I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15350

High School and Exam Scores: Does Their Predictive Validity for Academic Performance Vary with Programme Selectivity?

Pedro Luis Silva
Carla Sá
Ricardo Biscaia
Pedro N. Teixeira

JUNE 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15350

High School and Exam Scores: Does Their Predictive Validity for Academic Performance Vary with Programme Selectivity?

Pedro Luis Silva
CIPES, University of Porto and IZA

Ricardo Biscaia
CIPES and University of Porto

Carla Sá
CIPES, EEG and University of Minho

Pedro N. Teixeira
CIPES, University of Porto and IZA

JUNE 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

High School and Exam Scores: Does Their Predictive Validity for Academic Performance Vary with Programme Selectivity?*

Students are admitted into higher education based on their past performance. This paper compares two measures of past cognitive skills: teacher and national exam scores. By using a nationwide dataset, we look at how the predictive power of teacher assessment and exam scores for selecting successful students may vary with the degree of selectivity of higher education programmes. We find that teacher scores predict students' performance in higher education more accurately, and its predictive power remains the same independently of the selectivity programme indicator considered. We found that national exam scores are noisier and only gain relevance for highly selective programmes. Furthermore, we explore national exams' volatility and institutional selectivity as potential mechanisms to justify the results. Our results provide solid policy hints on the role that high school scores and admission exams should have for access and performance in higher education.

JEL Classification: I23, I21, I20, J24

Keywords: admission exams, teacher scores, higher education, selectivity

Corresponding author:
Pedro Luis Silva
Centre for Research in Higher Education Policies
Rua 1.o Dezembro 399
4450-227 Matosinhos
Porto
Portugal
E-mail: pedro.luis.silva@cipes.up.pt

* We thank Steve Desjardins and Pilar Beneito for valuable comments. The paper has also benefited from the comments of participants at the LESE conference in Lisbon, and at the seminar(s) of University of Valencia. We thank the Direção-Geral do Ensino Superior (DGES) and the Direção-Geral de Estatísticas da Educação e Ciência (DGEEC) for the access to the two data sets used in this paper. The data sets were merged at the premises of DGEEC and cannot be disclosed to anyone who has not signed a confidentiality agreement with DGEEC. We acknowledge financial support from the Fundação para a Ciência e Tecnologia (FCT), grant PTDC/CED-EDG/5530/2020, UIB/0757/2020, and UIDB/03182/2020. The authors have no relevant or material financial interests that relate to the research described in this paper. All omissions and errors are our own.

1 Introduction

The expansion of higher education (HE) has created important challenges on the side of institutions and the need to find the best mechanisms to select those most suitable to the demands of each programme (combination of institution/degree). Higher education institutions (HEIs) compete with each other for students, especially for the most talented ones. Their quality and the quantity of the programmes they offer are often judged by the ability of the students they attract. Recognizing that student recruitment is essential to increasing excellence, institutions are paying growing attention to the design of recruitment strategies and admission criteria. Furthermore, mismatches between students' ability and the requirements of the HE programmes or institutions may worsen student academic and future labour market outcomes (Dillon and Smith, 2017; Campbell et al., 2021).

Institution quality and selectivity are normally measured by the average achievement of the admitted students. Prior achievement, such as high school and admission exam scores, are seen as predictors of future student academic performance. Such widely used selection mechanisms have played a signalling role in identifying best matching options. Standardized tests are an important selection method since it is an efficient tool for student evaluation. System-wide, it is also effective to control grade inflation in the high school system, as it provides a reference standpoint to compare the high school grading system. The main disadvantage of such method is that it is a one-shot approach. All the students' knowledge is assessed only a few times, if not once. Therefore, this approach is subject to variables other than knowledge, such as the ability of the student to perform well under pressure. High school scores result from continuous testing of students throughout high school and, therefore, might result in a better assessment. This approach is validated by the finding that high school scores are a relevant predictor for HE success, both measured by the Grade Point Average (GPA) (Zwick, 2019; Silva et al., 2020) or by completion rate (Westrick et al., 2015; Silva, 2022). Hence, institutions would guarantee better students (on average) for their programmes by using high school scores as a criterion.

This paper looks at how the predictive power of teacher assessment and exam scores for selecting successful students may vary with the degree of selectivity of HE programmes. To do so, we develop a measure of the programme's degree of selectivity, a continuous

measure based on the number of applicants to the programme in the top 10 percentile of all students admitted to HE. The programme selectivity indicator is then a measure that considers information revealed by students' preferences and evaluates candidates' quality simultaneously. The construction of this indicator allows us to identify tiers of selectivity in HE. The higher the indicator, the more high ability students the programme receives and the more selective it is.

We rely on longitudinal, administrative data for all students attending HE in Portugal, from 2013 to 2018. We are one of the first studies to use this longitudinal data set to cover the population of HE students in Portugal, which improves on research that uses relatively small samples of students or universities. The data set provides a rich set of covariates, such as student demographics, socioeconomic background and performance in HE. In our paper, success in HE is measured by the number of completed credits in the first year, a count variable, which we model using a negative binomial model. For all HEIs, we observe students' admission and subsequent performance, having information about individual characteristics as well.

At first, we find that high school scores seem to be a robust predictor of students' academic performance at university. Thus, the predictive power of high school scores on future student performance is stable and sometimes decreases slightly in selectivity levels. However, when looking at programmes' selectivity, we conclude that high-stakes exams are relevant given exam scores prediction of university performance increases as long as the selectivity increases.

We provide evidence that justifies why admission exams might better predict students' ability for more selective programmes. First, we measure exams' volatility, showing that the admission exams are more volatile for those enrolled in less-selective institutions. As a result, the admission exam has less predictive power on performance for less selective programmes. Secondly, we identify institutions' characteristics that might explain their degree of selectivity. For example, more selective institutions rely on more strict admission criteria and allocate a higher weight to the admission exams rather than the high school score. Finally, we discuss how students' socioeconomic backgrounds may affect our results. We discuss possible mechanisms in which parents' socioeconomic status can affect students' performance in high school and thus affect our results.

These findings contribute to several related strands of literature regarding the stu-

dents' performance, selection and selectivity in HE. On the selection side, this paper adds to the discussion on which screening devices should a HEI consider when selecting their students in order to assure the best match. Prior work shows that combining high school scores and high stake assessments is the best way of selecting students (Westrick et al., 2015; Zwick, 2019; Silva et al., 2020). Nevertheless, both measures have pros and cons. High school scores have the huge disadvantage of comparing students, since grading standards are bound to vary significantly between schools (Atkinson and Geiser, 2009). At the same time, high school scores are the result of continuous testing of students throughout high school, and therefore might result in a better assessment. In turn, high-stake assessments have the natural advantage of not being subject to different grading standards.

Additionally, this paper also relates to the existing research on studying students' returns to attending elite universities. Dillon and Smith (2017) and Walker and Zhu (2018) analyse the mismatch between students' ability and institution quality. They find that the most informed students pursue higher institution quality rather than a close match between institution quality and their abilities. Thus, returns to elite universities (Brewer and Eide, 1999; Li et al., 2012; Anelli, 2020) vary across institutions largely due to student selectivity. We observe that more selective institutions rely in different screening devices to select their students, even if the gain of it is relatively low.

This paper also contributes to the literature on the predictive power of tests scores on the implications for admission policies. For instance, although high-stakes assessments influence students' decision to apply to HE (Papay et al., 2016), Bettinger et al. (2013) proposes to reduce the number of ACT components to improve college admissions. At the same time, other papers show the predictive power of past performance, such as high school transcripts (e.g., Belfield and Crosta, 2012; Cyrenne and Chan, 2012; Dooley et al., 2012), are relevant to college admission as well. Our results also corroborate that high-stakes assessments are a predictive tool of students' future academic performance.

The paper proceeds as follows. In the next section we will focus on the Portuguese context, namely by providing some background of HE system. This will be followed by the presentation of the data and methodology in sections 3 and 4. Section 5 presents the main results of our empirical analysis and in section 6 we will discuss some of the mechanisms that might explain our results. Section 7 concludes.

2 Background: access and entry requirements

HE in Portugal is currently composed of universities and polytechnics, both public and private. Each year, the government sets the number of vacancies available for each programme (combination of institution/degree). Recently, the *Numerus Clausus* system has not constrained the overall access to HE due to the combined effect of the increase in the system capacity and the decrease in demand. However, *numeri clausi* can still impose caps on prestigious programmes and institutions, where demand clearly surpasses the vacancies available (OECD, 2019).

