UNIVERSITA' DEGLI STUDI DI PADOVA

**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI
"M.FANNO"**

**CORSO DI LAUREA MAGISTRALE IN
ECONOMICS AND FINANCE**

TESI DI LAUREA

"PREDICTIVE POWER OF THE ADMISSION TEST ON COLLEGE PERFORMANCE"

RELATORE:

CH.MO PROF. Rocco Lorenzo

LAUREANDO/A: Candiba Ana

MATRICOLA N. 1183784

ANNO ACCADEMICO 2019 – 2020

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

The candidate declares that the present work is original and has not already been submitted, totally or in part, for the purposes of attaining an academic degree in other Italian or foreign universities. The candidate also declares that all the materials used during the preparation of the thesis have been explicitly indicated in the text and in the section "Bibliographical references" and that any textual citations can be identified through an explicit reference to the original publication.

Firma dello studente

____

Contents

List of figures

List of tables

Introduction

One way through which universities can select their future students is by the implementation of admission tests. These are standardized tests that verify the knowledge of applicants in different fields and have the aim of selecting the students that are most likely to be academically successful. Therefore, if the demand of applications is too high for a university to bear, then an admission test will help select some students, that is the ones who obtained the best test results. During the selection process universities can also take into account other factors, in addition to the admission test, such as the high school final mark or the high school grade point average. Furthermore, especially in the United States, most admission rules use other indicators of student preparedness, such as recommendation letters, extracurricular activities and essays (Rothstein, 2004).

Each university and college around the world sets its own admission standards. In the United States, for example, high school students usually are required to take a SAT or an ACT test which are national standardized tests comprised of different sections related to school subjects (reading, English, mathematics and science) that give universities a measure through which they can analyse the performance of students together with the high school GPA.

In Italy, instead, each university decides its own admission rules at a local level: some universities will allow free access, others will require students to undergo a test specifically prepared for that undergraduate program. There are some exceptional degrees (e.g. medicine, animal medicine, architecture, and others), however, that are managed from a national point of view by the MIUR (Ministry of Education, Universities and Research). In this case the MIUR decides each year the number of places offered based on the labour demand and on the availability of places of each university[1].

For what concerns the undergraduate program of Economics held in the University of Padova, the admission rules are decided locally.

The aim of this work is to analyse the admission test that must be taken in order to be admitted to the bachelor's degree of Economics in the University of Padova. I will examine how the

[1] The admission rules from a national point of view are regulated by a specific law: Legge 2 Agosto 1999, n.264 "*Norme in materia di accessi ai corsi universitari*".

admission test is used and what is its predictiveness for what concerns the academic results and success of future students.

The implementation of admission tests for the selection of future students is considered to be one of the best ways of predicting the students that are most likely to be academically successful. But do admission tests really predict who will be the best future students? And to which extend?

My work is divided as follows: Chapter 1 presents an overview on the literature that studies the predictiveness of the admission test. Chapter 2 describes the admission rules that were implemented by the department of Economics of the University of Padova during the four academic years (2013-2016) that I will analyse. Chapter 3 describes the data and the variables employed in this dissertation as well as the empirical methodology. Chapter 4 shows the results and the discussion.

Chapter 1

Literature review

There is a vast literature that tries to answer to the question: How well the admission test can predict the students' future performance in college? Most of the literature examines the predictiveness of standardized tests in the United States: the SAT and the ACT are the tests most taken into consideration. These tests are similar from a point of view of the total duration, which is about 3 hours (almost 4 hours if also considering the optional Writing part of the tests), but nonetheless there are some differences between them. The ACT is formed by four parts (English, Mathematics, Reading and Science) whereas the SAT has three sections (Writing, Reading and Mathematics). The grade range also differs: the composite score of the ACT lies between 0 and 36, while regarding the SAT it ranges from 0 to 1600. The SAT is developed and managed by the College Board, which is a non-profit organization that released the first SAT test in 1926. Since then the SAT has been redesigned throughout the years and its name and scoring have been changed several times. Over the years the College Board has revised the SAT test trying the improve its predictiveness and has released many reports that study the SAT validity and its changes and revisions. The results of one of the most recent publications (Westrick, 2019), based on data from more than 223,000 students across 171 colleges, show that the SAT is strongly predictive of college performance: students with higher SAT scores are more likely to have higher grades in college. The SAT scores are also shown to be predictive of second year retention. Furthermore, it is argued that the SAT is as effective as the high school grade point average (HSGPA) in predicting students' success in college but nonetheless the best way of predicting future performance is by combining the SAT and the HSGPA. These results are confirmed by other publications by the College Board (Bridgeman, 2000; Kobrin, 2008; Marini, 2019).

In regard to the ACT test, Sawyer (2013) found that test scores are more useful than HSGPA in case of highly selective universities, but the opposite occurs in situations involving low selectivity in admission.

Other studies, while analysing the predictiveness of the test scores, have taken into consideration other factors like the socioeconomic status of students and the self-regulatory competencies. A meta-analysis (Westrick, 2015) examined the relationship of the ACT scores, high school grades, and the socioeconomic status (SES) with the students' performance during the first three years of college. Traditionally, in the literature, the key indicator of the college success has been considered to be the first-year grade point average (FYGPA), whereas Westrick et al. put their attention also on the academic performance and the retention beyond the first year. They found that ACT scores and high school GPA are positively correlated with the first year GPA and that generally they have much stronger relationship with the 1st-year GPA than does the socioeconomic status.[2] For what concerns the second year they found that both ACT scores and high school GPA were valid predictors of students' academic performance and the mean correlations were as strong as the corresponding mean correlations with the first year GPA. Finally, it was established that retention during the second and third year was better predicted by the academic performance during the previous years.

Other factors that could influence the college performance are the self-regulatory competencies that include self-control and grit, the tendency to sustain passion and perseverance toward challenging goals (Galla, 2019). It was argued by Galla that, while both the test scores and the high school grades were found to be important in predicting the performance of future students, these measures were relevant for different reasons: "the incremental predictive validity of high school grades for college graduation was explained by self-regulation, whereas the incremental predictive validity of SAT scores for college graduation was explained by cognitive ability". It appears that universities seek to develop at least two types of expertise in their students: the cognitive skills and the soft skills, that comprise also self-regulation (Stemler, 2012). Given that the SAT test does not capture the element of self-regulation, it is important for universities to consider also the GPA when deciding who to admit. In fact, Galla found that HSGPA predict college graduation better than admission test scores.

A critique to the traditional method of analysing the predictiveness of the admission tests and the HSGPA is that the traditional validity studies do not take into account other variables that may predict college performance, such as the individual and high school demographic variables (Rothstein 2004). Rothstein, using data from the University of California, shows that background characteristics of students and high schools are strong predictors of SATs and argues that "the exclusion of student background characteristics from prediction models inflates the SAT's apparent validity". In addition, Hoffman and Lowitzki (2005) show that, when it comes to minority students, high school grades are stronger predictors of college performance

[2] The estimated mean correlation of the first year GPA with the ACT scores was 0.51, while the one with the high school GPA was 0.58. In contrast, the estimated mean correlation between the socioeconomic status and the first year GPA was 0.24. (Westrick, 2015)

than are test scores. In fact, if the merit of students is defined only by the standardized tests (SAT), the result of this decision will be a lower diversity in college, whereas defining merit using also high school grades is more compatible with diversity in universities (Alon, 2007).

Given that the SAT and ACT tests are composed of different sections, it may be argued that some of these subtests are more effective in predicting college performance than others. Bettinger et al. (2013) show that even though the ACT composite score has a strong positive correlation with the college outcomes, “this overall correlation masks an important pattern: Mathematics and English scores are much more tightly correlated with college success than are Reading and Science scores.”

Chapter 2

Admission test for bachelor’s degree in Economics (Unipd)

My dissertation analyses the predictiveness of the admission test for the bachelor’s degree in economics of the University of Padova. The whole process of the admission is managed locally: the university decides the number of places available, the fields and the questions of the test, the number of questions, the test duration and whether to take into account also the high school final grade for the creation of the final ranking. All these details are disclosed on the Calls for application[3] (Avviso di Ammissione) that each year the department of economics releases for the new students. The data used in this work comprise 4 cohorts of bachelor students enrolled between 2013 and 2016. During this period, the department set 450 available places for the EU citizens and non-EU citizens regularly living in Italy, and 10 for the non-EU citizens living abroad, with the only exception of the 2013/2014 academic year. As a matter of fact, in 2013 two distinct undergraduate programs were offered by the department of Economics, International Economics (degree class L33) and Economics and Management (degree class L18)[4]. These two programs were cancelled in 2014 and from that year onward a new undergraduate program (Economics, degree class L18) was offered. For the year 2013 I will be considering only the students enrolled in the Economics and Management program[5], for which there were available 225 places for the EU citizens and non-EU citizens regularly living in Italy, and 5 for the non-EU citizens living abroad. During the four years considered, some changes have been implemented to the admission rules. Hereafter I will present the admission standards decided by the department of Economics in the University of Padova.

---

[3] All the Calls for application can be found on the website of the department of Economics of the University of Padova at the following link:
https://www.economia.unipd.it/avvisi-di-ammissione-e-materiale-utile-anni-precedenti

[4] The degree programs of the same level that satisfy a distinctive and specific set of characteristics and essential educational objectives belong to the same degree class as defined by the Ministry of Education, University and Research (MIUR).
Source: D.M. 22 Ottobre 2004, n. 270 “*Modifiche al regolamento recante norme concernenti l'autonomia didattica degli atenei, approvato con decreto del Ministro dell'universita' e della ricerca scientifica e tecnologica 3 novembre 1999, n. 509”* issued by the MIUR.

[5] In this way the degree class (L18) remains the same throughout the four years.

The department of Economics has provided the candidates with two admission tests each year. The candidates could undertake the test in spring (in April or in May) or in summer (August). If a candidate did not pass the spring selection, they had the possibility to undertake again the admission test held in August. The structure of the test remained the same throughout these four years, it consisted of 80 multiple choice questions to answer in 80 minutes. The test was divided into four sections with the aim of verifying different abilities: the verbal and memory abilities, the analytical and quantitative abilities and the general knowledge[6].Thirty-five questions regarded the verbal ability, the memory skills and the general knowledge, whereas the remaining forty-five questions were meant to verify the analytical and quantitative abilities.
The final score of the test was computed assigning:

- 1 point for each correct answer
- -0,5 points for the wrong answers
- 0 points for the blank answers

For the creation of the final ranking, also the high school grades were taken into account, but they were given different weights during this four-year period. The final score was expressed on a 100-point scale. In 2013 and 2014 fifty points were reserved to the high school grades, the other fifty points could be obtained through the admission test[7]. For what concerns the high school performance, in spring it was measured computing the mean of the grades, obtained by the candidates during the fourth year of the secondary school, of some selected high school subjects[8]. Whereas during the summer selection the finial high school grade (Voto di Maturità) was taken into account[9] as an indicator of the high school performance. During the spring selections of 2015 and 2016 the final score was computed using only the admission test

[6] The verbal abilities included the reading and the lexical comprehension, while the memory skills referred to the capability of reading, remembering, and using certain information from a text. The analytical and quantitative abilities comprised logical thinking abilities and mathematical knowledge. The general knowledge part meant to verify the knowledge about contemporary history, current economic topics, the organization of the State and national and international institutions. These details are explained in the Calls for application of the years 2013, 2014, 2015 and 2016.

[7] The maximum points a candidate could obtain in the admission test was 80. This score was normalised as follows:

- 50 points if the test score was above 70
- 0 points if the test score was below 0
- $\left(50 * \frac{TS}{70}\right)$ if the score was comprised between 0 and 70, where TS represents the test score.

[8] The selected subjects were Italian language, history, mathematics, a foreign language and two more free choice school subjects.

[9] During the spring selection the normalised mean of the school grades was computed as follows: $(50 * \frac{GPA-6}{4})$
During the summer selection the final mark was normalised as follows:

- 50 points were assigned to the candidates who obtained the maximum final grade (100)
- 0 points were given in case of minimum final grade (60)
- For the intermediate grade, the following formula was used: $(50 * \frac{FG-60}{100-60})$, where FG is the final high school grade.

results[10], whereas in August, thirty points of the final score were attributed to the high school mark[11] and the remaining seventy to the admission test score[12]. Table 1 summarizes how the 100 points of the final ranking score was divided between the Admission test and the high school grade point average (HSGPA) throughout the four year.

Table 1: How the final ranking score is created

	2013		2014		2015		2016	
	April	**August**	**May**	**August**	**April**	**August**	**April**	**August**
Admission test	50	50	50	50	100	70	100	70
HSGPA	50	50	50	50	0	30	0	30

The structure and score of the admission test did not change during these four years, whereas the way of computing the final ranking score differed throughout the years and even between the spring and summer selection of the same year for what concerns 2015 and 2016.

In the Calls for application it is also explained how the allocation and reallocation of the available places are managed. During the spring selection a first ranking of the admitted candidates is released. If not all the admitted candidates decide to enrol, a second ranking aimed to fulfil the vacant places is released. After the second reallocation, if still some places remain open, they will be added to the places made available during the summer selection. In August, instead, after the first and the second allocation, that work exactly as in spring, there are further reallocations, usually four or five, that try to fulfil the remaining open places. Candidates that wish to be included in these last rankings must make a request. Thus, these last rankings are formed only by considering the candidates that applied for it. All the candidates can apply regardless of their final score.

---

[10] The final score was computed as follows:

- if the admission test score was below 0 than the final score was set to 0.
- if the admission test score was above 0 than the final score was: $(100 * \frac{TS}{80})$, where TS represents the admission test score.

[11] The high school mark was normalised as follows:

- 30 points if the candidate obtained the maximum high school grade
- 0 points if the candidate obtained the minimum high school grade
- $(30 * \frac{FM-60}{100-60})$ for the intermediate grades

[12] The admission test score was normalised as follows:

- 70 points if the test score was above 70
- 0 points if the test score was below 0
- The actual result of the test score if it was comprised between 0 and 70

Chapter 3

Empirical analysis

3.1 Data

The datasets employed in this dissertation were provided by the department of Economics of the University of Padova. I was provided with four datasets, which are:

1. A dataset containing information about the candidates who undertook the admission test (their gender, date of birth, final score, admission test score, subtests scores, type of secondary school attended ad final high school grade).
2. A dataset with personal data about the enrolled students (gender, date of birth, citizenship, region and province of residence, province of birth, high school graduation year, high school final grade, type and location of secondary school).
3. A dataset containing the grades of all the exams and the date when they were successfully undertaken by the enrolled students.
4. A dataset containing information about the graduation of the enrolled students (degree score, arithmetic and weighted averages of the exams' grades and date of graduation).

3.1.1 Who are the applicants?

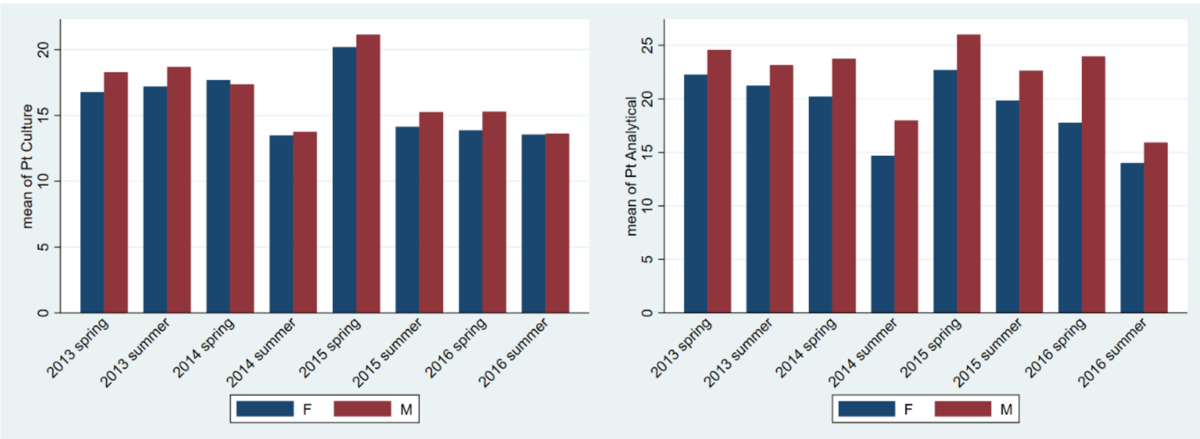
The first dataset contains information about the candidates that undertook the admission test for the Economics and Management degree (2013) and for the Economics degree (2014-2016). Table 2 contains details about the number of candidates that took the admission test each year and the percentage of males and females. The average number of candidates that took the test is 935 during each selection for a total number of applicants of 7477. Though the percentage of females is persistently slightly lower than the percentage of males, they remain constant near the 50% threshold throughout the four years. This Table also shows the average high school final grades of the candidates who took the admission test in summer. For what concerns the candidates who took the test in the spring of 2013 and 2014, the table presents the mean of the grades obtained during the fourth year of the secondary school. There is no available data for the candidates who undertook the test in the spring of 2015 and 2016 because, as pointed out before, during those selections the high school performance was not taken into account for the creation of the final ranking score. This information is also presented separately for males and females. Considering the whole sample of candidates, the average high school final mark of the four years is 79,35. It is interesting noticing that females perform better than males in high school, in fact both their final grade and GPA of the fourth year of the secondary school are, on average, higher compared to the average performance of males. The opposite is true, instead, for what concerns the results of the admission test. In fact, males tend to obtain on average higher total scores than females. Figures 1 shows the results achieved by males and females during each selection that took place throughout the four years. It confirms that males outperform females and gain on average four more points in the admission test results. The same trend can be observed also for what concerns the subtests, represented in Figure 2. The first subtest, aimed to assess the verbal abilities, memory abilities and general knowledge, shows better results for males, even though the difference between the two means is rather small, almost null or even positive in some cases. There is a clearer picture if we look at the second subtest proposed in Figure 2, that is the one aimed to measure the analytical and quantitative abilities of the candidates. Here males tend to perform better and on average gain three additional points compared to females.

