Knowing who you are: The Effect of Feedback on Short and Long Term Outcomes[1]

Sofoklis Goulas
Department of Economics
UNC at Chapel Hill
goulas@unc.edu

Rigissa Megalokonomou
Department of Economics
University of Queensland
r.megalokonomou@uq.edu.au

Abstract

We study the effect of disclosing relative performance information (feedback) on students' performance in high-school, on subsequent university enrollment, and on expected subsequent earnings. We exploit a large-scale natural experiment in which students in some cohorts receive information about their relative performance within their schools and across the nation. Using unique primary data, we find an asymmetric response to feedback: high-achieving students improve their final-year performance by 0.15 of a standard deviation, whereas the final-year performance of low-achieving students drops by 0.3 of a standard deviation. The results are more pronounced for females, indicating greater sensitivity to feedback. We also document the long-term effects of feedback: high-achieving students reduce their repetition rate for the national exams; they enroll into university departments that are more selective by 0.15 of a standard deviation and their expected annual earnings increase by 0.17 of a standard deviation. By contrast, the results for low-achieving students are negative. We provide suggestive evidence that feedback encourages students from low-income neighborhoods to enroll in university and to study in higher-quality programs, which may, in the long run, reduce income inequality.

Keywords: feedback, relative performance, university admission, rank, gender differences, income inequality

[1]We are deeply indebted to Fabian Waldinger for his invaluable advice and continuous support. Special thanks to Sascha O. Becker, Dan Bernhard and Clement de Chaisemartin for their guidance in this project. We also thank Robert Akerlof, Manuel Bagues, David Card, Saraswata Chaudhuri, Mirko Draca, Caroline Hoxby, Lutz Hendricks, Kala Krishna, Victor Lavy, Robin Naylor and Roland Rathelot. We also thank Diana Egerton-Warburton for very helpful work and also participants of the Warwick AMES, 30th Annual Congress of the European Economic Association, XXIV Meeting of the Economics of Education Association, 2nd Applied Econometrics IAAE, AUEB Conferences in Economics and Econometrics, 6th Annual Meeting of the French Economic Association, 18th IZA European Summer School in Labor Economics, 2014 and 2015 Conferences of the Eastern Economic Association, LEER Workshop on Education Economics, 2015 Royal Economic Society Conference, 2014 International Atlantic Economic Association, Warwick CWIP and seminar participants in Essex, Bogazici, Norwegian School of Economics, North Carolina at Chapel Hill, Queensland and Warwick for useful comments and suggestions. All errors and omissions remained are our own responsibility. Financial support from IAAE is gratefully acknowledged.

1 Introduction

Knowing how one's characteristics compare to those of others is important in many settings of decision making. Humans are social beings, and it is natural to make comparisons in terms of characteristics and abilities in given tasks (?). The theory of social comparison is about "our quest to know ourselves, about the search for self-relevant information and how people gain self knowledge and discover reality about themselves" (?). One measure that provides social comparison information is ordinal rank in addition to relative or absolute information. In the workplace, recent evidence suggests that an individual is influenced not only by his relative or absolute income but also by the ranked-ordered position of his wage within a reference group. So one's ordinal rank is used when people make comparisons with others (?, ?). Social comparison is also an indispensable part of bonding among adolescents. In education, it is rare for teachers or principals to provide relative performance information to students. Thus, there is little understanding of the effects of providing information on students' ranks, because there is limited variation in the nature of such information that students receive-thus precluding causal analyses. However, knowing one's ordinal rank may affect education investment decisions and, hence, future academic and labor market outcomes.

Improving pupil attainments has been and continues to be an important issue for policy makers and academics alike. In an effort to improve students' grades, education policies have focused on improving a wide array of school inputs, among them, reducing class size (?, ?), improving the quality of teachers (?, ?, ?), extending the term length (?) and improving the quality of a student's peer group (?, ?, ?). Disclosing social comparison information (such as rankings), and manipulating the availability of this information would be significantly less costly than the above-mentioned interventions.

This paper presents empirical evidence about whether providing high school students with information about their rank on externally marked, high-stake exams affects future performance in similar exams. Our analysis relies on the fact that different cohorts are subject to different policies regarding the provision of social comparison information.