HEIs select their students based on scores in admission exams and high school performance, which are mandatory requirements to gain admission to a HE programme. In the last two years of secondary education, students must take at least 4 subject-specific national exams: two in the 11th grade and two in the 12th grade. The government imposes some boundaries on admission requirements that HEIs must respect. The application score of each candidate is a weighted average of the high school score and the scores on the subject-specific national exams taken at the end of secondary education. However, the government imposes that each HEI should allocate minimum weights of 50% and 35%, respectively, to the high school and the admission exams' scores. The institution distributes the remaining 15% to either one or both admission criteria. Different programmes require different exams and might use different weights. Therefore, public HEIs have limited instruments to select their students within such a centralized access system. Although the majority of programmes consider combinations of one or two national exams, more selective programmes tend to give a higher weight to the admission exams (maximum of 50%) while low quality programs tend to give a lower weight (minimum of 35%).

Nevertheless, each institution/programme can decide which national exams to be consider as admission exams. They often define two or more possible combination of admission exams, but any combination must always consider the field-specific exam of each degree when defined by the government (for instance, Mathematics is the mandatory field-specific exam for economics and engineering degrees).[1] In each combination, the number of exams considered is, in general, either one or two; only medicine programmes require three exams. If considering more than one exam, they must have equal weight.

[1]For a discussion on the nature of these admission exams see Silva (2022).

Access to HE can be gained through several alternative pathways. Within the National Entrance Competition (*Concurso Nacional de Acesso*), the General Access Regime (GAR) is a national competition to all public HE programmes. About 83% of university students and 64% of students in polytechnics take this route, which makes it the main pathway to HE (OECD, 2019). Applications are made through a centralized admission process managed by Direção Geral do Ensino Superior (DGES; Directorate-General for Higher Education). On the other hand, candidates have to go through an institutional contest when applying to private institution programmes. There is still a number of special contests and regimens that target specific groups of candidates, for whom the general regime is not appropriate (e.g. mature students aged above 23 years old, international students).

After observing each programme's admission criteria and her national exam results, each applicant in the GAR can rank up to six programmes to which she applies. Students have an incentive to report a set of preferences (or at least a subset) that they judge feasible, based on the information of the entry scores in previous years. Given that they observe everything, there is an incentive for students to reveal their truthful rank of preferences, conditional on the national exams and high school results. Applicants know that they will be allocated to their higher feasible stated preference. The government solves the matching problem using a Deferred Acceptance mechanism (DA; Gale and Shapley (1962)). Through an iterative algorithm, the government finds a match between applicants' preferences and each programme's capacity. In the end, each applicant can gain admission in one programme only. If she does not accept it or is not offered a place, she will need to re-apply. There are three rounds each year. In the first round all the places are in the contest. Second and third rounds only happen for programmes not filling all places in previous rounds. In 2021, about 90% of the places were filled in the first round. The minimum admission score for each programme is that of the last student who was offered a place.

3 Data and variables

3.1 Dataset

We use two datasets from two different sources: i) Applications to public HEIs (DGES, 2019); ii) Students' performance at HE (DGEEC, 2019).

The application dataset contains all application data from DGES (Directorate-General for Higher Education). For each candidate, it contains her list of ranked preferences. Each preference listed has information on the degree and institution, the corresponding application score, high school score, admission exams' scores, and placement result. Moreover, for each programme, the dataset also gives information regarding the number of vacancies, admission requirements, admission score threshold and field of study. We have data for all applicants to public HE.

The dataset on students' performance at HE is provided by the Direção-Geral de Estatísticas da Educação e Ciência (DGEEC; Directorate-General for Education Statistics). This dataset is a snapshot of all students enrolled in HE. For each student in each year, the dataset retrieves the programme of enrollment, the field of study and academic performance. This dataset also reports student socio-economic characteristics, such as parents' education and job status. The latest dataset includes both public and private sectors.[2] Given that we only observe the applications to the public sector, we merge these two datasets for the public sector and find a one-to-one merge for 96% of the students admitted to first year of HE.[3]

Our combined dataset consists of six cohorts, from 2013-2014 to 2018-2019, including those students whom we can observe both the first-time application and the performance at HE by the end of their first year. We only consider students that applied under the General Access Regime (GAR; the centralized process of application to public HEIs)[4], in the first round of applications.

[2]Public HEIs represent around 80% of the system (According to DGEEC (2019), of the first-year students enrolled in 2016/2017 for the first time, 83% were in a public institution.)

[3]The link between the application and performance dataset was made by the first author at the premises of the Ministry of Science, Technology and Higher Education in Portugal.

[4]There is a reduced number of students that under GAR may obtain admission through special regimes, such as sons of diplomats, refugees, regional preferences, for instance. We exclude these students from our analysis as GAR accounts for approximately 70% of the candidates to Portuguese public HE and is the standard way of assessing HE. Finally, we only consider individuals that applied to public HEIs in Portugal Mainland.

3.2 Programme selectivity indicator

The definition of programme selectivity is central to our analysis. Based on student revealed preferences, we compute an indicator to identify tiers of selctivity in the HE system. That is crucial to determine the potential impacts of changes in HE. Portela et al. (2008) proposed a programme/institution strength measure (s_j) computed as the ratio between the total number of applicants choosing the programme j as a first option and the total number of available places for the programme. Nevertheless, the strength index does not consider how good candidates choosing a given programme as the first choice are. To account for applicants quality, we compute an adapted version of that measure which we designate as selectivity indicator[5] and is computed as

$$SI_j = \frac{fop90_j}{p_j} \qquad \text{(selectivity indicator)} \tag{1}$$

where $fop90_j$ is the total number of students who: (i) rank programme j as their first choice; and (ii) have an application score above the 90th percentile of all students admitted to HE in that year. p_j represents the number of vacancies available at the programme j. If SI_j is greater than zero, that means that programme j was able to attract some outstanding students. The higher the score, the more high ability students the programme receives and the more selective it is. Those students are in the top 10% of HE admitted students. SI_j increases as more outstanding students are admitted to programme j in a particular year.

[Insert Table 1]

SI_j is an indicator of the programme's selectivity level, a continuous variable that ranks programmes according to their degree of selectivity. For some programmes, this classification differs substantially from the conclusion driven by the more spartan strength index s_j. For instance, according to Table 1, the economics degree at the Universities of Minho and Coimbra has a strength index higher than one and an almost zero programme selectivity indicator. In other words, a high percentage of students ranked that programme as their first choice (economics/U Coimbra or economics/U Minho), but most

[5]This is adapted from the "excellence index", introduced by DGES (2020), and that it is an extension of the strength index. Nevertheless, they consider a programme as excellent if the admission score is higher or equal to 170 out of 200. We are going to use the percentiles instead of an absolute value.

of those students entering in their first option are not part of the top 10% students in the system. The strength index s_j is high for those programmes only because a good number of students wanted to attend that programme (i.e. due to location, prestige), but it does not necessarily translate into high-ability candidates. The programme selectivity indicator (SI_j) is then a measure that considers information revealed by students' preferences and evaluates candidates' quality at the same time.[6]

[Insert Figure 1]

The selectivity indicator gives us the indication of which programmes can be considered as the most selective ones. According to Figure 1, less than one-third of the programmes have a selectivity indicator above zero. Moreover, there are very few programmes with an indicator above 0.5 (see Tables A1 and A2 of the appendix). Thus, there is a concentration of programmes at the bottom of the selectivity indicator's distribution. In terms of students, more selective programmes are able to capture a high ability pool of incumbents. Although 50% of the students gain admission to a programme with a non-null selectivity level (see Table 2), only approximately 10% of the students are admitted to the top 5% selective programmes (see Figure 1).

[Insert Table 2]

According to Table 2, if we divide students into two types of programmes, those with a zero selectivity programme indicator and those with a non-null selectivity level, individual characteristics are overall the same across the different types. Nevertheless, when it comes to parents' education, we can see that more educated parents can allocate their children to more selective programmes. We are going to explore this channel later on in the paper.

3.3 Descriptive statistics

Table 3 presents the descriptive statistics for our analysis dataset.

[6] We performed a sensitivity analysis with different percentiles of the application score of all students admitted. Namely, we run a robustness check with 85th, 80th and 75th percentiles on the definitions of SI_j and we defined SI_j^{90}, SI_j^{85}, SI_j^{80} and SI_j^{75}. The destribution of the programmes based on their degree of selectivity has remained the same for most of the programmes. Additionally there is a strong correlation between different definitions of programme selectivity indicator ($corr(SI_j^{90}, SI_j^{85}) = 0.9744$; $corr(SI_j^{90}, SI_j^{80}) = 0.9373$; $corr(SI_j^{90}, SIj^{75}) = 0.8971$).