Table 2: Number of candidates and their high school final grade, 2013-2016

	2013		**2014**		**2015**		**2016**		**Total**
	April	**August**	**May**	**August**	**April**	**August**	**April**	**August**	
Number of candidates	791	871	1088	1013	946	880	996	892	**7477**
Female	46,3%	46,2%	48,9%	47,2%	43,8%	46,3%	45,6%	48,1%	**46,5%**
Male	53,7%	53,9%	51,1%	52,8%	56,2%	53,7%	54,4%	51,9%	**53,5%**
GPA	7,58	81,26	7,5	78,88		78,54		78,71	**79,35**
Female GPA	7,76	83,09	7,68	81,61		81,44		82,20	**82,08**
Male GPA	7,41	79,69	7,32	76,46		76,04		75,45	**76,91**

Figure 1: Mean of the Admission Test score for Males and Females, 2013-2016

Figure 2: Mean of the Reading and Math subtests scores for Males and Females, 2013-2016



The last subject that I want to analyse here refers to the secondary school that the candidates attended before college. To reach a comprehensive outlook, I divided the types of high school into eight categories, following the MIUR indications[13]:

1. Classical Lyceum (Liceo Classico)
2. Scientific Lyceum (Liceo Scientifico)
3. Linguistic Lyceum (Liceo Linguistico)
4. Human Sciences Lyceum (Liceo delle Scienze Umane)
5. Economic Institute (Istituto tecnico, settore economico)

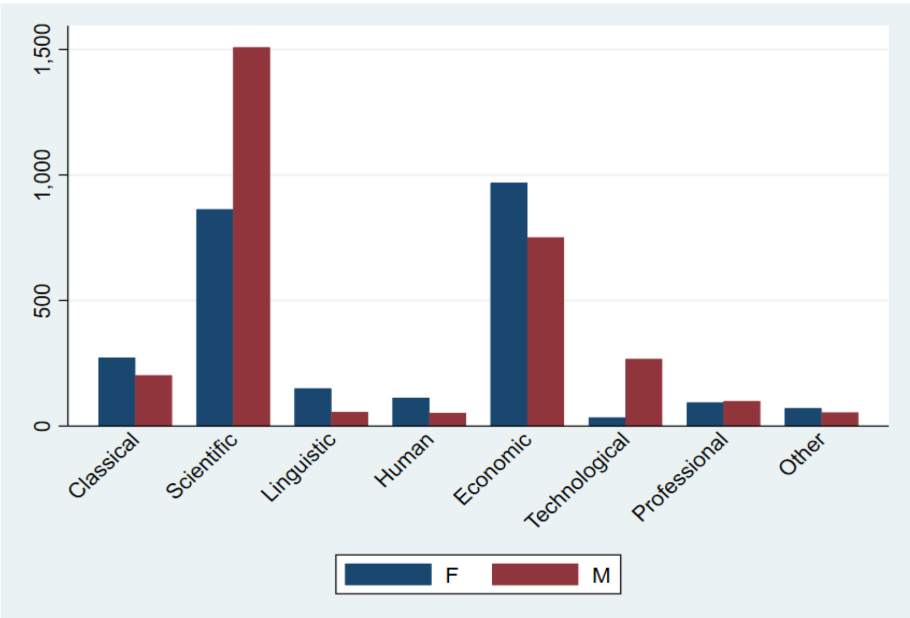
[13] https://www.miur.gov.it/web/guest/scuola-secondaria-di-secondo-grado

6. Technological Institute (Istituto tecnico, settore tecnologico)
7. Professional Institute (Istituto Professionale)
8. Other

Observing the data available, which comprises the whole four-year period[14], two secondary schools stand out among the others: the scientific lyceum and the economic institute. In fact, more than 70% of the total candidates[15] come from these two high schools, more precisely roughly 43% from the scientific lyceum and 30% from the economic institution. These two schools are followed by the classical lyceum (more than 8% of candidates), the technological institution (5%) and ultimately the rest of the schools, which represent about 3% of the candidates each. Another distinction can be made observing the gender. In fact, when looking at the whole sample, as seen in Table 2, there is a fair distribution for what concerns the gender, while if we analyse separately the subsamples regarding high school, some distinctions can be made. In particular, almost 65% of the candidates that come from a scientific lyceum are males, whereas regarding the technological institution, an even higher percentage (almost 90%) of candidates are males. The revers occurs for what concerns the rest of the secondary schools, although the differences between the number of male and female candidates here are smaller.

Figure 3 puts together the list of the secondary schools and the number of candidates, males and females, that attended each of these schools.

Figure 3: Number of candidates per secondary school, divided by males and females



[14] Only for the 2013 and 2014 spring selections there were no available data regarding the high schools attended by the candidates.

[15] The candidates, for whom data regarding high school was available, were 5556 for the period 2013-2016.

3.1.2 Who are the enrolled students?

In 2013 the department of Economics made available 230 places for the Economics and Management degree, whereas from 2014 to 2016 there were 460 available places each year. The students that actually enrolled were 220 for the first year, 442, 428 and 441 respectively for the years 2014, 2015 and 2016. Table 3 shows the number of students that enrolled during each selection. It's clear that, even though there are no differences in the number of available places between the spring and the summer selection, less students tend to enrol in the spring selection. This can be explained by the fact that not all the candidates who undertake the admission test are actually willing to enrol afterwards and, given that in spring after the release of the first ranking there is only one reallocation, some places always remain vacant, whereas in summer this risk is minimized by the addition of several reallocations. This means that in summer there are more chances for the candidates that were not immediately admitted to be reallocated.

The total number of students that enrolled in the four-year period is 1531, of which females represent 45,8% of the sample and males 54,2%. But looking only at the total averages is not enough because there are some striking differences between selections. In general, the percentages remain around 50% for both males and females with the exception of the spring selections of the last two years, in which females were only 31,5% (2015) and 25% (2016) of the enrolled students. As I explained in chapter 2, during these two selections the final ranking score was computed only by taking into account the results of the admission test. It is also important to recall from the paragraph above that males outperform females in the admission test, whereas the opposite occurs for what concerns the final high school grade. These two factors together can explain why the percentage of female students that enrolled during these two selections are so low. In fact, if the high school grade is left out of the final ranking score, females seem to be disadvantaged given their worse performance in the admission test.

For what concerns the final high school grades, the average of the whole sample is 87,64, which is 8 points higher than the average of the final high school grade of the candidates' sample. Furthermore, females are confirmed to perform better than males obtaining, on average, 5 more points than males.

Regarding the results of the admission test, a pattern similar to the candidates' can be noticed. In fact, although the average score obtained by the enrolled students from the admission test during the four years (45,7 points) is higher than the average score of the total number of candidates (37 points), also in this case males tend to perform better than females. This is true for both the analytical and the verbal subtests. Figures 4 and 5 show these results.

The enrolled students arrive mostly from a scientific lyceum (50,7%), an economic institute (30%) and a classical lyceum (9%). The distribution of males and females here is very similar to the distribution of the candidates' sample. Figure 6 shows the scholastic background of the enrolled students.

Table 3: Number of enrolled students and their high school final grade, 2013-2016

	2013		2014		2015		2016		Total
	April	August	May	August	April	August	April	August	
Enrolled students	74	146	180	262	184	244	180	261	**1531**
Female	55,4%	52,7%	52,2%	50,4%	31,5%	48,0%	25,0%	51,0%	**45,8%**
Male	44,6%	47,3%	47,8%	49,6%	68,5%	52,0%	75,0%	49,0%	**54,2%**
GPA	97,42	95,15	92,28	88,16	83,29	85,20	81,72	86,49	**87,64**
Female GPA	97,88	96,07	92,87	90,05	86,60	88,58	86,22	89,36	**90,62**
Male GPA	96,88	94,14	91,63	86,25	81,76	82,06	80,21	83,48	**85,17**

Figure 4: Mean of the Admission Test score for the enrolled Males and Females, 2013-2016

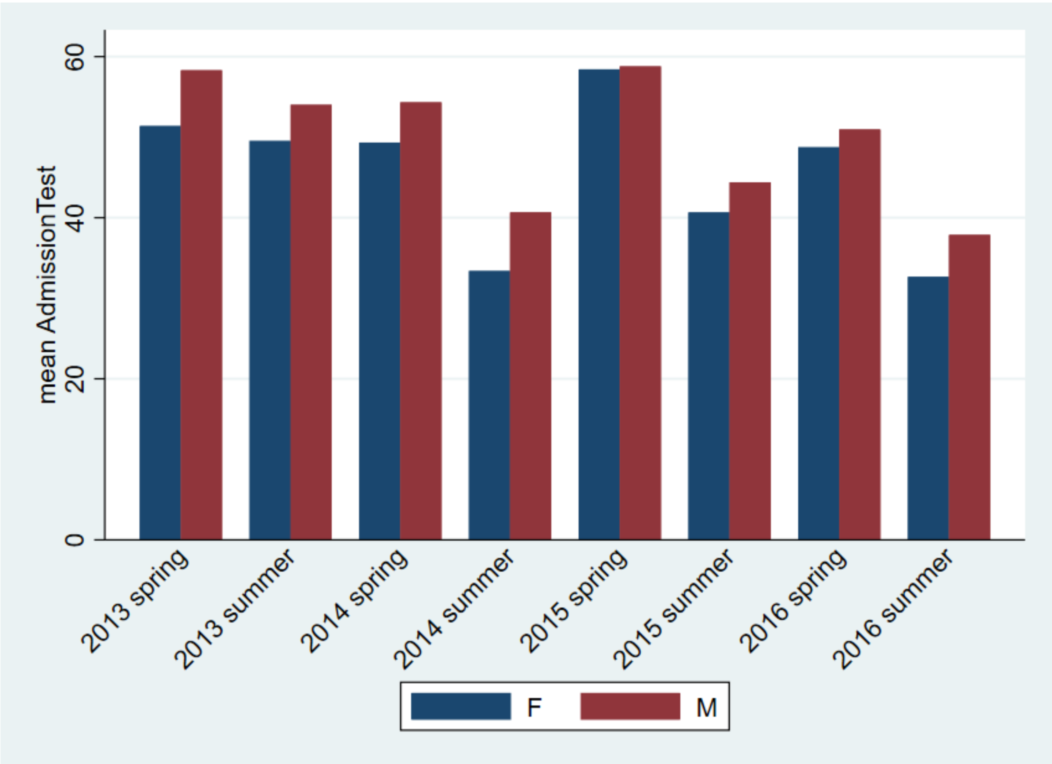


Figure 5: Mean of the Reading and Math subtests scores for the enrolled Males and Females, 2013-2016

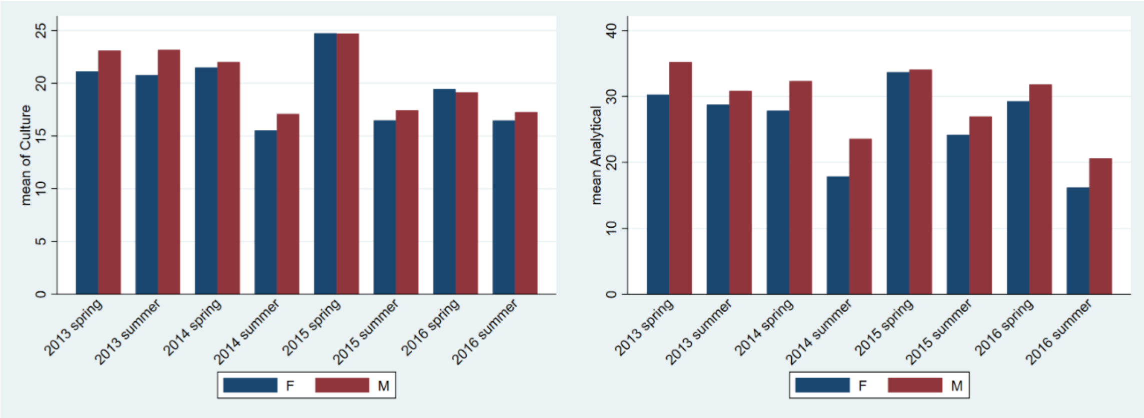
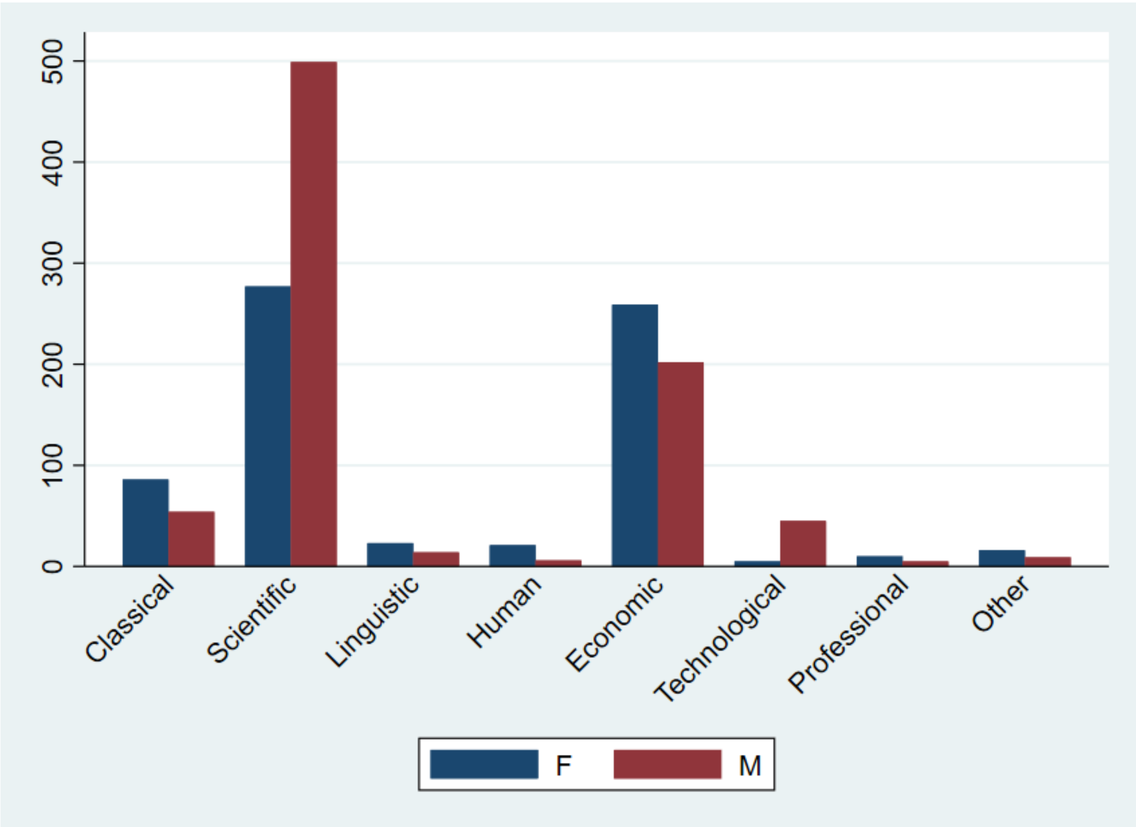


Figure 6: Number of enrolled students per secondary school, divided by males and females



3.1.2.1 Correlation between admission test and high school final grade

In the paragraphs above I showed what are, on average, the applicants' and the enrolled students' high school final grade (HSFG) and admission test (AT) scores. In this section I correlate these two relevant variables and show the differences between the whole population (applicants) and the subpopulation (enrolled students).

Table 4 shows the correlations between the HSFG and the admission test score for the two groups of population for each selection that took place in the four years under analysis. Concerning the enrolled students, there were no problems in doing the correlations because all the required information was available in the datasets. For the candidates there were some problems regarding the spring selections. In fact, most of the applicants who undertake the admission test in spring has not graduated from high school yet, thus the high school final grade is not available for these candidates. For the first two cohorts (2013 and 2014) the mean of the grades, of some selected subjects, obtained in the fourth year of the secondary school were taken into account as a proxy of the high school performance when calculating the final score. For the last two cohorts, instead, the high school performance was completely left out of the creation of the final score. So, the rankings in this case were formed only based on the admission test results. For this reason, it was not possible to recover any information regarding the school performance for the candidates of the spring selections of 2015 and 2016.

Therefore, for the enrolled students and the summer candidates I computed the correlation between the high school final grade and the admission test. For the candidates of the first two spring selections, instead of the HSFG I considered the mean of the grades of the fourth year of secondary school whereas for the last two spring selections it was not possible to compute any correlation.

The results show a positive and mostly stable correlation (around 30%) for the candidates' sample. However, this correlation becomes negative in the subsample of the enrolled students, especially for the first two cohorts (where the admission test and the HSFG contribute equally to the formation of final score). This outcome, although surprising, is just the result of the selection that the department of economics has to implement. In order to be admitted a candidate must obtain a total final score higher than a certain threshold[16], however there are several ways to achieve such threshold. A candidate could do well enough in both the admission test and the HSFG and consequently be admitted; or they could do very well only in one of the two variables considered and still be admitted (of course if the score of the other variable is not too low)

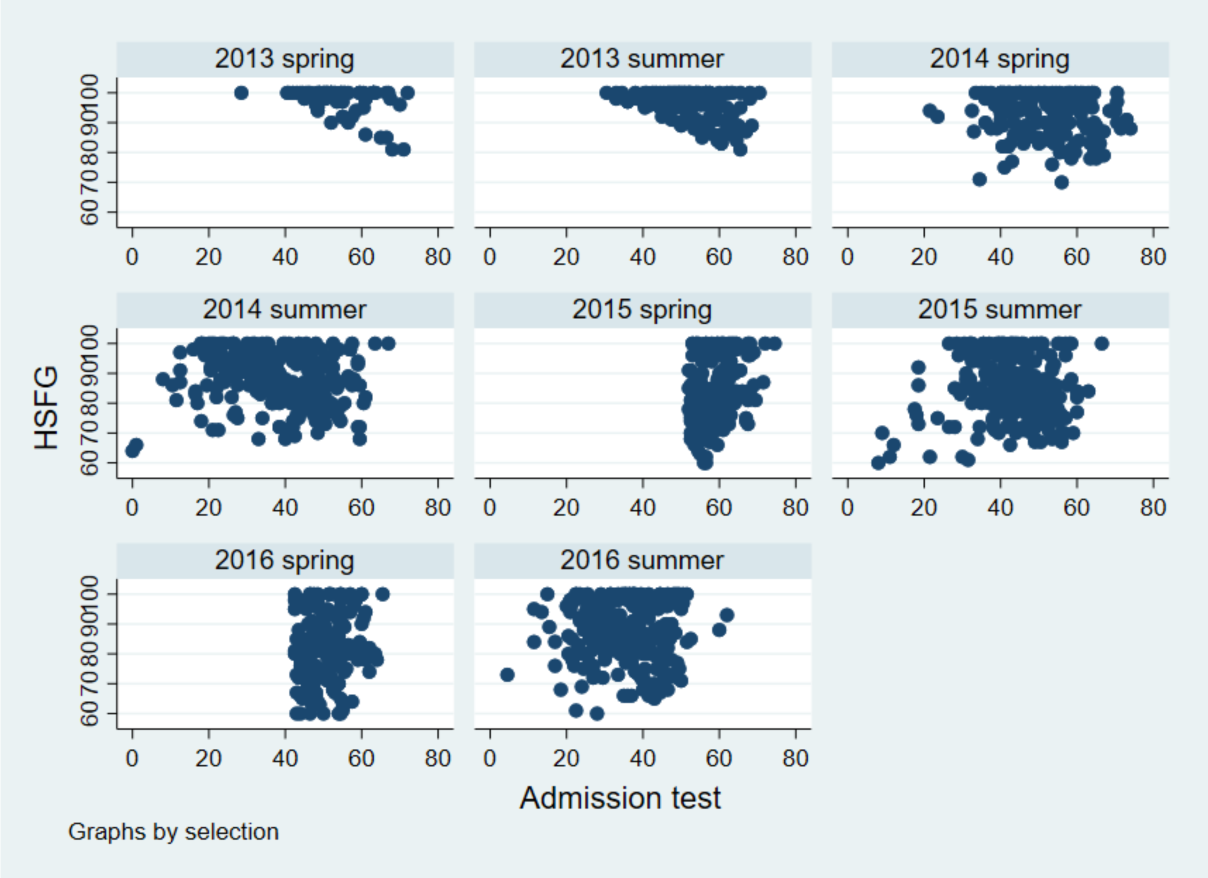
[16] Each year the threshold changes depending on the objective difficulty of the admission test and on the inner abilities and skills of the candidates.

because what really matters is just the total final score. In fact, the selection rule for the first two cohorts could be written as: $0{,}5 * AT + 0{,}5 * HSFG > T$ where T is the minimum threshold that candidates must reach in order to be admitted. Thus, it could happen that the correlation regarding the enrolled students turns out to be negative, especially when the threshold is set very high, that is when very few applicants are admitted. This is the case for the 2013 cohort that is characterized, as I will show in the next paragraph, by the lowest probability for the candidates of being admitted. This cohort indeed presents, for both selections, the highest negative correlations for the enrolled students (around -40%). The 2014 cohort also presents negative correlations, even though relatively smaller (around -15%), and this could be because of the lower threshold level with respect to 2013. For what concerns the last two cohorts, in spring we can see a positive correlation due to the fact that in these selections the admission rule was different: the candidates were admitted only based on their admission test score. Thus, there was only one way of reaching the minimum threshold required and this led to a more homogeneous group of enrolled students. In summer, instead, the correlations return to be smaller, if not negative. Again, the reason is to be found in the way the final score was formed (70 points came from the admission test and 30 points from the high school final grade).
Figure 6 shows the relationship between the admission test and the high school final grade for each selection, confirming the above argumentation.