We exploit a large-scale, natural experiment that took place in 2005, when Greece adopted policies that altered the education testing regime, and eliminated information on rank that has been available to students nationwide. Until 2005, all students had to take national exams in two consecutive grades: one year before high-school graduation (eleventh grade), and the year of high-school graduation (twelfth grade). Results for all students were published nationally. From the results, students could calculate their rank within the school and across the nation, thus enabling comparison within the school and nationwide.

After 2005, the penultimate-year (eleventh grade) national exams were abolished and replaced with school-level exams. As a result, penultimate-year students take exams on the same subjects as before, but only receive report cards with their own exam results; they no longer receive information about their penultimate-year *relative* performance in relation to others in their school or across the nation. These cohorts -like the previous ones- continue to take twelfth-grade, national exams that will determine their post-secondary placement. However, they no longer receive information regarding previous relative performance in similar exams (national exams).

We define a feedback regime as receiving information regarding one's performance in comparison to peers in school and nationwide. In the feedback regime that existed until 2005, eleventh-grade students could compare themselves to others allowing for social comparison. Thus, we find the effect of the feedback on students' short-term (subsequent exam performance) and long-term outcomes (repetition of national exams one year after graduation, university placement and expected annual earnings) by comparing students of the same prior performance (tenth-grade performance) across regimes. In other words, we compare students who do receive information on rank in the eleventh grade and those who do not receive this information. We then use two sets of subjects. The main analysis is conducted using those subjects on which students receive the relative performance information in the eleventh grade. The second group of subjects, the counterfactual group, is a group of subjects on which students do not receive any relative performance information in the eleventh grade in either regime. Using new data on school performance, school quality and national exams for university admission, we test the hypothesis that students' final-year exam (twelfth-grade) performance is independent of the feedback regime.

When students knew their performance in the eleventh-year exams, they could "translate" their hours of effort into their exam result. For a given level of effort exerted by their peers, they discover how much effort they need to put into the final-year exam to rank accordingly. Knowing their relative performance could affect investment decisions such as the amount of effort students decide to exert in their final year of school. Students' performance in the final-year (twelfth-grade) national exams is the most important determinant for university admission in this setting.

Our first finding is that high-achieving students perform better in the final-year national (externally graded) exams when they receive feedback. Providing relative performance information the year before, improves the next period's exam performance of the better students by 0.2 standard deviations and their relative national rank by 4-6 percentiles. This is of comparable magnitude to being taught by a teacher 1.5-2 standard deviations above the average (**?**, **?**), or to reducing the class size by 15 percent. (**?**, **?**). Additionally, we find evidence that the performance of students in the lower percentiles deteriorates when feedback is provided. In particular, their consecutive-year performance declines by 0.3 standard deviations, and their national rank decreases by 6-8 percentiles.

Our second finding reports the responses of males and females to feedback at different parts of the ability distribution. High-achieving students of both genders respond positively to positive feedback, and low-achieving students of both genders respond negatively to negative feedback. However, females seem to be considerably more sensitive to feedback at all parts of the ability distribution compared to males. Our results are consistent with the existing literature regarding the gender differential response to performance due to initial different levels of self-confidence (**?**), or competitiveness of the exam (**?**).

Our third finding is that the provision of feedback changes the matching of students to university departments. First, we rank all university departments (programs) based on selectiveness and we construct a program list from the most-selective programs (e.g. engineering and medicine) and to the least-selective (e.g. geo-technology and environmental studies). We find that feedback provision corresponds to high-achieving students moving up the program selectiveness ladder by 30 positions, which is 0.15 of a standard deviation. On the other hand, low-achieving students move down the program selectiveness ladder by 35 positions, which is 0.18 of a standard deviation. Using the national Hellenic Labor Force Survey information, we find the annual earnings of older people in each occupation, and we map them to university departments. When the social comparison information is disclosed, we find that high-achieving students experience an increase in expected earnings by 0.13 standard deviations. Further, feedback provision for low-achieving students imposes a decrease in expected earnings of 0.23 standard deviations.

Additionally, we find evidence to suggest that feedback encourages students from low-income neighborhoods to enroll in university and, as such, alters the socio-economic composition of students who are admitted to the top programs. More students from low-income neighborhoods gain admission to the most selective programs with the highest expected earnings after graduation (such as engineering and law), when feedback information is provided. This implies that feedback encourages social elevation for students from low income families.

