes over time we may take a more optimistic view: what we observe is an overall reduction of dropout rates and a progressive increase in the timely graduation probabilities across matriculation cohorts. The observed improvements in student trajectories are not mainly due to compositional effects, but instead, to changes in the behavior of the actors involved (students, faculty members, management).

Notwithstanding this positive result, the outlook for students from the vocational track is grimmer. At the aggregate level, they have experienced an increase in the dropout probability over time. Changes in student characteristics are largely responsible for this deterioration (contrary to the general trend, vocational track students have become older in recent matriculation cohorts, and age at enrolment has a large negative effect on academic careers). Still, while ceteris paribus their graduation probability has increased, they have not shared with the others a reduction in their chances to drop out. A reason could be their greater exposure to the economic crisis.

Two major contextual factors might have driven the observed 'behavioral' changes: the season of university reforms enforced since 2004 and the economic downturn. We have highlighted possible channels through which they could have affected student outcomes. Yet, identification of their effects is difficult and beyond the scope of the present work. To this aim, further research exploiting differences across territories and organizational units (institutions, schools, departments) is needed.

Let us go back to the question raised in the introduction of this paper: is there evidence that the Italian university system is becoming more inclusive and effective? We already pointed out that, according to aggregate national data on enrolment, the system is *not* becoming more inclusiveness at access, as the transition rates from high-school to university are not increasing (and the share of students with an upper secondary degree is rather stable). We have also argued that, instead, there are visible signs of an improvement in effectiveness, as the share of students who attain the degree has been rising steadily in the observed time span. To address the issue of inclusiveness at completion, conceived as effectiveness "for all", we rely on our analyses of the changes occurred in student trajectories. Our findings depict a blurred picture. We might speak of increasing effectiveness "for the majority", but not for all, because students from the vocational track have not experienced a reduction in the probability to leave university before attaining the degree.

Leaving equity considerations aside, we might pragmatically make some back-of-the envelope calculations aimed at delivering rough projections of the share of young individuals with a higher education degree. Assuming that the share of students attaining an upper secondary degree remains stable at 80% and the transition rate from upper secondary to tertiary education goes back up to 56%, the share of young individual entering university would reach 45%. Considering as degree attainment probability the average within-6-years graduation probability (56%) observed for our 2008 cohort, we would obtain a 25% share of youngsters with tertiary degree, a slightly lower figure than the current reported share of 27% for year 2017 (OECD 2018). Even if we assumed a very large improvement – that on average students reach the graduation probability estimated for median mark students from traditional lyceums enrolled at age 19 (73%) – at equilibrium the share of graduates would remain at 33%.

We may conclude that, in addition to making further large improvements in the system effectiveness there is also the need to sharply increase the share of students entering higher education. Even under the optimistic scenario of a 73% degree-attainment probability, to reach the share of 40% graduates, the transition rate from high school to university needs to increase from 56% to 69%. In the current system, this is possible only if participation rates of the students from all school types increase substantially. We may conclude that unless institutional reforms or powerful policy interventions aimed at fostering study progression of disadvantaged students are put in place, there seems to be little chance to raise the share of young people with tertiary education to the level of most other developed countries within the next future.

References

Agasisti T. (2009) Market forces and competition in university systems: theoretical reflections and empirical evidence from Italy, *International Review of Applied Economics* 23, 463 – 483.

Allison P. (2018) For Causal Analysis of Competing Risks, Don't use Fine & Gray's subdistribution method, https://statisticalhorizons.com/for-causal-analysis-of-competing-risks

Ambrogi F., F. Biganzoli, P. Boracchia (2009) Estimating crude cumulative incidences through multinomial logit regression on discrete cause-specific hazards, *Computational Statistics and Data Analysis*, 53, 2767–2779

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

Aina C. (2013) Parental background and university dropout in Italy, *Higher Education*, 65, 4, 437-456

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

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

ANVUR (2018) Rapporto biennale sullo stato del sistema universitario e della ricerca 2018, Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca, Roma.