Table 4: Correlation between the admission test and the high school final grade for the candidates and the enrolled students, 2013-2016

		candidates	enrolled
2013	spring	32,75%	-37,99%
	summer	36,57%	-43,03%
2014	spring	32,94%	-14,60%
	summer	28,88%	-16,58%
2015	spring		26,72%
	summer	22,35%	1,32%
2016	spring		16,83%
	summer	31,86%	-6,97%

Figure 7: Correlation between admission test and high school final grade, by selection, 2013-2016



3.1.2.2 Success rate

In Table 5 it is described the rate of success of the candidates, considering the whole sample and distinguishing between males and females. The rate of success measures the percentage of candidates that were admitted each year among the ones that tried the admission test, and it is computed as follows:

$$Rate\ of\ success\ = \frac{number\ of\ enrolled\ students}{number\ of\ candidates} * 100$$

The rate of success for males and females is:

$$Rate\ of\ success_i = \frac{number\ of\ enrolled\ students_i}{number\ of\ candidates_i} * 100$$

Where i represents either males or females.

The average rate of success is 20,4% but there are some differences that can be noticed between the spring and the summer selection. Each year in spring the success rate is smaller compared

to the success rate of the summer selection. As I explained above, the reasons for this may be the different reallocation system implemented by the department of economics. The average success rate of the spring selection is 15,9%, whereas for the summer selection it is 24,9% with a difference of 9 percentage points. It is also interesting noticing the gap between the success rate of females and males. Females have a higher success rate throughout the four years with the only exception of the spring selections of the years 2015 and 2016. This may be due to the fact that during these two selections, as I pointed out before, the final ranking was formed considering only the admission test score. Females tend to obtain higher grades in secondary school and this can explain why, if the high school GPA is taken into account for the formation of the final score, females tend to reach a higher success rate on average. On the other hand, when the final high school grade is omitted than females' success rate decreases sharply.
Table 5 shows the success rate of the candidates in each selection of the four years and, in the last row, the difference between the success rate of females and males. This difference is almost always positive, it takes negative values only during April of 2015 and 2016 where the success rate for females was respectively 10% and 15% lower than the males' success rate.

Table 5: Success rate of the candidates, 2013-2016

	2013		2014		2015		2016		Total
	April	August	May	August	April	August	April	August	
Success rate	9,4%	16,8%	16,5%	25,9%	19,5%	27,7%	18,1%	29,3%	20,4%
Female success rate	11,2%	19,2%	17,7%	27,6%	14,0%	28,8%	9,9%	31%	19,9%
Male success rate	7,8%	14,7%	15,5%	24,3%	23,7%	26,9%	24,9%	27,7%	20,7%
Δ	3,4%	4,4%	2,2%	3,3%	-9,7%	1,9%	-15%	3,4%	

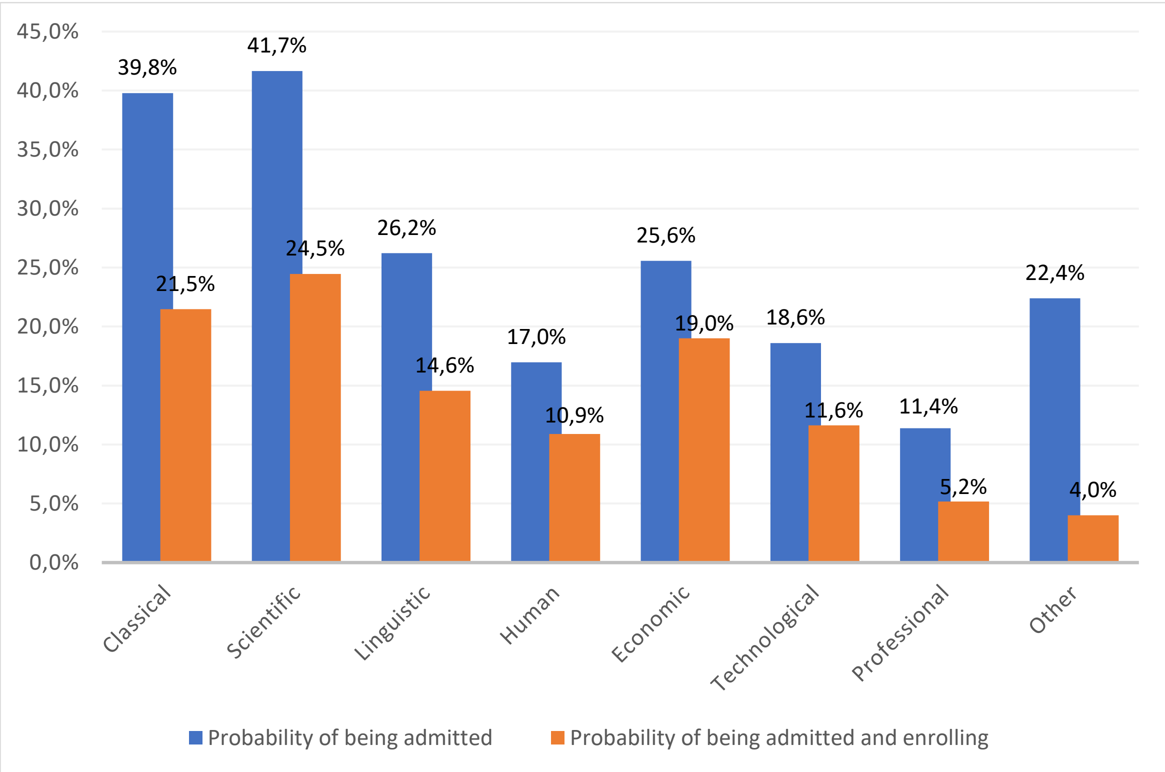
The scholastic background is also relevant for the determination of the success rate. In figure 8 the blue bar represents the probability[17] of the candidates of being admitted (regardless of their subsequent decision to actually enrol or not) in the first and second ranking during the whole 4-year period[18] for each high school type. Candidates coming from a scientific and a classical lyceum have higher probabilities of being admitted (respectively: 41,7% and 39,8%). They are

[17] Computed as follows: $\frac{number\ of\ admitted\ candidates\ _i}{number\ of\ total\ candidates_i} * 100$, where i is the type of secondary school.

[18] The scholastic background is available for all the selections except for the spring selections of 2013 and 2014.

followed by the linguistic lyceum (26,2%), the economic institute (25,6%), the technological institute (18,6%), the human sciences lyceum (17%) and finally the professional institute students (11,4%). The orange bar shows, instead, the probability of the candidates of being admitted in the first and second ranking and consequently enrolling. As before, the highest probability is obtained by the candidates that come from a scientific lyceum (24,5%), the classical lyceum and the economic institute students follow behind with a success rate of respectively 21,5% and 19%. Candidates coming from other schools (linguistic lyceum, human science lyceum, technological institute and professional institute), on the other hand, have lower probabilities of being admitted and enrolled (in this case the success rate ranges from 7,3% to 16%). Looking at the difference between the two probabilities it is clear that not all the admitted candidates decide to enrol. In particular, this difference is more pronounced for the students arriving from a classical, scientific or linguistic lyceum, which means that even though they have the highest success rate, they also have the highest probability of choosing not to enrol.

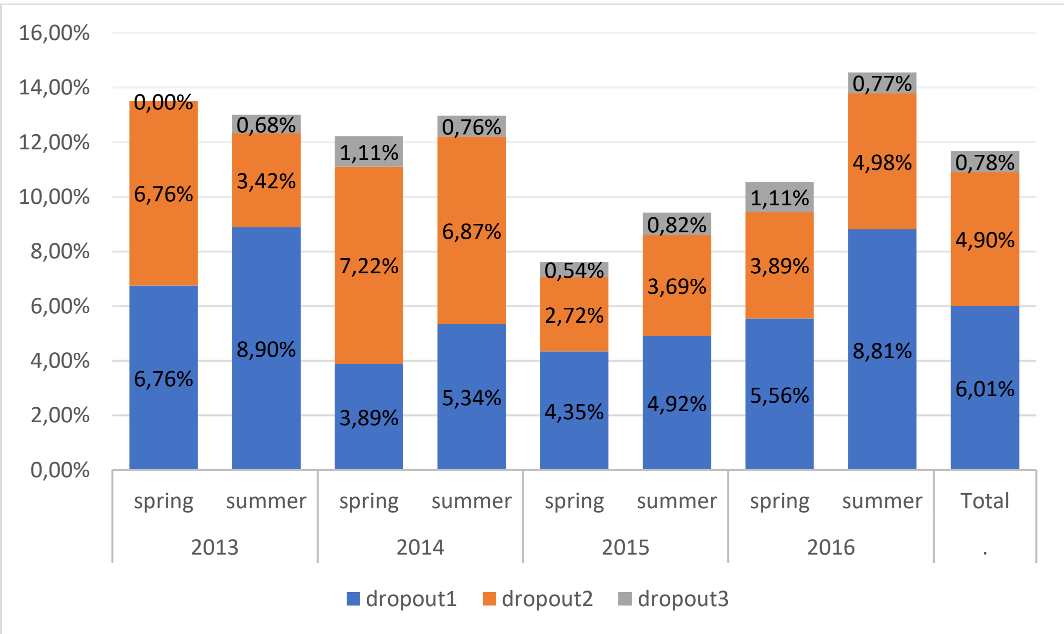
Figure 8: Success rate of students based on their scholastic background



3.1.2.3 Dropout rate

It is always desirable by universities to have low rates of dropout, which would mean investing their resources on students that are really willing to complete their academic career. Thus, given its importance, I analysed the dropout rate[19] during the first three years of college for the cohorts from 2013 to 2016. Figure 9 shows the percentage of students that dropped out of college during their first three years distinguishing between those who did the admission test in spring from those who did it in summer. From a general point of view it is clear that students tend to drop out more during the first two years of college, with an average dropout rate of 6% during the first year and 5% during the second year, while for the last year it falls below 1%. Observing only the first year of college it seems that, for each cohort, students who did the admission test in spring tend to dropout less than those who did it in summer. This pattern, however, is not present in the second and third years of college. Nonetheless, also when considering the three years altogether it seems that, except for the 2013 cohort, the dropout rate is lower if the admission test was taken in spring. On average 11,7% of students drop out of college during the first three years. The highest dropout rate (14,6%) was reached by the students of the 2016 cohort who did the admission test in summer, whereas the lowest (7,6%) was achieved in 2015 by the students who passed the admission test in spring.

Figure 9: Dropout rate of students, 2013-2016

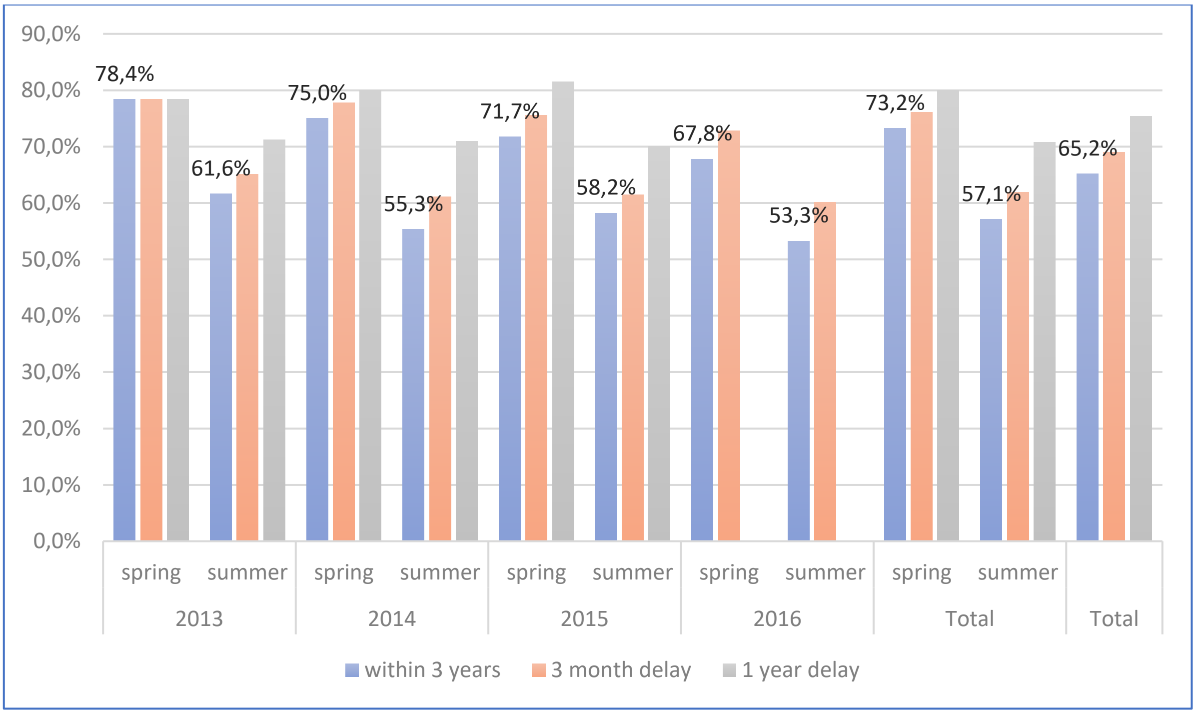


[19] The dropout rate was computed as: $\frac{(number\ of\ students\ that\ dropped\ out)_{ij}}{(number\ of\ enrolled\ students\ in\ year1)_i} *$, where i represents in which selection students undertook the admission test and j is the year during which students dropped out.

3.1.2.4 Graduation rate

At the department of Economics of the University of Padova, graduating in time means to graduate by December of the third year of a student's academic career. Following this definition, on average 65,2% of the enrolled students manage to graduate on time. This percentage rises above 75% if we also consider the students that graduate with a one-year delay. Figure 10 shows the differences between the four cohorts and, within each cohort, between the groups of students that did the admission test in spring and in summer. For each subgroup it is shown the percentage of students that graduated within three years (blue bar), with a small delay[20] (orange bar) and within four years[21] (grey bar). For the 2016 cohort there is available data only until March 2020, so it was not possible to compute the one-year delay graduation rate. The data show that the first subgroup, the one that did the admission test in spring, has higher graduation rates than the second subgroup and this is true for all the four cohorts. In fact, on average, the graduation rate of the first subgroup is above 73% whereas for the second subgroup it drops to 57,1%.

Figure 10: Graduation rate of the enrolled students by cohort, 2013-2016

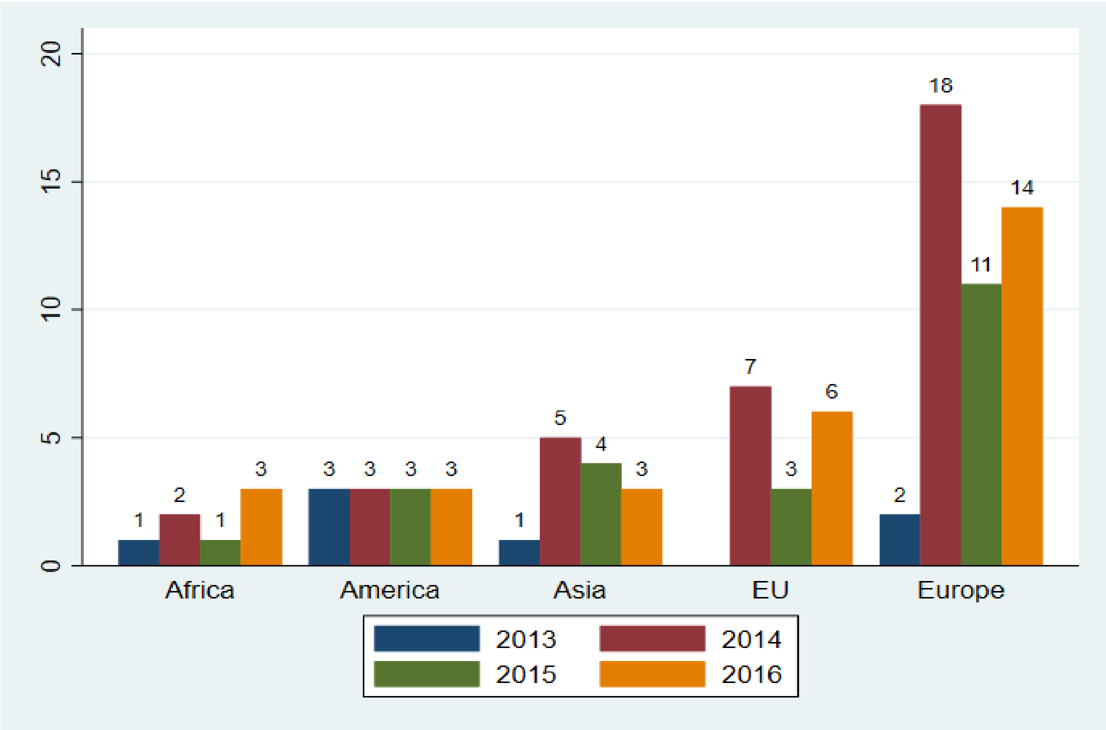


[20] Not later than the first call (Sessione di laurea) of their fourth year of college, that usually takes place in March or April.
[21] By December of their fourth year of college.