This paper makes two main contributions. First, this is the first large-scale study that documents the long-term effects of providing relative performance information in an educational setting. In particular, we document the direction and size of the effect of feedback on long-term outcomes such as repetition rates for the national exams, students' post-secondary placement, and expected earnings. We contribute to the literature by providing evidence that knowing one's ranking within the senior high school or nationwide has long-lasting effects, and changes an individual's career path.

Secondly, we explore how information transparency-particularly about a student's rank-affects educational outcomes. We accomplish this by making use of a unique information treatment. Disclosing information about ranks is not a standard practice for teachers or principals. Thus, the information treatment that we study in this paper is rare. We exploit a special setting in which high-school students receive explicit information about their relative position in at least two reference groups: school and country. Although students may generally observe their own perspicacity, they do not generally observe everyone's performance in the school and the country to deduce their ranking. Thus, we are able to separately identify the effect of knowing someone's ranking in each of these two groups.

This study has important policy implications. First, we provide evidence that a low-cost instrument -such as providing information on one's rank- has the potential to affect students' educational achievements. Our estimates of impact are at the lower end of those from the current literature on improving school inputs. Nonetheless, all the interventions studied so far (improving teacher quality, reducing class size, enhancing peer-group quality etc.) are significantly more costly than manipulating the availability of social comparison information. Thus, information on rank can be considered a new factor in the education production function[1].

Second, our findings imply that when the relative performance information is dis-

[1]Other determinants of the education production function that have been studied include: studies on class size (?, ?, ?), teachers' training and certification (?, ?), quality of teacher (?, ?), tracking (?), peer effects (?, ?), non-cognitive skills (?), classroom instructional time (?).

closed, it can be an important, additional factor in terms of the school choice. Our findings imply that being in a school with higher-achieving peers might not always be optimal for students. That is, students benefit from going to schools in which they are among the high-performing students (i.e. schools with more students with relatively lower performance levels). Making a school choice on the basis of rank only is unlikely to be correct, given that there are many other factors in the education production function.

In the recent years, the economic literature has shown increasing interest in the effect of feedback on feedback on exam performance.[2] **?** examine the effect of feedback information on students' future absolute performance using data for university students registered to departments with different feedback policies. In that study, feedback is defined as the knowledge of someone's absolute performance in the midterm exam in period one, and before students exert effort on their essay in period two. The authors find that the effect of feedback is positive for all students, and more pronounced for more-able students. Their study refers to feedback involving one's own absolute performance. Our study refers to the provision of feedback regarding relative performance.

The paper most closely related to ours is a study by **?**. The authors examine the effect of relative performance feedback on students' future absolute performance. They exploit a natural experiment that took place in a high school, in which for one year only, students received information about the average class score in addition to information about their own performance. Their findings suggest that feedback improves the performance of all students in the subsequent test. They do not find differential effects by gender along the ability distribution. A key difference between their paper and ours in the sample size. They use a small sample of one high school; by contrast, we use what is in many dimensions a nationally representative sample of 134 senior high schools. Another important difference is that **?** investigate the effect of providing information about someone's relative position within the class only. We contribute to the literature by exploiting an explicit information treatment in which the social comparison information refers to reference groups broader than the class, i.e. the school and the nation.

[2] The relative feedback information has been studied in the tournament literature. Some studies find that relative performance information has a positive effect for all participants in both tournaments and piece-rate payment schemes (**?**). On the other hand, some other studies find mixed results. **?** uses data on furniture sellers' effort, and finds that feedback has negative effects on the low-performing employees.)

More recently, **?** examine the effect of one's primary school ordinal rank on future exam performance. In their setting, students figure out their rank within their class from repeated interaction with their classmates. They find that being highly ranked in primary school has large and robust effects on secondary school achievement, with boys being more affected than girls. Our setting differs in that students receive explicit information regarding their rank within the school and nationwide, which facilitates the policy recommendations.

An interesting question is whether the effects are driven by students, parents or teachers or some combination of them. It is almost impossible to disentangle whether the effects are coming from the students or the parents. However, we can rule out the two mechanisms; neither sorting into schools by parents nor sorting into classes by teachers explains the effects. We discuss several possible mechanisms that cannot be fully excluded, and, thus could explain our findings. They are: 1) priors, 2) information about school quality, 3) parental investment, 4) practice and 5) learning about one's ability.

The paper is organized as follows. Section 2 provides a brief description of the institutional setting and the data. Section 3 sets out our empirical strategy. Section 4 presents the main results on short and long-term outcomes and discusses heterogeneous feedback effects by prior performance and gender. Section 5 discusses the potential mechanisms. Section 6 discusses the threats to identification and reports further robustness checks. In Section 7 we conclude and discuss policy implications.