[Insert Table 3]

We observe that 70% of the students are admitted to either their first or second-ranked preference. Students recognize that they will be admitted to their highest feasible preference, and so, on average, they try to not waste preferences with scenarios that are way out of their possibilities. The majority of students at HE studied at a public high school, and 58% of them are female. However, only 30% of the students has a mother/father with a HE degree, and one-fourth of the students receive a maintenance grant. This indicates that a large proportion of students admitted to HE seems to be from a disadvantageous background. Half of the students are admitted to a programme with a positive selectivity indicator, which are mainly concentrated in social sciences, engineering, and Health, which represent 15% of the programmes.

4 Methodology

In this section, we present our estimation strategy. Our main goal is to evaluate whether the predictive validity of teacher assessment and exam scores for selecting successful students may vary with the degree of Selectivity of higher education programmes. We focus on student's performance during the first year of HE. We consider as outcome (y_i) the number of credits obtained through the *European Credit Transfer System* (ECTS) by the end of the first academic year.

We estimate a fixed effects negative binomial model based on the following arguments. Firstly, the number of ECTS is a positive integer and its distribution is skewed to the right, which implies a count model type (either poisson or negative binomial, see Cameron and Trivedi (2005)). Since the over-dispersion test rejects the null hypothesis of absence of over-dispersion, the negative binomial model is appropriate. Secondly, institutions enjoy a certain level of autonomy and freedom when organizing each programme study plan, allocating their teaching staff and defining the teaching and evaluation methods. All combined may make the average number of completed ECTS specific to each programme (level of degree/institution combination), which justifies estimating a fixed-effects negative binomial model to account for possible unobserved heterogeneity.[7]

[7]Similar studies use negative binomial model when measuring demand and performance in higher education. For a review see Cardoso et al. (2008), Portela et al. (2009), Hilmer and Hilmer (2011), Akee et al. (2014), and Aina et al. (2018) .

The number of ECTS, y_i, is modeled by means of a negative binomial distribution:

$$P(y_i|\mathbf{x_i}) = \frac{\Gamma(\alpha^{-1}+y_i)}{\Gamma(\alpha^{-1})\Gamma(y_i+1)}\left(\frac{\alpha^{-1}}{\alpha^{-1}+u_i}\right)^{\alpha^{-1}}\left(\frac{u_i}{u_i+\alpha^{-1}}\right), \tag{2}$$

where $u_i = E(n_i|\mathbf{x_i}) = exp(\mathbf{x_i}'\beta)$, α is a constant, and $\mathbf{x_i'}$ the vector of controls, including the three main variables of interest: the high school score, the mean of the admission exams' scores and the programme selectivity indicator. We have tested for endogeneity of the programme selctivity measure, as follows: first, we have estimated the model by OLS and computed the predicted residuals; second, we estimate the negative binomial model with the predicted residuals as an additional explanatory variable; and finally, based on a significance test on its coefficient, we have rejected the endogeneity hypothesis.

We control for individual characteristics such as gender, type of high school attended, and whether she is a local student. Later in the paper we also consider parent's education as a control. Additionally, we control for cohort, field, institution and programme characteristics, as well as preference fixed effects. Preference is decomposed into dummies indicating the order with which each student applied to the programme that she was admitted to.[8] Finally, we also include later in the paper high school fixed effects as a robustness check.

5 Results

Table 4 presents a summary of the main results for all our different estimations. First, we start our analysis by considering the reduced form proposed in column (1). Then, in column (2), we include the control variables, and in column (3), we add the programme fixed effects. Later, we consider the programme selectivity indicator in our analysis in columns (4) to (6).[9] Finally, given that we are estimating a negative binomial model, we only present the average marginal effects.

We observe that high school scores have a higher effect on student performance at university than the admission exams, irrespective of the model that we use. In particular,

[8] After modeling the number of ECTS, we estimate the marginal effects. See Merkaj et al. (2020) for a full derivation of the negative binomial model.

[9] The programme selectivity indicator is collinear with programme fixed effects. For that reason, we drop the programme fixed effects and introduce institution and field fixed effects as a proxy for programme effects.

within the same programme, we observe that, on average, a student with an extra point on her high school score completes 0.280 more ECTS in the first year (column 4). However, an extra point on the admission exams only translates into 0.072 more ECTS, on average, in the first year. Thus, the scores attained in high school are more reliable predictors of future performance when compared to scores obtained in the admission exams.

Nevertheless, it is also true that this prediction varies according to the type of programme considered. For example, according to column (5), students tend to complete fewer ECTS within the first year on more selective programmes. In particular, when the programme selectivity indicator increases by 0.1, students complete 0.48 fewer ECTS within the first year. Thus, on average, students in more selective programmes find it more challenging to complete the first year, which is not surprising.

[Insert Table 4]

In columns (6), (8) and (10), we include the interaction terms between the programme selectivity indicator and the two measures of past performance. In order to assess whether the marginal effects for our main variables of interest differ depending on the selectiveness nature of the programmes, we estimate the average marginal effects at different percentiles of the programme selectivity indicator for the specification used in column (10). Results are presented in Figure 2. We observe that for the less selective institutions, the high school score is more relevant to explain the number of ECTS completed than the admission exams. However, when we look at the top 5% of selective programmes, the conclusion is different. The national exams gain precision in predicting future student performance at the top percentiles of the distribution. From Figure 2, we can also infer that, on average, the effect of high school score on the ECTS completion is more stable, and sometimes decreasing with the increase in levels of selectivity.

The national exams' prediction of university performance increases as long as our measure of selectivity increases. That is to say that standardized tests gain relevance when we consider the most selective programmes. Those programmes admit a pool of high ability students, since the admission cutoff is usually very high. Thus, students need to perform exceptionally well in high school and the national exams. In that case, the national exam becomes a more relevant predictor of future academic performance than the high school score. In the next section, we provide and explore some reasons for why that would be the case.

[Insert Figure 2]

6 Discussion

We provide four possible mechanisms and explanations for the heterogeneity present on the results, even though data availability only allows us to test empirically three of them. We present these explanations as the *competition effect*, the *volatility effect*, the *admission criteria effect* and the *parents' education effect*.

Our untested explanation is the *competition effect*. It is based on the irrelevance of the admission exams for students, which may vary with the selectivity level of the programmes they are applying for. When applying for highly selective programmes, students have no room to fail, and they will be mostly well prepared for the exam. Therefore, the exam result will be most likely related to their capacities at full performance and preparation levels.

On the other hand, if students are applying for less competitive positions or are less ambitious in terms of their HE choices – for instance, if they would instead enrol by proximity rather than their perception of programme quality or reputation – they face the admission exam less seriously. In this scenario, their performance might reflect less of their knowledge, but rather their (relative) lack of preparation or interest in the admission exam, since they will be admitted to one of their most preferred programmes anyway. Therefore, in the context of low-selective programmes, the admission exams' scores might reflect relatively less the student actual capabilities, when comparing to those students that were able to enroll in high-selective programmes. This helps explain why the admission exams might be a good predictor of student capacity, but more so for more selective programmes, where the exam scores are potentially more meaningful.[10]

6.1 *Volatility effect*

The volatility of the results of admission exams might contribute to its better explanatory power only at the more selective level. To participate in the first round of the General Access Regime, the students must take the admission exam once in that same year.

[10] We are unable to test this effect, since student preferences are only revealed after taking the admission exam. It is not possible to capture in our dataset the students that would have applied for highly selective programmes if their exam scores were better.

It means that the admission exam is a one-shot procedure, and therefore its result is more dependent on variables other than the knowledge or skills of the student, such as pressure or any other idiosyncratic factor that may happen during the exam. This does not say much regarding those who entered at highly selective institutions, since these have had certainly higher admission scores (and high school scores too) and are therefore more homogeneous in terms of their scores. However, students that entered institutions with lower requirements might consist of different types of students, namely: those who had significant high school scores but did not have a high performance in the exams; or students that have similarly graded exams when comparing to their high school scores. Therefore, the effect of the admission exams would be blurred in less selective institutions. These institutions combine students that suffered from those idiosyncratic effects or not, while more selective institutions will have much fewer students suffering from those effects, because if they did, their probability of being admitted to a high-selective institution would decrease significantly. Therefore, the results of the admission exams will be more volatile for those enrolled in less selective institutions, giving the admission exam less predictive power on performance for those institutions.

In order to assess the existence of the *volatility effect*, we have displayed in Figure 4 the distribution of the scores in admission exams and high school. As expected, the scores of students that access more selective institutions are higher than those who did not enter, both in the high school and the exams. However, it is clear in both samples that the admission exams are more volatile than the high school score, reflecting its one-shot nature. Since from Figure 4 it is not visible that the distribution of students who were admitted to more selective programmes is less volatile when comparing with students attending the less selective ones, we resort to Figure 5, where we computed the Coefficient of Variation[11] for both high school score and admission exams for different cohorts of elitism in the HE System. Since the coefficient of variation is an aggregate measure, and can not be obtained per observation, it cannot be included on the main regression for our research question.