3.1.2.5 Where do students come from?

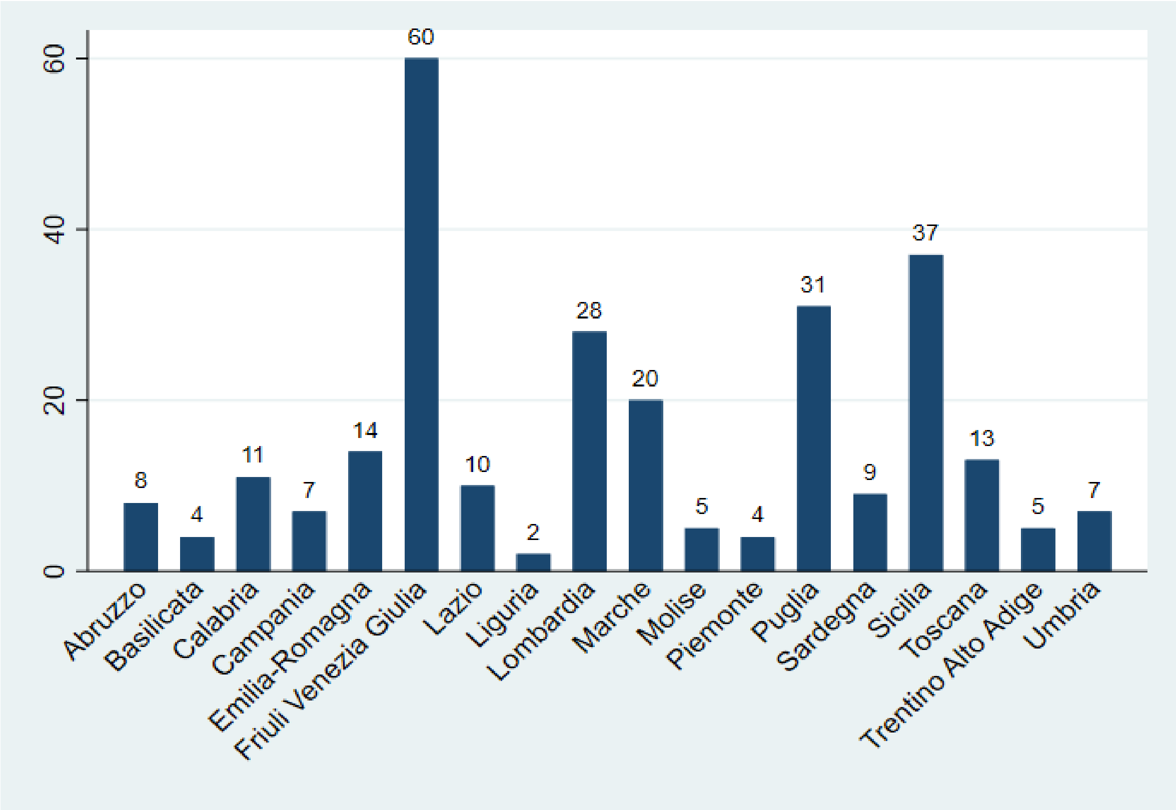
Considering all the four cohorts together, that is all the 1530 students, the great majority (93,9%) come from Italy, the remaining 6% (that is 93 students) have foreign origins. Figure 11 contains information about the number of enrolled students of the four cohorts born outside of Italy. 45 students come from European countries outside the EU, especially from Albania (18) and Moldova (19). Also those arriving from Romania (10) and from China (11) can be considered relatively big groups of students.

Figure 11: Number of students with foreign origins, 2013-2016



For what concerns the region of origin in Italy of the enrolled students, there is available data for 1521 students of which 1246 (82%) come from Veneto. The remaining students are distributed through the other regions as shown in figure 12. The regions that are more represented in this group of students are Friuli Venezia Giulia (60 students), Sicilia (37), Puglia (31) and Lombardia (28).

Figure 12: Regions of origin of the enrolled students, excluding Veneto

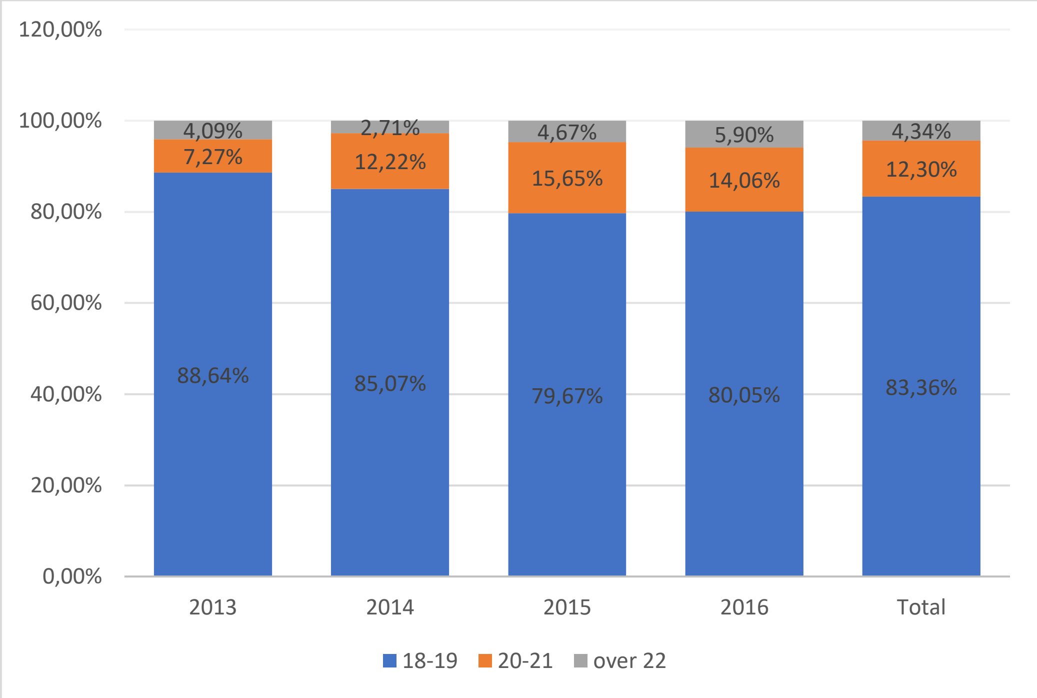


3.1.2.6 Students' age

Most students begin their academic carrier immediately after the end of high school, which means at the age of 18 or 19. In fact the youngest students represent the 83% of the total number of students enrolled in the four years. Students that enrol at the age of 20 or 21 are on average the 12% of the sample while the remaining 4% of students are older.

Over the four years, the share of the youngest students decreased slightly going from an 88% share in 2013 to an 80% in 2016, whereas the students of age 20-21 increased from a 7% share in 2013 to 14% in 2016. The students older than 22 years remained roughly constant throughout the four years. These results are shown in figure 13.

Figure 13: Students' age at enrolment, by cohort



3.2 Variables description

This work has the aim of analysing to what extend the admission test can predict the academic performance of future students at the department of Economics in Padova.
I will present here the outcome, treatment and control variables that I used for my analysis.

3.2.1 Dependent variables

To assess the performance of students during their academic career I computed the weighted grade point average for each of the three years.
First I computed the weighted mean of grades obtained during one year, and then I weighted this average with a multiplier.
For the computation of the weighted mean I considered all the exams undertaken during the winter, the summer and the autumn exam sessions of a year. For the third year, given that students should finish their exams in the summer session, I did not consider the exams taken later in the autumn session.
I computed the multiplier dividing the number of credits obtained by a student in their first year by the number of credits that the student should have obtained during that year. This value is equal to zero if no exams were passed and 1 if all the required exams were undertaken

successfully. For the second (third) year of college, I computed the multiplier dividing the number of credits that a student obtained during their first two (three) years by the number of total credits that the student should have obtained in that period of time. Also in this case I take into account the credits obtained in all three sessions for the first two years and the credits obtained till summer for the last year.

Using this multiplier means assigning a value of zero to all the exams that were not passed during a year, which is a very severe criterion but, given that I apply this method to all the students, it is not discriminatory and affects equally everyone.

Both the weighted mean of grades and the number of credits obtained in a year are indicators of college performance, thus by using also the multiplier I obtain a value that is more informative than the single indicators.

Thus, the value that I used as an indicator of performance is:

$$GPA_{ji} = weighted\ mean\ of\ grades_{ji} * \frac{credits\ obtained_{ji}}{credits\ required_{i}}$$

where GPA_{ji} is the weighted grade point average obtained in year i by student j.

In addition to the weighted GPA, I also considered the final grade obtained by the students at the end of their academic career as an indicator of performance.

3.2.2 Treatment and control variables

I will consider four independent variables in my analysis. The first one is the admission test, that ranges from 0 to 80 points, of which I want to find the predictive ability of the college performance.

I will also consider the final score, which is, as I explained before, the score candidates achieve as a result of both the admission test and the high school grades that are weighted differently throughout the years. The final score can range from 0 to 100.

Finally I will unpack the admission test into its two subtests (the first one measures the ability in reading and general knowledge, the second one assesses the analytical and quantitative abilities) and consider them separately as independent variables to see if one of them has a higher predictive power then the other.

In the attempt of clearing the effect of the independent variables, I will also use some controls. The high school final grade (HSFG) which ranges from 60 to 100 and the age of the enrolled students are continuous control variables. I will also consider some categorical variables:

- Female, a dummy that takes value of 1 if sex is female, 0 otherwise.

- Lyceum, a dummy that takes value of 1 if the students attended a lyceum (classical, scientific, linguistic or human sciences) before enrolling in college, 0 if they attended any other school.
- South, a dummy that takes value of 1 if the students come from a region of the south of Italy, 0 if they come from the centre or north of Italy.

The battery of control variables is not very abundant, but unfortunately it was not possible to retrieve other useful variables from the datasets in my possession.

3.3 Methodology

The aim of the work is to assess the predictive ability of the admission test on the college performance, but before doing so some notes must be made.

Since the sample of students is quite variegate, especially when observing separately the group of students that did the admission test in spring and the one that did it in summer, as I showed in the paragraphs above, when doing the analysis and regressions I will also distinguish between these two subgroups, in addition to the distinction between cohorts.

I also want to point out that the analysis on the 2013 cohort must be taken with caution for a series of reasons. First, the number of observations is halved with respect to the other cohorts and second, as I explained in chapter 2, the degree program and therefore the courses offered are not the same as for the other cohorts. This could lead to a problem in the comparison between cohorts and thus all the results attained for the 2013 cohort will be added in the Annex.

In order to have an homogeneous sample of students, in my analysis I will consider only the bundle of exams that were undertaken during the first three years of college, and in any case no later than the summer session of the third year. This means that all the exams that were successfully passed in later sessions will be omitted and therefore, for some students, the information employed in my analysis will be partial.

It is also important to notice that the number of students that enrol in the first year of college is not the same number that arrives at the third year and graduates. Some students decide to dropout, some others remain enrolled but do not undertake any exams, they just pay tuition fees, others decide to change university. Thereby each year the number of observations decreases slightly.

Thus, for those who decided to enrol but for some reason did not continue their studies, the data regarding their academic performance is missing and, of course, it is not possible evaluating the predictive ability of the admission test on these students' performance. It is only possible to see if the admission test can predict the probability of dropping out of college.

Since there is a group of students that dropped out, and a group of students that continued their academic career, I will use a two-part model to evaluate, first, if the admission test can predict the probability of dropping out of students and, second, the predictive ability of the admission test on the remaining students' performance.

For the first part I exploit a Probit model, which is a nonlinear regression where the dependent variable is a binary variable (Stock and Watson, 2011). As dependent variable I use the dummy Dropout that takes value of 1 if the student dropped out during the first, the second or the third year and 0 otherwise and construct the following model:

$$\Pr\left(Dropout_{ji} = 1 \,\middle|\, AT_j, X_j\right) = \varphi(\alpha + \beta_1 AT_j + \beta_2 X_j)$$

Where j represents each student and i is one of the three years during which a student could drop out, whereas AT is the admission test and X contains the control variables (HSFG, Lyceum, Female, Age, South).

The second part of the model aims to estimate, with an OLS regression, how well can the admission test predict the academic performance of students, considering only the sample of students that do not drop out. In addition to the admission test, I will also assess the predictiveness of the total final score and the two subtests (Reading and Math). As dependent variables I will consider the GPA of the first three years, computed as shown in the paragraph above, and the final grade that students obtain at the end of their career.

The OLS regression is:

$$Outcome_{ji} = \alpha + \beta_1 Explanatory\ variable_j + \beta_2 X_j + \varepsilon_{ji}$$

Where $Outcome_{ji}$ represents the GPA of the three years and the final grade of student j that belongs to one of the two subgroups of the four cohorts. The independent variable can represent either the admission test, the final score or the two subtest results of student j and X is the bundle of control variables.

Since I am interested in the ability of the independent variables, especially the admission test, of predicting the academic performance, analysing only the OLS results is not enough. To see how well a model predicts the variation in the dependent variable, the coefficient of multiple determination (R^2) must be given particular attention. But in a multiple regression model the R^2 represents the variance explained by the whole sample of explanatory variables. Thus, I will

decompose the R^2 to see to which extent each explanatory variable contributes to the variance of the dependent variables.
The decomposition of the R^2 is computed as follows:

$$R^2 = \sum_{j=1}^{k} a'_j r_{yx_j}$$

Where a'_j is the standardized regression coefficient[22] of the j-th explanatory variable and r_{yx_j} is the simple correlation coefficient between y, the dependent variable, and x_j (Borcard, 2002).
It must be acknowledged that the contribution of each explanatory variable is not independent from the influence of the other variables. Given that the explanatory variables are not orthogonal to each other, but they are more or less intercorrelated, each independent variable explains a portion of the R^2 that to some extent overlaps with the portions explained by the other variables. In other words, the explanatory variables are able to influence each other and do part of the job of the others. This means that, even though the decomposition process provides a general overview of the contributions of the explanatory variables, these contributions are a result not only of each independent variable but also of the intercorrelation between these variables.
For this reason, alongside with the decomposition of the R^2, I will also present in this work the partial r^2 of the independent variables, that is the coefficient of partial determination. The partial r^2 measures the mutual relationship between two variables, the dependent variable (y) and an explanatory variable (x_1), while all the other variables (X_j) are held constant with respect to both variables (y and x_1) (Borcard, 2002). The partial r^2 can also be explained in terms of residuals, in fact it can be obtained by looking at the r^2 of the regression of the residuals of y with respect to X_j on the residuals of x_1 with respect to X_j. By doing so the effect of X_j on y and x_1 is completely eradicated and all that remains is the relationship between y and x_1.
To sum up, the three steps to obtain the partial r^2 are:

1. Regress y on X_j and keep the residuals.
2. Regress x_1 on X_j and keep the residuals.
3. Regress the residuals of y on the residuals of x_1 and look at the r^2 of this regression.

The r^2 found in the last regression estimates the proportion of the unexplained variation of y that becomes explained when adding x_1 to the model (Borcard, 2002).

[22] Computed by multiplying the unstandardized regression coefficient of the explanatory variable by the ratio of the standard deviations of the explanatory variable and the dependent variable.

Chapter 4

Results

In this chapter I present the results of my analysis concerning the cohorts from 2014 to 2016. As I said before, I leave the 2013 cohort results out, given the fewer number of observations and the different structure of the degree program for this cohort.

4.1 Probit results

The probit regression estimates the probability of dropping out during one of the first three years of college. For this analysis I consider each cohort as a whole, in fact distinguishing between the two subgroups of students (those who did the admission test in spring and those who did it in summer) is pointless given the modest number of students that actually drop out each year. The main independent variable that I consider here is the admission test, the other explanatory variables are the high school final grade, age and the dummies for school, sex and region of origin. Tables 6, 7 and 8 show the probit regressions for the cohorts 2014, 2015 and 2016. The three columns represent the probability of dropping out during the first three years of college. It seems that neither the admission test nor the high school final grade can predict the probability of dropping out. The coefficients of these explanatory variables are, in fact, mostly negative, but very close to zero and not statistically significant.

Table 6: Probit results for cohort 2014

	(1)	(2)	(3)
	Dropout1	Dropout2	Dropout3
AT	-.011	-.004	-.016
	(.009)	(.009)	(.022)
HSFG	-.001	-.001	.09
	(.014)	(.012)	(.061)
Lyceum	-.048	-.082	-.371
	(.269)	(.238)	(.573)
Age	.131**	-.292	-1.034
	(.054)	(.249)	-1.583
Female	.048	.167	-.867

	(.226)	(.197)	(.538)
South	-.006	-.299	
	(.358)	(.366)	
_cons	-4.53**	6.128	16.333
	-1.866	-6.419	-39.478
Obs.,	436	416	341
R-squared	.z	.z	.z

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 7: Probit results for cohort 2015

	(1)	(2)	(3)
	Dropout1	Dropout2	Dropout3
AT	-.019	.009	.015
	(.012)	(.014)	(.031)
HSFG	-.02*	.002	-.006
	(.012)	(.014)	(.031)
Lyceum	.485	-.438	-.597
	(.345)	(.336)	(.726)
Age	.078***	.015	-.605
	(.029)	(.056)	-1.064
Female	-.364	.182	.175
	(.271)	(.265)	(.528)
South	.035	.868**	
	(.496)	(.383)	
_cons	-1.274	-2.684	12.058
	-1.321	-2.017	-26.114
Obs.,	426	406	370
R-squared	.z	.z	.z

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 8: Probit results for cohort 2016

	(1)	(2)	(3)
	Dropout1	Dropout2	Dropout3
AT	-.014	-.025**	-.016
	(.011)	(.012)	(.024)
HSFG	-.036***	.007	-.035
	(.011)	(.012)	(.025)
Lyceum	-.09	.133	-.902
	(.232)	(.265)	(.585)
Age	.048	.044	-.23

	(.042)	(.045)	(.294)
Female	.481**	-.363	-.324
	(.237)	(.267)	(.537)
South	.339	.504	
	(.559)	(.47)	
_cons	.738	-2.281	7.102
	-1.453	-1.639	-7.872
Obs.,	438	407	376
R-squared	.z	.z	.z

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

4.2 OLS results

In this section I present the results of the OLS regressions for the grade point average (GPA) of the first three years of college and for the final grade obtained at the end of the academic career. As I already stressed out before, I only consider the academic career of the first three years of college and leave out all the observations regarding exams that students took and passed later on. For what concerns the final grade, for each cohort I keep the observations regarding the students that graduated during one of the three calls available during the third year of college (July, October and December) and during the first call of the fourth year (March/April). This choice is straightforward because, regarding the final grade, there is available data only until the call of March 2020. For the 2016 cohort this is the fourth and last call available. Thus, in order to make the results of the different cohorts comparable, I consider four calls for graduation also for the other cohorts.

