

A Work Project, presented as part of the requirements for the Award of a Double Degree
Master's degree in Economics from the Nova School of Business and Economics, and in
Economics and Finance from LUISS Guido Carli

DIFFERENT ASSESSMENT METHODS AND PARENTS' EDUCATION

PEDRO MARIA TELES DE ALEGRIA BARAHONA DE LEMOS

Work project carried out under the supervision of:

Advisors: Ana Balcão Reis & Maria Savona

Co-advisors: Catarina Ângelo & Daniela Di Cagno

October/November 2022

Abstract

This study analyses whether teacher and test-based assessments differ systematically between low and high socio-economic status (SES) students throughout the Portuguese education system. As an indicator of the students' SES, we use the level of education of the students' parents. We use nationwide data on students and their assessments in 16 subjects in the 11th and 12th grades for a 10-year period ranging from 2009 to 2018. The analysis follows linear specifications to test whether there are significant socio-economic gaps between both grading schemes. The results suggest that, when comparing students with similar exam scores, the gap between teacher score and exam score increases for students from high-SES, suggesting that for students with the same exam score, low-SES students have a lower teacher grade compared to their high-SES peers. Thus, exams seem to be relatively more favourable for students from low-SES and teacher assessment more favourable for students from high-SES.

Keywords: Economics of Education, student achievement, assessment methods, socio-economic gap, parents' education.

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

Contents

- 1 Introduction** 4
- 2 Literature review** 5
- 3 Institutional setting** 8
- 4 Data and descriptive statistics** 10
 - 4.1 Composition effect 18
- 5 Regression analysis** 23
 - 5.1 Specification 23
 - 5.2 Results 24
 - 5.2.1 General regressions 24
 - 5.2.2 Regressions by exam score level 25
 - 5.2.3 Regressions by exam score level and gender 28
- 6 Concluding remarks** 31
- References** 34
- Appendices** 38
 - A Biennial and triennial subjects** 38
 - B Statistics using only the level of mothers' education** 39
 - C Analysis for three levels of parents' education** 41

D Analysis for six levels of parents' education	47
E Analysis using social support as proxy	51

1. Introduction

This study analyses whether different grading practices generate a gap in student assessment between low and high socio-economic (SES) students. To address this issue, we look for systematic assessment gaps between the teacher-given grade and exam scores for students from low- and high-SES, using nationwide data on Portuguese public school, from 2009 to 2018. As a proxy for the students' SES, we use the level of their parents' education. We follow the same analysis as Lindahl (2007), Marcenaro Gutiérrez and Vignoles (2015), and Ângelo and Balcão Reis (2021) in the methodology and approach to find a significant assessment gap, but focusing on the students' SES instead of the students' gender, given by the level of education of the students' parents. Ribeiro (2021) conducts a similar study but considering a different indicator of SES: receiving social support.

The possible existence of assessment gaps is very relevant, especially in contexts where these types of evaluation are the only measure of students' performance when applying to university (as is the case of Portugal). The findings of Ângelo and Balcão Reis (2021) suggest that teacher assessment is relatively more favourable to girls, whilst exam assessment is relatively more favourable to boys. Bearing in mind that there is this gap between male and female students, we examine if furthermore there is also a gap between students from low- and high-SES.

Our findings suggest that when comparing students with similar exam scores, the gap between teacher score and exam score increases for students from high-SES, that is, for students with the same exam score, low-SES have a lower teacher grade compared to their high-SES peers. Thus, exams seem to be relatively more favourable for students from low-SES and teacher assessment more favourable for students from high-SES.

The content of this study is organised as follows. [Section 2](#) presents a review of literature on

education and socio-economic status, and the impact of different grading schemes in student assessment. [Section 3](#) offers an overview of the Portuguese institutional setting. [Section 4](#) presents the data used with some descriptive statistics. [Section 5](#) introduces the regression analysis model to be estimated and the main results. Lastly, [Section 6](#) concludes.

2. Literature review

In 1966 the Coleman Report ordered by the US Government reported an influence of student background factors on achievement. Coleman et al. (1966) found that student background factors¹ account for about 10-25% of the variance in individual achievement. Ever since this report, the literature has confirmed that students' academic achievement is strongly correlated with their socio-economic status (SES).

Students' assessment varies according to the educational system and can vary within school or grade. Usually students will have part of their assessment measured throughout the year by their school professor, and part by an external entity, usually at the national level. Both of these assessment systems present advantages and disadvantages. Calsamiglia and Loviglio (2019) point that the teacher assessment should be more complete, since the teacher knows the student, and can evaluate different types of capacities and competences for a long period, which does not happen with one-shot evaluations like exams, that happen on one day and cannot evaluate everything a student might or not know. However, teachers may be more prone to biases and evidence suggests that teacher assessment of students' performance is done in relative terms

¹ The authors measured eight background factors: urbanism of background, parents' education, structural integrity of the home, smallness of family (number of brothers and sisters), items in home (telephone, record player, refrigerator, automobile, and vacuum cleaner), reading material in home (dictionary, encyclopaedia, daily newspaper, magazines, and books), parents' interest (whether parents talk about school, and whether parents read to children when they were small), and parents' educational desires (how good a student do parents' want their child to be, how far in school do they want their child to go, and attendance at PTA).

to other students in the class or school (Martínez, Stecher, and Borko, 2009; Calsamiglia and Loviglio, 2019). On this point, exams seem more objective, since they are equal to all students, and do not depend on the professor that grades them, since usually the evaluation rules are clearly defined and students are graded anonymously, although some groups might react better to high-stakes than others. When choosing one grading scheme instead of the other, one must bear in mind that different assessment types favour different types of students. Thus, the optimum assessment should probably use both grading schemes.

Many empirical studies have compared teacher assessment of pupil achievement versus exam assessment, an approach we follow in this study (Martínez, Stecher, and Borko, 2009; Marcenaro Gutiérrez and Vignoles, 2015; Guez, Peyre, and Ramus, 2020; Ângelo and Balcão Reis, 2021). By comparing results from different assessment types of the same student, some authors have found that some groups are systematically under-evaluated by their teachers compared to their results on exams or standardised test scores. This has been shown for boys (Lindahl, 2007; Lavy, 2008; Marcenaro Gutiérrez and Vignoles, 2015; Guez, Peyre, and Ramus, 2020; Ângelo and Balcão Reis, 2021), students with a migration background (Triventi, 2020; Gortázar, Martínez de Lafuente, and Vega Bayo, 2022), students from a lower caste in India (Hanna and Linden, 2012), black students (Orr, 2003; Botelho, Madeira, and Rangel, 2015), ethnic minorities (Burgess and Greaves, 2013). We will extend this literature by looking at the impact of SES on the assessment gap.

The impact of SES on students' grading has also been subject of study in the literature. Auwarter and Aruguete (2008) conclude that teachers are likely to develop negative attitudes towards low-SES students in general, but especially boys. Gortázar, Martínez de Lafuente, and Vega Bayo (2022) also find systematic teachers' under-assessment in low-SES students. Orr (2003) finds that wealth can affect academic achievement through its effect on the amount of

cultural capital to which a child is exposed, thus creating a socio-economic gap amongst students namely between black and white students, since the former are generally poorer in the American context.

Whether these gaps are related to the teachers' characteristics has been suggested by some. Lavy (2008) suggests it is the teacher's gender that influences the gender gap, a result not obtained by Krkovic et al. (2014), neither by Lindahl (2016). Ângelo and Balcão Reis (2022), on the other hand, find that having a male teacher tends to increase assessment gap (between teacher assessment and exam score) for all students. Others suggest teachers with a migrant background are more prone to reward non-native students (Lindahl, 2016). Bar and Zussman (2012) find evidence that student grades are linked to the political orientation of professors in the university context.

Apart from the teachers' characteristics, stakes of the assessment, writing skills, and self-discipline of the student may also influence the gender gap (Cornwell, Mustard, and van Parys, 2013; Guez, Peyre, and Ramus, 2020). Also, since boys are more prone to misbehaviour than girls (Marcenaro Gutiérrez, López Agudo, and Ropero García, 2017) this can contribute for being given lower scores by their teachers. van Bergen et al. (2017) link children's basic reading skills to several aspects of the home literacy environment, but suggest that most seem to be masked genetic effects. Furthermore, Marcenaro Gutiérrez and López Agudo (2016) find that agreement of parental and students' expectations presents a positive influence on students' achievement and the likelihood of selecting a high school track. These parental expectations have been found to be dependent of family socio-economic background. We will study the differential impact of student SES on two types of assessment, teacher grades and exam scores.

3. Institutional setting

Since 2009, students in Portugal are required to attend school from the year they turn 6 to the year they turn 18, which corresponds to 12 years of compulsory education². Grades are divided in four cycles of studies:

- **1st cycle:** 1st to 4th grade (ages 5/6 to 9/10);
- **2nd cycle:** 5th and 6th grade (ages 9/10 to 11/12);
- **3rd cycle:** 7th to 9th grade (ages 11/12 to 14/15);
- **Secondary:** 10th to 12th grade (ages 14/15 to 17/18).

At the end of the 9th grade, students can choose between following the academic path, which is for those who want to pursue a university degree, or the vocational one. In the year 2018—the last year we cover in this study—there were approximately 58% students in the academic path, the one we will focus.

The academic path is divided in four areas and each student must choose one: Sciences and Technologies; Socio-Economic Sciences; Languages and Humanities; and Visual Arts. For simplicity, we will refer to these areas as Sciences, Economics, Humanities, and Arts, respectively.

The subjects can be biennial (*i.e.*, from the 10th to the 11th grade) or triennial (*i.e.*, from the 10th to the 12th grade). All students must choose two biennial subjects specific to their area, and each area has a mandatory triennial subject. The list of these subjects is presented in [Appendix A](#). Apart from those, all areas have Philosophy, which is biennial, and Portuguese, which is triennial. There are also other subjects but since those are not subject to exams we will

² Although students are required to attend 12 years of compulsory education, they are not required to finish the 12th grade. If, during the mandatory years of education, a student failed one or more grades, when the 12 years of mandatory education end, he can choose to drop out of school, independently of the grade attended.

disregard them. We adopted a criterium of a minimum of 5,000 observations in each exam, and thus we will only consider 16 subjects.

Students are evaluated by their teachers in each subject during the whole academic year. The evaluation includes written tests (usually six per year), homework, oral presentations, class participation, student behaviour, laboratory sessions (in the Sciences course), amongst others. The weight of these elements differs between schools and subjects.

If the subject is biennial, the subject's final score will be the average of the teacher grade of the 10th and 11th grades; if it is triennial, the average of the 10th, 11th, and 12th grades. The teacher grades are given in a 0-20 scale. The minimum score to pass is 10. The exams are graded in a 0-200 scale³.

In addition, in some subjects there is a mandatory national exam. Students can take the national exams as internal or external students. If they take as internal students, the teacher grade only accounts for 70% of the final score, being the remaining 30% determined by the exam score. If they take as external ones, then the exam score accounts for 100% of their subject's final score, and there is no teacher assessment on the subject. To be able to take the exam as internal students, students must attend classes in the school they are taking the exam and must obtain a teacher grade of at least 10⁴.

Students are required to take four exams: two in the 11th grade, and two more on the 12th. It is mandatory for one of the exams of the 11th grade to be one of the two biennial subjects that the student chose; but regarding the second mandatory exam, students can choose to either take the exam of the second chosen biennial, or to take the Philosophy exam. In the 12th grade, students

³ The score is then divided by 10 and rounded to the nearest whole number to compute the subject's final score.

⁴ If students do not obtain this minimum grade of 10, they can then choose between repeating the academic year, or to try to have at least 9.5 in the exam, which they take as external students. It is also possible for an internal student to freely reject his teacher grade and thus take the exam as an external student; in this case there is no turning back: if the student does not obtain a minimum exam score of 9.5, he fails the subject.

must take the exam of their triennial mandatory subject as well as the Portuguese exam, which is the only exam that all students in the academic track must take.

Exams happen after the end of classes in June, after the publication of teachers' grades. Depending on the subject, they can include multiple choice, open questions, short answers, true or false, amongst others. Exams take between 90 and 150 minutes, with each exam having an extra 30 minutes of tolerance that students can use. The evaluation of the exam is anonymous: the student is not identifiable, and the exam is graded by a teacher from a different school than the one the student is attending. The exams are graded on absolute or criterion-referenced evaluation that is made available online by an agency of the Ministry of Education.

It is possible for a student to resit for an exam in July either to pass the subject—due to failure on the June season—, or to improve the exam score. We will only consider exams of the June season since the percentage of students who undertake the second season shows no relation with their parents' level of education—approximately 1/3 of all students, regardless of their parents' level of education, resit in the July season.

Lastly, higher education access is determined by a weighted average of all secondary subjects and the national exam scores, where the weight of the latter varies between 40% and 50% of the application score, depending on the university and degree the student is applying to. Universities can ask for one or two exams for the application⁵.

4. Data and descriptive statistics

This study uses administrative data from a database managed by the Portuguese Ministry of Education. This database includes all the students' exam scores with some information of the

⁵ The only exception is Medicine which requires three exams.

student: age, gender, area of studies, teacher grade (if the student is intern), if the student benefits from social support, the level of parents' education and employment situation; as well as the school attended and teacher characteristics, etc.

We note that our database only has data from public schools, which means that students from private schools are not considered in this study. We focus mainly in the level of parents' education as an indicator of socio-economic status. Our independent variable will be the difference between the teacher score and the exam score, *TeacherScore-ExamScore*, implying that we will only consider subjects that require students to take an exam. Our focus will be on how that variable changes according to the level of parents' education.

The data about parents' education is self reported and has 10 levels: no education, 1st cycle, 2nd cycle, 3rd cycle, secondary, postgraduate diploma, bachelor's degree, undergraduate degree, master's degree, and doctorate. We aggregate these 10 levels in 3 levels of education:

- **Level 1:** Basic education (from no education to 9th grade);
- **Level 2:** Secondary education (secondary);
- **Level 3:** Higher education (from post-graduation to doctorate).

We use for each student the maximum level of his parents' education which can be either the mother's or the father's education. To check if the results are very different from using only the mother's education, we run a descriptive statistics on the grades using only the mother's education which is presented in [Appendix B](#). We found no significant differences.

Our database has 303,620 observations with no information about neither of the parents' education. These make approximately 20% of all observations, which we will have to disregard. Therefore, we remain with 1,367,753 observations.