2 Institutional Setting and Data

2.1 Institutional Setting

All universities in Greece are public and the Ministry of Education manages the admission procedure. Access to tertiary education is based on the "admission grade". The admission grade in both regimes is a weighted average of the grades a student gets in the national exams (70% weight) and the school grades (30% weight). National exams for specific subjects take place on specific dates every year. The questions are the same for all students and the exams are externally marked. The school grade for every subject is the average of the term grades. Only final-year students can participate in the university admission procedure. All students are examined on five general or core subjects, plus several "speciality" elective subjects chosen at the beginning of the twelfth year.

First, students take the final-year exams and then their admission grades are announced. Then students apply by submitting to the Ministry of Education a list, in order of preference, of university departments to which they would like to be admitted in that year. Admission is made to a specific university department. The student's ordering of university departments is crucial: once a student gains admission to a university department in a higher place in his preference list, he cannot be admitted to any departments below that position. This means that students have to be very careful in constructing their preference list. The only way a student can avoid the university admission procedure is by not submitting a list of preferences. Then, each department admits the best students who have included this department in their preference list. All students are compared to each other according to their admission grades and every successful candidate is admitted to the first department in his list where there is an available place, and every student with a higher admission grade has already been allocated. The rest of the students are denied admission for that year.

At the end of this process, every department announces the grade of the last student it admitted in that year. This grade is considered to be the "cut-off grade" in that year. More selective/prestigious departments have higher cut-off grades and students are aware of the cut-off grades of the previous years when they construct their preference lists. The ranking of university departments according to their cut-off grades appears to stay largely unchanged, year after year, and this represents the students' evaluations for these departments.

The admission grade of a student in the non-feedback regime depends entirely on students' performance in the twelfth grade.[3] In the feedback regime, students' performance in the eleventh grade could take some weight (30%) in the calculation of the admission grade, but only if their eleventh grade performance exceeds that of their final year exams.[4] Again, the overall performance of a student in each grade is a combination of the national exams (70%) and the school grades (30%). The eleventh-grade material in not included/tested in the twelfth-grade exams. The results of the penultimate-year exams are not used in any other way in the university admission procedure. So, stu-

[3]It is a combination of the national exams (70%) and the school grades (30%).

[4] In this case, the overall performance of a student in the twelfth grade takes a weight of 70% and the overall performance of a student in the eleventh grade takes a weight of 30% in the calculation of the admission grade.

dents have incentives to perform well in the eleventh-grade national exams, but that performance is not sufficient to secure a specific university placement. Given that the number of seats is pre-determined, a specific score will not guarantee admission to a specific university department, because demand for that university department also comes into play.

2.2 How does feedback work?

Knowing one's own relative performance might affect the amount of effort a student exerts with regard to a certain objective. In the context of our study, the student's objective is to maximize his or her score and/or rank at the end of high school.

Consider a student in the treated group. In the world of this experiment, students compete with each other over access to a limited number of university places. At the end of the penultimate year, students take standardized exams in some subjects with external examiners and at least two anonymous external graders per subject.

Then two mechanisms are in action. First, everyone's results within the school become public knowledge: the names and detailed grades are displayed at the entrance of every school. This provides students with information about how well they can do given a specific level of effort, when national exams come around again. This means that students could calculate their distance from the school's average score, and their relative position within their school. Second, the names, details about national exam scores, and the cohort's average national exam score are published in the newspaper. This means that each student could calculate her distance from the national cohort's average score and derive her relative national rank.

We believe that students in the feedback regime calculate their eleventh grade rank within the school and nationally given the importance of their performance in the senior year exams. Knowing a student's national rank provides them with information about the competition in that year. Each year the newspaper reports the following: cohort's average national exam score, the cohort's minimum and maximum score, the score that corresponds to each decile and comparisons with last year's statistics. For each student the following is reported: student's first name, surname and father's name, score given by the first and the second examiners (Figure 2.3 and 2.4). This is published separately for each subject. The score given by each examiner ranges from 0 to 100. If the difference between the score given by the first and the second examiners is not greater than 13/100, then the final score is the average score between the scores given by the first and the second examiners. Otherwise, the final score is the average between the highest two scores given by any examiners. The raw final score used a 1-to-20 scale that we transformed into z-scores to facilitate the interpretation of the results.