For the OLS regression I distinguish between the students that did the admission test in spring and in summer.

As main independent variables, besides the admission test, I use the subtest scores (Math and Reading) and the total final score.

The results presented below show a relatively low R^2, in fact it ranges between 0.08 and 0.33, which means that all the explanatory variables together can explain more or less 20% of the variation of the dependent variables.

4.2.1 Admission test

Tables 9, 10 and 11 show the results of the OLS regressions when the main independent variable is the admission test. It seems that both the admission test and the high school final grade are positive and statistically significant which means that obtaining a higher score at the admission test and having a higher high school final grade lead to a higher GPA at college. This is true for all the cohorts and for both subgroups of students (spring and summer). The other explanatory variables have rarely significant coefficients. Age seems to have a negative impact on GPA, meaning that older students tend to perform worse than younger students at college. The school dummy, instead, seems to have a slightly positive effect, meaning that students that come from a lyceum tend to be more successful in college.

Table 9: OLS results for cohort 2014, Admission test

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
AT	0.086*	0.167***	0.122***	0.110***	0.124***	0.114***	0.296***	0.183**
	(0.049)	(0.043)	(0.040)	(0.041)	(0.041)	(0.037)	(0.083)	(0.075)
HSFG	0.205***	0.185***	0.194***	0.212***	0.151***	0.233***	0.461***	0.356***
	(0.063)	(0.054)	(0.051)	(0.051)	(0.052)	(0.047)	(0.104)	(0.100)
Lyceum	0.869	-0.141	1.713**	0.616	1.398	0.793	0.424	2.454
	-1.039	(0.973)	(0.831)	(0.915)	(0.853)	(0.825)	-1.798	-1.670
Age	-0.969	-0.456	0.733	-1.173**	0.712	-0.370	0.514	-1.368
	(0.664)	(0.498)	-1.126	(0.463)	-1.151	(0.480)	-2.320	-1.049
Female	-1.036	-1.444*	0.509	-0.451	0.551	-0.168	-1.313	-0.297
	(0.845)	(0.835)	(0.692)	(0.798)	(0.715)	(0.720)	-1.461	-1.418
South	-0.704	0.357	1.773	-0.233	2.032	-1.057	2.424	-1.853
	-1.710	-1.215	-1.453	-1.109	-1.489	-1.008	-2.922	-2.005
_cons	23.639	8.604	-20.710	26.256*	-14.287	6.766	31.599	92.100***
	-17.580	-14.536	-28.977	-13.663	-29.650	-13.806	-59.485	-30.349
Obs.,	170	232	154	206	152	200	140	158
R-squared	0.110	0.136	0.194	0.160	0.155	0.162	0.226	0.152

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 10: OLS results for cohort 2015, Admission test

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
AT	0.208**	0.235***	0.145*	0.226***	0.096	0.253***	0.359***	0.412***
	(0.086)	(0.051)	(0.085)	(0.049)	(0.079)	(0.048)	(0.132)	(0.100)
HSFG	0.136***	0.149***	0.197***	0.197***	0.153***	0.206***	0.398***	0.257***
	(0.044)	(0.044)	(0.044)	(0.044)	(0.041)	(0.044)	(0.071)	(0.083)
Lyceum	2.415*	-0.745	3.166**	0.688	1.569	-0.225	6.319***	1.629
	-1.225	-1.023	-1.224	(0.979)	-1.166	(0.967)	-2.049	-1.769
Age	-0.162	-0.434**	-0.341	-0.471**	-0.279	-0.468**	0.468	-0.545
	(0.342)	(0.179)	(0.326)	(0.191)	(0.313)	(0.187)	-1.291	(0.705)
Female	0.477	0.364	-0.150	0.291	0.044	0.205	-0.739	1.015
	(0.836)	(0.851)	(0.822)	(0.807)	(0.767)	(0.801)	-1.298	-1.472
South	2.271	-3.773**	0.994	-1.839	0.885	-1.664	0.790	-1.688
	-2.033	-1.518	-1.947	-1.663	-1.803	-1.628	-2.939	-2.930
_cons	0.329	8.699	3.485	5.138	11.924	6.090	30.476	70.577***
	-9.493	-6.744	-9.108	-6.835	-8.684	-6.742	-32.160	-20.084
Obs.,	174	213	164	194	160	190	139	149
R-squared	0.164	0.224	0.214	0.275	0.146	0.303	0.321	0.189

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 11: OLS results for cohort 2016, Admission test

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
AT	0.096	0.286***	0.007	0.256***	-0.025	0.175***	0.210	0.496***
	(0.092)	(0.053)	(0.082)	(0.051)	(0.081)	(0.048)	(0.142)	(0.090)
HSFG	0.253***	0.234***	0.208***	0.191***	0.148***	0.158***	0.332***	0.219**
	(0.050)	(0.050)	(0.045)	(0.048)	(0.045)	(0.047)	(0.079)	(0.088)
Lyceum	1.602	1.619*	2.120**	1.137	1.667	2.260**	3.284*	2.488
	-1.201	(0.970)	-1.064	(0.921)	-1.047	(0.874)	-1.826	-1.655
Age	-0.990**	-0.070	-1.361*	-0.082	-1.592*	-0.267	-3.010	-0.047
	(0.432)	(0.212)	(0.822)	(0.228)	(0.807)	(0.214)	-2.309	(0.620)
Female	0.153	0.926	0.162	-0.393	0.102	-0.738	0.604	3.193**
	-1.131	(0.957)	-1.027	(0.898)	-1.013	(0.869)	-1.702	-1.577
South		-3.372*		-1.606		-1.714		1.302
		-1.924		-1.963		-1.843		-3.488
_cons	16.684	-11.716*	34.359*	-3.697	49.344**	8.745	129.165**	57.566***
	-11.946	-6.883	-20.145	-7.023	-19.800	-6.649	-53.330	-16.368
Obs.,	167	223	158	199	155	193	131	156
R-squared	0.226	0.232	0.185	0.215	0.122	0.184	0.192	0.249

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

4.2.2 Subtests (Math and Reading)

If instead of observing the total score of the admission test we unpack it into its two subtests, new findings might be pursued. In fact, when considering as main independent variables the subtest aimed to assess the analytical and quantitative abilities (Math) separately from the subtest aimed to verify the general knowledge and reading abilities (Reading), the regressions show clearly different results.

The Math subtest presents positive and significant coefficients throughout the three cohorts and for both subgroups of students for all the dependent variables considered. This cannot be said for the Reading subtest which presents mostly non-significant effects. The Reading subtest seems to be less relevant for the determination of the future students' success in college.

The regression results of the Math and Reading subtests are shown in tables 12-17.

Table 12: OLS results for cohort 2014, Math

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
Math	0.201***	0.278***	0.217***	0.134***	0.228***	0.101**	0.427***	0.252***
	(0.068)	(0.052)	(0.055)	(0.052)	(0.056)	(0.047)	(0.116)	(0.094)
HSFG	0.221***	0.203***	0.215***	0.216***	0.173***	0.234***	0.508***	0.363***
	(0.061)	(0.052)	(0.049)	(0.052)	(0.051)	(0.048)	(0.104)	(0.100)
Lyceum	0.336	-0.157	1.455*	0.939	1.094	1.348*	0.463	2.991*
	-1.022	(0.906)	(0.813)	(0.874)	(0.831)	(0.797)	-1.779	-1.562
Age	-0.824	-0.408	0.895	-1.149**	0.898	-0.370	0.910	-1.397
	(0.656)	(0.485)	-1.105	(0.464)	-1.127	(0.487)	-2.324	-1.042
Female	-0.705	-1.063	0.784	-0.421	0.864	-0.319	-0.934	-0.049
	(0.839)	(0.818)	(0.688)	(0.803)	(0.709)	(0.732)	-1.481	-1.427
South	-0.347	0.339	1.903	-0.308	2.226	-1.251	2.101	-1.907
	-1.657	-1.180	-1.406	-1.107	-1.435	-1.018	-2.874	-1.993
_cons	17.146	6.026	-26.925	26.366*	-21.371	8.619	19.583	93.301***
	-17.494	-14.125	-28.515	-13.681	-29.071	-14.019	-59.784	-29.972
Obs.,	170	232	154	206	152	200	140	158
R-squared	0.140	0.181	0.225	0.159	0.193	0.140	0.231	0.158

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 13: OLS results for cohort 2015, Math

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
Math	0.240**	0.383***	0.163	0.333***	0.206**	0.373***	0.531***	0.500***
	(0.105)	(0.068)	(0.105)	(0.067)	(0.097)	(0.066)	(0.171)	(0.136)
HSFG	0.149***	0.154***	0.207***	0.207***	0.148***	0.219***	0.396***	0.278***
	(0.043)	(0.043)	(0.042)	(0.043)	(0.040)	(0.043)	(0.070)	(0.084)
Lyceum	2.258*	-0.635	3.045**	0.967	1.058	0.123	5.095**	2.726
	-1.250	(0.950)	-1.254	(0.931)	-1.192	(0.917)	-2.132	-1.703
Age	-0.177	-0.329*	-0.349	-0.360*	-0.349	-0.343*	0.250	-0.593
	(0.345)	(0.177)	(0.329)	(0.191)	(0.312)	(0.186)	-1.288	(0.712)
Female	0.452	0.544	-0.157	0.361	0.139	0.278	-0.684	0.943
	(0.837)	(0.830)	(0.823)	(0.801)	(0.761)	(0.794)	-1.287	-1.493
South	2.381	-3.502**	1.042	-1.963	1.180	-1.821	1.398	-2.611
	-2.043	-1.486	-1.956	-1.644	-1.793	-1.607	-2.937	-2.927
_cons	3.834	5.600	5.920	2.366	13.159	2.755	40.043	73.989***
	-9.291	-6.668	-8.917	-6.884	-8.427	-6.793	-31.868	-20.203
Obs.,	174	213	164	194	160	190	139	149
R-squared	0.161	0.259	0.211	0.286	0.162	0.316	0.332	0.171

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 14: OLS results for cohort 2016, Math

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
Math	0.202**	0.500***	0.142	0.384***	0.043	0.217***	0.338**	0.683***
	(0.093)	(0.074)	(0.088)	(0.074)	(0.088)	(0.071)	(0.153)	(0.130)
HSFG	0.261***	0.287***	0.202***	0.234***	0.141***	0.189***	0.361***	0.300***
	(0.046)	(0.049)	(0.042)	(0.047)	(0.042)	(0.046)	(0.075)	(0.088)
Lyceum	1.584	1.862**	1.851*	1.593*	1.506	2.749***	3.423*	3.385**
	-1.165	(0.910)	-1.038	(0.885)	-1.032	(0.849)	-1.786	-1.624
Age	-0.926**	-0.149	-1.349	-0.158	-1.587*	-0.321	-3.114	-0.062
	(0.429)	(0.203)	(0.815)	(0.227)	(0.806)	(0.216)	-2.285	(0.625)
Female	0.383	1.273	0.628	-0.351	0.357	-0.956	0.927	2.840*
	-1.100	(0.925)	-1.011	(0.895)	-1.006	(0.873)	-1.680	-1.573
South		-3.654*		-2.256		-2.221		0.075
		-1.861		-1.948		-1.853		-3.508
_cons	13.031	-13.813**	30.607	-3.930	47.296**	9.537	128.917**	55.881***
	-11.809	-6.647	-19.970	-6.997	-19.800	-6.721	-52.740	-16.540
Obs.,	167	223	158	199	155	193	131	156
R-squared	0.243	0.279	0.199	0.221	0.123	0.168	0.209	0.238

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 15: OLS results for cohort 2014, Reading

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
Reading	-0.055	-0.033	0.038	0.078	0.023	0.165**	0.263	0.081
	(0.090)	(0.078)	(0.074)	(0.073)	(0.076)	(0.066)	(0.161)	(0.141)
HSFG	0.224***	0.158***	0.206***	0.200***	0.166***	0.216***	0.459***	0.323***
	(0.064)	(0.055)	(0.053)	(0.052)	(0.055)	(0.047)	(0.110)	(0.101)
Lyceum	1.800*	1.685*	2.670***	1.343	2.366***	1.173	2.166	3.884**
	(0.973)	(0.953)	(0.803)	(0.893)	(0.826)	(0.806)	-1.774	-1.668
Age	-0.974	-0.536	0.357	-1.225***	0.317	-0.475	-0.310	-1.673
	(0.671)	(0.514)	-1.153	(0.470)	-1.179	(0.483)	-2.393	-1.062
Female	-1.414*	-2.257***	0.014	-0.941	0.025	-0.615	-2.489*	-0.927
	(0.831)	(0.837)	(0.693)	(0.784)	(0.715)	(0.703)	-1.468	-1.420
South	-1.844	-0.119	0.510	-0.510	0.687	-1.196	0.124	-2.321
	-1.687	-1.253	-1.457	-1.120	-1.491	-1.014	-2.940	-2.034
_cons	27.351	19.275	-7.190	31.465**	-0.113	12.611	61.742	107.946***
	-17.580	-14.826	-29.514	-13.700	-30.182	-13.731	-60.961	-30.270
Obs.,	170	232	154	206	152	200	140	158
R-squared	0.095	0.078	0.144	0.135	0.102	0.147	0.169	0.121

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 16: OLS results for cohort 2015, Reading

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
Reading	0.093	0.094	0.074	0.173*	-0.073	0.200**	0.102	0.382**
	(0.123)	(0.097)	(0.121)	(0.094)	(0.114)	(0.093)	(0.198)	(0.176)
HSFG	0.163***	0.179***	0.217***	0.223***	0.176***	0.232***	0.453***	0.258***
	(0.043)	(0.046)	(0.042)	(0.046)	(0.040)	(0.046)	(0.070)	(0.087)
Lyceum	3.125**	1.265	3.680***	2.254**	1.840	1.463	7.613***	3.518**
	-1.210	-1.009	-1.200	(0.965)	-1.143	(0.970)	-2.072	-1.762
Age	-0.018	-0.500***	-0.241	-0.536***	-0.201	-0.544***	0.880	-0.593
	(0.342)	(0.188)	(0.323)	(0.202)	(0.308)	(0.200)	-1.316	(0.734)
Female	0.334	-0.549	-0.280	-0.446	-0.066	-0.619	-1.144	-0.244
	(0.847)	(0.868)	(0.824)	(0.824)	(0.765)	(0.829)	-1.323	-1.485
South	1.818	-4.391***	0.631	-2.623	0.658	-2.506	-0.087	-3.155
	-2.057	-1.584	-1.951	-1.730	-1.799	-1.717	-3.002	-3.021
_cons	3.895	15.593**	5.630	10.878	15.466*	12.806*	33.600	83.176***
	-9.642	-6.879	-9.202	-7.031	-8.772	-7.019	-33.263	-20.541
Obs.,	174	213	164	194	160	190	139	149
R-squared	0.138	0.149	0.201	0.207	0.140	0.217	0.284	0.122

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 17: OLS results for cohort 2016, Reading

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
Reading	-0.167	0.116	-0.205*	0.184**	-0.110	0.177**	-0.128	0.391**
	(0.120)	(0.087)	(0.109)	(0.083)	(0.108)	(0.077)	(0.186)	(0.150)
HSFG	0.293***	0.250***	0.240***	0.195***	0.160***	0.155***	0.389***	0.223**
	(0.049)	(0.055)	(0.045)	(0.051)	(0.045)	(0.049)	(0.081)	(0.096)
Lyceum	2.129*	3.130***	2.375**	2.277**	1.706*	2.841***	3.904**	4.344**
	-1.180	-1.003	-1.027	(0.939)	-1.021	(0.869)	-1.826	-1.733
Age	-0.957**	-0.228	-1.330	-0.134	-1.574*	-0.286	-2.866	0.044
	(0.432)	(0.225)	(0.813)	(0.241)	(0.805)	(0.219)	-2.322	(0.665)
Female	-0.329	-0.709	0.015	-1.690*	0.144	-1.549*	-0.187	0.844
	-1.081	(0.964)	(0.962)	(0.893)	(0.955)	(0.836)	-1.639	-1.606
South		-3.939*		-1.999		-1.833		1.192
		-2.042		-2.068		-1.887		-3.752
_cons	20.340*	-0.964	35.138*	3.344	48.746**	12.886*	133.937**	66.851***
	-11.441	-6.999	-19.759	-7.192	-19.576	-6.623	-53.641	-17.442
Obs.,	167	223	158	199	155	193	131	156
R-squared	0.230	0.135	0.204	0.133	0.128	0.150	0.181	0.136

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

4.2.3 Final score

As I already stressed out, the total final score (FS) is obtained combining the admission test score and the high school final grade (or the GPA of the fourth year of secondary school if the HSFG is not available). The share of these two variables changes throughout the cohorts: in 2014 the admission test and the high school grades contribute to the creation of the final score with 50 points each, in spring of 2015 and 2016 only the admission test is taken into account and in the summer of these two years 70 points are attributed to the admission test and the remaining 30 points to the HSFG.

Tables 18, 19 and 20 show the results of the regressions when the final score is considered as main independent variable. The results show positive and significant coefficients, especially for the summer subgroup of students.

It must be acknowledged that in these regressions I consider both the final score and the HSFG as independent variables but, given that the HSFG contributes - in different proportions during each selection - to the creation of the final score, these two variables are correlated and lead to a problem of multicollinearity. Tables 36, 37 and 38 in the Annex show the OLS regressions without controlling for the HSFG. It is interesting noticing that the R squared for the summer selections almost does not change, whereas in spring, if the HSFG is not considered, the R squared of the regressions becomes rather small, especially in 2015 and 2016. This means that in spring, when only the admission test is taken into account, the model can poorly predict the variation in the dependent variables.