Some important conclusions can be drawn from Figure 5. The result from Figure 4 is confirmed: independently of the selectivity level, the volatility of the admission exam score is relatively higher than that of the high school score. Then, as the selectivity

[11]The coefficient of variation measures the relative dispersion of scores around for its mean.

level increases, both admission exams and high school score volatility decreases, which can be attributed due to the nature of the distribution, highly concentrated on the right part of the [95,200] grade interval, since students that have entered in very high-selective programmes ($SI_j > 0.6$) have high scores both in the admission exam and in high school. The most important conclusion is on the *Volatility Effect*. As we hypothesized, the difference between the volatility in admission exams and high school exams would be decreasing for increasing degrees of selectivity. That is clear in Figure 5, where the difference is big between 0 and 0.6, but very small when the selectivity indicator is higher than 0.6.

[Insert Table 5]

[Insert Figure 4]

[Insert Figure 5]

6.2 *Admission criteria effect*

Other explanation for the predictive power of the admission exams for the performance in HE only for the most selective programmes relates to the institution's ability to select the required admission exams for each programme. As discussed previously, in Portugal, institutions are not able to do their own admission exams, or to attribute it more weight than the maximum of 50% allowed. However, there is still a significant range for admission exams' differentiation, as institutions can, if they are more strict, for instance, demand the exams of Mathematics as mandatory for access; or allow a reduced number of combinations between different admission exams. On the other hand, lower-tiered institutions can ease the access to their institution by allowing entrance with just the Portuguese exam, or allowing for a great variety of combinations to suit the candidate's best interests, namely, by allowing the student to select the exam in which he better performed on. Given the existing rules to choose the required exams, we would expect more selective programmes/HEIs to choose the exams in fields that are essential for the programme, whereas less selective institutions would choose exams that maximize the number of admitted students. The admission exam is expected to be more meaningful for the content of the programmes in the more selective institutions when compared to

the less selective institutions. This would therefore make the admission exam a weaker predictor of performance on non-selective institutions.

To test this hypothesis, we have estimated a regression, in which the dependent variable is the selectivity indicator (Table 6). Each observation is a programme, and we opted for a Tobit Regression due to the high number of zeros of the dependent variable. As explanatory variables, we use the information on whether the programme is more selective or not according to the limited selectivity options the institution has. The variable "1 Mandatory Exam" is a dummy variable that reflects whether the programme presents at least one exam that is repeated in all its combinations - which is a selectivity procedure given the students inability to avoid such mandatory exams. The "Number of different exams" is a variable that shows the number of different exams used for admission across all combinations. The following variables are categorical, as they refer to the number of combinations allowed, being *one combination* the omitted category. Another group of categorical variables distinguishes between the procedures where admission exams are worth 35%, 40%, 45% or 50%, with 35% being the omitted category. Proportion of females, of public school students, and of non-local students are included as controls.

[Insert Table 6]

The regression shows that part of the selectivity phenomenon can be explained by the selection practices of the institutions. Allowing for less entry combinations, as well as giving a higher weight for admission exams significantly improves the selectivity of the programme. While having a mandatory exam does not seem to have a significant effect after introducing field, institution and cohort fixed effects, the coefficient for the number of different exams is also positive. This reflects that when the program is assumed to have the number of different entrance exams combinations fixed, having more exams implies more selectivity, as the student is most likely forced to present results on more than one exam. For example, Medicine applicants have to take three exams to access the program: Mathematics, Biology and Chemistry. This reinforces the main result of this paper - that admission exams are a good predictor of academic performance, but mostly on the most selective programmes. The selectivity of the program plays a role in reinforcing the effect of the admission exams in the performance of students in more selective programmes.

6.3 *Parents' education effect*

As it can be seen in Table 2, students can be divided according to the selectivity of the programme they attend. The composition of students that belongs to each group of programmes is relatively similar, except for the difference in terms of the parents' educational level. This poses a potential problem as the selectivity indicator variable is correlated with the educational level of parents, and therefore our estimate of the effect of attending a more selective programme in students' performance might be confounded by this difference in the student composition in each group. However, including the variables on parental education would imply i) A reduction in the number of observations per year, as this information is not available for all students; and ii) significant heterogeneity problems, given the expected high correlation between parental education and the performance variables associated with high school, that is, the high school scores and the admission exam scores.

Our expectations between the relationship of parental education and the performance variables stem from different factors. Firstly, on the relationship between socioeconomic variables and performance, as high-endowed parents usually employ means for their sons and their daughters to perform well in general. This might include investing in extra-curricular activities or private tuition that have a positive relationship on student performance both on the long run but also on the one-shot admission exams (e.g. Smyth, 2009). Another path for this advantage to occur is on the selection of schools, since having a better socioeconomic background might imply accessing better or more performance-oriented high schools (e.g. Graddy and Stevens, 2005; Green et al., 2012). In Portugal, there is even the possibility that some schools might be inflating their high school scores due to the tight competition to access HE in the country (Nata et al., 2014). This reveals also that these students might have better access to the most selective HEIs as well. Secondly, better qualified parents - even after assuming equal socioeconomic conditions - will have a better perception of the value of HE studies for their children, and therefore are a direct factor for better performance across all education levels.

As a robustness check we present on Table A3 of the appendix the main regression of the paper, but adding as explanatory variables the dummy variables on whether the mother and father of the student have a HE degree, as well as the interactions between these two variables and the high school score and admission exams variables. As it can

be noted from the previous section, the results for the programme selectivity indicator variable, as well as for the internal score and national exams score exhibit similar patterns of significance. Therefore, we can conclude that our results are robust in relation to the inclusion or exclusion of a variable presenting potential heterogeneity problems.

7 Conclusion

In this paper, we assessed whether the performance on high school scores or admission exam scores at student level are good predictors for performance in HE. Namely, on the completion of credits (ECTS) in the first year of the programme. To do so, we conducted a negative binomial regression. Other than the necessary cohort and preference fixed effects, we tested the robustness of such relationship by making different regressions with a number of controls, as well as field, institution and programme fixed effects. Additionally, we have used a measure for student demand as a proxy for selectivity at programme level.

We have concluded that both the high school scores and admission exams are positively correlated with the performance of the student for all the model specifications we have tested, even though the effect of the admission exams is lower. The most novel conclusion that we find in our analysis is that when separating the effect of our two main variables by the degree of selectivity of the programme, we conclude that the high school score effect is more relevant in less selective programmes, and that the admission exams become more relevant in more selective programmes. This result indicates important differences between how performance is explained is more or less selective programmes.

We sought an explanation for this phenomenon, and we provide four possible mechanisms that justify our results. Nevertheless, we only prove three of them empirically. The first is the "Volatility Effect", in which we conclude that the one-shot nature of the admission exams, coupled with the selection issue that those who entered more selective programmes are those that have had a good admission exam, leaves us with the idea that the grades in admission exams of students in less selective programmes are more dependent from one-shot idiosyncrasies. Therefore, admission exam scores reflect less of the students' ability than those assessed in more selective programmes.

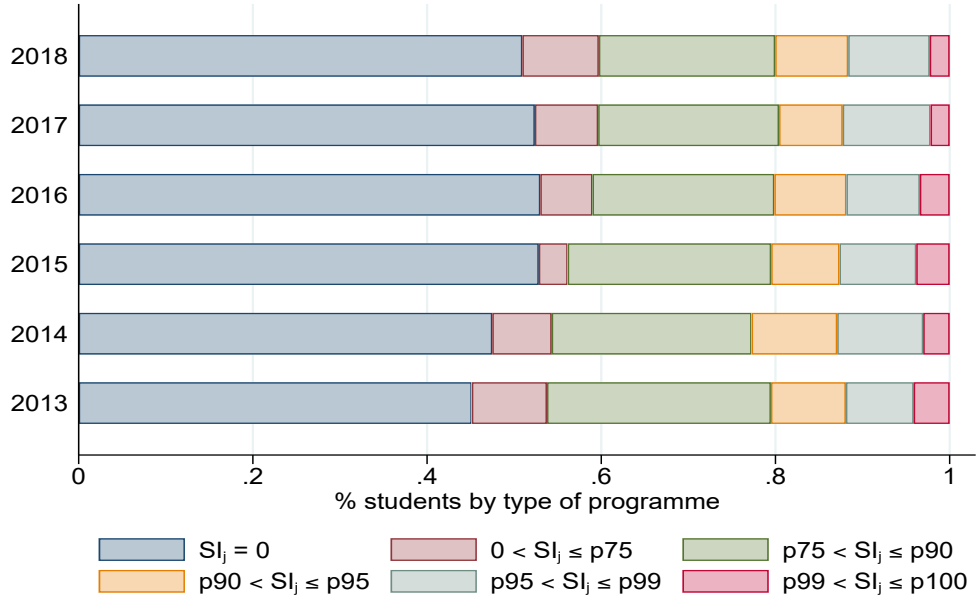
The second mechanism is the "institutional selectivity effect", which arises from the fact that more selective programmes might have different admission standards for their

admission exams. We observe that these programmes usually 1) allow less combinations of admission exams; 2) attribute a larger weight to the admission exams on the application process; and 3) ask for admission exams that are more closely related to the field of the programme. Therefore, the admission exams in more selective programmes are more meaningful because they signal knowledge more related to the programme content and its difficulty than in less selective programmes.