[Table 1](#) presents some descriptive statistics regarding students. Column (1) shows the variables of interest. Column (2) has the mean for all the observations, which includes students

with no information on their parents' education. Column (3) has the mean of the observations we will use, which only includes students' that provided information on their parents' education (approximately 80% of all students). Columns (4)-(6) have the mean for students per level of their parents' education. Columns (7)-(11) provide the correspondent standard deviation of columns (2)-(6). Column (12) provides a *T*-test of the differences between students with and without information about their parents' education. The period of analysis is from 2009 until 2018.

By comparing columns (2)-(3), the means of students with and without parents' education data, we see that students with no information on their parents' education are more likely to be from a lower socio-economic background since, despite the overall differences between the two populations being small, the internet (51.48% vs 70.56%) and computer at home (54.19% vs 75.98%) means are drastically inferior. Also, they are slightly older (18.00 vs 17.79), which means that they failed more years, and their grades are slightly lower. Furthermore, the percentage of students born in Portugal is smaller, which might be another indicator of a lower socio-economic background, since most students not born in Portugal are from CPLP countries (Community of Portuguese Language Countries), that are more likely to come from a lower socio-economic background.

Thus, we must acknowledge a small bias: the sample that we are going to study could have less low-SES students than those that actually exist. Since approximately 20% of all students do not provide information on their parents' education, we are thus disregarding 1 in every 5 students that are more likely to be from a lower socio-economic background. However, the differences appear to be very small, and thus we are confident that it will not affect our conclusions. The fact that there is almost no difference in the percentage receiving social support also reinforces this conclusion (26.36% vs 25.42%).

Looking at columns (4)-(6), the means show us what we expected beforehand: as the level of parents' education increases, students are more likely to have higher scores. We have less social support (41.17%, 23.66%, and 6.09%), and more internet at home (65.95%, 72.84%, and 74.64%). Furthermore, we see a decrease in age (17.88, 17.79, and 17.67), fewer students improving their grades (0.51%, 0.44%, and 0.31%), and higher grades in all subjects.

One interesting result is the decrease in the percentage of girls. This means that, as the level of parents' education increases, boys are more likely to pursue the academic track of secondary education. However, in lower socio-economic backgrounds, it appears that the percentage of boys that give up the academic track is much higher. There is also an increase in students born in Portugal with the level of parents' education, since most students not born in Portugal are from CPLP countries which are typically from lower socio-economic backgrounds, thus it is no surprising that the percentage increases as the level of parents' education increases (94.76%, 95.16%, and 96.47%).

[Table 2](#) presents descriptive statistics of teacher and exam scores. Column (1) indicates the subject. Column (2) has the number of observations. Column (3) shows the level of parents' education. Column (4) has the percentage of students for each level of parents' education. Columns (5)-(7) have the means of teacher score, exam score, and the difference *TeacherScore-ExamScore*. Column (8) provides a *T*-test⁶.

We note that the means of teacher score and exam score all increase with the level of parents' education. Regarding the mean teacher score, the biggest increase is from levels 2 to 3 of parents' education that exceed 1 point in 7/16 subjects: Descriptive Geometry (1.37), Mathematics A (1.27), Physics and Chemistry (1.27), Mathematics B (1.13), Portuguese Literature (1.11),

⁶ The *T*-test is for the difference in mean differences between students with a level 1 parents' education and students with a level 2 parents' education, and students with a level 2 parents' education and students with a level 3 parents' education.

Biology and Geology (1.10), and History of Culture and Arts (1.03). The differences between levels 1 and 2 are smaller and never exceed 0.5 points.

Concerning the mean exam score, the behaviour is similar to the mean teacher score: it increases less from levels 1 to 2, and more from levels 2 to 3. Between levels 1 and 2 the difference is bigger than the teacher score as we have 5/16 subjects that exceed 0.5 points of difference and 1 that exceeds 1: Descriptive Geometry (1.22), Mathematics B (0.95), Physics and Chemistry (0.67), Philosophy (0.63), Mathematics A (0.62), and Mathematics Applied to Social Sciences (0.54). Between levels 2 and 3—except for Drawing (0.51) and Spanish (0.63)—all subjects exceed 1 point and 1 exceeds 2 points, with other 3 approaching 2 points: Descriptive Geometry (2.79), Mathematics A (1.97), Physics and Chemistry (1.97), and Mathematics B (1.93).

By analysing column (7) we see that all the mean differences are positive, and all decrease with the level of parents' education⁷. All the differences are smaller from levels 2 to 3 than from levels 1 to 2, as the means. There is one exception that is significant: Portuguese (-0.14 from levels 1 to 2, and -0.08 from levels 2 to 3).

From this table's results, we see that the higher the level of parents' education, the higher the grades, either from the teacher or in the exam. This is no surprise, as it is already known from the literature. However, the table also seems to indicate that, as the level of parents' education increases, the difference between the teacher and exam score decreases. Thus, higher-SES students decrease less the exam score with respect to their teacher-assigned grade, which could indicate that, in relative terms, exams would be more favourable to these students. However, as we present in the next subsection, there is a composition effect, and therefore this result is misleading.

⁷ There is only one exception for Economics between levels 1 and 2 that increases from 2.64 to 2.67, however with the *T*-test we see that it is not significant.

Table 1: Means and standard deviations of students' variables for three levels of parents' education (2009-2018).

Variable (1)	Overall (2)	Mean				Standard Deviation					Mean differences ^a (12)
		W/ info (3)	Level 1 (4)	Level 2 (5)	Level 3 (6)	Overall (7)	W/ info (8)	Level 1 (9)	Level 2 (10)	Level 3 (11)	
Age	18.00	17.79	17.88	17.79	17.67	1.00	0.81	0.86	0.81	0.72	0.21*** [123.87]
Girls	0.57	0.58	0.62	0.57	0.52	0.50	0.49	0.48	0.49	0.50	-0.01*** [-7.55]
Born in Portugal	0.93	0.95	0.95	0.95	0.96	0.26	0.21	0.22	0.21	0.18	-0.03*** [-59.59]
Social support	0.26	0.25	0.41	0.24	0.06	0.44	0.44	0.49	0.43	0.24	0.01*** [10.81]
Internet at home	0.51	0.71	0.66	0.73	0.75	0.50	0.46	0.47	0.44	0.44	-0.19*** [-200.00]
Computer at home	0.54	0.76	0.75	0.77	0.76	0.50	0.43	0.43	0.42	0.43	-0.22*** [-250.00]
Improving exam grade	0.01	0.00	0.01	0.00	0.00	0.08	0.07	0.07	0.07	0.06	0.00*** [16.67]
Applying to university	0.99	1.00	0.99	1.00	1.00	0.08	0.07	0.08	0.06	0.04	-0.00*** [-10.86]
Area of studies:											
Sciences	0.51	0.58	0.53	0.58	0.65	0.50	0.49	0.50	0.49	0.48	-0.08*** [-76.77]
Economics	0.11	0.09	0.07	0.09	0.11	0.31	0.28	0.25	0.28	0.32	-0.29*** [-10.02]
Humanities	0.29	0.26	0.32	0.26	0.16	0.46	0.44	0.47	0.44	0.37	0.04*** [42.07]
Arts	0.09	0.07	0.07	0.07	0.07	0.29	0.26	0.26	0.26	0.26	0.02*** [32.57]
Subjects (teacher grade):											
Biology and Geology	13.39	13.74	13.19	13.50	14.60	2.39	2.48	2.29	2.35	2.57	-0.35*** [-25.29]
Descriptive Geometry	14.16	14.60	13.74	14.23	15.60	3.01	3.09	2.85	2.98	3.08	-0.44*** [-12.47]
Drawing	14.80	15.14	14.89	15.05	15.54	2.15	2.12	2.05	2.09	2.16	-0.34*** [-11.86]
Economics	13.87	14.16	13.74	13.87	14.71	2.45	2.52	2.42	2.45	2.55	-0.29*** [-10.02]
French	12.94	13.26	12.94	13.30	14.13	2.20	2.34	2.20	2.33	2.50	-0.33*** [-5.51]
Geography	12.95	13.20	12.91	13.12	13.90	2.18	2.24	2.12	2.19	2.40	-0.25*** [-16.82]
History	12.68	13.06	12.76	13.03	13.91	2.24	2.38	2.24	2.34	2.59	-0.38*** [-20.84]
History of Culture and Arts	13.00	13.24	12.78	13.09	14.12	2.34	2.47	2.27	2.35	2.67	-0.23*** [-5.72]
Mathematics A	12.82	13.50	12.80	13.17	14.44	2.66	2.85	2.59	2.68	2.97	-0.68*** [-44.69]
Mathematics Applied to Social Sciences	13.11	13.38	13.09	13.39	14.16	2.44	2.52	2.39	2.48	2.72	-0.28*** [-10.64]
Mathematics B	12.80	13.24	12.59	12.96	14.09	2.50	2.67	2.39	2.53	2.79	-0.44*** [-7.22]
Philosophy	13.53	13.80	13.42	13.64	14.57	2.27	2.31	2.16	2.25	2.42	-0.27*** [-10.15]
Physics and Chemistry	12.90	13.40	12.72	13.09	14.36	2.55	2.72	2.44	2.54	2.85	-0.50*** [-33.11]

Continued on next page

Variable (1)	Overall (2)	W/ info (3)	Mean			Standard Deviation					Mean differences ^a (12)
			Level 1 (4)	Level 2 (5)	Level 3 (6)	Overall (7)	W/ info (8)	Level 1 (9)	Level 2 (10)	Level 3 (11)	
Portuguese	12.97	13.44	12.95	13.26	14.25	2.18	2.31	2.11	2.20	2.43	-0.47*** [-50.68]
Portuguese Literature	12.85	13.21	12.90	13.14	14.25	2.25	2.34	2.17	2.28	2.60	-0.37*** [-7.70]
Spanish	14.94	15.12	15.03	15.06	15.50	2.22	2.17	2.10	2.18	2.31	-0.18*** [-4.65]
Number of observations	1,671,373	1,367,753	41.21%	27.73%	31.06%						

^a In column (12), the difference tested is column (3) minus column (2), *StudentsWithInformation-AllStudents*. The *T*-test is for mean in differences between these two population values and the *T*-statistic is presented in square brackets. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.

Table 2: Means and estimated standard errors of teachers' and exams' scores for three levels of parents' education (2009-2018).

Subject (1)	Obs. (2)	Educ. level (3)	% of students (4)	Mean teacher score (5)	Mean exam score (6)	Mean difference (7)	Difference in mean differences ^a (8)
Biology and Geology	193,877	1	38.81%	13.19 (0.01)	9.16 (0.01)	4.03 (0.01)	0.11*** [8.56] 0.39*** [28.62]
		2	28.11%	13.50 (0.01)	9.58 (0.01)	3.92 (0.01)	
		3	33.07%	14.60 (0.01)	11.07 (0.01)	3.53 (0.01)	
Descriptive Geometry	36,472	1	34.25%	13.74 (0.03)	8.86 (0.05)	4.89 (0.04)	0.74*** [13.31] 1.42*** [26.34]
		2	26.46%	14.23 (0.03)	10.08 (0.06)	4.15 (0.04)	
		3	39.29%	15.60 (0.03)	12.87 (0.05)	2.73 (0.03)	
Drawing	25,536	1	40.70%	14.89 (0.02)	12.36 (0.03)	2.53 (0.03)	0.21*** [4.90] 0.02 [0.52]
		2	27.46%	15.05 (0.03)	12.73 (0.03)	2.33 (0.03)	
		3	31.85%	15.54 (0.02)	13.24 (0.03)	2.30 (0.03)	
Economics	34,804	1	33.02%	13.74 (0.02)	11.10 (0.03)	2.64 (0.03)	-0.03 [-0.72] 0.47*** [12.40]
		2	27.50%	13.87 (0.03)	11.21 (0.04)	2.67 (0.03)	
		3	39.48%	14.71 (0.02)	12.52 (0.03)	2.19 (0.02)	
French	6,807	1	52.06%	12.94 (0.04)	11.60 (0.06)	1.34 (0.05)	0.12 [1.58] 0.34*** [3.69]
		2	29.65%	13.30 (0.05)	12.08 (0.07)	1.22 (0.06)	
		3	18.29%	14.13 (0.07)	13.25 (0.09)	0.88 (0.07)	
Geography	108,209	1	48.04%	12.91 (0.01)	10.56 (0.01)	2.35 (0.01)	0.23*** [13.78] 0.25*** [12.66]
		2	28.28%	13.12 (0.01)	11.00 (0.02)	2.12 (0.01)	
		3	23.68%	13.90 (0.01)	12.03 (0.02)	1.87 (0.01)	
History	82,533	1	52.28%	12.76 (0.01)	10.13 (0.02)	2.62 (0.01)	0.20*** [8.24] 0.30*** [10.14]
		2	27.95%	13.03 (0.02)	10.61 (0.02)	2.42 (0.02)	
		3	19.78%	13.91 (0.02)	11.79 (0.03)	2.12 (0.02)	
History of Culture and Arts	17,241	1	43.95%	12.78 (0.03)	9.53 (0.04)	3.25 (0.03)	0.02 [0.33] 0.37*** [6.10]
		2	28.89%	13.09 (0.03)	9.86 (0.05)	3.23 (0.04)	
		3	27.16%	14.12 (0.04)	11.26 (0.05)	2.87 (0.04)	
Mathematics A	203,682	1	36.49%	12.80 (0.01)	9.78 (0.02)	3.02 (0.01)	0.25*** [14.81] 0.69*** [41.78]
		2	27.23%	13.17 (0.01)	10.40 (0.02)	2.77 (0.01)	
		3	36.29%	14.44 (0.01)	12.37 (0.02)	2.07 (0.01)	
Mathematics Applied to Social Sciences	49,677	1	52.51%	13.09 (0.01)	10.30 (0.02)	2.78 (0.02)	0.24*** [7.35] 0.51*** [12.71]
		2	27.56%	13.39 (0.02)	10.84 (0.03)	2.55 (0.03)	
		3	19.93%	14.16 (0.03)	12.13 (0.04)	2.04 (0.03)	
Mathematics B	8,607	1	36.59%	12.59 (0.04)	9.72 (0.07)	2.87 (0.06)	0.58*** [6.41] 0.80*** [9.05]
		2	26.14%	12.96 (0.05)	10.67 (0.09)	2.29 (0.07)	
		3	37.27%	14.09 (0.05)	12.60 (0.07)	1.49 (0.06)	
Philosophy	41,410	1	41.99%	13.42 (0.02)	9.67 (0.03)	3.75 (0.02)	0.41*** [11.59] 0.72*** [18.66]
		2	30.84%	13.64 (0.02)	10.30 (0.03)	3.34 (0.03)	
		3	27.18%	14.57 (0.02)	11.95 (0.04)	2.62 (0.03)	
Physics and Chemistry	190,597	1	37.19%	12.72 (0.01)	8.20 (0.01)	4.52 (0.01)	0.31*** [21.35] 0.69*** [46.60]
		2	27.56%	13.09 (0.01)	8.87 (0.02)	4.21 (0.01)	
		3	35.25%	14.36 (0.01)	10.84 (0.02)	3.52 (0.01)	
Portuguese	343,089	1	41.45%	12.95 (0.01)	10.23 (0.01)	2.71 (0.01)	0.15*** [13.85] 0.08*** [7.23]
		2	27.46%	13.26 (0.01)	10.69 (0.01)	2.57 (0.01)	
		3	31.09%	14.25 (0.01)	11.76 (0.01)	2.49 (0.01)	

Continued on next page

Subject (1)	Obs. (2)	Educ. level (3)	% of students (4)	Mean teacher score (5)	Mean exam score (6)	Mean difference (7)	Difference in mean differences ^a (8)
Portuguese Literature	11,912	1	54.75%	12.90 (0.03)	10.55 (0.04)	2.35 (0.04)	0.23*** [3.45]
		2	26.80%	13.14 (0.04)	11.02 (0.06)	2.12 (0.05)	0.12 [1.41]
		3	18.45%	14.25 (0.06)	12.25 (0.07)	2.00 (0.06)	
Spanish	13,300	1	54.22%	15.03 (0.02)	13.34 (0.04)	1.69 (0.03)	0.11* [2.05]
		2	28.45%	15.06 (0.04)	13.49 (0.05)	1.58 (0.04)	0.25*** [3.55]
		3	17.33%	15.50 (0.05)	14.17 (0.06)	1.33 (0.05)	

^a In column (8), there are two *T*-tests for the difference in mean differences between *Level1Students-Level2Students*, and *Level2Students-Level3Students*, respectively. The *T*-statistic is presented in square brackets. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.