Table 18: OLS results for cohort 2014, Final score

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
FS	0.053	0.185***	0.055	0.138***	0.084*	0.155***	0.256***	0.256**
	(0.053)	(0.054)	(0.044)	(0.051)	(0.045)	(0.045)	(0.094)	(0.105)
HSFG	0.187***	-0.048	0.183***	0.042	0.125**	0.043	0.359***	0.036
	(0.069)	(0.081)	(0.057)	(0.077)	(0.058)	(0.069)	(0.117)	(0.154)
Lyceum	1.376	0.184	2.498***	0.696	2.038**	0.798	1.732	2.454
	(0.984)	(0.966)	(0.806)	(0.900)	(0.822)	(0.808)	-1.743	-1.670
Age	-1.020	-0.431	0.391	-1.144**	0.397	-0.320	-0.081	-1.368
	(0.669)	(0.502)	-1.148	(0.464)	-1.166	(0.478)	-2.353	-1.049
Female	-1.287	-1.567*	0.109	-0.472	0.205	-0.144	-1.975	-0.297
	(0.833)	(0.839)	(0.695)	(0.794)	(0.713)	(0.714)	-1.460	-1.418
South	-1.485	0.221	0.492	-0.309	0.820	-1.121	-0.323	-1.853
	-1.644	-1.222	-1.417	-1.105	-1.441	-1.000	-2.820	-2.005
_cons	27.238	23.179	-8.754	36.063***	-3.376	16.895	53.546	111.294***
	-17.545	-14.473	-29.401	-13.536	-29.856	-13.551	-60.005	-29.383
Obs.,	170	232	154	206	152	200	140	158
R-squared	0.098	0.122	0.151	0.161	0.123	0.170	0.197	0.152

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 19: OLS results for cohort 2015, Final score

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
FS	0.167**	0.235***	0.116*	0.226***	0.077	0.253***	0.287***	0.412***
	(0.069)	(0.051)	(0.068)	(0.049)	(0.063)	(0.048)	(0.106)	(0.100)
HSFG	0.136***	-0.027	0.197***	0.028	0.153***	0.017	0.398***	-0.051
	(0.044)	(0.064)	(0.044)	(0.063)	(0.041)	(0.062)	(0.071)	(0.115)
Lyceum	2.415*	-0.745	3.166**	0.688	1.569	-0.225	6.319***	1.629
	-1.225	-1.023	-1.224	(0.979)	-1.166	(0.967)	-2.049	-1.769
Age	-0.162	-0.434**	-0.341	-0.471**	-0.279	-0.468**	0.468	-0.545
	(0.342)	(0.179)	(0.326)	(0.191)	(0.313)	(0.187)	-1.291	(0.705)
Female	0.477	0.364	-0.150	0.291	0.044	0.205	-0.739	1.015
	(0.836)	(0.851)	(0.822)	(0.807)	(0.767)	(0.801)	-1.298	-1.472
South	2.271	-3.773**	0.994	-1.839	0.885	-1.664	0.790	-1.688
	-2.033	-1.518	-1.947	-1.663	-1.803	-1.628	-2.939	-2.930

_cons	0.329	19.253***	3.485	15.288**	11.924	17.461***	30.476	89.102***
	-9.493	-6.595	-9.108	-6.790	-8.684	-6.685	-32.160	-19.587
Obs.,	174	213	164	194	160	190	139	149
R-squared	0.164	0.224	0.214	0.275	0.146	0.303	0.321	0.189

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 20: OLS results for cohort 2016, Final score

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
FS	0.077	0.249***	0.006	0.227***	-0.020	0.153***	0.168	0.424***
	(0.073)	(0.050)	(0.066)	(0.048)	(0.065)	(0.045)	(0.113)	(0.083)
HSFG	0.253***	0.055	0.208***	0.027	0.148***	0.049	0.332***	-0.080
	(0.050)	(0.066)	(0.045)	(0.063)	(0.045)	(0.061)	(0.079)	(0.114)
Lyceum	1.602	1.772*	2.120**	1.242	1.667	2.346***	3.284*	2.649
	-1.201	(0.979)	-1.064	(0.925)	-1.047	(0.877)	-1.826	-1.678
Age	-0.990**	-0.107	-1.361*	-0.106	-1.592*	-0.284	-3.010	-0.034
	(0.432)	(0.213)	(0.822)	(0.230)	(0.807)	(0.215)	-2.309	(0.628)
Female	0.153	0.814	0.162	-0.456	0.102	-0.805	0.604	2.962*
	-1.131	(0.969)	-1.027	(0.906)	-1.013	(0.873)	-1.702	-1.595
South		-3.486*		-1.740		-1.817		1.100
		-1.942		-1.974		-1.850		-3.530
_cons	16.684	1.153	34.359*	7.625	49.344**	16.457**	129.165**	77.608***
	-11.946	-6.553	-20.145	-6.846	-19.800	-6.495	-53.330	-16.560
Obs.,	167	223	158	199	155	193	131	156
R-squared	0.226	0.217	0.185	0.205	0.122	0.177	0.192	0.230

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

4.3 Decomposition of R^2

As I mentioned before, the R squared of a multiple regression contains the partial contributions of all the explanatory variables. In this section I decompose the coefficient of multiple determination and present the contributions of each independent variable. The decomposition is made on the R squared of the regressions of the dependent variable GPA1, which is the grade point average of the first year of college, on the main independent variables (Admission test, Final score, Math and Reading) and the control variables (High school final grade, school type, age, gender and region of origin).

Tables 21, 22 and 23 show the detailed results of the decomposition for the 2014-2016 cohorts. On the left-hand side of the tables there are the results regarding the spring selection and on the other side the summer selection is considered. The first column of each side presents the raw regression coefficients of the explanatory variables of the regressions. The second column contains the standardized regression coefficients, and the third column presents the correlation between the independent variables and the explained variable, GPA1. In the last column it is shown the contribution of each explanatory variable to the R squared, while at the bottom of this column the whole R^2 is provided.

Table 24 summarizes the decomposition results of the three cohorts. Here are shown the averages of the contributions of each explanatory variable for the three cohorts. The table is divided into four sections, each referring to a main independent variable (Admission test, Final score, Math, and Reading). In fact, the first row contains the averages of these independent variables' contributions, while the rows below present the averages for the control variables. In the last row the mean of the R^2 of the regressions for the three cohorts is computed.

The results show that, for what concerns the Admission test, on average the contribution to the R squared is around 3% for the spring group of students and above 10% for the summer students. This means that in spring the admission test seems to contribute very little in explaining the variation of the dependent variable (GPA1). On the other hand, the admission test in summer explains a considerably larger portion of variance of GPA1. On average the gap between the two contributions is around 7%.

It is also worthwhile checking the results of the two subtests. The decomposition regarding the subtest aimed to verify the analytical and quantitative abilities (Math) reflects the results found for the admission test. In fact, on average the Math subtest contributes to the R^2 with a 4,3% share in spring and with a 14% share in summer, confirming a higher contribution of the subtest in summer. The average gap for the Math subtest between spring and summer is about 10%. For what concerns the second subtest, aimed to verify the verbal abilities and general knowledge, the contribution to the R^2 both in spring and in summer is rather small, very close to zero. In fact, it is, on average, equal to 0,07% in spring and to 0,75% in summer. This means that the contribution in explaining the variation of the dependent variable GPA1 of the Reading subtest is rather insignificant.

Regarding the other explanatory variables, the high school final grade (HSFG) is the most relevant in terms of contribution to the R^2, although its pattern is quite different from the admission test. In fact, the mean of its contributions is around 11% in spring and 6% in summer. This denotes that the high school final grade explains a larger share of the variation of GPA1 in spring than it does in summer, which is exactly the opposite of what happens with the admission test. Thus, these two explanatory variables, although both relevant, tend to go in opposite directions. The other independent variables, instead, contribute to the R^2 on a smaller scale. In fact, the share of the variation of the GPA1 explained by each of these variables is, on average, below 1%. Only the dummy for Age seems to contribute a little more, exceeding the 1% share.

Table 21: Decomposition of the R2 for the 2014 cohort

	spring				summer			
	raw coeff	stand coeff	corr	**R^2**	raw coeff	stand coeff	corr	**R^2**
AT	0,086	0,153	0,185	**2,84%**	0,167	0,299	0,257	**7,68%**
HSFG	0,205	0,266	0,198	**5,25%**	0,185	0,244	0,141	**3,45%**
Lyceum	0,869	0,074	0,069	**0,51%**	-0,141	-0,011	0,104	**-0,12%**
Age	-0,969	-0,108	-0,110	**1,19%**	-0,456	-0,059	-0,110	**0,65%**
Female	-1,036	-0,095	-0,124	**1,18%**	-1,444	-0,115	-0,153	**1,76%**
South	-0,704	-0,033	-0,002	**0,01%**	0,357	0,020	0,084	**0,17%**
				10,98%				**13,59%**
Final score	0,053	0,082	0,184	**1,51%**	0,185	0,334	0,310	**10,35%**
HSFG	0,187	0,243	0,198	**4,80%**	-0,048	-0,063	0,141	**-0,89%**
Lyceum	1,376	0,117	0,069	**0,81%**	0,184	0,015	0,104	**0,15%**
Age	-1,020	-0,114	-0,110	**1,25%**	-0,431	-0,055	-0,110	**0,61%**
Female	-1,287	-0,118	-0,124	**1,47%**	-1,567	-0,125	-0,153	**1,91%**
South	-1,485	-0,070	-0,002	**0,01%**	0,221	0,012	0,084	**0,10%**
				9,85%				**12,23%**
Math	0,201	0,263	0,240	**6,32%**	0,278	0,383	0,324	**12,44%**
HSFG	0,221	0,286	0,198	**5,65%**	0,203	0,268	0,141	**3,79%**
Lyceum	0,336	0,029	0,069	**0,20%**	-0,157	-0,012	0,104	**-0,13%**
Age	-0,824	-0,092	-0,110	**1,01%**	-0,408	-0,053	-0,110	**0,58%**
Female	-0,705	-0,065	-0,124	**0,81%**	-1,063	-0,085	-0,153	**1,30%**
South	-0,347	-0,016	-0,002	**0,00%**	0,339	0,019	0,084	**0,16%**
				13,98%				**18,13%**
Reading	-0,055	-0,048	0,019	**-0,09%**	-0,033	-0,029	0,014	**-0,04%**
HSFG	0,224	0,290	0,198	**5,72%**	0,158	0,208	0,141	**2,95%**
Lyceum	1,800	0,153	0,069	**1,06%**	1,685	0,134	0,104	**1,39%**
Age	-0,974	-0,109	-0,110	**1,19%**	-0,536	-0,069	-0,110	**0,76%**
Female	-1,414	-0,130	-0,124	**1,62%**	-2,257	-0,180	-0,153	**2,76%**
South	-1,844	-0,087	-0,002	**0,02%**	-0,119	-0,007	0,084	**-0,06%**
				9,51%				**7,75%**

Table 22: Decomposition of the R2 for the 2015 cohort

	spring				summer			
	raw coeff	stand coeff	corr	**R^2**	raw coeff	stand coeff	corr	**R^2**
AT	0,208	0,185	0,264	**4,87%**	0,235	0,341	0,311	**10,61%**
HSFG	0,136	0,256	0,301	**7,70%**	0,149	0,237	0,266	**6,30%**
Lyceum	2,415	0,150	0,123	**1,84%**	-0,745	-0,057	0,023	**-0,13%**
Age	-0,162	-0,034	-0,042	**0,14%**	-0,434	-0,155	-0,235	**3,65%**
Female	0,477	0,042	0,100	**0,42%**	0,364	0,028	0,014	**0,04%**
South	2,271	0,083	0,169	**1,41%**	-3,773	-0,159	-0,123	**1,95%**
				16,38%				**22,42%**
Final score	0,167	0,185	0,264	**4,87%**	0,235	0,433	0,417	**18,06%**
HSFG	0,136	0,256	0,301	**7,70%**	-0,027	-0,043	0,266	**-1,15%**
Lyceum	2,415	0,150	0,123	**1,84%**	-0,745	-0,057	0,023	**-0,13%**
Age	-0,162	-0,034	-0,042	**0,14%**	-0,434	-0,155	-0,235	**3,65%**
Female	0,477	0,042	0,100	**0,42%**	0,364	0,028	0,014	**0,04%**
South	2,271	0,083	0,169	**1,41%**	-3,773	-0,159	-0,123	**1,95%**
				16,38%				**22,42%**
Math	0,240	0,173	0,223	**3,87%**	0,383	0,394	0,376	**14,83%**
HSFG	0,149	0,281	0,301	**8,44%**	0,154	0,245	0,266	**6,52%**
Lyceum	2,258	0,140	0,123	**1,72%**	-0,635	-0,049	0,023	**-0,11%**
Age	-0,177	-0,038	-0,042	**0,16%**	-0,329	-0,118	-0,235	**2,78%**
Female	0,452	0,039	0,100	**0,39%**	0,544	0,042	0,014	**0,06%**
South	2,381	0,087	0,169	**1,48%**	-3,502	-0,148	-0,123	**1,81%**
				16,06%				**25,87%**
Reading	0,093	0,056	0,120	**0,68%**	0,094	0,069	0,089	**0,61%**
HSFG	0,163	0,308	0,301	**9,26%**	0,179	0,286	0,266	**7,60%**
Lyceum	3,125	0,194	0,123	**2,38%**	1,265	0,097	0,023	**0,22%**
Age	-0,018	-0,004	-0,042	**0,02%**	-0,500	-0,179	-0,235	**4,22%**
Female	0,334	0,029	0,100	**0,29%**	-0,549	-0,043	0,014	**-0,06%**
South	1,818	0,067	0,169	**1,13%**	-4,391	-0,185	-0,123	**2,27%**
				13,76%				**14,85%**

Table 23: Decomposition of the R2 for the 2016 cohort

	spring				summer			
	raw coeff	stand coeff	corr	**R^2**	raw coeff	stand coeff	corr	**R^2**
AT	0,096	0,079	0,188	**1,48%**	0,286	0,368	0,348	**12,81%**
HSFG	0,253	0,417	0,417	**17,39%**	0,234	0,316	0,250	**7,90%**
Lyceum	1,602	0,103	0,003	**0,03%**	1,619	0,118	0,122	**1,44%**
Age	-0,990	-0,162	-0,219	**3,56%**	-0,070	-0,021	-0,080	**0,16%**
Female	0,153	0,010	0,120	**0,12%**	0,926	0,068	0,075	**0,50%**
South	0,000		0,000	**0,00%**	-3,372	-0,111	-0,038	**0,42%**
				22,59%				**23,24%**

Final score	0,077	0,079	0,188	**1,48%**	0,249	0,398	0,432	**17,18%**
HSFG	0,253	0,417	0,417	**17,39%**	0,055	0,073	0,250	**1,84%**
Lyceum	1,602	0,103	0,003	**0,03%**	1,772	0,129	0,122	**1,58%**
Age	-0,990	-0,162	-0,219	**3,56%**	-0,107	-0,031	-0,080	**0,25%**
Female	0,153	0,010	0,120	**0,12%**	0,814	0,059	0,075	**0,44%**
South	0,000		0,000	**0,00%**	-3,486	-0,115	-0,038	**0,44%**
				22,59%				**21,72%**
Math	0,202	0,155	0,173	**2,67%**	0,500	0,440	0,343	**15,10%**
HSFG	0,261	0,430	0,417	**17,95%**	0,287	0,387	0,250	**9,68%**
Lyceum	1,584	0,102	0,003	**0,03%**	1,862	0,136	0,122	**1,66%**
Age	-0,926	-0,152	-0,219	**3,33%**	-0,149	-0,044	-0,080	**0,35%**
Female	0,383	0,026	0,120	**0,31%**	1,273	0,093	0,075	**0,69%**
South	0,000		0,000	**0,00%**	-3,654	-0,120	-0,038	**0,46%**
				24,29%				**27,94%**
Reading	-0,167	-0,102	0,037	**-0,38%**	0,116	0,091	0,186	**1,69%**
HSFG	0,293	0,483	0,417	**20,16%**	0,250	0,337	0,250	**8,43%**
Lyceum	2,129	0,137	0,003	**0,05%**	3,130	0,228	0,122	**2,79%**
Age	-0,957	-0,157	-0,219	**3,44%**	-0,228	-0,067	-0,080	**0,53%**
Female	-0,329	-0,022	0,120	**-0,27%**	-0,709	-0,052	0,075	**-0,39%**
South	0,000		0,000	**0,00%**	-3,939	-0,130	-0,038	**0,49%**
				23,00%				**13,55%**

Table 24: Averages of the decomposition 2014-2016

	AT		FS		Math		Reading	
	spring	summer	spring	summer	spring	summer	spring	summer
Main explanatory variable	3,06%	10,37%	2,62%	15,20%	4,29%	14,12%	0,07%	0,75%
HSFG	10,11%	5,88%	9,96%	-0,07%	10,68%	6,66%	11,71%	6,32%
Lyceum	0,80%	0,40%	0,89%	0,53%	0,65%	0,47%	1,16%	1,47%
Age	1,63%	1,49%	1,65%	1,50%	1,50%	1,23%	1,55%	1,84%
Female	0,57%	0,77%	0,67%	0,80%	0,50%	0,68%	0,55%	0,77%
South	0,47%	0,84%	0,47%	0,83%	0,49%	0,81%	0,38%	0,90%
average R^2	16,65%	19,75%	16,28%	18,79%	18,11%	23,98%	15,42%	12,05%

4.4 Partial r^2

The partial r squared allows to evaluate whether the addition of an explanatory variable to a regression model can be useful and to which extent.

In the paragraph above it was shown that the control variables Lyceum, Age, Female and South are not of great relevance in explaining the variation of GPA1. Therefore, in this last section I focus the attention only on the main explanatory variables (Admission test, Final score, Math and Reading) and on the high school final grade. As dependent variables I consider the grade point average of all three years and the final grade. For what concerns the regressions with the grade point average of the first year as the dependent variable, I find the partial r squared of all the main explanatory variables and the high school final grade. For the other dependent variables (GPA2, GPA3 and final grade) I just focus on the two independent variables that are of most interest, admission test and HSFG.

For the main independent variables, I find the partial r squared by keeping all the other controls (HSFG, lyceum, age, female, south) constant, as shown at the end of chapter 3. For the HSFG, instead, I find the partial r squared by keeping constant the controls and the admission test.

As for the rest of my analysis, I separate the group of students between spring and summer. Table 25 contains the results of this analysis for each dependent variable. For the regressions that consider GPA1 as the outcome variable, it can be noticed that the partial r squared of the admission test is on average lower in the spring selection (2%) with respect to the summer selection (9%), with a difference between the two selections of 7 percentage points. However, for the 2014 cohort this gap is rather moderate (4,5%) compared with the other two cohorts, where the gap is around 6% for the 2015 cohort and 11% for the 2016 cohort. The same pattern is found when looking at the Math's partial r squared. In fact, on average the spring selection presents a partial r squared of 3,7% whereas the summer selection is undoubtedly higher (almost 14%). Also in this case the 2014 cohort presents the lowest gap between selections. For what concerns the Reading subtest, the partial r squared is almost always below 1% with a very little difference between spring and summer, meaning that adding this variable to the model does not contribute much to the improvement of the model itself. The HSFG, instead, has a steadier and more regular partial r squared. The difference between selections is indeed small, being around 8% in spring and 6% in summer.