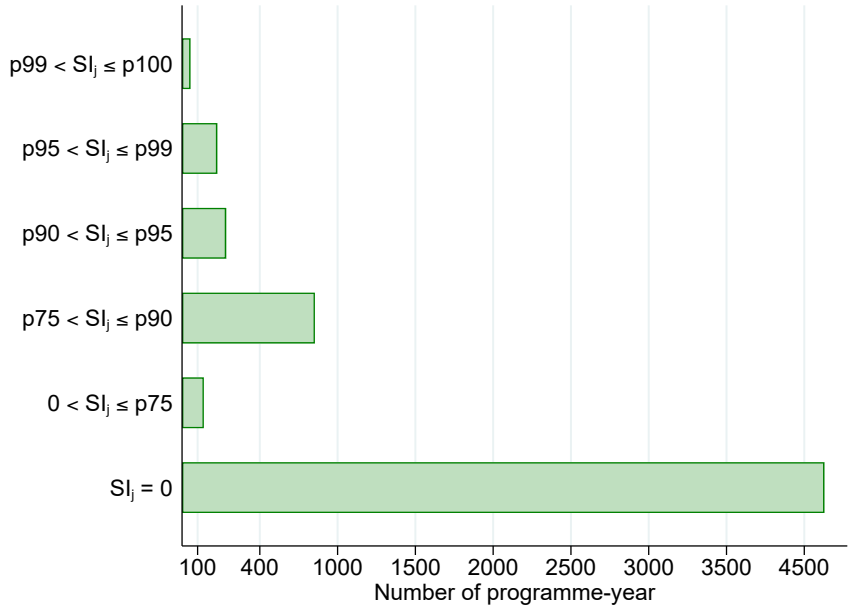
Thirdly, we discuss how the students' socioeconomic background can affect our results. We present different channels in which parent education can affect students' performance at high school and then bias our results. Namely, high-endowed parents usually have more resources to invest in students' extracurricular activities as well as in the preparation of students to the admission exams. Additionally, also the choice of high school might affect students' access to HE. Nevertheless, even controlling for parents' education, our results remain the same.

Finally, our results can provide some solid policy hints on the role of the admission criteria to HE. In light of these results, we can argue in favour of multiple criteria in higher education admission processes. As exam scores and high school grades are valid predictors of future performance and success in higher education, institutions should consider both when selecting students. However, the higher predictive power of exam grades in the most demanded programmes may justify that different weights on each criterion should apply to different programmes.