4.1. Composition effect

To verify if we have a composition effect, we present the following two tables. [Table 3](#) presents descriptive statistics of teacher and exam scores by subject and by exam score level. [Table 4](#) presents the percentage of students in each subject and exam score level by the level of parents' education.

Since the exam scores are graded from a scale from 0 to 200, we convert to a 1-5 scale using the criteria defined by the Ministry of Education: level 1 corresponds to the interval of scores [0; 35[, level 2 to [35; 95[, level 3 to [95; 135[, level 4 to [135; 175[, and level 5 to [175; 200].

Regarding [Table 3](#), column (1) indicates the subject; column (2) has the number of observations; column (3) has the exam score level; column (4) has the percentage of students for each exam score level; columns (5)-(7) have the means of teacher score, exam score, and the difference *TeacherScore-ExamScore*, respectively.

Since students are required to have at least 10 to be admitted to exam, the teacher score means are all positive and greater than 10. The means increase with the exam score level, which is also expected since students with greater exam score are more likely to have a higher teacher score. Column (7) shows that the mean difference decreases with the exam score level, that is, the higher the exam score level, the smaller the difference *TeacherScore-ExamScore*. This result

allows us to conclude that students with a higher exam score level decrease less the exam score with respect to their teacher-assigned grade.

In [Table 4](#), column (1) indicates the subject; column (2) has the number of observations; column (3) has the level of parents' education; column (4) has the percentage of students for each level of parents' education; columns (5)-(9) have the percentage of students by exam score level for each level of parents' education, respectively.

The percentage of students in each level of parents' education for the whole sample is given in [Table 1](#): 41.21% in level 1, 27.73% in level 2, and 31.06% in level 3. By looking at the percentages presented in [Table 4](#), we confirm that the students are not equally distributed across the exam scores. Students from the lower level of parents' education always decrease as the exam score level increases, whilst students from level 3 always increase. The percentage of students whose parents have an education level equal to 1 decreases with the exam score level, whilst the opposite occurs for students with parents with an education level of 3. No pattern, however, can be identified for students whose parents have an education level of 2. For this reason, and to increase the number of observations in each level of analysis, in the next section we will merge the levels 1 and 2 of parents' education.

[Table 3](#) shows that, as the exam score level increases, the difference *TeacherScore-ExamScore* decreases. [Table 4](#) shows that, as the exam level score increases, the percentage of students with more educated parents also increases. It may happen that the trend for the difference between teacher grade and exam score to decrease with parents' education observed in [Table 2](#) is just the result of a larger number of students with more educated parents with higher exam scores. So, the following analysis will take this into account.

Table 3: Means and estimated standard errors of teachers' and exams' scores for five levels of exam scores (2009-2018).

Subject (1)	Obs. (2)	Exam score level ^a (3)	% of students (4)	Mean teacher score (5)	Mean exam score (6)	Mean difference (7)
Biology and Geology	193,877	1	1.21%	11.17 (0.02)	2.63 (0.01)	8.54 (0.03)
		2	44.64%	12.13 (0.01)	6.93 (0.01)	5.20 (0.01)
		3	35.70%	14.26 (0.01)	11.12 (0.00)	3.15 (0.01)
		4	16.61%	16.63 (0.01)	14.93 (0.01)	1.70 (0.01)
		5	1.84%	18.30 (0.02)	18.07 (0.01)	0.24 (0.02)
Descriptive Geometry	36,472	1	10.50%	11.85 (0.03)	2.32 (0.01)	9.54 (0.03)
		2	33.32%	12.85 (0.02)	6.23 (0.02)	6.62 (0.02)
		3	20.74%	14.50 (0.03)	11.20 (0.01)	3.30 (0.03)
		4	17.30%	16.13 (0.03)	15.31 (0.01)	0.82 (0.03)
		5	18.15%	18.06 (0.03)	19.11 (0.01)	-1.06 (0.02)
Drawing	25,536	1	0.16%	13.32 (0.30)	1.94 (0.19)	11.38 (0.44)
		2	10.75%	13.66 (0.04)	7.74 (0.02)	5.92 (0.04)
		3	47.06%	14.65 (0.02)	11.52 (0.01)	3.13 (0.02)
		4	36.92%	15.92 (0.02)	15.03 (0.01)	0.89 (0.02)
		5	5.11%	17.19 (0.05)	18.22 (0.02)	-1.03 (0.05)
Economics	34,804	1	0.67%	11.88 (0.11)	2.70 (0.05)	9.18 (0.12)
		2	27.06%	12.34 (0.02)	7.13 (0.01)	5.21 (0.02)
		3	37.33%	13.69 (0.02)	11.25 (0.01)	2.45 (0.02)
		4	27.34%	15.71 (0.02)	15.19 (0.01)	0.52 (0.02)
		5	7.60%	17.57 (0.03)	18.30 (0.01)	-0.74 (0.03)
French	6,807	1	0.44%	11.33 (0.21)	2.42 (0.18)	8.91 (0.30)
		2	20.61%	11.67 (0.04)	7.39 (0.04)	4.28 (0.05)
		3	43.37%	12.59 (0.03)	11.41 (0.02)	1.18 (0.03)
		4	30.06%	14.67 (0.05)	15.13 (0.02)	-0.46 (0.05)
		5	5.52%	17.02 (0.09)	18.34 (0.03)	-1.32 (0.09)
Geography	108,209	1	0.19%	11.27 (0.10)	2.61 (0.05)	8.66 (0.11)
		2	26.70%	11.77 (0.01)	7.56 (0.01)	4.21 (0.01)
		3	51.61%	13.04 (0.01)	11.23 (0.00)	1.81 (0.01)
		4	20.32%	15.27 (0.01)	14.79 (0.01)	0.49 (0.01)
		5	1.18%	17.53 (0.05)	18.00 (0.01)	-0.47 (0.05)
History	82,533	1	1.59%	11.15 (0.04)	2.56 (0.02)	8.59 (0.04)
		2	34.26%	11.75 (0.01)	6.93 (0.01)	4.82 (0.01)
		3	40.25%	12.97 (0.01)	11.15 (0.01)	1.82 (0.01)
		4	20.73%	14.94 (0.02)	15.01 (0.01)	-0.06 (0.02)
		5	3.17%	16.97 (0.04)	18.25 (0.01)	-1.28 (0.04)
History of Culture and Arts	17,241	1	1.97%	11.69 (0.09)	2.56 (0.04)	9.13 (0.10)
		2	38.65%	12.02 (0.02)	6.92 (0.02)	5.11 (0.03)
		3	40.90%	13.34 (0.03)	11.07 (0.01)	2.27 (0.03)
		4	16.14%	15.47 (0.05)	14.95 (0.02)	0.52 (0.04)
		5	2.35%	17.25 (0.10)	18.27 (0.03)	-1.02 (0.09)
Mathematics A	203,682	1	4.13%	10.70 (0.01)	2.35 (0.01)	8.36 (0.01)
		2	34.38%	11.47 (0.01)	6.63 (0.01)	4.84 (0.01)
		3	29.16%	13.17 (0.01)	11.17 (0.00)	1.99 (0.01)
		4	22.24%	15.59 (0.01)	15.22 (0.01)	0.37 (0.01)
		5	10.10%	17.89 (0.01)	18.53 (0.01)	-0.64 (0.01)

Continued on next page

Subject (1)	Obs. (2)	Exam score level ^a (3)	% of students (4)	Mean teacher score (5)	Mean exam score (6)	Mean difference (7)
Mathematics Applied to Social Sciences	49,677	1	2.33%	11.05 (0.04)	2.35 (0.02)	8.70 (0.04)
		2	33.90%	11.91 (0.01)	6.95 (0.01)	4.96 (0.02)
		3	35.95%	13.23 (0.02)	11.21 (0.01)	2.03 (0.02)
		4	22.28%	15.16 (0.02)	15.07 (0.01)	0.09 (0.02)
		5	5.54%	17.21 (0.04)	18.36 (0.01)	-1.15 (0.04)
Mathematics B	8,607	1	3.39%	10.92 (0.08)	2.24 (0.05)	8.68 (0.09)
		2	31.81%	11.62 (0.03)	6.78 (0.03)	4.83 (0.04)
		3	32.51%	12.85 (0.04)	11.17 (0.02)	1.68 (0.04)
		4	23.78%	14.91 (0.05)	15.17 (0.03)	-0.26 (0.05)
		5	8.50%	17.11 (0.07)	18.42 (0.03)	-1.31 (0.07)
Philosophy	41,410	1	2.54%	11.96 (0.05)	2.57 (0.02)	9.40 (0.06)
		2	35.67%	12.46 (0.01)	6.78 (0.01)	5.68 (0.02)
		3	36.28%	13.80 (0.02)	11.17 (0.01)	2.63 (0.02)
		4	21.81%	15.62 (0.02)	15.03 (0.01)	0.59 (0.02)
		5	3.70%	17.30 (0.04)	18.24 (0.02)	-0.94 (0.04)
Physics and Chemistry	190,597	1	3.79%	10.80 (0.01)	2.65 (0.01)	8.16 (0.01)
		2	50.63%	11.84 (0.01)	6.54 (0.01)	5.30 (0.01)
		3	27.54%	14.23 (0.01)	11.12 (0.01)	3.10 (0.01)
		4	14.58%	16.68 (0.01)	15.10 (0.01)	1.58 (0.01)
		5	3.46%	18.52 (0.02)	18.33 (0.01)	0.19 (0.02)
Portuguese	343,088	1	0.51%	11.04 (0.03)	2.68 (0.01)	8.36 (0.03)
		2	30.47%	11.88 (0.00)	7.22 (0.00)	4.66 (0.01)
		3	45.50%	13.36 (0.00)	11.13 (0.00)	2.23 (0.00)
		4	21.48%	15.47 (0.01)	14.84 (0.00)	0.63 (0.01)
		5	2.04%	17.51 (0.02)	18.01 (0.01)	-0.50 (0.02)
Portuguese Literature	11,912	1	1.23%	11.26 (0.11)	2.40 (0.06)	8.86 (0.12)
		2	28.80%	11.96 (0.03)	7.07 (0.03)	4.89 (0.03)
		3	43.65%	13.05 (0.03)	11.17 (0.02)	1.88 (0.03)
		4	22.73%	14.68 (0.04)	14.94 (0.02)	-0.26 (0.04)
		5	3.58%	16.59 (0.11)	18.18 (0.03)	-1.58 (0.11)
Spanish	13,300	1	0.06%	11.50 (0.42)	2.02 (0.38)	9.47 (0.57)
		2	9.94%	13.22 (0.05)	7.68 (0.03)	5.55 (0.05)
		3	35.90%	14.30 (0.03)	11.58 (0.02)	2.72 (0.03)
		4	43.67%	15.75 (0.03)	15.32 (0.01)	0.43 (0.03)
		5	10.43%	17.15 (0.05)	18.38 (0.02)	-1.22 (0.05)

^a In column (3), exam score level 1 corresponds to the interval of scores [0; 35[, level 2 to [35; 95[, level 3 to [95; 135[, level 4 to [135; 175[, and level 5 to [175; 200].

Table 4: Total number of observations by level of parents' education and by exam grade (2009-2018).