Argentin G., M. Triventi (2011) Social inequality in higher education and labor market in a period of institutional reforms: Italy, 1992-2007, *Higher Education*, 61: 309-323

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

Azzolini D., A. Martini, E. Rettore, B. Romano, A. Schizzerotto, L. Vergolini (2018) Testing a social innovation in financial aid for low-income students: experimental evidence for Italy, *IZA Discussion Paper* No 11625

Austin P.C., J. P. Fine (2017) Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Statistics in Medicine* 36:4291-4400

Barone C., M. Triventi, G. Assarelli Explaining Social Inequalities in Access to University: A Test of Rational Choice Mechanisms in Italy, *European Sociological Review*, 2018, Vol. 34, No. 5, 554–569

Barro, R. J. (2001), Human Capital and Growth, *American Economic Review*, 91(2): 12–17

Berens J., Schneider K., Görtz S., Oster S., Burghoff J. (2018) Early Detection of Students at Risk. Predicting Student Dropouts Using Administrative Student Data and Machine Learning Methods, CESIFO WP 7259

Blundell R., L. Dearden, C. Meghir, B. Sianesi (1999), Human capital investment: the returns from education and training to the individual, the firm and the economy, *Fiscal Studies*, 20, no. 1, 1–23

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

Bound J., M. Lovenheim, S. Turner (2010) Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources, *American Economic Journal: Applied Economics*, 2(3), 129-57.

Bowen W., M. Chingos, M. McPherson (2009) *Crossing the Finish Line: Completing College at America's Public Universities*, Princeton University Press, Princeton

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

Cameron S. and Heckman J. (2001). The dynamics of educational attainment for black, Hispanic, and white males. *Journal of Political Economy,* 109, 455-499.

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

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

Carrieri V., M. D'Amato, R. Zotti (2015) On the causal effects of selective admission policies on students' performances: evidence from a quasi-experiment in a large Italian university. *Oxford Economic Papers*

Cattaneo M., P. Malighetti, M. Meoli, S. Paleari (2016) University spatial competition for students: the Italian case, *Regional Studies*

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

Cobalti, A., and A. Schizzerotto (1993) Inequality of educational opportunity in Italy, in Shavit Y. and H.P. Blossfeld *Persistent inequality. Changing educational attainment in thirteen countrie*s, Westview Press, Boulder, 155-176.

Collett D. (2015) *Modelling survival data in medical research*, Chapman and Hall/CRC, third edition

Collins, Randall. 1971. "Functional and Conflict Theories of Educational Stratification." *American Sociological Review* 36 (6): 1002.

Contini D., A. Scagni (2013) Social Origin Inequalities in Educational Careers in Italy. Performance or Decision Effects? in Jackson Michelle (eds), *Determined to Fail? Performance, Choice and Education*, Stanford University Press

Contini, D., Cugnata, F., and Scagni, A. (2018). Social selection in higher education. Enrolment, dropout and timely degree attainment in Italy. *Higher Education*, 75, 785–808.

Cox D. R. and D. Oakes (1990) *Analysis of survival data*, Chapman and Hall, Cambridge, UK

Declercq K., F. Verboven (2018) Enrollment and degree completion in higher education without admission standards, *Economics of Education Review*, 66, 223-244

Denning J. T., E. R. Eide, M. Warnick (2019) Why Have College Completion Rates Increased? EdWorkingPaper: 19-77 Annenberg Institute at Brown University

De Santis G., E. Pirani., M. Porcu (2019) *Rapporto sulla Popolazione. L'istruzione in Italia*, Il Mulino, Bologna

Delen D. (2010) A comparative analysis of machine learning techniques for student retention management, *Decision Support System*, 49, 498-506

Desjardins S., D.A. Ahlburg, B. P. McCall (2002) A temporal investigation of factors related to timely degree completion, *The Journal of Higher Education,* Vol. 73, No. 5

Desjardins S., Ahlburgb D. A., McCallc B. P. (2006) The effects of interrupted enrolment on graduation from college: Racial, income, and ability differences, *Economics of Education Review*, 25, 575–590

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.

Enea, M., Attanasio, M. (2016) An Association Model for Bivariate Data with Application to the Analysis of University Students' Success. *Journal of Applied Statistics*, 43 (1), 46-57

Enea, M. (2018) From South to North? Mobility of Southern Italian students at the transition from the first to the second level university degree. In C. Perna, M. Pratesi, A. Ruiz-Gazen (Eds). *Studies in Theoretical and Applied Statistics.* Springer

European Commission/EACEA/Eurydice (2018) The European Higher Education Area in 2018: Bologna Process Implementation Report. Luxembourg: Publications Office of the European Union.

Eurostat (2018) Education and training in the EU-facts and figures http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do

Fairlie (1999) The absence of the African-American owned business: an analysis of the dynamics of self-employment, *Journal of Labor Economics*, 17-1, 80-108

Gambetta D. (1987) *Were they pushed or did they jump? Individual decision mechanisms in education*. Cambridge University Press

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

Gloker D. (2011) The effect of student aid on the duration of study, *Economics of Education Review*, 30, 177–190

Guetto, R., L. Vergolini (2017). Educational expansion without equalization: a reappraisal of the 'Effectively Maintained Inequality' hypothesis in children's choice of the upper secondary track, *European Societies*, 19, 1

Hanushek E. A., L. Woessmann (2008) The Role of Cognitive Skills in Economic Development, *Journal of Economic Literature*, 46:3, 607–668

International Monetary Fund. European Dept. (2019), Italy: Selected Issues, Country Report No. 19/41.