8 Figures



(a) Students enrolled in selective programmes



(b) Number of selective programmes

Figure 1

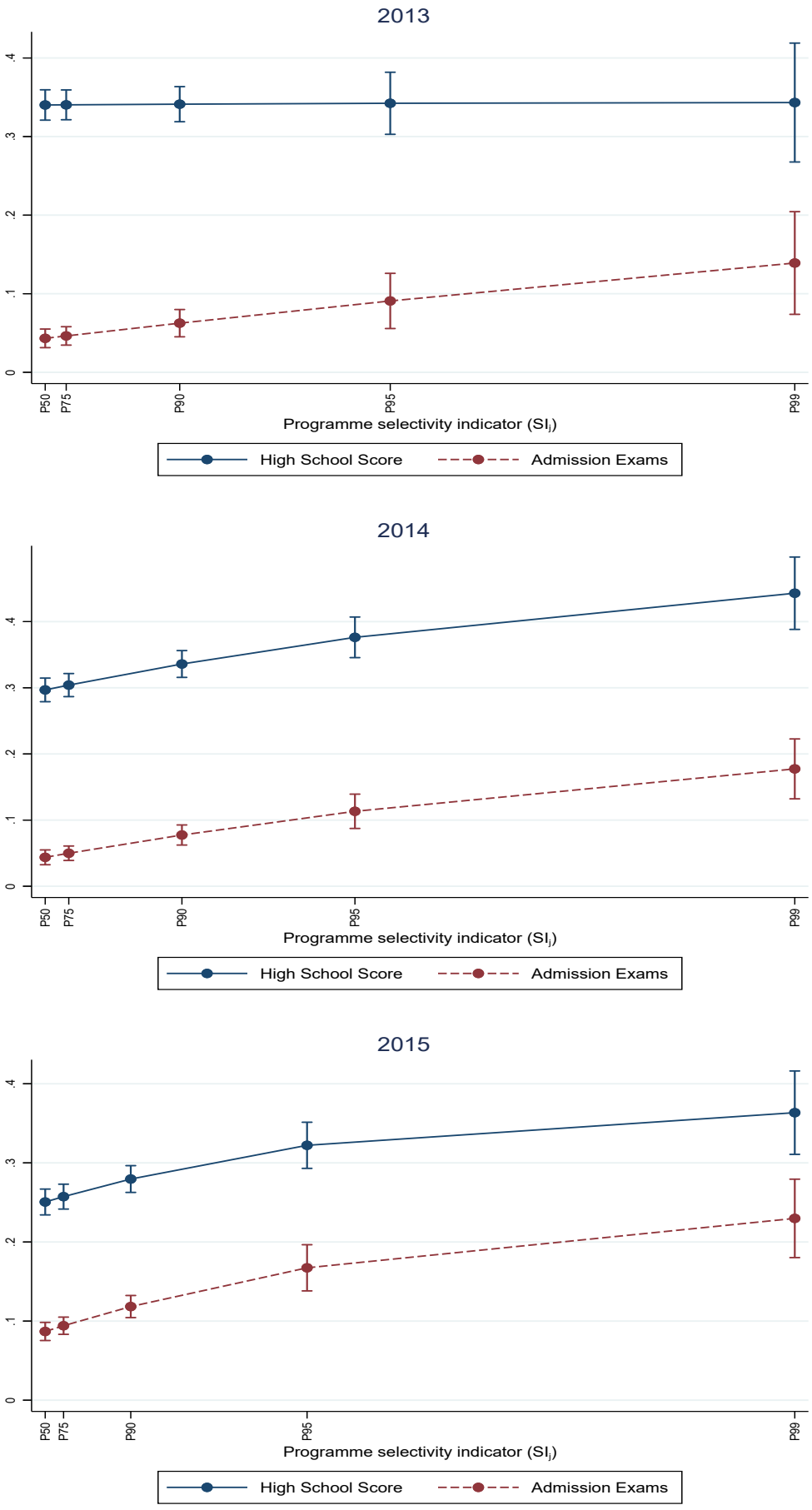


Figure 2: Marginal effects according to programme selectivity

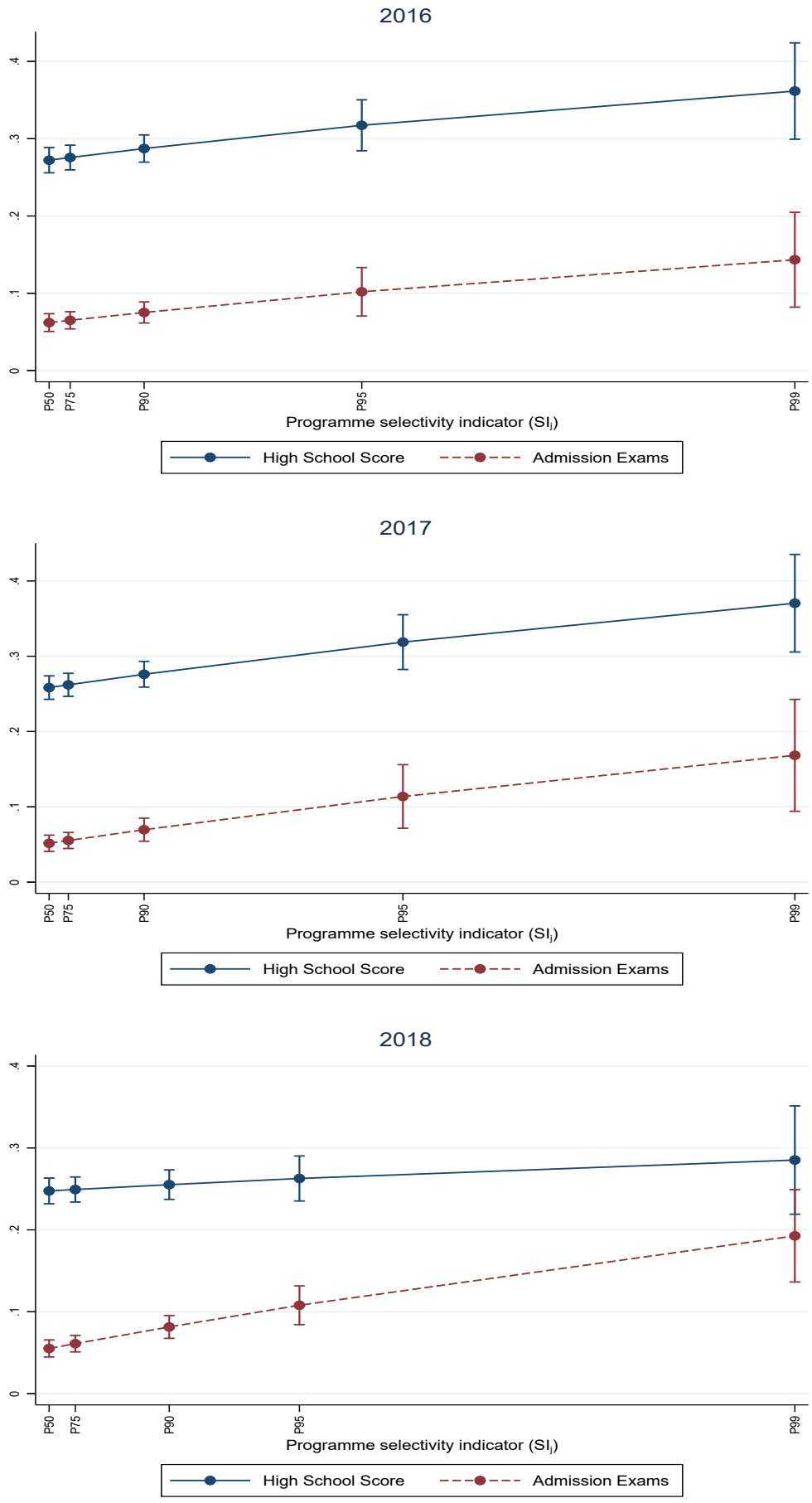


Figure 3: Marginal effects according to programme selectivity (continuation)

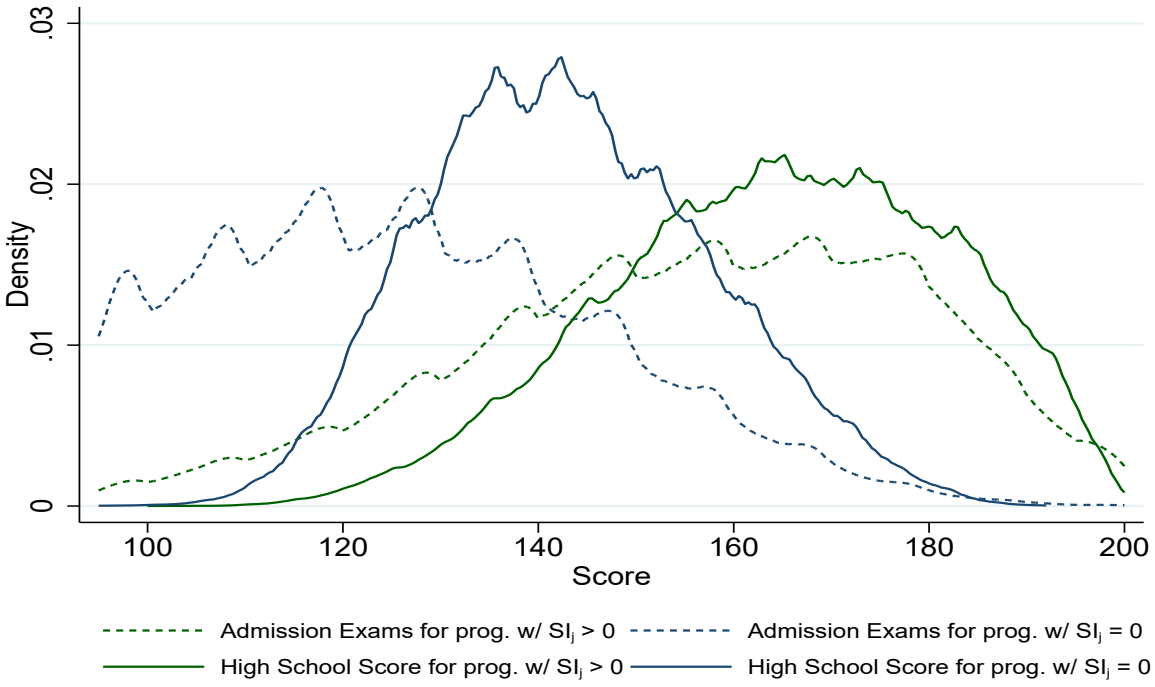


Figure 4: Distribution of admission exams and high school score

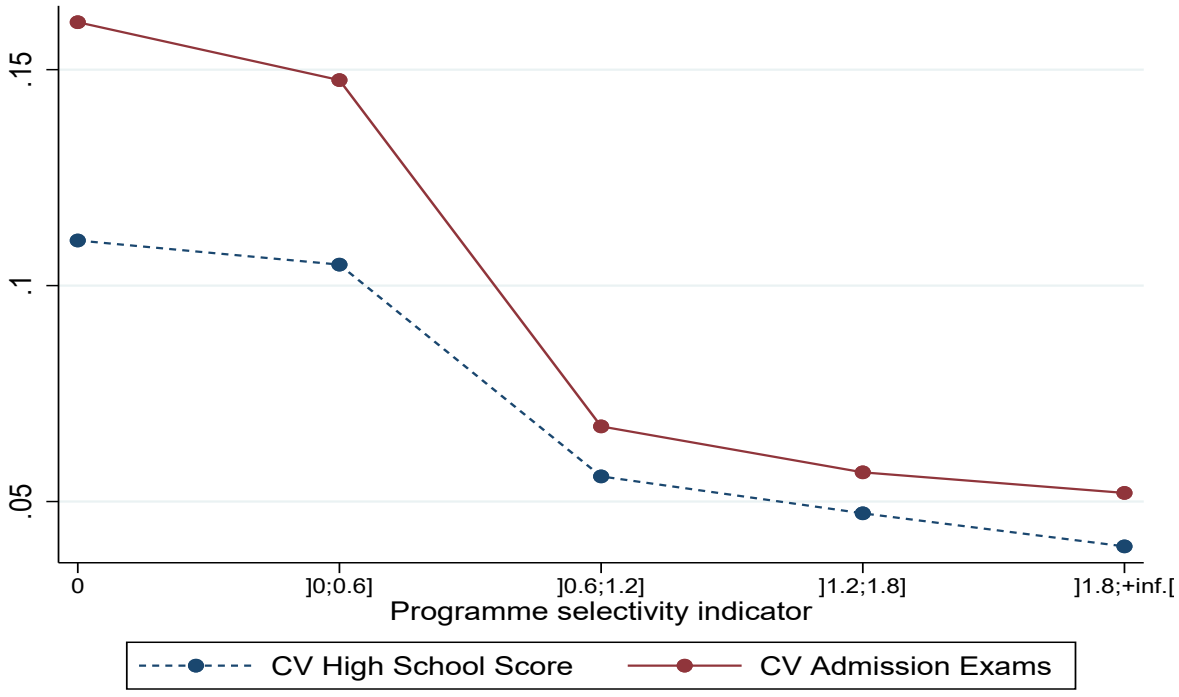


Figure 5: Coefficient of Variation

9 Tables

Table 1: Programme selectivity indicator for the economics degree in Portugal Mainland

		2015			2016		
Programme	Vacancies	Threshold	s_j	SI_j	Threshold	s_j	SI_j
Universidade Nova de Lisboa	223	173.5 (1)	2.43 (1)	0.71 (1)	169.5 (1)	2.04 (1)	0.53 (1)
Universidade do Porto	210	171.5 (2)	2.13 (2)	0.61 (2)	169.5 (1)	1.73 (2)	0.47 (2)
Universidade do Minho	79	163.6 (3)	1.00 (4)	0	164.6 (3)	1.06 (3)	0.05 (3)
ISCTE (Lisboa)	80	160.5 (4)	0.81 (6)	0	160.0 (4)	1.06 (4)	0.01 (7)
Universidade de Aveiro	50	156.4 (5)	0.66 (7)	0	154.0 (5)	1.02 (5)	0
Universidade de Lisboa	145	154.5 (6)	0.88 (5)	0.02 (4)	154.0 (5)	0.77 (7)	0.02 (5)
Universidade de Coimbra	154	149.0 (7)	1.42 (3)	0.03 (3)	145.0 (7)	0.92 (6)	0.03 (4)
Universidade de Trás os Montes	36	139.8 (8)	0.61 (9)	0	140.7 (8)	0.61 (9)	0
Universidade da Beira Interior	45	128.0 (9)	0.24 (10)	0	129.9 (9)	0.27 (10)	0.02 (6)
Universidade do Algarve	25	124.2 (10)	0.64 (8)	0	122.0 (10)	0.68 (8)	0
Universidade de Évora	37	118.8 (11)	0.16 (11)	0	120.8 (11)	0.27 (11)	0

Source: Authors calculations. Notes: Scores are measured in a scale between 0 and 200. We only consider the economics degree for the years 2015 and 2016.