When the grade point average of the second and the third year and the final grade are considered as the outcomes of the regressions, on a general basis the partial r squared of the admission test still remains higher in summer, but when observing the details of each cohort, it can be seen that this is true only for the 2015 and 2016 cohort. On the other hand, the 2014 cohort presents an opposite trend. In this case spring takes the lead with a higher partial r squared, but still the difference between selections remains rather small. One reason why the 2014 cohort is more balanced between selections could be the fact that in this year the admission rules were the same for both the spring and the summer selections (the final score was formed given a 50% weight to both the admission test and the high school final grade). Whereas in 2015 and 2016 in spring only the admission test was accounted for the creation of the final score, instead in summer the admission test and the HSFG were given respectively a 70% and a 30% share in the final score. For what concerns the HSFG, its partial r squared has a less clear pattern. It still seems to be higher in spring but, as before, the gap between selections is very tiny and the situation even reverses when it comes to the regressions where the output variable is the GPA3. Thus, the bottom line is that the HSFG seems to be more stable across cohorts and also across selections. To sum up, it seems that the admission test is coherent across cohorts but tends into opposite directions when distinguishing between selections. Instead, the HSFG is overall more reliable and stable.

Table 25: Partial r2 of AT, HSFG, FS, Math and Reading using as dependent variable GPA1, GPA2, GPA3 and final grade

	2014		2015		2016		Mean	
	spring	summer	spring	summer	spring	summer	**spring**	**summer**
GPA1								
r2 AT	1,84%	6,33%	3,37%	9,24%	0,67%	11,87%	**1,96%**	**9,15%**
r2 HSFG	6,12%	4,79%	5,44%	4,99%	13,79%	8,96%	**8,45%**	**6,25%**
r2 FS	0,60%	4,86%	3,37%	9,24%	0,67%	10,15%	**1,55%**	**8,08%**
r2 Math	5,13%	11,24%	3,00%	13,26%	2,84%	17,11%	**3,66%**	**13,87%**
r2 Reading	0,23%	0,08%	0,34%	0,45%	1,18%	0,82%	**0,58%**	**0,45%**
GPA2								
r2 AT	5,87%	3,43%	1,82%	10,17%	0,00%	11,68%	**2,56%**	**8,43%**
r2 HSFG	8,96%	7,49%	11,36%	9,26%	11,93%	7,64%	**10,75%**	**8,13%**
GPA3								
r2 AT	5,73%	4,74%	0,97%	13,05%	0,06%	6,60%	**2,25%**	**8,13%**
r2 HSFG	5,44%	10,80%	8,19%	10,34%	6,62%	5,83%	**6,75%**	**8,99%**
FG								
r2 AT	8,69%	3,77%	5,30%	10,58%	1,73%	16,87%	**5,24%**	**10,41%**
r2 HSFG	12,78%	7,74%	19,27%	6,32%	12,31%	3,98%	**14,79%**	**6,01%**

Table 26 describes how the partial r squared changes when females and males are analysed separately. For this evaluation I consider only the case where the grade point average of the first year is the outcome variable and I find the partial r squared for the admission test and the high school final grade. The results show little difference between males and females. The partial r squared of the admission test is slightly lower for females with respect to males, whereas the opposite occurs for what concerns the high school final grade. Thus, it seems that when the admission test is added, the regression model improves slightly more when males are considered. For females, instead, the same happens with the HSFG. But from a wider point of view, the difference between males and females still remains narrow.

Table 26: Partial r2 of AT and HSFG using GPA1 as dependent variable, for females and males

	2014		2015		2016		Mean	
	spring	summer	spring	summer	spring	summer	**spring**	**summer**
female								
r2 AT	1,68%	8,18%	2,03%	6,18%	0,75%	11,86%	**1,49%**	**8,74%**
r2 HSFG	9,39%	2,62%	9,26%	8,76%	6,79%	13,60%	**8,48%**	**8,33%**
male								
r2 AT	2,28%	5,08%	4,54%	12,94%	0,48%	10,91%	**2,43%**	**9,64%**
r2 HSFG	3,47%	7,42%	3,08%	1,96%	16,50%	5,43%	**7,68%**	**4,94%**

4.5 Discussion

In my analysis I find some interesting, rather surprising and sometimes not so straightforward results. First of all, I find a positive correlation between the admission test and the high school final grade in the candidates' sample, but a mostly negative correlation in the enrolled students' sample. Furthermore, I find that the regression models employed in my analysis can explain on average 20% of the variation of the dependent variables. Both the admission test and the high school final grade are found to be relevant in explaining the variation of the outcome variables. The same applies for the Math subtest, while the Reading subtest seems negligible.

Since I am not interested in the causal effect of the explanatory variables but in understanding how well they can predict the dependent variables, I focus more on the decomposition and on the partial r squared. The decomposition of the R^2 of the regressions highlights the differences that occur between the admission test and the high school final grade. In particular, the findings show a distinct discrepancy between these two variables when observing the results concerning the spring and the summer selections, for all the cohorts involved. In spring the admission test is able to explain a rather small portion of the variation of the dependent variables, whereas in summer this portion increases sharply. The opposite occurs for what concerns the high school final grade, although it remains more stable across selections. Similar remarks can be made also when looking at the partial r squared. In fact, the additional explanatory power of the admission test tends to be very small in spring and increases decidedly in summer. All the three cohorts present this pattern, however in 2014 it is less evident and sometimes even reversed. The HSFG, instead, has a higher additional explanatory power in spring but again the gap between selections here tends to be modest.

The reason for these divergences could depend on the way the admission test and the high school final grade reflect the true ability, which we cannot observe, of students. Of course, these two variables cannot represent the ability perfectly: there will be a certain noise that interferes with their capability of reflecting the true ability of students.

We could imagine a model where both the admission test and the HSFG are equal to the true ability of students plus a certain error[23]. Looking at the results, it can be argued that the error embodied in the admission test may have a higher variance with respect to the error embodied in the high school final grade. This would mean that in the admission test there is a strong noise attributed to the error. Therefore, even though both variables reflect the true ability of students[24], the admission test seems to reflect it in a vaguer way.

The error term, and also the correlation between the error and the true ability, affects the coefficients of the admission test and of the high school final grade in the regression models. The regression coefficients, in turn, affect the contribution that a variable gives to the R^2.

Therefore, given that the error and its correlation with the ability vary across selections, also the regression coefficients, the contribution to the R^2 and the partial r^2 will be affected and will consequently present discrepancies across selections, especially for what concerns the admission test.

Thus, from my analysis I can conclude that the best way of predicting the performance of students is by employing in a regression model both the admission test and the high school final grade. Sometimes the admission test will contribute more to the prediction of the students' performance, other times the opposite will occur.

[23] Assuming it is independent and with mean 0.

[24] we can see this by looking at the correlation between the admission test and the high school final grade in the candidates' sample. The correlation is positive ad around 30%.

Conclusion

Higher education is an important step in the life of the young adults. Deciding to pursue an academic career means deciding to invest financial resource, time and effort, keeping in mind also the opportunity costs (Anchor et al., 2011). This decision must be a well-informed choice in order not to waste such resources. The choice to enrol in an academic faculty can be driven by various factors, but mainly: either trying to secure a better job placement in the future, with higher earnings, or just wanting to learn and to expand the personal knowledge. On the other side, the universities, who offer such knowledge, also have to deal with limited resources. Therefore, the two players involved in this game must find an equilibrium. Especially when the demand for enrolment is too high the universities must find a way to choose who to accept and who to reject. The most straightforward way is by implementing an admission test, which is an objective test that measures the abilities of candidates in specific fields of study. Also other measures may be taken into account, like the high school final grade. Therefore, universities decide what admission rules to implement, keeping in mind that the ultimate aim is admitting only the most prominent students, that is the students that will most probably achieve a successful academic career.

Since one of the admission rules that most universities implement is indeed the admission test, one cannot help but wonder if the admission test can really predict who the best students will be and, if so, how accurately can it predict such result.

To answer these questions, I analysed four groups of students belonging to four cohorts (from 2013 to 2016) that enrolled at the department of economics of the university of Padova.

The admission rule for this department consisted mostly in admitting students based on their score obtained in the admission test and based on the final grade achieved in high school. As I mentioned many times, the admission rule was not the same for all the cohorts, and even within cohorts it changed sometimes between spring and summer selections. This led to some differences in the results when distinguishing between selections.

Since the admission test and the high school final grade were usually considered in the admission rules, first of all I tried to correlate them to see the nature of their relationship. I found a positive correlation (30%) for the whole sample of candidates and a mostly negative relationship within the subsample of enrolled students. This outcome is indeed the result of the admission rules implemented. Nevertheless, both the admission test and the HSFG reflect a certain portion of the true ability of students. Thus, in my analysis I employed these two variables, and some other controls, as the explanatory variables of the academic performance of enrolled students. The academic performance was measured through the weighted grade point averages of the first three years of college and through the final grade obtained at the end of the career.

First, I measured, through a probit regression model, the ability of the admission test and HSFG to predict the probability of dropping out of college during one of the first three years, finding little evidence of such capacity.

Therefore, I focused my analysis on the OLS regression and, in particular, on the R squared, since the aim of this study was not to find a causal relationship between the explanatory variables and the academic performance but instead the predictive ability of these variables. I repeated the analysis for each selection that took place in the four years. My results show that the regression models can predict around 20% of the academic performance's variation. The admission test and the HSFG are the main two variables that generally present positive and statistically significant coefficients. However, in order to evaluate their contribution to the R^2, I decomposed it and found some interesting results. The admission test seems to contribute consistently to the summer selections and very little in spring. Whereas the opposite occurs for the high school final grade, even though the gap between selections is less emphasised. The partial r squared confirms these findings for what concerns the GPA of the first year of college, the results seem to be vaguer for the other key indicators of performance. Finally, when unpacking the admission test into its two subtests (Math and Reading) I found that the Math subtest has more predictive power, whereas the Reading subtest seems to be irrelevant. These results are in line with those found by Bettinger et al. (2013) regarding the composite score of the ACT.

The interpretation of these results is not so straightforward. I imagine that, even though both the AT and the HSFG reflect the true ability of students, the differences between selections are due to the noise (an error) embodied in the variables. Especially for the admission test the error seems to have a large variance which leads the AT in opposite directions throughout selections. The noise embodied in the HSFG, instead, seems to have a smaller variance, and consequently also the variable is more stable. This seems plausible since the admission test, although objective, tries to measure the ability of students in just 80 minutes. There are a lot of variables that do not depend on the personal ability and that can affect positively or negatively the output of the test. Feeling bad the day of the test or simply being too anxious[25] can lead to a distortive

[25] Several papers have shown that admission tests tend to underpredict females' academic success (Connor, 1992; Hancock, 1999; Gillborn, 2000; Silverstein, 2000; Froese, 2001). It is not clear yet why this happens, but

final result. On the other hand, the high school final grade represents the ability that students proved to have during the five years of secondary school. Thus, it might seem more reliable, even though it is less objective. In fact, it depends on the type of secondary school attended by students, on the "generosity" of teachers regarding grades, on the location of the secondary school and so on. Despite all these reasons, the HSFG still seems more stable for what concerns my results. The bottom line, in any case, is that the best way of predicting the future performance of students is by employing both the admission test and the high school final grade because by doing so the probability of choosing the best students increases, since during some selections the admission test will predict the performance better, during other selections the HSFG will do the job. Besides these two variables, other factors could be taken into account. Following the example of the admission rules implemented by most American universities, also extracurricular activities, recommendation letters and essays could be considered for a more reliable admission rule.

My analysis is of course not without limitations. It must be highlighted that, unfortunately, we can only observe the performance of the admitted and enrolled students. However, what we really would like to find out is if the enrolled students perform systematically better than those who are left out by the admission rule. This is of course not possible since those who are not admitted cannot pursue this academic career. An attempt to address this problem was implemented by Migliaretti et al. (2017) who analysed the cohort of medical students at the university of Turin who enrolled in the 2014-2015 academic year. They found that the test scores were indeed able to predict the academic success in the first year of college.

I also want to point out that in my analysis I could not employ many explanatory variables that potentially could affect to some extent the academic performance, like the socio-economic status of students, their attendance in class, the family income and their parents' background education.

Finally, I want to underline that, even though the students are admitted on the basis of an impartial admission rule, they can decide whether to enrol or not. The data indeed show that a quite large number of admitted students (above 100 for each selection) decide not to enrol. This may as well affect the predictiveness of the admission test.

Nevertheless, my results are in line with the literature regarding the predictiveness of the admission test and the high school final grade.

I want to make a final remark regarding further research. I think it could be interesting to analyse the predictive ability of the admission test for the most recent cohorts, starting from the

---

one of the reasons could be the fact that females tend to be more anxious and therefore perform worse under pressure (Saygin, 2020).

academic year of 2019-2020, enrolled in the economics degree program at the university of Padova since the admission rule for these cohorts changed completely. In fact, the admission test consists now of 36 questions (including a Math, a Reading and a Logical subtest) to answer in 90 minutes. Thus, the structure of the test is quite different from the one I analysed, and it could be interesting to make a confrontation between them.

The right for education is a fundamental right for students but since universities do not have unlimited resources some necessary choices must be made. The admission test remains the most objective, straightforward and easiest way of selecting students, however it seems that if the high school final grade is also employed the ability of predicting future academic success improves.

Acknowledgement

I would like to thank the IT staff of the department of Economics of the university of Padova. I want to thank especially Pierfrancesco Consolo who helped me collect and organize the data I needed for my analyses and Antonio Trabucco who has shown great patience and made it possible for me to work on my thesis allowing me remotely access to the university's computers, since these trying and uncertain times have made it impossible to go to the department in person.

References

Alon, S., & Tienda, M. (2007). Diversity, opportunity, and the shifting meritocracy in higher education. *American Sociological Review*, *72*(4), 487-511.

Anchor, J. R., Fišerová, J., Maršíková, K., & Urbánek, V. (2011). Student expectations of the financial returns to higher education in the Czech Republic and England: Evidence from business schools. *Economics of Education Review*, *30*(4), 673-681.

Bettinger, E. P., Evans, B. J., & Pope, D. G. (2013). Improving college performance and retention the easy way: Unpacking the ACT exam. *American Economic Journal: Economic Policy*, *5*(2), 26-52.

Borcard, D. (2002). *Partial R-square, contribution and fraction*. Working paper, University of Montreal.

Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). Predictions of freshman grade-point average from the revised and recentered SAT® I: Reasoning Test. *ETS Research Report Series*, *2000*(1), i-16.

Connor, K., & Vargyas, E. J. (1992). The legal implications of gender bias in standardized testing. *Berkeley Women's LJ*, *7*, 13.

D.M. 22 Ottobre 2004, n. 270 "*Modifiche al regolamento recante norme concernenti l'autonomia didattica degli atenei, approvato con decreto del Ministro dell'universita' e della ricerca scientifica e tecnologica 3 novembre 1999, n. 509*"

Froese-Germain, B. (2001). Standardized testing+ high-stakes decisions= educational inequity. *Interchange*, *32*(2), 111-130.

Galla, B. M., Shulman, E. P., Plummer, B. D., Gardner, M., Hutt, S. J., Goyer, J. P., ... & Duckworth, A. L. (2019). Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *American Educational Research Journal*, *56*(6), 2077-2115.

Gillborn, D., & Mirza, H. S. (2000). Educational Inequality: Mapping Race, Class and Gender. A Synthesis of Research Evidence.

Hancock, T. (1999). The gender difference: Validity of standardized admission tests in predicting MBA performance. *Journal of Education for Business*, *75*(2), 91-93.

Hoffman, J. L., & Lowitzki, K. E. (2005). Predicting college success with high school grades and test scores: Limitations for minority students. *The review of higher education*, *28*(4), 455-474.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). Validity of the SAT® for Predicting First-Year College Grade Point Average. Research Report No. 2008-5. *College Board*.

Legge 2 agosto 1999, n.264 "*Norme in materia di accessi ai corsi universitari"*.

Marini, J. P., Westrick, P. A., Young, L., Ng, H., Shmueli, D., & Shaw, E. J. (2019). Differential Validity and Prediction of the SAT®: Examining First-Year Grades and Retention to the Second Year. *College Board*.

Migliaretti, G., Bozzaro, S., Siliquini, R., Stura, I., Costa, G., & Cavallo, F. (2017). Is the admission test for a course in medicine a good predictor of academic performance? A case− control experience at the school of medicine of Turin. *BMJ open*, *7*(11).

Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics*, *121*(1-2), 297-317.

Sawyer, R. (2013). Beyond correlations: Usefulness of high school GPA and test scores in making college admissions decisions. *Applied measurement in education*, *26*(2), 89-112.

Saygin, P. O. (2020). Gender bias in standardized tests: evidence from a centralized college admissions system. *Empirical Economics*, *59*(2), 1037-1065.

Silverstein, A. L. (2000). Standardized tests: The continuation of gender bias in higher education. *Hofstra L. Rev.*, *29*, 669.

Stemler, S. E. (2012). What should university admissions tests predict?. *Educational Psychologist*, *47*(1), 5-17.

Stock, J. H., & Watson, M. W. (2011). Introduction to Econometrics. 3rd international edition.

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT® scores, high school grades, and SES. *Educational Assessment*, *20*(1), 23-45.

Westrick, P. A., Marini, J. P., Young, L., Ng, H., Shmueli, D., & Shaw, E. J. (2019). Validity of the SAT® for Predicting First-Year Grades and Retention to the Second Year. *College Board*.

Web references

Ministero dell'Istruzione, dell'Università e della Ricerca. (n.d.). Scuola secondaria di secondo grado. Retrieved from https://www.miur.gov.it/web/guest/scuola-secondaria-di-secondo-grado

Università degli studi di Padova. (2013). Avviso di ammissione, anno accademico 2013-2014 Laurea Triennale in Economia e Management SELEZIONE PRIMAVERILE. Retrieved from https://www.economia.unipd.it/sites/decon.unipd.it/files/AvvisoAmmissioneEconomiaManagement1314_primavera.pdf

Università degli studi di Padova. (2013). Avviso di ammissione, anno accademico 2013-2014 Laurea in Economia e Management SELEZIONE ESTIVA. Retrieved from https://www.economia.unipd.it/sites/decon.unipd.it/files/AvvisoAmmissioneEconomiaManagement1314.pdf

Università degli studi di Padova. (2014). Avviso di ammissione, anno accademico 2014-2015 Laurea in Economia - Selezione Primaverile. Retrieved from https://www.economia.unipd.it/sites/economia.unipd.it/files/Economia%20primaverile%202014%20definitivo_2.pdf

Università degli studi di Padova. (2014). Avviso di ammissione, anno accademico 2014-2015 Laurea in Economia - SELEZIONE ESTIVA. Retrieved from https://www.economia.unipd.it/sites/economia.unipd.it/files/Economia%20-selezione%20estiva%20-%20definitivo_2.pdf

Università degli studi di Padova. (2015). Avviso di ammissione, anno accademico 2015-2016 Laurea in Economia - Selezione Primaverile. Retrieved from https://www.economia.unipd.it/sites/economia.unipd.it/files/2015_Economia_Selezione%20primaverile.pdf

Università degli studi di Padova. (2015). Avviso di ammissione, anno accademico 2015-2016 Laurea in Economia - SELEZIONE ESTIVA. Retrieved from https://www.economia.unipd.it/sites/economia.unipd.it/files/Economia_Avviso_ammissione_Agosto2015.pdf

Università degli studi di Padova. (2016). Avviso di ammissione, anno accademico 2016-2017 Laurea in Economia - Selezione Primaverile. Retrieved from https://www.economia.unipd.it/sites/economia.unipd.it/files/def%20Economia%20selezione%20primaverile%202016_0.pdf

Università degli studi di Padova. (2016). Avviso di ammissione, anno accademico 2016-2017 Corso di Laurea in Economia - SELEZIONE ESTIVA. Retrieved from https://www.economia.unipd.it/sites/economia.unipd.it/files/2016%20Economia%20selezione%20estiva_0.pdf

Università degli studi di Padova. (2019). Avviso di ammissione, anno accademico 2019-2020 Selezione Primaverile e Selezione Estiva per l'ammissione al Corso di Laurea in Economia. Retrieved from https://www.unipd.it/sites/unipd.it/files/2019/2019%20Economia%20TOLC-E%20V6%20web.pdf

Annex

Regression results for the 2013 cohort.