Ishitani (2006) Studying Attrition and Degree Completion Behavior among First-Generation College Students in the United States, *The Journal of Higher Education*, Vol. 77, No. 5, pp. 861-885

Jackson M. (2013), *Determined to Succeed? Performance, Choice and Education*, Stanford University Press

Jenkins S. P. (2005) Survival Analysis https://www.iser.essex.ac.uk/files/teaching/stephenj/ec968/pdfs/ec968lnotesv6.pdf

Lancaster T. (1990) *The econometric analysis of transition data*, Econometric Society Monographs, Cambridge University Press

Larsen M.R., H. B. Sommersel, M.S. Larsen (2103) *Evidence on dropout phenomena at universities*, Danish Clearinghouse for Educational Research

Lassibille G., L.Navarro Gomez (2008) Why do higher education students drop out? Evidence from Spain, *Education Economics*, 16, 89-105

Manski C. F. (1989) Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon *Economics of Education Review*, 8, 4, 305-312

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

OECD (2017a) *Enhancing Higher Education System Performance*. Report on benchmarking higher education system performance: conceptual framework and data, OECD Publishing

OECD (2017b) *State of Higher Education* 2015-16, OECD Publishing

OECD (2018) *Education at a Glance* 2018: OECD Indicators, OECD Publishing

Orr D., A. Usher, C. Haj, G. Atherton, I Gaenta (2017) Study on the impact of admission systems on higher education outcomes. Volume 1: comparative report, European Commission.

Ortiz, E. A., C. Dehon, C. (2013). Roads to success in the Belgian French community's higher education system: predictors of dropout Bruxelles. *Research in. Higher Education* 54, 693–723.

Panichella N. and M. Triventi (2014) "Social inequalities in the choice of secondary school: Long-term trends during educational expansion and reforms in Italy." *European Societies* 16.5, 666-693.

Romito, M, Pilutti S., D. Contini (2020) Why do students leave university? Qualitative research at an Italian higher education institution, *European Journal of Education*, forthcoming

Schizzerotto A., C. Barone (2006). *Sociologia dell'istruzione*. Il Mulino, Bologna

Schnepf, S. V. (2014). Do Tertiary Dropout Students Really Not Succeed in European Labor Markets? Bonn: Institute for the Study of Labor, *IZA Discussion Paper* 8015

Scott, M.A., B.B. Kennedy (2005). Pitfalls in pathways: Some perspectives on competing risks event history analysis in education research. *Journal of Educational and Behavioral Statistics*, *30*(4), 413-442.

Smith J. P., R. A. Naylor (2001) Dropping out of university: a statistical analysis of the probability of withdrawal for UK university students, *Journal of the Royal Statistical Society A*, 164, 2, 389-405

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

Stinebrickner R., T. Stinebrickner (2013) Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model, NBER Working Paper 1894

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

Vergolini, L., N. Zanini (2015). Away, but not too far from home. The effects of financial aid on university enrolment decisions. *Economics of Education Review*, *49*, 91-109

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

ADDITIONAL DESCRIPTIVE STATISTICS

Table A1. Student characteristics at enrolment (BA degrees), by school type (cohorts 2004 and 2010)