Students use this information to calculate their national rank. Given that the names are alphabetically sorted, calculating a student's rank using even a single newspaper's scores is already a good indicator of a student's national rank. Calculating the school rank is much easier given that the average school cohort consists of 79 students.

Consider a student in the control group. During the penultimate year of senior high school, he chooses an effort level to prepare for his exams, which are now given only at the school level. Within the school, teachers coordinate to cover the same material, and usually give the same exam questions. Before the summer break in the penultimate year, our student takes exams on the same five subjects, and receives a written report from school with his own grades. When he reaches twelfth grade, he has access to the same material, study guides and past exam papers as any student in the treated group. However, he is unaware of how his schoolmates and his cohort did relative to him in the penultimate year final exams. Table 3 reports the summary statistics of the variables of interest across the two regimes. Some of the differences seem to be significantly different from zero but they are either very small or economically non-meaningful. The exact timing is presented in Figure 2.

2.3 Data Collection

To study the effect of disclosing rankings, we need a prior measure of performance that is not affected by the feedback provision, i.e. students' tenth-grade performance. Data on students' performance in the tenth grade are not centrally collected[5], and can only be found in the school archives. We visited 134 senior high schools across the country and constructed a database of student performance in every subject throughout senior high school. Our novel dataset combines information from various sources:

1. We obtained administrative data from the Hellenic Ministry of Education regarding the performance of all students in the twelfth-grade national exams from 2003 to 2009. This dataset contains student level information about gender, national

[5] The Ministry of Education collects data on students' scores that are used in the calculation of the admission grade.

and school exam results in each subject nationally examined in twelfth grade, the senior high school attended, year of birth, graduation year from senior high school, and speciality subjects chosen at the beginning of twelfth grade. It also contains university admission-related information, such as the university department where each student gained admission, number of applications made to university departments, and the reported ordinal preference position of the admitted university department in the student's preference list. The dataset refers to the period 2003-2009, and gives us information about 435,589 students.

2. Because the Ministry does not collect information on students' tenth-grade performance, we collected this information directly from the schools.[6] More specifically, we have physically visited and collected data from more than 147 public, experimental[7] and private schools from cities and the countryside. The final sample includes 134 schools which corresponds to 10 % of the school population. We exclude evening schools[8] from our analysis because they differ[9] in many aspects from the other types of schools.[10]This dataset includes the following information: year of birth, indicators for gender, indicators for class, graduation year, school and/or national exam results in the tenth, eleventh and twelfth grade in all subjects, speciality chosen at the beginning of the eleventh and twelfth grade and a unique, individual student identification that stays the same throughout senior high school. We have had short interviews with the principal of every school in our sample to find out about any effects potentially affecting our outcomes of interest. Inter alia, principals were asked about the size and history of the school, facilities, attrition and teacher quality.

We match the twelfth-grade, school-level data with the administrative data using

[6]The tenth-grade performance data are recorded in each school's archives either in their computers or in their history books. In most schools the data for all the years were extracted from their computers. There were cases-especially for the data referring to the first years of our sample period- where we photocopied pages from the history books in schools' storage area.

[7]Experimental schools are public schools where admission in these schools is based on a randomized lottery.

[8]Which are public schools that offer evening lessons in order to target employed students.

[9]University cut-offs differ for students graduating from evening schools compared to any other type of school. Including these schools in the analysis provides similar results, only varying at the second decimal point. Contact authors for further results

[10]We also exclude schools that had at least one year school cohort size smaller than ten students because these small schools may be atypical in some dimensions. Results including those schools are very similar. Contact authors for further results.

the following combination of information: year of birth, gender, high school attended, graduation year, speciality chosen at the beginning of twelfth grade, and school as well as national exam scores in each subject examined at the national level. The matching between the dataset provided by the Ministry of Education and the school datasets was very satisfactory[11] providing us with a complete senior high-school performance history for 45,746 students, which is our sample size.

3. We obtained average household income information for 2009 for every postcode in the country from the Ministry of Economy and Finance. We employ this as a proxy for neighborhood income.

4. We obtained postcode data on urban density information from the Ministry of Internal Affairs. Urban areas are those with more than 20,000 inhabitants.