Subject (1)	Obs. (2)	Educ. level (3)	% of students (4)	Exam score level ^a				
				1 (5)	2 (6)	3 (7)	4 (8)	5 (9)
Biology and Geology	193,877	1	38.81%	51.62%	46.07%	36.24%	26.31%	16.98%
		2	28.11%	30.29%	30.14%	28.29%	23.36%	17.14%
		3	33.07%	18.09%	23.79%	35.47%	50.34%	65.88%
Descriptive Geometry	36,472	1	34.25%	52.06%	43.11%	33.02%	25.24%	17.67%
		2	26.46%	28.84%	29.37%	27.52%	25.21%	19.73%
		3	39.29%	19.10%	27.52%	39.47%	49.56%	62.60%
Drawing	25,536	1	40.70%	47.50%	48.89%	43.62%	36.05%	29.89%
		2	27.46%	20.00%	27.58%	27.73%	27.46%	24.83%
		3	31.85%	32.50%	23.53%	28.65%	36.49%	45.29%
Economics	34,804	1	33.02%	38.79%	38.39%	34.83%	28.21%	21.85%
		2	27.50%	34.48%	31.33%	28.45%	24.15%	20.64%
		3	39.48%	26.72%	30.28%	36.72%	47.64%	57.50%
French	6,807	1	52.06%	83.33%	60.01%	53.79%	46.92%	34.31%
		2	29.65%	10.00%	28.72%	30.76%	29.23%	28.19%
		3	18.29%	6.67%	11.26%	15.45%	23.85%	37.50%
Geography	108,209	1	48.04%	74.88%	57.07%	48.20%	36.71%	27.74%
		2	28.28%	19.43%	28.06%	29.06%	26.96%	22.88%
		3	23.68%	5.69%	14.87%	22.74%	36.32%	49.37%
History	82,533	1	52.28%	65.32%	58.39%	52.39%	43.77%	33.82%
		2	27.95%	25.15%	27.66%	28.45%	28.11%	24.92%
		3	19.78%	9.53%	13.95%	19.15%	28.12%	41.26%
History of Culture and Arts	17,241	1	43.95%	55.46%	49.92%	43.53%	32.31%	23.65%
		2	28.89%	30.68%	30.66%	28.73%	25.77%	22.41%
		3	27.16%	13.86%	19.42%	27.74%	41.91%	53.94%
Mathematics A	203,682	1	36.49%	52.65%	44.53%	36.20%	28.66%	20.54%
		2	27.23%	28.26%	29.90%	28.50%	24.71%	19.57%
		3	36.29%	19.10%	25.57%	35.29%	46.62%	59.89%
Mathematics Applied to Social Sciences	49,677	1	52.51%	63.49%	59.15%	52.47%	45.39%	36.17%
		2	27.56%	23.79%	27.13%	28.52%	27.76%	24.66%
		3	19.93%	12.72%	13.72%	19.00%	26.85%	39.18%
Mathematics B	8,607	1	36.59%	53.08%	48.76%	35.31%	26.67%	17.08%
		2	26.14%	30.14%	27.61%	28.23%	23.01%	19.81%
		3	37.27%	16.78%	23.63%	36.45%	50.32%	63.11%
Philosophy	41,410	1	41.99%	57.66%	50.30%	41.13%	31.22%	22.91%
		2	30.84%	29.21%	32.15%	31.60%	28.95%	23.04%
		3	27.18%	13.13%	17.55%	27.28%	39.83%	54.05%
Physics and Chemistry	190,597	1	37.19%	55.00%	44.46%	31.63%	22.94%	15.47%
		2	27.56%	28.04%	29.88%	27.49%	22.24%	16.04%
		3	35.25%	16.96%	25.65%	40.87%	54.82%	68.49%
Portuguese	343,088	1	41.45%	57.49%	49.80%	41.35%	31.29%	21.64%
		2	27.46%	25.94%	28.42%	28.04%	25.54%	21.00%
		3	31.09%	16.57%	21.78%	30.61%	43.16%	57.36%

Continued on next page

Subject (1)	Obs. (2)	Educ. level (3)	% of students (4)	Exam score level ^a				
				1 (5)	2 (6)	3 (7)	4 (8)	5 (9)
Portuguese Literature	11,912	1	54.75%	65.99%	62.72%	55.11%	46.05%	37.70%
		2	26.80%	24.49%	25.79%	27.52%	27.03%	25.29%
		3	18.45%	9.52%	11.48%	17.37%	26.92%	37.00%
Spanish	13,300	1	54.22%	50.00%	57.72%	57.28%	52.44%	47.80%
		2	28.45%	37.50%	29.88%	28.02%	28.75%	27.25%
		3	17.33%	12.50%	12.41%	14.70%	18.80%	24.95%

^a The exam score level 1 corresponds to the interval of scores [0; 35[, level 2 to [35; 95[, level 3 to [95; 135[, level 4 to [135; 175[, and level 5 to [175; 200].

5. Regression analysis

5.1. Specification

In the previous section, an analysis of descriptive statistics for our database suggested that high-SES students could, in relative terms, benefit from taking exams. The results were remarkably consistent throughout all the 16 subjects studied. In this section, a regression analysis will be carried out to better understand the socio-economic gap in assessments, when controlling for observable students' characteristics. We will see that a regression that includes all observations, that is, at the whole distribution of grades, is consistent with the results from the previous section; however, when we perform the analysis by level of exam score, we see that our previous conclusions were deceitful, due to the unbalanced distribution of students along the exam score levels.

We assume a linear specification for the gap equation to estimate the following model:

$$G_{it} = \alpha + \beta E_i + \gamma X_{it} + \mu_t + \epsilon_{it}, \quad (1)$$

where G_{it} is the gap between teacher and exam assessment (*TeacherScore-ExamScore*) for student i in year t . G_{it} is assumed to be a function of the level of education of the students'

parents, E ($E = 0$ no parent with higher education, and $E = 1$ at least one parent with higher education). The model considers a vector of co-variables X_{it} , which includes the student's age, area of studies, and a set of dummy variables controlling for the nationality (Portuguese or not), gender, if the exam is taken with the purpose of applying to university, and if the student is improving a previous exam score. The model also includes year fixed effects, μ_t . The model is estimated using ordinary least squares pooling all observations from the available repeated cross sections.

We are only considering two levels of parents' education: parents with no higher education, and at least one parent with higher education. We conducted this analysis also using the three levels of parents' education considered in [Section 4](#), which is presented in [Appendix C](#). In the latter case, two coefficients were estimated in the regression instead of one: one for each level above the base level (basic education). The results are coherent with the ones presented in this section, however there are fewer significant coefficients⁸.

5.2. Results

5.2.1. General regressions

[Table 5](#) presents the OLS estimation results of the model specification described in [Equation \(1\)](#) for the 16 different subjects. Column (1) indicates the subject; column (2) has the number of observations; and column (3) has the estimated coefficient of interest, β . The standard errors were corrected for clustering at the school level. The period of analysis is from 2009 until 2018.

⁸ Moreover, we tried to conduct the same analysis by aggregating the parents' education in six levels. Unfortunately, due to insufficient observations, we were only able to present the descriptive statistics and the regressions for the whole distribution of grades for the six level analysis. The results for the whole distribution of grades are remarkably consistent with the results for the three level analysis: each level increase of parents' education shows a better student performance, and a decrease in the difference *TeacherScore-ExamScore*. Regarding the results from the exam score level, it was not possible to confirm. The results are presented in [Appendix D](#).

All the coefficients from the regressions are negative and significant. This means that increasing the parents' education to higher education reduces the assessment gap. Since all coefficients are significant and negative, we would be prone to believe that the level of education decreases the difference between teacher and exam score. However, as we said, this is deceitful, because there is a composition effect.

Table 5: Linear regressions for two levels of parents' education (2009-2018).

Subject (1)	Observations (2)	β^a (3)
Biology and Geology	193,877	-0.432*** (0.028)
Descriptive Geometry	36,472	-1.412*** (0.087)
Drawing	25,536	-0.165** (0.056)
Economics	34,804	-0.463*** (0.068)
French	6,807	-0.336** (0.107)
Geography	108,209	-0.233*** (0.038)
History	82,533	-0.421*** (0.048)
History of Culture and Arts	17,241	-0.402*** (0.077)
Mathematics A	203,682	-0.802*** (0.050)
Mathematics Applied to Social Sciences	49,677	-0.613*** (0.057)
Mathematics B	8,607	-0.965*** (0.083)
Philosophy	41,410	-0.827*** (0.062)
Physics and Chemistry	190,597	-0.745*** (0.042)
Portuguese	343,088	-0.159*** (0.027)
Portuguese Literature	11,912	-0.278** (0.103)
Spanish	13,300	-0.265*** (0.071)

^a Dependent variable is the assessment gap, *TeacherScore-ExamScore*. Standard errors in parentheses are corrected for clustering at the school level. Only the coefficient of interest, β , is presented. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.

5.2.2. Regressions by exam score level

We now make the same regression using the model specified in [Equation \(1\)](#) but for each exam score level. We use the conversion scale presented in [Section 4](#). As previously referred, the

minimum passing grade is 9.5, thus, exam score levels 1 and 2 correspond to students who failed in the exam.

Table 6 presents the OLS estimation results of the model specifications described in Equation (1) for the 16 different subjects by exam score level. Column (1) indicates the subject; column (2) has the number of observations; columns (3)-(7) have the estimated coefficient of interest, β , with their respective standard errors (corrected for clustering at the school level).

The results contradict our first regression analysis. In fact, all significant coefficients (50/80) are positive, which clearly contradicts Table 5 for the whole distribution of grades. Thus, when comparing similar students, meaning students on the same range of exam scores, instead of the whole grading distribution, we obtain an opposite effect: the level of parents' education seems to *increase* the difference between teacher and exam score. The results range from 0.074 in the exam score level 2 of Biology and Geology to 2.203 in the exam score level 1 of Geography.

In light of the results from Section 4, regarding the composition effect, we can be sure that the results from Table 2 and Table 5 are misleading, as students are not evenly distributed. We must, then, consider the results from Table 6, which presents the regressions by the exam score level. Therefore, when comparing similar students, we confirm that low-SES students decrease less the exam score with respect to their teacher-assigned grade. As a consequence, when comparing students with similar scores, the results suggest that exams are relatively more favourable for low-SES students, whilst teacher grading seems to be more favourable for high-SES students.

Table 6: Linear regressions by exam score level with a scale 1-5 for two levels of parents' education (2009-2018).

Subject (1)	Observations (2)	Exam score level ^a				
		1 (3)	2 (4)	3 (5)	4 (6)	5 (7)
Biology and Geology	193,877	0.195* (0.082)	0.074** (0.023)	0.233*** (0.027)	0.220*** (0.029)	0.141** (0.046)
Descriptive Geometry	36,472	0.193* (0.092)	-0.040 (0.064)	0.249** (0.072)	0.122 (0.065)	0.183** (0.062)
Drawing	25,536	-1.019 (0.963)	0.051 (0.101)	0.143** (0.050)	0.274*** (0.056)	0.370** (0.111)
Economics	34,804	0.040 (0.360)	0.107 (0.064)	0.165* (0.067)	0.145* (0.060)	0.133* (0.065)
French	6,807	0.228 (1.747)	-0.008 (0.166)	0.148 (0.107)	0.436*** (0.121)	0.269 (0.171)
Geography	108,209	2.203* (0.856)	-0.000 (0.040)	0.119** (0.036)	0.287*** (0.043)	0.185* (0.092)
History of Culture and Arts	17,241	0.098 (0.301)	0.076 (0.078)	0.495*** (0.077)	0.652*** (0.112)	0.522** (0.190)
History	82,533	0.011 (0.150)	0.091* (0.046)	0.298*** (0.040)	0.482*** (0.053)	0.387*** (0.081)
Mathematics Applied to SS	49,677	0.451** (0.138)	0.023 (0.054)	0.129* (0.053)	0.183*** (0.052)	0.358*** (0.079)
Mathematics A	203,682	0.131** (0.046)	-0.023 (0.027)	0.207*** (0.031)	0.268*** (0.039)	0.259*** (0.038)
Mathematics B	8,607	0.579* (0.238)	-0.017 (0.103)	0.056 (0.087)	0.212* (0.101)	-0.016 (0.130)
Philosophy	41,410	0.131 (0.166)	-0.015 (0.050)	0.206*** (0.049)	0.235*** (0.054)	0.387*** (0.089)
Physics and Chemistry	190,597	0.118** (0.042)	-0.040 (0.026)	0.103** (0.034)	0.069 (0.039)	0.044 (0.039)
Portuguese Literature	11,912	0.361 (0.429)	0.212 (0.131)	0.515*** (0.085)	0.683*** (0.112)	0.948*** (0.219)
Portuguese	343,088	0.268** (0.089)	0.182*** (0.023)	0.424*** (0.027)	0.540*** (0.028)	0.564*** (0.044)
Spanish	13,300	—	-0.185 (0.173)	-0.087 (0.084)	0.040 (0.071)	0.350** (0.113)

^a Dependent variable is the assessment gap, *TeacherScore-ExamScore*. Only the coefficient of interest, β , is presented. Standard errors are presented in parentheses (corrected for clustering at the school level). The exam score level 1 corresponds to the interval of scores [0; 35[, level 2 to [35; 95[, level 3 to [95; 135[, level 4 to [135; 175[, and level 5 to [175; 200]. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.

To verify the robustness of our results, we conducted the same regressions of the last two subsections but using social support information instead of the level of parents' education. We use the same [Equation \(1\)](#) but E is now a dummy for social support ($E = 0$ for students who do not benefit from social support, and $E = 1$ for students who benefit from social support).

The results are coherent with our results, as it had already been shown by Ribeiro (2021). Therefore, both the level of parents' education and social support indicate the same results: low-SES students tend to have a smaller gap between teacher and exam score. Nonetheless, we note that using social support as a proxy for students' SES instead of the level of parents' education diminishes the number of low-SES students from 68.94% to 25.42%. The results are presented in [Appendix E](#).

5.2.3. Regressions by exam score level and gender

Given the results obtained by Ângelo and Balcão Reis (2021), which showed the existence of an assessment gap between girls and boys in the Portuguese educational system, we conducted the regression by exam score level also by gender.

[Table 7](#) presents the OLS estimation results of the model specifications described in [Equation \(1\)](#) for the 16 different subjects by exam score level. Column (1) indicates the subject; column (2) has the students' gender; column (3) has the percentage of students of each gender; columns (4)-(8) have the estimated coefficient of interest, β , with their respective standard errors (corrected for clustering at the school level).

When looking at the gender regressions, we see that virtually all significant coefficients are positive, thus we conclude that the socio-economic gap is also present when we conduct the analysis for girls and boys separately. Additionally, more than half of the coefficients are significant for girls (46/80), but only approximately 1/3 are significant for boys (29/80), suggesting

that the socio-economic gap effect seems to be more observable for girls. Furthermore, almost all coefficients seem to be bigger for girls than boys, which would indicate that, for students with the same exam score level, girls have a higher teacher grade, compared to boys. This would be in line with the results from the literature⁹.

Finally, we conducted also these regressions by gender using SASE as a proxy for students' SES instead of the level of their parents' education. The results confirm again that using SASE as a proxy yields the same behaviour as using the level of parents' education. The results are presented in [Appendix E](#).

Table 7: Linear regressions by exam score level with a scale 1-5 and by gender for two levels of parents' education (2009-2018).

Subject (1)	Gender (2)	% of students (3)	Exam score level ^a				
			1 (4)	2 (5)	3 (6)	4 (7)	5 (8)
Biology and Geology	Girls	54.78%	0.249 (0.129)	0.093** (0.029)	0.270*** (0.031)	0.230*** (0.034)	0.145* (0.064)
	Boys	45.22%	0.132 (0.105)	0.052* (0.026)	0.187*** (0.029)	0.204*** (0.034)	0.140* (0.062)
Descriptive Geometry	Girls	48.45%	0.069 (0.116)	0.023 (0.082)	0.318** (0.101)	0.265* (0.104)	0.238** (0.081)
	Boys	51.55%	0.383* (0.152)	-0.110 (0.083)	0.183* (0.086)	0.003 (0.079)	0.149 (0.078)
Drawing	Girls	66.97%	0.999 (2.644)	0.133 (0.119)	0.230*** (0.063)	0.353*** (0.061)	0.353** (0.129)
	Boys	33.03%	-2.134 (1.414)	-0.102 (0.165)	-0.029 (0.075)	0.082 (0.087)	0.412* (0.186)
Economics	Girls	47.88%	0.420 (0.553)	0.118 (0.095)	0.197* (0.085)	0.198** (0.075)	0.207* (0.082)
	Boys	52.12%	-0.331 (0.464)	0.098 (0.069)	0.143 (0.074)	0.088 (0.067)	0.059 (0.092)
French	Girls	71.25%	-1.616 (0.880)	-0.003 (0.188)	0.137 (0.135)	0.476*** (0.128)	0.226 (0.174)
	Boys	28.75%	—	-0.017 (0.285)	0.156 (0.137)	0.329 (0.226)	0.275 (0.473)

Continued on next page

⁹ However, we do not test for significance, and therefore we can only confirm that the socio-economic gap is found for both boys and girls.