Table 2: Descriptive statistics by programme selectivity

	Programmes with $SI = 0$		Programmes with $SI > 0$	
	Mean	Std. Dev.	Mean	Std. Dev.
No. Individuals	89,664		88,252	
Age	18.52	1.89	18.03	1.24
Admission Exams	125.91	20.27	154.8	23.28
High school GPA	141.07	15.58	163.2	17.99
Female (share)	.59		.57	
Public high school (share)	.92		.90	
Non-local student (share)	.30		.31	
Mother has HE (share)	.25		.45	
Father has HE (share)	.18		.36	
Applied and received to a maintenance grant (share)	.30		.22	
Applied and did not received a maintenance grant (share)	.06		.06	
ECTs accumulated by the end of the 1st year	52.09	16.17	54.74	14.83

Source: Authors calculations.

Table 3: Descriptive statistics, analysis population

A. Data Structure			
Initial year	2013/2014		
Final year	2018/2019		
Number of years	6		
Number of Individuals	177,916		
	Mean	Std. Dev.	N
B. Individuals			
Age	18.28	1.62	177,916
Female (share)	.58		177,916
Admission Exams	140.24	26.17	177,916
High school GPA	152.05	20.14	177,916
Public high school (share)	.91		151,635
Non-local student (share)	.31		177,916
Mother has HE (share)	.35		170,639
Father has HE (share)	.27		167,393
Applied and received to a maintenance grant (share)	.26		177,916
Applied and did not received a maintenance grant (share)	.06		177,916
C. Placement			
Degree of placement (no. individuals)			
Bachelor			140,645
Integrated Master			37,271
Preferences of placement (share)			
1st	.58		177,916
2nd	.21		177,916
3rd	.10		177,916
4th	.05		177,916
5th	.03		177,916
6th	.02		177,916
Admission score (0-200)	147.2	20.17	177,916
Admitted to a programme with $SI > 0$ (share)	0.50		177,916
Admitted to a programme with $SI > 0$, by field (share)			
Education	0.01		177,916
Arts & Humanities	0.12		177,916
Social Sciences	0.21		177,916
Sciences	0.11		177,916
Engineering	0.20		177,916
Agriculture	0.01		177,916
Health	0.14		177,916
Services	0.02		177,916
D. Outcomes			
ECTs accumulated by the end of the 1st year	54.05	14.54	152,332

Source: Author's calculations. Notes: Scores are measured in a scale between 0 and 200. We only consider individuals that applied to public HEIs in Portugal Mainland under the GAR in the 1st round of applications

Table 4: Baseline results (average marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
No. of ECTS by the end of the 1st year (Negative Binomial)										
High school score	0.146***	0.136***	0.238***	0.280***	0.164***	0.185***	0.226***	0.234***	0.268***	0.280***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Admission exams	0.029***	0.037***	0.089***	0.072***	0.052***	0.084***	0.072***	0.087***	0.058***	0.069***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
SI_j					-4.839***	-18.451***	-1.755***	-9.040***	-1.108***	-8.243***
					(0.110)	(0.682)	(0.149)	(0.671)	(0.153)	(0.674)
Female		4.114***	2.386***	2.213***	4.059***	4.056***	2.590***	2.594***	2.422***	2.427***
		(0.083)	(0.080)	(0.079)	(0.083)	(0.083)	(0.082)	(0.082)	(0.082)	(0.082)
Public high school		1.207***	1.358***	2.668*	1.225***	1.223***	1.484***	1.482***	2.274	2.269
		(0.141)	(0.121)	(1.374)	(0.141)	(0.140)	(0.126)	(0.126)	(1.416)	(1.415)
Non_local student		2.834***	-0.496***	-0.252**	2.794***	2.807***	-0.848***	-0.825***	-0.645***	-0.619***
		(0.090)	(0.091)	(0.099)	(0.089)	(0.089)	(0.093)	(0.093)	(0.101)	(0.101)
N	130,371	130,371	130,371	130,371	130,371	130,371	130,371	130,371	130,371	130,371
Pseudo R2										
High School Score × SI_j						✓		✓		✓
Admission Exams × SI_j						✓		✓		✓
High School FE				✓					✓	✓
Programme FE			✓	✓						
Field FE (3 dig)							✓	✓	✓	✓
Institution FE							✓	✓	✓	✓
Cohort and Preference FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Robust standard errors are in parentheses. *,** and *** represents statistical significance from 10%, 5% and 1% respectively.

Table 5: Admission exams volatility effect (OLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Admission Exams Score							
High school score	0.869***	0.875***	0.858***	0.326***	0.384***	0.635***	0.649***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)
SI_j						15.700***	239.762***
						(0.219)	(4.566)
High School Score × SI_j							-1.197***
							(0.024)
Cohort FE							
2014/2015			3.237***	2.713***	2.723***	2.954***	3.013***
			(0.203)	(0.166)	(0.162)	(0.171)	(0.168)
2015/2016 (elections)			9.317***	9.302***	9.344***	9.183***	9.455***
			(0.194)	(0.158)	(0.155)	(0.166)	(0.162)
2016/2017			7.339***	7.722***	7.701***	7.340***	7.669***
			(0.194)	(0.158)	(0.155)	(0.165)	(0.162)
2017/2018			7.552***	9.466***	9.287***	8.267***	8.606***
			(0.190)	(0.155)	(0.153)	(0.161)	(0.158)
2018/2019			4.097***	6.478***	6.261***	5.063***	5.359***
			(0.190)	(0.156)	(0.154)	(0.161)	(0.158)
Constant	7.907***	6.378***	5.602***	86.683***	77.271***	25.076	25.059***
	(0.400)	(0.453)	(0.479)	(0.771)	(1.400)	(48318.743)	(1.823)
N	130,371	130,371	130,371	130,371	130,371	130,371	130,371
R^2	0.413	0.417	0.439	0.658	0.676	0.608	0.621
Programme FE				✓	✓		
Field FE (3 dig)						✓	✓
Institution FE						✓	✓
High School FE					✓	✓	✓
Controls		✓	✓	✓	✓	✓	✓
Preference FE		✓	✓	✓	✓	✓	✓

Notes: Robust standard errors are in parentheses. *,** and *** represents statistical significance from 10%, 5% and 1% respectively. Controls: female, public high school, non-local student.

Table 6: Admission criteria effect

Selectivity Programme Indicator (SI_j) - Tobit model	(1)	(2)	(3)
1 Mandatory Exam	-0.055***	-0.072***	-0.018
	(0.017)	(0.017)	(0.017)
No. of different exams	0.123***	0.120***	0.116***
	(0.022)	(0.022)	(0.020)
No. of entrance exam combinations allowed			
1 Two	-0.190***	-0.186***	-0.149***
	(0.033)	(0.033)	(0.031)
1 Three	-0.345***	-0.341***	-0.265***
	(0.043)	(0.043)	(0.039)
1 Four	-0.432***	-0.418***	-0.373***
	(0.067)	(0.066)	(0.059)
1 Five	-0.567***	-0.563***	-0.501***
	(0.099)	(0.099)	(0.088)
1 Six	-0.496***	-0.478***	-0.425***
	(0.120)	(0.119)	(0.107)
Admission Exams weight			
1 40%	0.318***	0.318***	0.029
	(0.021)	(0.022)	(0.029)
1 45%	0.382***	0.380***	0.228***
	(0.068)	(0.069)	(0.056)
1 50%	0.479***	0.480***	0.194***
	(0.028)	(0.028)	(0.025)
Share of Females		0.031	0.036
		(0.023)	(0.025)
Share of Public High School		0.259***	0.159***
		(0.042)	(0.042)
Share of Non local students		0.053***	0.513***
		(0.018)	(0.059)
Constant	-0.570***	-0.790***	-0.692***
	(0.047)	(0.061)	(0.071)
N	5,687	5,687	5,687
Pseudo R^2	0.27	0.28	0.40
Field FE (3 dig)			✓
Institution FE			✓
Cohort FE			✓

Notes: Robust standard errors are in parentheses. *,** and *** represents statistical significance from 10%, 5% and 1% respectively.