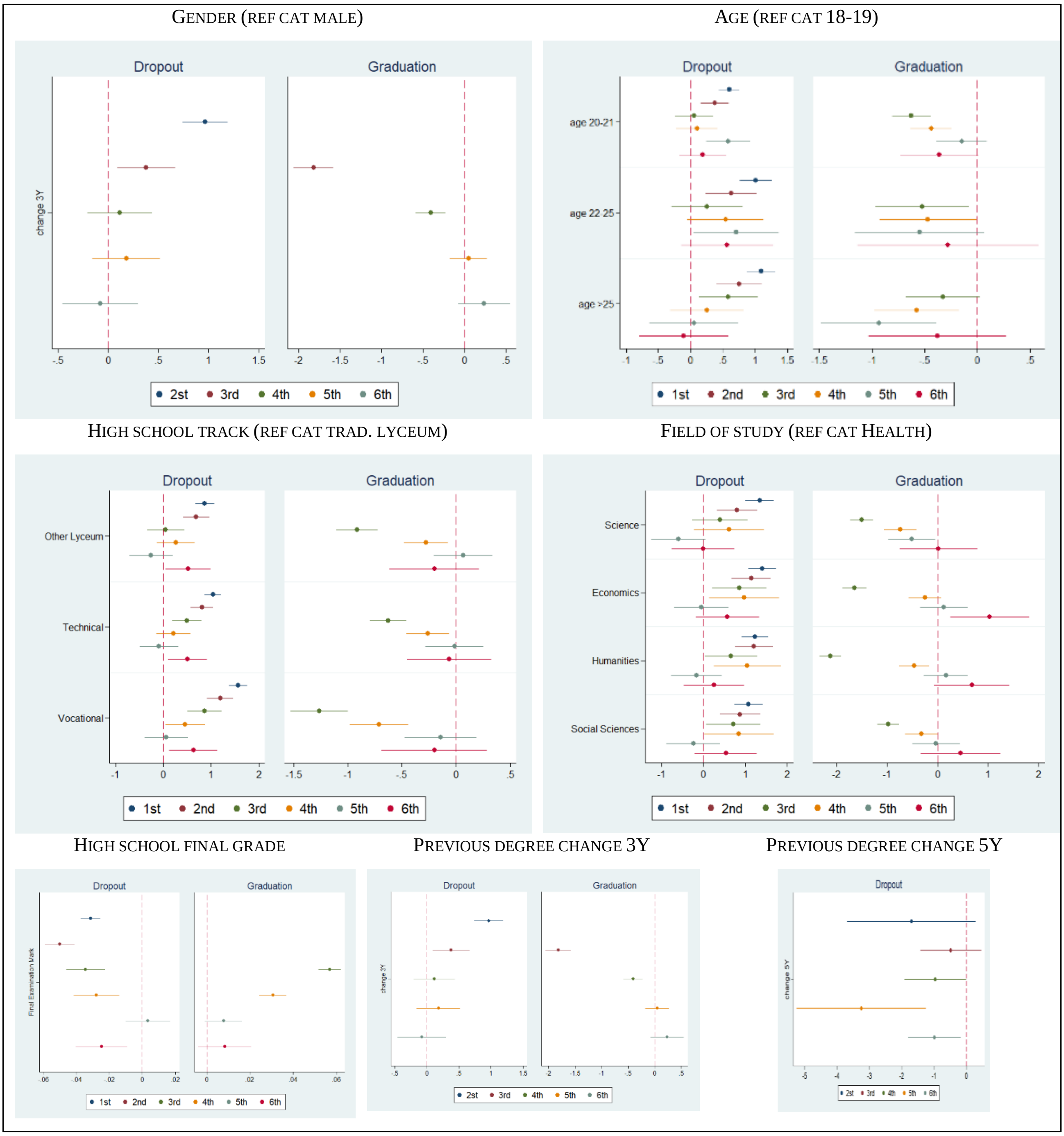
Cohort 2004						
	% matriculated age:			% females	Average final high school mark	% over total
	18-19	20-21	22+			
Traditional Lyceum	82	14	4	61	78.47	47.7
Other Lyceum	62	20	18	87	78.69	9.8
Technical	54	21	25	44	77.42	23.4
Vocational	61	24	15	73	79.48	19.1
TOTAL	69	18	13	62	78.44	100
Cohort 2010						
	% matriculated age:			% females	Average final high school mark	% over total
	18-19	20-21	22+			
Traditional Lyceum	83	15	2	61	76.16	58.3
Other Lyceum	68	24	8	87	78.02	14.8
Technical	58	26	16	49	76.16	17.2
Vocational	47	33	20	61	76.27	9.7
TOTAL	73	20	7	63	76.45	100

NOTE. 3-years law students excluded from this table and from the analysis because institutional changes limit the comparability over time

STEP-BY-STEP MODEL ESTIMATES

Figure A1 displays 'step-by-step' model estimates for the 2008 matriculation cohort by year after enrolment (years 1-6 for dropout, years 3-6 for graduation).

Figure A1. Step-by-step model estimates



NOTE. We report coefficients' point estimates and confidence intervals (bars) for the options corresponding to destination states. For the change to longer degrees, we do not repot estimates relative to the graduation option, because chances to graduate in the first 4 years are nil.

GOODNESS OF FIT

Table A2. Cumulative observed and predicted dropout/graduation probabilities (CIF), by academic year (cohort 2008)

	Dropout						Graduation			
Year	**1**	**2**	**3**	**4**	**5**	**6**	**3**	**4**	**5**	**6**
Observed	15.34%	21.95%	25.54%	28.39%	30.89%	33.10%	27.99%	43.91%	51.23%	54.24%
Predicted	15.65%	22.23%	25.86%	28.76%	31.27%	33.58%	27.75%	43.72%	51.16%	54.37%