Subject (1)	Gender (2)	% of students (3)	Exam score level ^a				
			1 (4)	2 (5)	3 (6)	4 (7)	5 (8)
Geography	Girls	61.62%	2.416 (1.509)	-0.015 (0.049)	0.157*** (0.042)	0.321*** (0.048)	0.248 (0.127)
	Boys	38.38%	2.146*** (0.522)	0.028 (0.052)	0.078 (0.040)	0.246*** (0.056)	0.133 (0.132)
History of Culture and Arts	Girls	68.36%	-0.020 (0.353)	0.110 (0.092)	0.567*** (0.090)	0.715*** (0.117)	0.411* (0.207)
	Boys	31.64%	0.101 (0.480)	-0.001 (0.119)	0.326** (0.111)	0.474* (0.189)	0.930* (0.392)
History	Girls	70.24%	-0.112 (0.171)	0.118* (0.051)	0.313*** (0.047)	0.475*** (0.059)	0.399*** (0.092)
	Boys	29.76%	0.469 (0.386)	0.029 (0.070)	0.267*** (0.054)	0.491*** (0.074)	0.368** (0.140)
Mathematics Applied to Social Sciences	Girls	69.75%	0.484* (0.196)	0.053 (0.066)	0.160* (0.067)	0.210** (0.060)	0.390*** (0.085)
	Boys	30.25%	0.286 (0.228)	-0.039 (0.074)	0.066 (0.071)	0.105 (0.091)	0.290 (0.153)
Mathematics A	Girls	53.36%	0.236*** (0.063)	0.009 (0.035)	0.259*** (0.038)	0.294*** (0.045)	0.277*** (0.039)
	Boys	46.64%	0.059 (0.052)	-0.056 (0.029)	0.149*** (0.034)	0.232*** (0.042)	0.233*** (0.049)
Mathematics B	Girls	65.05%	0.873** (0.325)	0.029 (0.122)	0.041 (0.110)	0.311* (0.124)	0.152 (0.152)
	Boys	34.95%	0.095 (0.407)	-0.078 (0.173)	0.078 (0.132)	0.004 (0.167)	-0.500 (0.276)
Philosophy	Girls	64.99%	0.067 (0.246)	0.068 (0.062)	0.234*** (0.055)	0.255*** (0.056)	0.401*** (0.107)
	Boys	35.01%	0.212 (0.219)	-0.112 (0.070)	0.161* (0.067)	0.182* (0.091)	0.349* (0.174)
Physics and Chemistry	Girls	51.23%	0.159* (0.073)	-0.010 (0.034)	0.166*** (0.039)	0.076 (0.046)	-0.004 (0.052)
	Boys	48.77%	0.082 (0.053)	-0.067* (0.026)	0.035 (0.037)	0.057 (0.043)	0.087 (0.048)
Portuguese Literature	Girls	75.09%	0.181 (0.626)	0.220 (0.171)	0.500*** (0.098)	0.701*** (0.121)	0.994*** (0.240)
	Boys	24.91%	0.202 (0.508)	0.212 (0.182)	0.574** (0.173)	0.599* (0.235)	0.395 (0.442)
Portuguese	Girls	58.09%	0.050 (0.169)	0.219*** (0.027)	0.450*** (0.029)	0.550*** (0.028)	0.582*** (0.051)
	Boys	41.91%	0.387*** (0.105)	0.162*** (0.026)	0.395*** (0.029)	0.521*** (0.040)	0.501*** (0.076)

Continued on next page

Subject (1)	Gender (2)	% of students (3)	Exam score level ^a				
			1 (4)	2 (5)	3 (6)	4 (7)	5 (8)
Spanish	Girls	69.53%	—	-0.142 (0.226)	-0.117 (0.096)	-0.024 (0.078)	0.494*** (0.112)
	Boys	30.47%	—	-0.225 (0.262)	-0.032 (0.133)	0.188 (0.126)	-0.158 (0.238)

^a Dependent variable is the assessment gap, *TeacherScore-ExamScore*. Only the coefficient of interest, β , is presented. Standard errors are presented in parentheses (corrected for clustering at the school level). The exam score level 1 corresponds to the interval of scores [0; 35[, level 2 to [35; 95[, level 3 to [95; 135[, level 4 to [135; 175[, and level 5 to [175; 200]. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.

6. Concluding remarks

The purpose of this study has been to analyse whether the difference between teacher-based assessment and exam-based assessment at Portuguese public schools is associated with students' SES, measured by the level of their parents' education.

Because low-SES students are not evenly distributed throughout the scores, and to be able to compare students with similar exam scores, we conducted our analysis by exam score level. Overall, all students are given worse grades in exams than by their teachers, this is already known in the Portuguese context (Ângelo and Balcão Reis, 2021; Ribeiro, 2021). But what is most notable is that, using a linear model, we find evidence that, for students with the same exam score level, low-SES students tend to have a smaller gap between their teacher score and their exam score. Thus, evidence suggests that exams are relatively more favourable for low-SES students. On the other hand, teacher assessment seems to be relatively more favourable for high-SES students.

These results are very relevant and raise important questions such as: are low-SES students being under-evaluated by their teachers? Is there a teacher bias against low-SES students? Or do students from low socio-economic backgrounds lack stimuli from home that the schools are not

being able to provide? Unfortunately, all of these questions lay outside the scope of this study. As a consequence, we must acknowledge that we cannot identify the mechanisms behind these discrepancies. It might be the case that students from low-SES are more prone to misbehaviour than students from high-SES, or are more prone to be absent from classes, or are less keen to attend school than their high-SES peers. It has been shown in the literature that parental expectations are dependent on family socio-economic background, namely the level of income of the household, and that agreement of parental and students' expectations presents a positive influence on students' achievement (Marcenaro Gutiérrez and López Agudo, 2016). All of these factors might contribute to the assessment gap found.

Despite these limitations, our findings suggest two implications. Firstly, there are obviously systematic differences between teacher assessment and exam assessment for all students, regardless of socio-economic factors. Although it is possible, as some literature suggests, that those differences arise from potential teachers' biases, it might also be the case that teachers assess different aspects of students' performance which are not evaluated in exams, and could be more likely to assess non-cognitive skills such as student behaviour and motivation. Secondly, it seems that students' SES systematically generates a gap between low- and high-SES students, and this gap must be considered when drawing public policies. Teachers should be aware that, for students with the same exam score level, low-SES students are systematically being under-evaluated by their teachers in comparison to their high-SES peers.

In addition, just as the findings from Ângelo and Balcão Reis (2021) suggest that exams favour boys and teacher grading favours girls, our evidence points that, for similar students, exams favour low-SES students and teacher grading high-SES students. We find that our results are valid when we further take gender into account, since the socio-economic gap is present for both boys and girls.

Lastly, since teacher and exam scores are the only measure of a student's performance when applying to university, the existence of this assessment gap between low- and high-SES students should raise a concern as to what schools could do to overcome this issue. In this regard, the current no-mandatory exam policy that arose amidst the Covid crisis in 2020 is of our concern. Exams have ceased to be mandatory and students will only take the exams they need to pursue a university degree. Therefore, since 2020, the exam scores no longer account for the final subject grade. Based on our evidence, we are concerned that suppressing mandatory exams possibly increases even further the assessment gap between low- and high-SES students, especially when applying to university.

References

- Ângelo, C., and A. Balcão Reis. 2021. "Gender gaps in different grading systems." *Education Economics* 29 (1): 105–119. <https://doi.org/10.1080/09645292.2020.1853681>.
- Ângelo, C., and A. Balcão Reis. 2022. "Gender gaps in different assessment systems: The role of teacher gender." *Nova SBE Working Paper Series 2022* (640). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4048290.
- Auwarter, A. E., and M. S. Aruguete. 2008. "Effects of Student Gender and Socioeconomic Status on Teacher Perceptions." *The Journal of Educational Research* 101 (4): 242–246. <https://doi.org/10.3200/JOER.101.4.243-246>.
- Bar, T., and A. Zussman. 2012. "Partisan Grading." *American Economic Journal: Applied Economics* 4 (1): 30–48. <https://doi.org/10.1257/app.4.1.30>.
- Botelho, F., R. A. Madeira, and M. A. Rangel. 2015. "Racial Discrimination in Grading: Evidence from Brazil." *American Economic Journal: Applied Economics* 7 (4): 37–52. <https://doi.org/10.1257/app.20140352>.
- Burgess, S., and E. Greaves. 2013. "Test Scores, Subjective Assessment, and Stereotyping of Ethnic Minorities." *Journal of Labor Economics* 31 (3): 535–576. <https://doi.org/10.1086/669340>.
- Calsamiglia, C., and A. Loviglio. 2019. "Grading on a curve: When having good peers is not good." *Economics of Education Review* 73. <https://doi.org/10.1016/j.econedurev.2019.101916>.

- Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York. 1966. *Equality of Educational Opportunity*. Technical report. U.S. Department of Health, Education, and Welfare. <https://eric.ed.gov/?id=ED012275>.
- Cornwell, C., D. B. Mustard, and J. van Parys. 2013. “Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments.” *The Journal of Human Resources* 48 (1): 236–264. <https://doi.org/doi:10.1353/jhr.2013.0002>.
- Gortázar, L., D. Martínez de Lafuente, and A. Vega Bayo. 2022. “Comparing teacher and external assessments: Are boys, immigrants, and poorer students undergraded?” *Teaching and Teacher Education* 115. <https://doi.org/10.1016/j.tate.2022.103725>.
- Guez, A., H. Peyre, and F. Ramus. 2020. “Sex differences in academic achievement are modulated by evaluation type.” *Learning and Individual Differences* 83-84. <https://doi.org/10.1016/j.lindif.2020.101935>.
- Hanna, R. N., and L. L. Linden. 2012. “Discrimination in Grading.” *American Economic Journal: Economic Policy* 4 (4): 146–168. <https://doi.org/10.1257/pol.4.4.146>.
- Krkovic, K., S. Greiff, S. Kupiainen, M.-P. Vainikainen, and J. Hautamäki. 2014. “Teacher evaluation of student ability: what roles do teacher gender, student gender, and their interaction play?” *Educational Research* 56 (2): 244–257. <https://doi.org/10.1080/00131881.2014.898909>.
- Lavy, V. 2008. “Do gender stereotypes reduce girls’ or boys’ human capital outcomes? Evidence from a natural experiment.” *Journal of Public Economics* 92 (10-11): 2083–2105. <https://doi.org/10.1016/j.jpubeco.2008.02.009>.
- Lindahl, E. 2007. “Comparing teachers’ assessments and national test results: Evidence from Sweden.” *Working Paper* 2007 (24). <http://hdl.handle.net/10419/78672>.

- Lindahl, E. 2016. “Are teacher assessments biased? – evidence from Sweden.” *Education Economics* 24 (2): 224–238. <https://doi.org/10.1080/09645292.2015.1014882>.
- Marcenaro Gutiérrez, O. D., and L. A. López Agudo. 2016. “Testing the impact on educational achievement of expectations.” *XXIII Encuentro de Economía Pública*, <http://hdl.handle.net/10630/11005>.
- Marcenaro Gutiérrez, O. D., L. A. López Agudo, and M. Á. Ropero García. 2017. “Gender Differences in Adolescents’ Academic Achievement.” *YOUNG* 26 (3): 250–270. <https://doi.org/10.1177/1103308817715163>.
- Marcenaro Gutiérrez, O. D., and A. Vignoles. 2015. “A comparison of teacher and test-based assessment for Spanish primary and secondary students.” *Educational Research* 57 (1): 1–21. <https://doi.org/10.1080/00131881.2014.983720>.
- Martínez, J. F., B. Stecher, and H. Borko. 2009. “Classroom Assessment Practices, Teacher Judgments, and Student Achievement in Mathematics: Evidence from the ECLS.” *Educational Assessment* 14 (2): 78–102. <https://doi.org/10.1080/10627190903039429>.
- Orr, A. J. 2003. “Black-White Differences in Achievement: The Importance of Wealth.” *Sociology of Education* 76 (4): 281–304. <https://doi.org/10.2307/1519867>.
- Ribeiro, L. 2021. “The socioeconomic gap between teacher and test-based assessments.” Master’s thesis, Nova School of Business and Economics. <http://hdl.handle.net/10362/133144>.
- Triventi, M. 2020. “Are Children of Immigrants Graded Less Generously by their Teachers than Natives, and Why? Evidence from Student Population Data in Italy.” *International Migration Review* 54 (3): 765–795. <https://doi.org/10.1177/0197918319878104>.

van Bergen, E., T. van Zuijen, D. Bishop, and P. F. de Jong. 2017. “Why Are Home Literacy Environment and Children’s Reading Skills Associated? What Parental Skills Reveal.” *Reading Research Quarterly* 52 (2): 147–160. <https://doi.org/10.1002/rrq.160>.

Appendices

A. Biennial and triennial subjects

The biennial and triennial subjects of each course are presented in the following list. Note that students will only choose two of the biennial subjects.

❖ Sciences:

- Biennials: Physics and Chemistry; Biology and Geology; Descriptive Geometry.
- Triennial: Mathematics A

❖ Economics:

- Biennials: Economics; Geography; History B.
- Triennial: Mathematics A.

❖ Humanities:

- Biennials: Geography; Latin; Portuguese Literature; Mathematics Applied to Social Sciences; French/Spanish/German.
- Triennial: History A.

❖ Arts:

- Biennials: Descriptive Geometry; Mathematics B; History of Culture and Arts.
- Triennial: Drawing.

We do not include the subjects German, History B, and Latin in our study since we adopted a criterium of a minimum of 5,000 observations per subject.

B. Statistics using only the level of mothers' education

Table 8: Means and estimated standard errors of teachers' and exams' scores for three levels of mothers' education (2009-2018).