References

Aina, C., E. Baici, G. Casalone, and F. Pastore (2018). The economics of university dropouts and delayed graduation: a survey. *IZA Discussion Papers* (11421).

Akee, R. Q., T. J. Halliday, and S. Kwak (2014). Investigating the effects of furloughing public school teachers on juvenile crime in hawaii. *Economics of Education Review 42*, 1–11.

Anelli, M. (2020). The returns to elite university education: A quasi-experimental analysis. *Journal of the European Economic Association 18*(6), 2824–2868.

Atkinson, R. C. and S. Geiser (2009). Reflections on a century of college admissions tests. *Educational Researcher 38*(9), 665–676.

Belfield, C. R. and P. M. Crosta (2012). Predicting success in college: The importance of placement tests and high school transcripts. ccrc working paper no. 42. *Community College Research Center, Columbia University*.

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

Brewer, D. J. and E. R. Eide (1999). Does it pay to attend an elite private college? *The Journal of Human Resources 34*(1), 104–123.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge university press.

Campbell, S., L. Macmillan, R. Murphy, and G. Wyness (2021). Matching in the dark? inequalities in student to degree match. *Journal of Labor Economics*.

Cardoso, A. R., M. Portela, C. Sá, and F. Alexandre (2008). Demand for higher education programs: the impact of the bologna process. *CESifo Economic Studies 54*(2), 229–247.

Cyrenne, P. and A. Chan (2012). High school grades and university performance: A case study. *Economics of Education Review 31*(5), 524–542.

DGEEC (2019). Registro de alunos inscritos e diplomados no ensino superior (RAIDES) [Registry of student enrollment and graduation from higher education]. Ministry of Science, Technology and Higher Education, and Ministry of Education. Portugal.

DGES (2019). Concurso nacional de acesso [National competition to access higher education], Ministry of Science, Technology and Higher Education. Portugal.

DGES (2020). Relatório do grupo de trabalho sobre o acesso ao ensino superior [report of the working group on access to higher education]. Report, https://wwwcdn.dges.gov.pt/sites/default/files/relatorio.grupo_de_trabalho.acesso_ao_ensino_superior.2020.pdf.

Dillon, E. W. and J. A. Smith (2017). Determinants of the match between student ability and college quality. *Journal of Labor Economics 35*(1), 45–66.

Dooley, M. D., A. A. Payne, and A. L. Robb (2012). Persistence and academic success in university. *Canadian Public Policy 38*(3), 315–339.

Gale, D. and L. S. Shapley (1962). College admissions and the stability of marriage. *The American Mathematical Monthly 69*(1), 9–15.

Graddy, K. and M. Stevens (2005). The impact of school resources on student performance: A study of private schools in the united kingdom. *ILR Review 58*(3), 435–451.

Green, F., S. Machin, R. Murphy, and Y. Zhu (2012). The changing economic advantage from private schools. *Economica 79*(316), 658–679.

Hilmer, M. J. and C. E. Hilmer (2011). Is it where you go or who you know? on the relationship between students, ph. d. program quality, dissertation advisor prominence, and early career publishing success. *Economics of Education Review 30*(5), 991–996.

Li, H., L. Meng, X. Shi, and B. Wu (2012). Does attending elite colleges pay in china? *Journal of Comparative Economics 40*(1), 78–88.

Merkaj, E., R. Lucchetti, and F. Fiorillo (2020). The role of local leaders in regional development funding: Evidence from an elite survey. *Journal of Regional Science 60*(4), 712–737.

Nata, G., M. J. Pereira, and T. Neves (2014). Unfairness in access to higher education: a 11 year comparison of grade inflation by private and public secondary schools in portugal. *Higher Education 68*(6), 851–874.

OECD (2019). *OECD Review of Higher Education, Research and Innovation: Portugal*. OECD Publishing, Paris, https://doi.org/10.1787/9789264308138-en.

Papay, J. P., R. J. Murnane, and J. B. Willett (2016). The impact of test score labels on human-capital investment decisions. *Journal of Human Resources 51*(2), 357–388.

Portela, M., N. Areal, C. Sá, F. Alexandre, J. Cerejeira, A. Carvalho, and A. Rodrigues (2008). Evaluating student allocation in the portuguese public higher education system. *Higher Education 56*(2), 185–203.

Portela, M., C. Sá, F. Alexandre, and A. R. Cardoso (2009). Perceptions of the bologna process: what do students' choices reveal? *Higher Education 58*(4), 465–474.

Silva, P. L. (2022). Specialists or all-rounders: How best to select university students? *IZA Discussion Papers* (15271).

Silva, P. L., L. C. Nunes, C. Seabra, A. Balcao Reis, and M. Alves (2020). Student selection and performance in higher education: admission exams vs. high school scores. *Education Economics 28*(5), 437–454.

Smyth, E. (2009). Buying your way into college? private tuition and the transition to higher education in ireland. *Oxford Review of Education 35*(1), 1–22.

Walker, I. and Y. Zhu (2018). University selectivity and the relative returns to higher education: Evidence from the uk. *Labour Economics 53*, 230–249.

Westrick, P. A., H. Le, S. B. Robbins, J. M. Radunzel, and F. L. Schmidt (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.

Zwick, R. (2019). Assessment in american higher education: The role of admissions tests. *The ANNALS of the American Academy of Political and Social Science 683*(1), 130–148.

A Appendix: Data

Table A1: Distribution of the programme seletivity indicator (SI_j)

	2013/2014			2014/2015			2015/2016		
	Max SI_j	No. programmes	No. Individuals	Max SI_j	No. programmes	No. Individuals	Max SI_j	No. programmes	No. Individuals
$SI_j = 0$	0	677	12,277	0	678	12,818	0	692	16,173
$0 < SI_j \leq$ p75	0.018	29	2,347	0.014	21	1,840	0.012	10	1,016
p75 $< SI_j \leq$ p90	0.070	147	6,993	0.067	140	6,194	0.063	140	7,153
p90 $< SI_j \leq$ p95	0.200	49	2,342	0.183	43	2,655	0.169	46	2,402
p95 $< SI_j \leq$ p99	1.007	35	2,116	0.925	38	2,665	0.933	38	2,698
p99 $< SI_j \leq$ Max	1.619	10	1,214	1.616	9	819	2.176	9	1,174
N	-	947	27,289	-	929	26,991	-	935	30,616

Source: Authors' calculations. Notes: Scores are measured in a scale between 0 and 200. We only consider individuals that applied to public HEIs under the GAR in the 1st round of applications.

Table A2: Distribution of the programme seletivity indicator (SI_j) - continuation

	2016/2017			2017/2018			2018/2019		
	Max SI_j	No. programmes	No. Individuals	Max SI_j	No. programmes	No. Individuals	Max SI_j	No. programmes	No. Individuals
$SI_j = 0$	0	698	16,180	0	696	16,522	0	688	15,674
$0 < SI_j \leq$ p75	0.013	19	1,827	0.016	22	2,296	0.020	39	2,701
p75 $< SI_j \leq$ p90	0.073	139	6,374	0.077	142	6,559	0.086	146	6,229
p90 $< SI_j \leq$ p95	0.150	49	2,527	0.183	48	2,309	0.219	48	2,580
p95 $< SI_j \leq$ p99	1.033	37	2,573	1.171	39	3,186	1.122	39	2,880
p99 $< SI_j \leq$ Max	2.259	10	1,045	2.553	9	765	2.313	10	773
N	-	952	30,526	-	956	31,637	-	970	30,837

Source: Authors' calculations. Notes: Scores are measured in a scale between 0 and 200. We only consider individuals that applied to public HEIs under the GAR in the 1st round of applications.

Table A3: Parents' education (average marginal effects)

	(1)	(2)	(3)	(4)
High school score	0.236***	0.235***	0.274***	0.274***
	(0.003)	(0.003)	(0.004)	(0.004)
Admission exams	0.074***	0.074***	0.058***	0.059***
	(0.002)	(0.002)	(0.002)	(0.002)
SI_j	-1.563***	-1.868***	-1.044***	-1.414***
	(0.156)	(0.158)	(0.160)	(0.162)
N	122,291	122,291	122,291	122,291
Pseudo R2				
Mother with HE	✓	✓	✓	✓
Father with HE	✓	✓	✓	✓
Mother with HE × SI_j		✓		✓
Father with HE × SI_j		✓		✓
High School FE			✓	✓
Controls	✓	✓	✓	✓
Field FE (3 dig)	✓	✓	✓	✓
Institution FE	✓	✓	✓	✓
Cohort and Preference FE	✓	✓	✓	✓

Notes: Robust standard errors are in parentheses. *,** and *** represents statistical significance from 10%, 5% and 1% respectively. Controls: female, public high school and non-local student.