Table 27: Probit results for cohort, 2013

	(1)	(2)
	Dropout1	Dropout2
AT	-.051***	.008
	(.019)	(.024)
HSFG	.003	-.021
	(.029)	(.035)
Lyceum	.676**	.637
	(.342)	(.485)
Age	-.002	.014
	(.099)	(.1)
Female	-.632**	.75**
	(.293)	(.38)
South	-.984*	.414
	(.54)	(.472)
_cons	.916	-1.429
	-4.237	(4.85)
Obs.,	217	199
R-squared	.z	.z

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 28: OLS results for cohort 2013, Admission test

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
AT	0.097	0.274***	0.111	0.244***	0.091	0.201***	-0.039	0.516***
	(0.078)	(0.079)	(0.079)	(0.069)	(0.073)	(0.061)	(0.141)	(0.103)
HSFG	0.419***	0.327**	0.239*	0.314***	0.207	0.316***	0.544**	0.469***
	(0.130)	(0.127)	(0.134)	(0.112)	(0.125)	(0.100)	(0.228)	(0.170)
Lyceum	0.326	0.067	0.978	1.288	0.956	1.232	2.904	-0.418
	-1.302	-1.331	-1.349	-1.206	-1.273	-1.062	-2.349	-1.851
Age	-0.390	-1.733***	-2.102	-1.291***	-1.719	-0.806**	-2.735	-5.097*
	-1.557	(0.390)	-1.615	(0.335)	-1.505	(0.324)	-2.817	-2.686
Female	-2.121*	-0.946	0.662	0.170	1.050	1.016	-1.259	-1.547
	-1.148	-1.173	-1.168	-1.044	-1.103	(0.916)	-2.138	-1.549
South	-6.192***	-2.179	-5.758***	0.144	-7.851***	0.424	-11.141***	-1.314
	-1.688	-1.900	-1.769	-1.860	-1.646	-1.616	-3.568	-2.635
_cons	-10.958	21.445	49.516	13.187	45.720	3.868	124.999	164.114**
	-45.175	-17.104	-46.347	-14.767	-43.164	-13.795	-80.770	-73.549
Obs.,	68	120	62	110	61	105	57	94
R-squared	0.334	0.23	0.261	0.24	0.391	0.22	0.232	0.295

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 29: OLS results for cohort 2013, Math

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
Math	0.114	0.425***	0.143	0.324***	0.121	0.246***	-0.116	0.599***
	(0.101)	(0.102)	(0.111)	(0.092)	(0.103)	(0.083)	(0.200)	(0.151)
HSFG	0.400***	0.385***	0.213	0.358***	0.186	0.343***	0.555**	0.530***
	(0.128)	(0.127)	(0.133)	(0.116)	(0.124)	(0.104)	(0.225)	(0.184)
Lyceum	0.180	-0.999	0.688	0.793	0.692	0.949	3.487	-0.827
	-1.378	-1.379	-1.491	-1.271	-1.397	-1.138	-2.592	-2.033
Age	-0.299	-1.576***	-1.966	-1.178***	-1.594	-0.753**	-2.878	-4.703*
	-1.568	(0.378)	-1.632	(0.332)	-1.519	(0.326)	-2.824	-2.799
Female	-2.261**	-1.601	0.595	-0.620	1.022	0.389	-1.582	-3.205**
	-1.120	-1.126	-1.168	-1.020	-1.103	(0.905)	-2.158	-1.577
South	-6.336***	-3.028*	-5.944***	-0.921	-7.981***	-0.487	-11.179***	-3.800
	-1.669	-1.809	-1.745	-1.800	-1.623	-1.578	-3.519	-2.641
_cons	-9.781	14.539	50.135	9.996	45.741	3.640	129.072	158.453**
	-45.239	-17.021	-46.538	-15.122	-43.297	-14.129	-80.523	-76.941
Obs.,	68	120	62	110	61	105	57	94
R-squared	0.332	0.262	0.257	0.238	0.389	0.203	0.236	0.233

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 30: OLS results for cohort 2013, Reading

	GPA1		GPA2		GPA3		FG	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	spring	summer	spring	summer	spring	summer	spring	summer
Reading	0.072	0.067	0.112	0.166	0.085	0.176*	0.052	0.579***
	(0.127)	(0.131)	(0.137)	(0.117)	(0.128)	(0.101)	(0.238)	(0.185)
HSFG	0.399***	0.215*	0.237*	0.215*	0.203	0.234**	0.564**	0.252
	(0.131)	(0.129)	(0.138)	(0.115)	(0.128)	(0.101)	(0.233)	(0.178)
Lyceum	0.945	1.763	1.837	2.914**	1.647	2.591**	2.674	2.810
	-1.226	-1.303	-1.235	-1.167	-1.167	-1.016	-2.158	-1.855
Age	-0.550	-1.576***	-2.337	-1.214***	-1.911	-0.746**	-2.732	-4.832*
	-1.571	(0.413)	-1.632	(0.353)	-1.518	(0.336)	-2.819	-2.888
Female	-2.561**	-1.568	0.176	-0.112	0.634	0.918	-0.900	-1.376
	-1.089	-1.254	-1.102	-1.134	-1.037	(0.982)	-1.955	-1.731
South	-6.688***	-3.578*	-6.241***	-0.874	-8.267***	-0.254	-10.816***	-2.458
	-1.644	-2.006	-1.739	-1.965	-1.615	-1.686	-3.569	-2.858
_cons	-1.226	40.223**	59.136	28.775*	53.883	15.846	119.547	190.015**
	-44.757	-17.014	-46.135	-14.696	-42.878	-13.688	-79.917	-78.921
Obs.,	68	120	62	110	61	105	57	94
R-squared	0.321	0.15	0.243	0.163	0.379	0.159	0.231	0.185

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 31: OLS results for cohort 2013, Final score

	GPA1		GPA2		GPA3		FG	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	spring	summer	spring	summer	spring	summer	spring	summer
FS	0.219**	0.384***	0.047	0.342***	-0.041	0.282***	0.023	0.725***
	(0.096)	(0.111)	(0.103)	(0.097)	(0.095)	(0.085)	(0.176)	(0.145)
HSFG	0.293**	-0.152	0.191	-0.113	0.201	-0.037	0.542**	-0.436*
	(0.130)	(0.162)	(0.142)	(0.145)	(0.131)	(0.129)	(0.238)	(0.219)
Lyceum	0.457	0.068	1.657	1.289	1.732	1.233	2.592	-0.418
	-1.195	-1.330	-1.274	-1.206	-1.199	-1.062	-2.189	-1.850
Age	-0.108	-1.733***	-2.109	-1.292***	-1.998	-0.807**	-2.615	-5.099*
	-1.520	(0.390)	-1.674	(0.335)	-1.555	(0.324)	-2.892	-2.685
Female	-2.387**	-0.948	0.021	0.167	0.429	1.014	-0.948	-1.551
	-1.025	-1.172	-1.082	-1.044	-1.021	(0.916)	-1.940	-1.549
South	-6.217***	-2.181	-6.383***	0.141	-8.680***	0.422	-10.861***	-1.317
	-1.582	-1.900	-1.747	-1.860	-1.621	-1.616	-3.591	-2.634
_cons	-17.027	50.191***	56.901	38.804***	61.478	24.983*	118.189	218.318***
	-43.629	-16.385	-47.790	-14.212	-44.385	-13.302	-82.837	-73.616
Obs.,	68	120	62	110	61	105	57	94
R-squared	0.371	0.23	0.237	0.24	0.376	0.22	0.231	0.296

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 32: Decomposition of the R2 for the 2013 cohort

	spring				summer			
	raw coeff	stand coeff	corr	**R^2**	raw coeff	stand coeff	corr	**R^2**
AT	0,097	0,166	0,197	**3,26%**	0,274	0,355	0,207	**7,33%**
HSFG	0,419	0,370	0,262	**9,69%**	0,327	0,248	0,098	**2,42%**
Lyceum	0,326	0,032	-0,086	**-0,27%**	0,067	0,005	0,041	**0,02%**
Age	-0,390	-0,028	0,060	**-0,17%**	-1,733	-0,381	-0,317	**12,06%**
Female	-2,121	-0,220	-0,186	**4,10%**	-0,946	-0,071	-0,018	**0,13%**
South	-6,192	-0,438	-0,384	**16,82%**	-2,179	-0,106	-0,097	**1,03%**
				33,43%				**22,98%**
Final score	0,219	0,250	0,372	**9,28%**	0,384	0,379	0,287	**10,87%**
HSFG	0,293	0,259	0,262	**6,79%**	-0,152	-0,115	0,098	**-1,12%**
Lyceum	0,457	0,045	-0,086	**-0,38%**	0,068	0,005	0,041	**0,02%**
Age	-0,108	-0,008	0,060	**-0,05%**	-1,733	-0,381	-0,317	**12,06%**
Female	-2,387	-0,248	-0,186	**4,62%**	-0,948	-0,071	-0,018	**0,13%**
South	-6,217	-0,439	-0,384	**16,88%**	-2,181	-0,106	-0,097	**1,03%**
				37,14%				**22,99%**
Math	0,114	0,152	0,169	**2,57%**	0,425	0,435	0,255	**11,07%**
HSFG	0,400	0,354	0,262	**9,28%**	0,385	0,291	0,098	**2,84%**
Lyceum	0,180	0,018	-0,086	**-0,15%**	-0,999	-0,072	0,041	**-0,30%**
Age	-0,299	-0,021	0,060	**-0,13%**	-1,576	-0,347	-0,317	**10,97%**
Female	-2,261	-0,235	-0,186	**4,37%**	-1,601	-0,121	-0,018	**0,22%**
South	-6,336	-0,448	-0,384	**17,21%**	-3,028	-0,148	-0,097	**1,42%**
				33,15%				**26,23%**
Reading	0,072	0,064	0,121	**0,78%**	0,067	0,049	0,010	**0,05%**
HSFG	0,399	0,352	0,262	**9,24%**	0,215	0,162	0,098	**1,58%**
Lyceum	0,945	0,092	-0,086	**-0,79%**	1,763	0,128	0,041	**0,53%**
Age	-0,550	-0,039	0,060	**-0,24%**	-1,576	-0,347	-0,317	**10,97%**
Female	-2,561	-0,266	-0,186	**4,95%**	-1,568	-0,118	-0,018	**0,22%**
South	-6,688	-0,473	-0,384	**18,16%**	-3,578	-0,174	-0,097	**1,68%**
				32,10%				**15,03%**

Table 33: Partial r2 of AT, HSFG, FS, Math and Reading using as dependent variable GPA1, GPA2, GPA3 and final grade, 2013

	2013	
	spring	summer
GPA1		
r2 AT	2,46%	9,56%
r2 HSFG	14,31%	5,51%
r2 FS	7,85%	9,57%
r2 Math	2,04%	13,34%
r2 Reading	0,52%	0,23%
GPA2		
r2 AT	3,42%	10,77%
r2 HSFG	5,18%	6,95%
GPA3		
r2 AT	2,72%	10,00%
r2 HSFG	4,71%	9,31%
FG		
r2 AT	0,15%	22,26%
r2 HSFG	10,24%	7,96%

Table 34: Partial r2 of AT and HSFG using GPA1 as dependent variable, for females and males, 2013

	2013	
	spring	summer
female		
pR^2 AT	0,37%	13,76%
pR^2 HSFG	26,96%	5,23%
male		
pR^2 AT	8,80%	7,14%
pR^2 HSFG	12,28%	5,53%

OLS regressions when the main independent variable is the final score and the high school final grade is omitted from the control variables.

Table 35: OLS results for cohort 2013, Final score (without HSFG)

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
FS	0.287***	0.180***	0.086	0.157***	0.004	0.153***	0.151	0.410***
	(0.093)	(0.056)	(0.099)	(0.049)	(0.091)	(0.042)	(0.172)	(0.079)
Lyceum	-0.514	0.547	0.855	1.640	1.077	1.279	0.818	2.258
	-1.131	-1.181	-1.179	-1.057	-1.109	(0.920)	-2.095	-1.591
Age	-0.626	-1.381***	-1.537	-0.908***	-2.215	-0.528*	-2.938	-1.159
	-1.474	(0.376)	-1.597	(0.329)	-1.475	(0.311)	-2.807	-1.413
Female	-2.279**	-1.008	0.371	0.178	0.604	1.133	-0.297	-0.957
	-1.042	-1.156	-1.075	-1.039	-1.008	(0.892)	-1.960	-1.532
South	-5.651***	-2.785	-5.661***	-0.288	-8.184***	0.344	-9.855***	-2.650
	-1.594	-1.775	-1.713	-1.728	-1.579	-1.467	-3.637	-2.486
_cons	20.510	42.864***	58.058	32.839***	83.647**	24.618**	170.710**	97.261**
	-40.434	-11.436	-44.229	-10.019	-40.894	-9.386	-78.178	-39.944
Obs.,	69	122	63	112	62	107	58	104
R-squared	0.323	0.199	0.198	0.187	0.351	0.191	0.156	0.321

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 36: OLS results for cohort 2014, Final score (without HSFG)

	GPA1		GPA2		GPA3		FG	
	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer	(1) spring	(2) summer
FS	0.113**	0.161***	0.113***	0.159***	0.124***	0.176***	0.390***	0.330***
	(0.050)	(0.036)	(0.041)	(0.034)	(0.041)	(0.030)	(0.087)	(0.060)
Lyceum	0.292	0.483	1.504*	0.448	1.372*	0.545	0.636	1.046
	(0.918)	(0.823)	(0.767)	(0.776)	(0.771)	(0.699)	-1.620	-1.347
Age	-0.967	-0.408	0.338	-1.167**	0.367	-0.339	-0.414	-1.657*
	(0.681)	(0.500)	-1.184	(0.461)	-1.180	(0.477)	-2.363	(0.899)
Female	-1.029	-1.718**	0.370	-0.333	0.405	-0.005	-1.097	-0.869
	(0.843)	(0.799)	(0.712)	(0.751)	(0.716)	(0.677)	-1.475	-1.316
South	-0.311	0.034	1.634	-0.137	1.608	-0.933	2.109	-1.255
	-1.618	-1.179	-1.414	-1.057	-1.410	(0.953)	-2.825	-1.882
_cons	39.692**	19.817	5.971	39.123***	6.442	19.957	85.876	116.816***
	-17.267	-13.299	-29.948	-12.295	-29.868	-12.617	-59.888	-23.867
Obs.,	170	232	154	206	152	200	144	183
R-squared	0.057	0.121	0.091	0.160	0.095	0.169	0.149	0.194

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 37: OLS results for cohort 2015, Final score (without HSFG)

	GPA1		GPA2		GPA3		FG	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	spring	summer	spring	summer	spring	summer	spring	summer
FS	0.237***	0.220***	0.224***	0.243***	0.157**	0.263***	0.538***	0.420***
	(0.067)	(0.034)	(0.067)	(0.033)	(0.061)	(0.033)	(0.114)	(0.059)
Lyceum	1.171	-0.488	1.285	0.438	0.023	-0.372	1.957	1.751
	-1.187	(0.832)	-1.217	(0.786)	-1.132	(0.779)	-2.099	-1.371
Age	-0.271	-0.423**	-0.497	-0.483**	-0.400	-0.475**	-0.926	-0.220
	(0.349)	(0.177)	(0.344)	(0.187)	(0.324)	(0.183)	(0.598)	(0.671)
Female	1.064	0.219	0.760	0.404	0.728	0.274	1.461	0.747
	(0.835)	(0.786)	(0.843)	(0.748)	(0.774)	(0.739)	-1.427	-1.307
South	4.162**	-3.897***	3.777*	-1.660	3.025*	-1.558	6.720**	-2.686
	-1.989	-1.488	-1.954	-1.594	-1.776	-1.559	-3.207	-2.727
_cons	10.019	17.527***	17.023*	17.004***	22.841***	18.481***	81.248***	75.262***
	-9.192	-5.302	-9.100	-5.400	-8.497	-5.299	-15.417	-17.590
Obs.,	174	214	164	195	160	191	150	171
R-squared	0.116	0.228	0.113	0.280	0.070	0.308	0.177	0.265

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 38: OLS results for cohort 2016, Final score (without HSFG)

	GPA1		GPA2		GPA3		FG	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	spring	summer	spring	summer	spring	summer	spring	summer
FS	0.202***	0.276***	0.117*	0.240***	0.059	0.177***	0.318***	0.387***
	(0.074)	(0.038)	(0.065)	(0.036)	(0.063)	(0.034)	(0.115)	(0.065)
Lyceum	-0.906	1.373	0.108	1.040	0.313	1.994***	0.269	3.282**
	-1.176	(0.849)	-1.031	(0.796)	(0.994)	(0.756)	-1.785	-1.416
Age	-1.247***	-0.080	-1.969**	-0.099	-2.020**	-0.272	-2.881	-0.040
	(0.461)	(0.211)	(0.863)	(0.229)	(0.822)	(0.214)	-2.455	(0.627)
Female	2.230*	1.197	1.966*	-0.274	1.381	-0.461	3.024*	2.445*
	-1.132	(0.849)	-1.009	(0.799)	(0.965)	(0.757)	-1.703	-1.415
South		-3.152*		-1.557		-1.493		0.535
		-1.897		-1.923		-1.802		-3.432
_cons	37.040***	3.758	59.717***	9.070	67.222***	19.078***	146.443**	72.772***
	-12.086	-5.728	-20.601	-5.949	-19.648	-5.585	-56.553	-15.056
Obs.,	167	223	158	199	155	193	131	156
R-squared	0.102	0.215	0.073	0.204	0.059	0.174	0.078	0.228

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*