Subject	Obs.	Education level	% of students	Mean teacher score	Mean exam score	Mean difference
Biology and Geology	189,526	1	45.09%	13.23 (0.01)	9.22 (0.01)	4.01 (0.01)
		2	26.19%	13.59 (0.01)	9.69 (0.02)	3.90 (0.01)
		3	28.72%	14.70 (0.01)	11.19 (0.02)	3.50 (0.01)
Descriptive Geometry	35,412	1	40.76%	13.78 (0.02)	9.00 (0.04)	4.79 (0.03)
		2	25.23%	14.43 (0.03)	10.44 (0.06)	3.99 (0.04)
		3	34.01%	15.72 (0.03)	13.11 (0.05)	2.61 (0.04)
Drawing	24,793	1	47.30%	14.91 (0.02)	12.40 (0.03)	2.51 (0.03)
		2	25.72%	15.10 (0.03)	12.79 (0.04)	2.31 (0.03)
		3	26.98%	15.60 (0.03)	13.29 (0.03)	2.31 (0.03)
Economics	33,679	1	38.98%	13.76 (0.02)	11.16 (0.03)	2.61 (0.03)
		2	26.92%	13.97 (0.03)	11.32 (0.04)	2.65 (0.03)
		3	34.10%	14.79 (0.02)	12.62 (0.04)	2.17 (0.03)
French	6,613	1	59.47%	12.97 (0.04)	11.63 (0.05)	1.34 (0.04)
		2	26.13%	13.37 (0.06)	12.19 (0.08)	1.18 (0.06)
		3	14.40%	14.32 (0.08)	13.45 (0.11)	0.87 (0.08)
Geography	105,179	1	54.45%	12.94 (0.01)	10.60 (0.01)	2.33 (0.01)
		2	25.66%	13.18 (0.01)	11.10 (0.02)	2.08 (0.01)
		3	19.89%	14.01 (0.02)	12.14 (0.02)	1.87 (0.02)
History	80,370	1	58.81%	12.78 (0.01)	10.19 (0.02)	2.59 (0.01)
		2	24.89%	13.11 (0.02)	10.69 (0.03)	2.42 (0.02)
		3	16.30%	14.03 (0.02)	11.93 (0.03)	2.10 (0.03)
History of Culture and Arts	16,698	1	50.62%	12.81 (0.02)	9.56 (0.04)	3.26 (0.03)
		2	26.51%	13.17 (0.04)	9.99 (0.05)	3.18 (0.05)
		3	22.88%	14.26 (0.04)	11.40 (0.06)	2.86 (0.05)
Mathematics A	199,162	1	42.87%	12.85 (0.01)	9.88 (0.02)	2.97 (0.01)
		2	25.60%	13.28 (0.01)	10.59 (0.02)	2.70 (0.01)
		3	31.53%	14.56 (0.01)	12.51 (0.02)	2.05 (0.01)
Mathematics Applied to Social Sciences	48,402	1	58.84%	13.11 (0.01)	10.35 (0.02)	2.75 (0.02)
		2	24.55%	13.47 (0.02)	10.97 (0.04)	2.50 (0.03)
		3	16.60%	14.27 (0.03)	12.27 (0.04)	2.00 (0.03)
Mathematics B	8,359	1	42.95%	12.64 (0.04)	9.86 (0.07)	2.79 (0.05)
		2	24.82%	13.08 (0.06)	10.90 (0.09)	2.19 (0.07)
		3	32.23%	14.19 (0.05)	12.72 (0.08)	1.47 (0.06)
Philosophy	40,274	1	48.04%	13.44 (0.02)	9.74 (0.03)	3.70 (0.02)
		2	28.68%	13.72 (0.02)	10.43 (0.03)	3.29 (0.03)
		3	23.28%	14.67 (0.03)	12.12 (0.04)	2.55 (0.03)

Continued on next page

Subject	Obs.	Education level	% of students	Mean teacher score	Mean exam score	Mean difference
Physics and Chemistry	186,325	1	43.46%	12.76 (0.01)	8.28 (0.01)	4.48 (0.01)
		2	25.87%	13.21 (0.01)	9.05 (0.02)	4.15 (0.01)
		3	30.67%	14.48 (0.01)	11.00 (0.02)	3.47 (0.01)
Portuguese	334,820	1	47.92%	12.99 (0.01)	10.30 (0.01)	2.70 (0.01)
		2	25.44%	13.34 (0.01)	10.79 (0.01)	2.55 (0.01)
		3	26.64%	14.35 (0.01)	11.87 (0.01)	2.48 (0.01)
Portuguese Literature	11,590	1	61.07%	12.90 (0.03)	10.58 (0.04)	2.32 (0.04)
		2	23.48%	13.29 (0.04)	11.15 (0.07)	2.15 (0.06)
		3	15.45%	14.36 (0.06)	12.46 (0.08)	1.90 (0.07)
Spanish	12,885	1	61.23%	15.04 (0.02)	13.38 (0.04)	1.66 (0.03)
		2	24.59%	15.07 (0.04)	13.48 (0.06)	1.58 (0.05)
		3	14.19%	15.61 (0.05)	14.30 (0.07)	1.31 (0.06)

Note: Standard errors presented in parentheses are corrected for clustering at the school level.

C. Analysis for three levels of parents' education

Table 9: Linear regressions for three levels of parents' education (2009-2018).

Subject	Observations	Education level	OLS ^a
Biology and Geology	193,877	2	-0.121*** (0.023)
		3	-0.484*** (0.035)
Descriptive Geometry	36,472	2	-0.492*** (0.089)
		3	-1.636*** (0.111)
Drawing	25,536	2	-0.160** (0.048)
		3	-0.231*** (0.065)
Economics	34,804	2	-0.098 (0.055)
		3	-0.509*** (0.084)
French	6,807	2	-0.149 (0.093)
		3	-0.390** (0.122)
Geography	108,209	2	-0.148*** (0.029)
		3	-0.292*** (0.045)
History	82,533	2	-0.315*** (0.039)
		3	-0.534*** (0.056)
History of Culture and Arts	17,241	2	-0.076 (0.071)
		3	-0.433*** (0.092)
Mathematics A	203,682	2	-0.263*** (0.039)
		3	-0.918*** (0.061)
Mathematics Applied to Social Sciences	49,677	2	-0.261*** (0.048)
		3	-0.705*** (0.067)
Mathematics B	8,607	2	-0.496*** (0.082)
		3	-1.176*** (0.099)
Philosophy	41,410	2	-0.347*** (0.053)
		3	-0.978*** (0.074)
Physics and Chemistry	190,597	2	-0.227*** (0.031)
		3	-0.846*** (0.051)
Portuguese	343,088	2	-0.119*** (0.021)
		3	-0.209*** (0.032)
Portuguese Literature	11,912	2	-0.247* (0.103)
		3	-0.361** (0.121)
Spanish	13,300	2	-0.137* (0.060)
		3	-0.313*** (0.082)

^a Dependent variable is the assessment gap, *TeacherScore-ExamScore*. Standard errors in parentheses are corrected for clustering at the school level. Only the coefficient of interest, β , is presented. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.

Table 10: Linear regressions by exam score level with a scale 1-5 and by gender for three levels of parents' education (2009-2018).

Subject	Gender	Obs.	Education level	Exam score level ^a				
				1	2	3	4	5
Biology and Geology	All	193,877	2	-0.059 (0.062)	0.010 (0.021)	0.044 (0.024)	0.009 (0.029)	0.005 (0.072)
			3	0.171* (0.085)	0.078** (0.028)	0.253*** (0.033)	0.224*** (0.036)	0.144* (0.063)
	Girls	54.78%	2	-0.148 (0.077)	0.038 (0.024)	0.088** (0.028)	0.019 (0.035)	-0.056 (0.097)
			3	0.195 (0.133)	0.108** (0.033)	0.308*** (0.036)	0.239*** (0.041)	0.116 (0.083)
	Boys	45.22%	2	0.089 (0.107)	-0.025 (0.026)	-0.013 (0.030)	-0.004 (0.041)	0.073 (0.106)
			3	0.170 (0.113)	0.040 (0.031)	0.181*** (0.035)	0.201*** (0.042)	0.177* (0.084)
Descriptive Geometry	All	36,472	2	0.004 (0.095)	-0.126* (0.062)	0.024 (0.084)	-0.095 (0.094)	0.049 (0.080)
			3	0.194 (0.105)	-0.092 (0.075)	0.260** (0.091)	0.073 (0.082)	0.209* (0.083)
	Girls	48.45%	2	0.135 (0.111)	-0.226** (0.084)	0.044 (0.105)	-0.032 (0.138)	0.320* (0.132)
			3	0.116 (0.122)	-0.068 (0.095)	0.338** (0.120)	0.249 (0.128)	0.406** (0.119)
	Boys	51.55%	2	-0.147 (0.133)	-0.014 (0.083)	0.002 (0.113)	-0.157 (0.118)	-0.129 (0.101)
			3	0.323 (0.173)	-0.116 (0.095)	0.183 (0.110)	-0.081 (0.106)	0.079 (0.094)
Drawing	All	25,536	2	-0.701 (0.734)	-0.277** (0.096)	-0.009 (0.049)	0.054 (0.056)	-0.119 (0.122)
			3	-1.135 (1.011)	-0.051 (0.111)	0.140* (0.057)	0.297*** (0.067)	0.315* (0.127)
	Girls	66.97%	2	0.286 (1.284)	-0.274* (0.134)	0.032 (0.055)	0.043 (0.060)	0.066 (0.142)
			3	1.042 (2.879)	0.031 (0.132)	0.243*** (0.067)	0.372*** (0.073)	0.384* (0.149)
	Boys	33.03%	2	-0.239 (1.431)	-0.327* (0.143)	-0.081 (0.082)	0.084 (0.107)	-0.534* (0.220)
			3	-2.145 (1.472)	-0.221 (0.169)	-0.060 (0.086)	0.117 (0.102)	0.165 (0.203)

Continued on next page

Subject	Gender	Obs.	Education level	Exam score level ^a				
				1	2	3	4	5
Economics	All	34,804	2	0.205 (0.272)	0.041 (0.062)	0.017 (0.054)	-0.063 (0.065)	0.096 (0.093)
			3	0.142 (0.398)	0.126 (0.077)	0.173* (0.081)	0.115 (0.076)	0.181* (0.080)
	Girls	47.88%	2	0.256 (0.421)	0.006 (0.084)	0.051 (0.074)	-0.071 (0.088)	0.042 (0.118)
			3	0.535 (0.570)	0.121 (0.109)	0.219* (0.098)	0.166 (0.089)	0.228* (0.099)
	Boys	52.12%	2	-0.191 (0.348)	0.074 (0.072)	-0.022 (0.064)	-0.052 (0.085)	0.123 (0.130)
			3	-0.436 (0.544)	0.136 (0.082)	0.132 (0.089)	0.062 (0.086)	0.121 (0.107)
French	All	6,807	2	0.412 (0.586)	0.083 (0.121)	0.114 (0.086)	0.053 (0.105)	-0.045 (0.213)
			3	0.214 (1.806)	0.019 (0.179)	0.189 (0.112)	0.456** (0.132)	0.248 (0.183)
	Girls	71.25%	2	0.439 (0.682)	0.003 (0.148)	0.184 (0.108)	0.042 (0.127)	-0.086 (0.235)
			3	-1.601 (0.932)	-0.002 (0.197)	0.204 (0.143)	0.492** (0.142)	0.190 (0.175)
	Boys	28.75%	2	—	0.258 (0.186)	-0.042 (0.137)	0.084 (0.195)	0.864 (0.721)
			3	—	0.079 (0.309)	0.141 (0.147)	0.363 (0.252)	0.762 (0.666)
Geography	All	108,209	2	0.343 (0.271)	-0.052 (0.030)	-0.034 (0.028)	0.036 (0.039)	0.046 (0.121)
			3	2.279* (0.868)	-0.019 (0.044)	0.105* (0.043)	0.303*** (0.051)	0.207 (0.110)
	Girls	61.62%	2	0.457 (0.341)	-0.046 (0.036)	-0.009 (0.034)	0.063 (0.046)	-0.130 (0.157)
			3	2.493 (1.522)	-0.030 (0.053)	0.154** (0.050)	0.347*** (0.055)	0.186 (0.142)
	Boys	38.38%	2	-0.097 (0.484)	-0.060 (0.041)	-0.073* (0.032)	-0.001 (0.057)	0.260 (0.187)
			3	2.117*** (0.534)	0.004 (0.056)	0.046 (0.046)	0.245*** (0.067)	0.248 (0.162)
History	All	82,533	2	0.037 (0.093)	-0.065 (0.034)	-0.007 (0.029)	0.023 (0.046)	0.269** (0.100)
			3	0.022 (0.155)	0.070 (0.051)	0.296*** (0.044)	0.491*** (0.060)	0.502*** (0.097)
	Girls	70.24%	2	0.133 (0.109)	-0.077 (0.039)	0.000 (0.035)	0.036 (0.052)	0.225* (0.114)
			3	-0.074 (0.177)	0.094 (0.056)	0.313*** (0.051)	0.490*** (0.065)	0.496*** (0.109)
	Boys	29.76%	2	-0.216 (0.177)	-0.039 (0.047)	-0.033 (0.047)	-0.016 (0.078)	0.395* (0.173)
			3	0.397 (0.390)	0.015 (0.075)	0.255*** (0.058)	0.485*** (0.081)	0.535*** (0.164)

Continued on next page

Subject	Gender	Obs.	Education level	Exam score level ^a					
				1	2	3	4	5	
History of Culture and Arts	All	17,241	2	0.487*	0.011	0.141	0.269*	0.093	
				(0.222)	(0.062)	(0.075)	(0.123)	(0.297)	
				3	0.275	0.081	0.552***	0.774***	0.568*
					(0.310)	(0.084)	(0.090)	(0.133)	(0.256)
	Girls	68.36%		2	0.491	0.050	0.175	0.174	0.292
					(0.314)	(0.073)	(0.092)	(0.134)	(0.322)
			3	0.164	0.130	0.638***	0.795***	0.566*	
				(0.382)	(0.100)	(0.106)	(0.136)	(0.283)	
Boys	31.64%		2	0.669*	-0.062	0.073	0.568*	-0.621	
				(0.330)	(0.099)	(0.102)	(0.252)	(0.625)	
			3	0.293	-0.025	0.355**	0.717**	0.659	
				(0.480)	(0.126)	(0.119)	(0.219)	(0.479)	
Mathematics A	All	203,682	2	0.062*	-0.035	0.040	0.076*	-0.024	
				(0.031)	(0.026)	(0.027)	(0.034)	(0.044)	
				3	0.154**	-0.038	0.225***	0.304***	0.247***
					(0.048)	(0.034)	(0.038)	(0.048)	(0.049)
	Girls	53.36%		2	0.029	-0.034	0.092**	0.083*	0.009
					(0.048)	(0.031)	(0.034)	(0.042)	(0.050)
			3	0.246***	-0.004	0.298***	0.331***	0.282***	
				(0.067)	(0.042)	(0.044)	(0.055)	(0.051)	
Boys	46.64%		2	0.088*	-0.045	-0.037	0.042	-0.073	
				(0.042)	(0.030)	(0.033)	(0.043)	(0.062)	
			3	0.094	-0.076*	0.131**	0.253***	0.196**	
				(0.054)	(0.035)	(0.042)	(0.051)	(0.063)	
Mathematics Applied to Social Sciences	All	49,677	2	0.076	-0.028	0.027	-0.009	0.119	
				(0.119)	(0.042)	(0.041)	(0.051)	(0.095)	
				3	0.472**	0.014	0.139*	0.179**	0.407***
					(0.147)	(0.060)	(0.059)	(0.058)	(0.089)
	Girls	69.75%		2	-0.031	0.020	0.056	0.015	0.116
					(0.142)	(0.051)	(0.048)	(0.058)	(0.113)
			3	0.474*	0.059	0.179*	0.216**	0.437***	
				(0.209)	(0.071)	(0.073)	(0.067)	(0.096)	
Boys	30.25%		2	0.287	-0.130	-0.041	-0.072	0.133	
				(0.197)	(0.067)	(0.066)	(0.093)	(0.179)	
			3	0.361	-0.083	0.051	0.077	0.345*	
				(0.242)	(0.080)	(0.077)	(0.093)	(0.174)	
Mathematics B	All	8,607	2	0.145	-0.168	-0.166	0.094	-0.064	
				(0.229)	(0.087)	(0.103)	(0.134)	(0.232)	
				3	0.631**	-0.079	-0.019	0.256*	-0.051
					(0.235)	(0.110)	(0.102)	(0.125)	(0.159)
	Girls	65.05%		2	0.167	-0.078	0.037	0.205	-0.129
					(0.265)	(0.116)	(0.132)	(0.172)	(0.285)
			3	0.937**	-0.001	0.058	0.409**	0.082	
				(0.318)	(0.126)	(0.127)	(0.150)	(0.199)	
Boys	34.95%		2	0.091	-0.300*	-0.480**	-0.073	0.054	
				(0.496)	(0.149)	(0.153)	(0.235)	(0.407)	
			3	0.122	-0.189	-0.130	-0.028	-0.470	
				(0.423)	(0.182)	(0.148)	(0.210)	(0.341)	

Continued on next page

Subject	Gender	Obs.	Education level	Exam score level ^a				
				1	2	3	4	5
Philosophy	All	41,410	2	-0.116 (0.138)	-0.140** (0.047)	0.048 (0.043)	0.052 (0.060)	0.118 (0.124)
			3	0.090 (0.176)	-0.072 (0.058)	0.228*** (0.056)	0.260*** (0.066)	0.447*** (0.105)
	Girls	64.99%	2	-0.152 (0.191)	-0.141* (0.056)	0.048 (0.049)	0.074 (0.065)	0.127 (0.141)
			3	0.019 (0.256)	0.014 (0.068)	0.255*** (0.062)	0.289*** (0.070)	0.466*** (0.123)
	Boys	35.01%	2	-0.113 (0.202)	-0.146* (0.063)	0.051 (0.067)	-0.011 (0.113)	0.024 (0.284)
			3	0.170 (0.229)	-0.175* (0.080)	0.185* (0.078)	0.176 (0.110)	0.361 (0.229)
Physics and Chemistry	All	190,597	2	0.078* (0.034)	-0.066** (0.023)	-0.035 (0.030)	-0.075 (0.039)	-0.028 (0.057)
			3	0.147** (0.044)	-0.068* (0.033)	0.086* (0.043)	0.032 (0.047)	0.029 (0.052)
	Girls	51.23%	2	0.094 (0.051)	-0.063* (0.028)	0.013 (0.037)	-0.053 (0.045)	0.015 (0.074)
			3	0.190** (0.073)	-0.035 (0.040)	0.172*** (0.048)	0.051 (0.055)	0.004 (0.065)
	Boys	48.77%	2	0.052 (0.046)	-0.073** (0.024)	-0.102** (0.037)	-0.102 (0.054)	-0.067 (0.085)
			3	0.102 (0.057)	-0.100** (0.032)	-0.016 (0.046)	0.004 (0.053)	0.051 (0.070)
Portuguese	All	343,088	2	-0.015 (0.074)	0.019 (0.019)	0.094*** (0.019)	0.120*** (0.026)	-0.006 (0.054)
			3	0.262** (0.091)	0.190*** (0.027)	0.464*** (0.031)	0.596*** (0.034)	0.561*** (0.055)
	Girls	58.09%	2	-0.039 (0.117)	0.045 (0.025)	0.133*** (0.022)	0.141*** (0.027)	0.010 (0.062)
			3	0.038 (0.170)	0.235*** (0.032)	0.502*** (0.033)	0.613*** (0.033)	0.587*** (0.063)
	Boys	41.91%	2	0.013 (0.091)	-0.004 (0.021)	0.030 (0.023)	0.069 (0.042)	-0.058 (0.120)
			3	0.392*** (0.109)	0.160*** (0.030)	0.409*** (0.034)	0.554*** (0.049)	0.473*** (0.098)
Portuguese Literature	All	11,912	2	-0.135 (0.362)	0.025 (0.100)	0.046 (0.081)	0.128 (0.109)	-0.377 (0.333)
			3	0.325 (0.462)	0.220 (0.144)	0.531*** (0.097)	0.732*** (0.125)	0.795** (0.256)
	Girls	75.09%	2	-0.106 (0.480)	-0.007 (0.117)	0.065 (0.095)	0.123 (0.117)	-0.351 (0.342)
			3	0.164 (0.655)	0.218 (0.183)	0.521*** (0.113)	0.747*** (0.132)	0.851** (0.261)
	Boys	24.91%	2	-0.159 (0.400)	0.056 (0.135)	0.012 (0.123)	0.141 (0.259)	-0.910 (0.763)
			3	0.139 (0.516)	0.230 (0.200)	0.578** (0.178)	0.658* (0.267)	0.017 (0.493)

Continued on next page

Subject	Gender	Obs.	Education level	Exam score level ^a				
				1	2	3	4	5
Spanish	All	13,300	2	—	-0.141 (0.134)	0.029 (0.073)	-0.148* (0.061)	0.185 (0.098)
			3	—	-0.233 (0.175)	-0.077 (0.094)	-0.013 (0.079)	0.418** (0.126)
	Girls	69.53%	2	—	-0.094 (0.178)	0.049 (0.091)	-0.136* (0.065)	0.205 (0.106)
			3	—	-0.172 (0.231)	-0.100 (0.108)	-0.073 (0.084)	0.570*** (0.128)
	Boys	30.47%	2	—	-0.209 (0.182)	0.008 (0.100)	-0.187 (0.113)	0.070 (0.247)
			3	—	-0.303 (0.265)	-0.029 (0.139)	0.118 (0.132)	-0.129 (0.265)

^a Dependent variable is the assessment gap, *TeacherScore-ExamScore*. Standard errors in parentheses are corrected for clustering at the school level. Only the coefficient of interest, β , is presented. The exam score level 1 corresponds to the interval of scores [0; 35[, level 2 to [35; 95[, level 3 to [95; 135[, level 4 to [135; 175[, and level 5 to [175; 200]. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.

D. Analysis for six levels of parents' education

We aggregate the 10 levels in 6 levels of education:

- **Level 1:** Basic education (from no education to 4th grade);
- **Level 2:** 6th grade;
- **Level 3:** 9th grade;
- **Level 4:** Secondary education (secondary);
- **Level 5:** Low higher education (post-graduation, bachelor's, or undergraduate);
- **Level 6:** High higher education (master or doctorate).

Table 11: Means and estimated standard errors of teachers' and exams' scores for six levels of parents' education (2009-2018).

Variable	Mean							Standard Deviation						
	Overall	1	2	3	4	5	6	Overall	1	2	3	4	5	6
Age	17.79	18.01	17.85	17.85	17.79	17.68	17.64	0.81	0.94	0.84	0.84	0.81	0.73	0.70
Girls	0.58	0.65	0.63	0.61	0.57	0.52	0.51	0.49	0.48	0.48	0.49	0.49	0.50	0.50
Born in Portugal	0.95	0.93	0.95	0.95	0.95	0.97	0.96	0.21	0.25	0.21	0.22	0.21	0.18	0.19
Social support	0.25	0.48	0.45	0.36	0.24	0.06	0.04	0.44	0.50	0.50	0.48	0.43	0.24	0.20
Internet at home	0.71	0.59	0.63	0.70	0.73	0.75	0.75	0.46	0.49	0.48	0.46	0.44	0.44	0.43
Computer at home	0.76	0.72	0.75	0.77	0.77	0.76	0.76	0.43	0.45	0.44	0.42	0.42	0.43	0.43
Improving exam grade	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.07	0.08	0.07	0.07	0.07	0.06	0.05
Applying to university	1.00	0.99	0.99	0.99	1.00	1.00	1.00	0.07	0.11	0.08	0.07	0.06	0.04	0.04
Area of studies:														
Sciences	0.58	0.50	0.54	0.54	0.58	0.65	0.67	0.49	0.50	0.50	0.50	0.49	0.48	0.47
Economics	0.09	0.06	0.06	0.07	0.09	0.11	0.12	0.28	0.24	0.24	0.26	0.28	0.32	0.32
Humanities	0.26	0.37	0.33	0.31	0.26	0.16	0.14	0.44	0.48	0.47	0.46	0.44	0.37	0.35
Arts	0.07	0.07	0.07	0.08	0.07	0.07	0.08	0.26	0.26	0.25	0.26	0.26	0.26	0.27
Number of observations	1,367,753	6.72%	14.99%	19.50%	27.73%	27.09%	3.97%							

Table 12: Means and estimated standard errors of teachers' and exams' scores for six levels of parents' education (2009-2018).

Subject	Obs.	Education level	% of students	Mean teacher score	Mean exam score	Mean difference
Biology and Geology	193,877	1	5.78%	13.03 (0.02)	8.91 (0.03)	4.12 (0.02)
		2	14.32%	13.20 (0.01)	9.15 (0.02)	4.05 (0.01)
		3	18.72%	13.23 (0.01)	9.24 (0.02)	3.99 (0.01)
		4	28.11%	13.50 (0.01)	9.58 (0.01)	3.92 (0.01)
		5	28.91%	14.53 (0.01)	10.97 (0.02)	3.55 (0.01)
		6	4.17%	15.09 (0.03)	11.71 (0.04)	3.37 (0.03)
Descriptive Geometry	36,472	1	5.11%	13.47 (0.06)	8.21 (0.12)	5.26 (0.09)
		2	11.91%	13.74 (0.04)	8.68 (0.08)	5.06 (0.06)
		3	17.23%	13.83 (0.04)	9.17 (0.07)	4.66 (0.05)
		4	26.46%	14.23 (0.03)	10.08 (0.06)	4.15 (0.04)
		5	33.62%	15.53 (0.03)	12.70 (0.05)	2.82 (0.04)
		6	5.67%	16.04 (0.07)	13.89 (0.12)	2.15 (0.09)
Drawing	25,536	1	6.86%	14.84 (0.05)	12.15 (0.07)	2.69 (0.06)
		2	14.02%	14.93 (0.03)	12.39 (0.05)	2.55 (0.05)
		3	19.82%	14.88 (0.03)	12.41 (0.04)	2.47 (0.04)
		4	27.46%	15.05 (0.03)	12.73 (0.03)	2.33 (0.03)
		5	27.75%	15.50 (0.03)	13.19 (0.03)	2.31 (0.03)
		6	4.10%	15.78 (0.07)	13.52 (0.09)	2.26 (0.08)
Economics	34,804	1	5.15%	13.62 (0.06)	11.30 (0.08)	2.32 (0.07)
		2	11.00%	13.73 (0.04)	11.14 (0.06)	2.59 (0.05)
		3	16.88%	13.78 (0.03)	11.01 (0.05)	2.76 (0.04)
		4	27.50%	13.87 (0.03)	11.21 (0.04)	2.67 (0.03)
		5	34.45%	14.66 (0.02)	12.45 (0.03)	2.21 (0.03)
		6	5.03%	15.08 (0.06)	13.04 (0.09)	2.04 (0.07)
French	6,807	1	10.37%	12.81 (0.08)	11.34 (0.13)	1.47 (0.10)
		2	17.73%	12.89 (0.06)	11.50 (0.10)	1.39 (0.08)
		3	23.96%	13.03 (0.05)	11.78 (0.08)	1.25 (0.07)
		4	29.65%	13.30 (0.05)	12.08 (0.07)	1.22 (0.06)
		5	16.07%	14.10 (0.07)	13.21 (0.10)	0.89 (0.07)
		6	2.22%	14.36 (0.22)	13.57 (0.29)	0.80 (0.19)
Geography	108,209	1	8.49%	12.78 (0.02)	10.30 (0.03)	2.48 (0.03)
		2	17.29%	12.94 (0.02)	10.56 (0.02)	2.38 (0.02)
		3	22.26%	12.94 (0.01)	10.66 (0.02)	2.28 (0.02)
		4	28.28%	13.12 (0.01)	11.00 (0.02)	2.12 (0.01)
		5	20.99%	13.86 (0.02)	11.98 (0.02)	1.88 (0.02)
		6	2.69%	14.23 (0.05)	12.42 (0.05)	1.81 (0.04)
History	82,533	1	9.89%	12.61 (0.02)	10.00 (0.04)	2.61 (0.03)
		2	19.25%	12.75 (0.02)	10.05 (0.03)	2.69 (0.02)
		3	23.13%	12.82 (0.02)	10.26 (0.03)	2.57 (0.02)
		4	27.95%	13.03 (0.02)	10.61 (0.02)	2.42 (0.02)
		5	17.57%	13.85 (0.02)	11.73 (0.03)	2.12 (0.02)
		6	2.20%	14.38 (0.06)	12.25 (0.09)	2.13 (0.07)
History of Culture and Arts	17,241	1	7.13%	12.70 (0.06)	9.55 (0.09)	3.15 (0.09)
		2	15.18%	12.80 (0.04)	9.48 (0.07)	3.32 (0.06)
		3	21.64%	12.79 (0.04)	9.55 (0.06)	3.24 (0.05)
		4	28.89%	13.09 (0.03)	9.86 (0.05)	3.23 (0.04)
		5	23.53%	14.04 (0.04)	11.15 (0.06)	2.89 (0.05)
		6	3.63%	14.64 (0.11)	11.95 (0.15)	2.69 (0.11)

Continued on next page

Subject	Obs.	Education level	% of students	Mean teacher score	Mean exam score	Mean difference
Mathematics A	203,682	1	5.57%	12.59 (0.02)	9.48 (0.04)	3.12 (0.03)
		2	13.41%	12.82 (0.02)	9.76 (0.03)	3.06 (0.02)
		3	17.50%	12.85 (0.01)	9.90 (0.02)	2.95 (0.02)
		4	27.23%	13.17 (0.01)	10.40 (0.02)	2.77 (0.01)
		5	31.46%	14.37 (0.01)	12.25 (0.02)	2.12 (0.01)
		6	4.83%	14.95 (0.03)	13.16 (0.05)	1.79 (0.03)
Mathematics Applied to Social Sciences	49,677	1	9.54%	12.86 (0.03)	9.95 (0.06)	2.91 (0.05)
		2	19.36%	13.10 (0.02)	10.36 (0.04)	2.74 (0.03)
		3	23.61%	13.17 (0.02)	10.40 (0.04)	2.77 (0.03)
		4	27.56%	13.39 (0.02)	10.84 (0.03)	2.55 (0.03)
		5	17.81%	14.12 (0.03)	12.04 (0.04)	2.08 (0.03)
		6	2.12%	14.54 (0.09)	12.86 (0.13)	1.68 (0.09)
Mathematics B	8,607	1	5.81%	12.41 (0.10)	9.24 (0.18)	3.16 (0.15)
		2	12.78%	12.61 (0.07)	9.60 (0.12)	3.01 (0.10)
		3	18.00%	12.64 (0.06)	9.95 (0.11)	2.68 (0.08)
		4	26.14%	12.96 (0.05)	10.67 (0.09)	2.29 (0.07)
		5	32.47%	14.03 (0.05)	12.53 (0.08)	1.50 (0.06)
		6	4.80%	14.50 (0.14)	13.06 (0.20)	1.45 (0.15)
Philosophy	41,410	1	5.81%	13.35 (0.04)	9.26 (0.08)	4.09 (0.07)
		2	14.93%	13.44 (0.03)	9.61 (0.05)	3.83 (0.04)
		3	21.24%	13.43 (0.02)	9.83 (0.04)	3.60 (0.03)
		4	30.84%	13.64 (0.02)	10.30 (0.03)	3.34 (0.03)
		5	23.58%	14.51 (0.02)	11.85 (0.04)	2.66 (0.03)
		6	3.59%	14.95 (0.06)	12.62 (0.10)	2.32 (0.07)
Physics and Chemistry	190,597	1	5.56%	12.46 (0.02)	7.74 (0.03)	4.72 (0.02)
		2	13.69%	12.72 (0.02)	8.15 (0.02)	4.57 (0.02)
		3	17.93%	12.79 (0.01)	8.38 (0.02)	4.42 (0.01)
		4	27.56%	13.09 (0.01)	8.87 (0.02)	4.21 (0.01)
		5	30.66%	14.27 (0.01)	10.70 (0.02)	3.57 (0.01)
		6	4.60%	14.93 (0.03)	11.72 (0.04)	3.21 (0.03)
Portuguese	343,089	1	6.98%	12.74 (0.01)	9.93 (0.02)	2.80 (0.02)
		2	15.05%	12.97 (0.01)	10.23 (0.01)	2.73 (0.01)
		3	19.42%	13.00 (0.01)	10.34 (0.01)	2.66 (0.01)
		4	27.46%	13.26 (0.01)	10.69 (0.01)	2.57 (0.01)
		5	27.07%	14.19 (0.01)	11.70 (0.01)	2.49 (0.01)
		6	4.02%	14.62 (0.02)	12.17 (0.03)	2.44 (0.02)
Portuguese Literature	11,912	1	10.23%	12.81 (0.06)	10.34 (0.10)	2.47 (0.09)
		2	20.67%	12.89 (0.04)	10.46 (0.07)	2.43 (0.06)
		3	23.85%	12.95 (0.04)	10.72 (0.06)	2.22 (0.06)
		4	26.80%	13.14 (0.04)	11.02 (0.06)	2.12 (0.05)
		5	16.12%	14.20 (0.06)	12.18 (0.08)	2.03 (0.07)
		6	2.33%	14.56 (0.15)	12.74 (0.20)	1.82 (0.16)
Spanish	13,300	1	9.50%	14.81 (0.06)	13.24 (0.09)	1.57 (0.08)
		2	20.26%	15.16 (0.04)	13.37 (0.06)	1.79 (0.05)
		3	24.45%	15.01 (0.04)	13.35 (0.05)	1.65 (0.05)
		4	28.45%	15.06 (0.04)	13.49 (0.05)	1.58 (0.04)
		5	15.68%	15.48 (0.05)	14.21 (0.07)	1.27 (0.06)
		6	1.65%	15.71 (0.15)	13.86 (0.21)	1.85 (0.18)

Note: Standard errors presented in parentheses are corrected for clustering at the school level.

E. Analysis using social support as proxy

We note that, for this analysis using social support as proxy, the signs of the results have to be read in the opposite sense as our prior analysis with the level of parents' education, since now for eq. (1) we have $E = 1$ for low-SES students, and $E = 0$ for high-SES students.

Table 13: Linear regressions using social support (2009-2018).

Subject	Observations	β^a
Biology and Geology	193,877	0.237*** (0.025)
Descriptive Geometry	36,472	0.974*** (0.083)
Drawing	25,536	0.221*** (0.056)
Economics	34,804	0.178** (0.058)
French	6,807	0.260** (0.085)
Geography	108,209	0.154*** (0.027)
History	82,533	0.266*** (0.038)
History of Culture and Arts	17,241	0.107 (0.068)
Mathematics A	203,682	0.552*** (0.043)
Mathematics Applied to Social Sciences	49,677	0.244*** (0.046)
Mathematics B	8,607	0.734*** (0.102)
Philosophy	41,410	0.419*** (0.048)
Physics and Chemistry	190,597	0.466*** (0.033)
Portuguese	343,088	0.123*** (0.023)
Portuguese Literature	11,912	0.139 (0.096)
Spanish	13,300	0.167** (0.052)

^a Dependent variable is the assessment gap, *TeacherScore-ExamScore*. Standard errors in parentheses are corrected for clustering at the school level. Only the coefficient of interest, β , is presented. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.

Table 14: Linear regressions by exam grade with a scale 1-5 and by student gender using social support (2009-2018).

Subject	Gender	Obs.	Exam score level ^a				
			1	2	3	4	5
Biology and Geology	All	193,877	-0.099 (0.057)	0.020 (0.021)	-0.044 (0.024)	-0.083** (0.030)	-0.118 (0.062)
	Girls	54.78%	-0.155* (0.076)	0.030 (0.025)	-0.070* (0.028)	-0.088* (0.035)	-0.060 (0.084)
	Boys	45.22%	0.013 (0.093)	0.007 (0.025)	-0.006 (0.029)	-0.075 (0.039)	-0.174* (0.086)
Descriptive Geometry	All	36,472	-0.041 (0.079)	0.224*** (0.058)	0.094 (0.078)	0.068 (0.081)	0.067 (0.079)
	Girls	48.45%	-0.109 (0.098)	0.181* (0.077)	0.007 (0.112)	-0.054 (0.118)	0.012 (0.129)
	Boys	51.55%	0.041 (0.123)	0.277*** (0.079)	0.181 (0.096)	0.172 (0.108)	0.087 (0.088)
Drawing	All	25,536	3.067*** (0.704)	-0.002 (0.101)	0.113* (0.051)	-0.024 (0.058)	0.133 (0.117)
	Girls	66.97%	2.356* (0.920)	-0.136 (0.133)	0.093 (0.058)	-0.078 (0.069)	0.079 (0.134)
	Boys	33.03%	3.680** (1.133)	0.277 (0.151)	0.148 (0.080)	0.091 (0.092)	0.251 (0.226)
Economics	All	34,804	-0.368 (0.253)	-0.099 (0.060)	-0.010 (0.054)	-0.026 (0.064)	-0.138 (0.076)
	Girls	47.88%	-0.160 (0.374)	-0.070 (0.081)	0.005 (0.068)	-0.054 (0.081)	-0.173 (0.091)
	Boys	52.12%	-0.529 (0.415)	-0.134 (0.071)	-0.021 (0.067)	0.001 (0.088)	-0.104 (0.118)
French	All	6,807	0.801 (0.518)	0.102 (0.117)	-0.023 (0.074)	-0.073 (0.111)	-0.178 (0.215)
	Girls	71.25%	0.812 (0.575)	0.123 (0.147)	0.023 (0.088)	-0.046 (0.129)	-0.303 (0.215)
	Boys	28.75%	2.195** (0.545)	0.119 (0.199)	-0.096 (0.118)	-0.174 (0.173)	0.568 (0.593)
Geography	All	108,209	-0.510* (0.225)	0.048 (0.028)	0.043 (0.026)	-0.110** (0.036)	-0.156 (0.117)
	Girls	61.62%	-0.547* (0.254)	0.077* (0.032)	0.046 (0.031)	-0.110* (0.044)	-0.144 (0.145)
	Boys	38.38%	-0.637 (0.440)	-0.025 (0.042)	0.034 (0.033)	-0.112* (0.056)	-0.181 (0.185)

Continued on next page

Subject	Gender	Obs.	Exam score level ^a				
			1	2	3	4	5
History of Culture and Arts	All	17,241	-0.011 (0.216)	-0.095 (0.062)	-0.288*** (0.071)	-0.373** (0.124)	-0.623* (0.262)
	Girls	68.36%	-0.129 (0.296)	-0.134 (0.071)	-0.351*** (0.084)	-0.475** (0.135)	-0.663* (0.309)
	Boys	31.64%	0.071 (0.313)	-0.023 (0.097)	-0.160 (0.105)	-0.039 (0.217)	-1.328* (0.608)
History	All	82,533	-0.082 (0.087)	0.070* (0.032)	0.039 (0.030)	-0.044 (0.041)	-0.176 (0.095)
	Girls	70.24%	-0.136 (0.104)	0.093** (0.035)	0.032 (0.035)	-0.021 (0.046)	-0.076 (0.097)
	Boys	29.76%	0.065 (0.190)	0.011 (0.050)	0.066 (0.044)	-0.098 (0.073)	-0.470** (0.177)
Mathematics Applied to Social Sciences	All	49,677	-0.088 (0.093)	0.019 (0.039)	0.004 (0.040)	0.007 (0.050)	-0.122 (0.085)
	Girls	69.75%	-0.083 (0.123)	0.050 (0.044)	-0.018 (0.048)	-0.009 (0.055)	-0.200* (0.093)
	Boys	30.25%	-0.103 (0.158)	-0.051 (0.062)	0.059 (0.063)	0.061 (0.093)	0.101 (0.195)
Mathematics A	All	203,682	-0.030 (0.036)	0.157*** (0.025)	0.036 (0.029)	-0.049 (0.036)	-0.070 (0.041)
	Girls	53.36%	-0.028 (0.051)	0.142*** (0.030)	-0.026 (0.033)	-0.082* (0.041)	-0.134** (0.046)
	Boys	46.64%	-0.039 (0.043)	0.178*** (0.032)	0.126** (0.038)	0.006 (0.047)	0.033 (0.057)
Mathematics B	All	8,607	-0.437* (0.174)	0.245** (0.083)	0.004 (0.098)	0.061 (0.137)	0.322 (0.172)
	Girls	65.05%	-0.335 (0.235)	0.196* (0.095)	-0.194 (0.127)	-0.077 (0.177)	0.013 (0.209)
	Boys	34.95%	-0.694* (0.340)	0.345* (0.146)	0.330* (0.146)	0.327 (0.195)	1.093*** (0.274)
Philosophy	All	41,410	0.072 (0.114)	0.083* (0.038)	-0.041 (0.038)	-0.006 (0.051)	-0.096 (0.097)
	Girls	64.99%	0.079 (0.168)	0.062 (0.045)	-0.029 (0.045)	-0.039 (0.060)	-0.168 (0.114)
	Boys	35.01%	0.072 (0.169)	0.122 (0.065)	-0.071 (0.067)	0.086 (0.097)	0.186 (0.223)
Physics and Chemistry	All	190,597	0.043 (0.033)	0.116*** (0.021)	0.045 (0.029)	0.033 (0.034)	-0.039 (0.050)
	Girls	51.23%	0.019 (0.043)	0.107*** (0.024)	0.014 (0.033)	0.019 (0.041)	-0.044 (0.065)
	Boys	48.77%	0.081 (0.049)	0.126*** (0.024)	0.087* (0.038)	0.053 (0.045)	-0.028 (0.070)

Continued on next page

Subject	Gender	Obs.	Exam score level ^a				
			1	2	3	4	5
Portuguese Literature	All	11,912	0.330 (0.250)	-0.119 (0.091)	-0.175** (0.064)	-0.390*** (0.108)	-0.330 (0.251)
	Girls	75.09%	0.109 (0.404)	-0.099 (0.105)	-0.206** (0.073)	-0.329** (0.113)	-0.493* (0.239)
	Boys	24.91%	0.436 (0.355)	-0.151 (0.144)	-0.070 (0.130)	-0.730** (0.264)	0.273 (0.672)
Portuguese	All	343,088	0.003 (0.066)	-0.026 (0.018)	-0.110*** (0.021)	-0.220*** (0.026)	-0.191** (0.056)
	Girls	58.09%	0.071 (0.097)	-0.024 (0.022)	-0.115*** (0.023)	-0.209*** (0.028)	-0.204** (0.063)
	Boys	41.91%	-0.049 (0.088)	-0.031 (0.022)	-0.101*** (0.024)	-0.250*** (0.040)	-0.116 (0.108)
Spanish	All	13,300	—	0.156 (0.093)	0.006 (0.063)	0.088 (0.052)	0.031 (0.117)
	Girls	69.53%	—	0.092 (0.129)	0.011 (0.078)	0.107 (0.061)	0.023 (0.111)
	Boys	30.47%	—	0.244 (0.147)	-0.009 (0.101)	0.037 (0.108)	0.143 (0.264)

^a Dependent variable is the assessment gap, *TeacherScore-ExamScore*. Only the coefficient of interest, β , is presented. Standard errors are presented in parentheses (corrected for clustering at the school level). The percentage of students in each exam score level is also presented. The exam score level 1 corresponds to the interval of scores [0; 35[, level 2 to [35; 95[, level 3 to [95; 135[, level 4 to [135; 175[, and level 5 to [175; 200]. *Significant at $p < .05$, **significant at $p < .01$, ***significant at $p < .001$.