5. We obtained the Labor Force Survey data for the year 2003 from the National Statistical Authority. We use quarterly data to create a variable that maps college occupations into annual earnings.[12] We do that if respondent's reported education is in the same field as her actual occupation in 2003. Respondents report their occupations with high precision.[13] The earnings data are grouped into ten bins indicating the ten national deciles with the highest frequency. We use the lowest bound of each bin[14] to construct a variable that measures minimum expected annual earnings for each occupation.

Every school follows the same curriculum, and students are assigned to public schools based on a school district system. This school district system assigns students to schools based on geographical distance. Students are alphabetically assigned to classes in tenth grade and do not change class throughout senior high school. Moreover, teachers are allocated to public schools based on geographical criteria and no quality criteria are taken into consideration in the process. Figure 1 presents the geographic position of each school included in the sample. The density of the school population in Athens is 32%- thus, many of the schools in our sample are located in Athens.

[11] 92 % of students were matched, with only 8 % missing due to the lack of values either in the school level data or the administrative Ministry level data.

[12] We also map college fields to occupations.

[13] 209 classified occupations are reported and respondent have to indicate which one is closest to their actual occupations.

[14] Multiplied by 12 months.

Table 1 presents descriptive statistics about the available variables in the sample in the twelfth grade. The variable "internal migration" takes the value one if the district of university department to which the student is admitted is different from the district of residence; the latter being proxied by the school district. Moreover, the variable "early enrollment" takes the value of one if the student enrols in the first grade before the age of six.[15]Interestingly, on average, 82 % of students gain admission to at least one university department. Given that there are no fixed cut-offs, if there is not much demand for a particular university department, the cut-off grade in that year is very low.

Table 2 reports the mean characteristics of the schools in our sample and the whole school population, and, thus , allows us to investigate whether our sample is representative. There are some variables for which there is a statistically significant difference between the 134 sample schools and the rest of the school population, and these differences are mainly related to the sampling methods that we used.[16] So, though the sample may not be fully representative of national responses, but it nonetheless looks very similar.

2.4 Test Scores

Our prior measure of performance is based on the overall students' performance in the tenth grade (GPA). The tenth grade GPA takes into account students' tenth grade performance in thirteen subjects. The performance in each subject is a weighted average of the final school-exam result and the performance of the student during the school year. Teachers receive guidance on how to mark the final tenth-grade school exam and test scores are not curved. We use the within school rank of each student based on the tenth grade GPA as a prior measure of performance.

Our main outcome variable is a student's twelfth grade rank in two reference groups (the school and nationally). These outcome variables -the within school and the national rank- take into account students' twelfth grade performance in five core-education subjects. Students take exams at the end of the twelfth grade in these core-education

[15] According to the law, this happens if the student is born in the first quarter of the calendar year.

[16]i.e. the share of schools in Athens in our sample is higher than the share of schools in Athens in the population. Furthermore, the share of private schools in the sample is 4 % smaller than the share of private schools in the population and the share of experimental schools in the sample is 4 % higher than the share of experimental schools in the population.

subjects. A student's performance in these five subjects is the most important determinant for the calculation of the high-school graduation grade under both regimes. Before 2005, these five subjects were all examined at a national level. From 2005 onwards, two subjects are examined at a national level whereas the other three subjects are examined at a school level. This change in the number of subjects examined at a national and school level happened in the same year as the abolition of feedback. We do various robustness checks to examine if this change affects our results. In particular, we use various outcome variables (the rank in each subject separately; the average rank in those subjects examined at the national level; or the average rank in the five core-education subjects) and the estimated effects follow the same patterns. We call the core-education subjects "incentivized", because performance in these subjects is taken into account in the calculation of the admission grade.

All schools in the sample offer three academic tracks in the twelfth grade. Each student has to choose the academic track that is relevant to the post-secondary degree they desire to pursue. Each track offers different subjects. Depending on the track students choose, they take national exams in four track-specific subjects in both regimes [17] We do not include the test scores in these four subjects in the main analysis because the choice of track is based on endogenous criteria, i.e.their perceived differential ability or preferences for a particular degree after high school graduation. Robustness checks show that the results remain almost unchanged when the track-specific subjects are taken into account.

In addition to the core-education subjects and the track-specific subjects, students take compulsory within-school exams in three subjects (Sociology, Religion and Modern Greek Literature) in both grades; eleventh and twelfth. Students take school exams at the end of the eleventh grade and each student receives a report card. This report card shows each student's own performance in these exams without providing information about the class or school average score. In the twelfth grade, students are examined again on these subjects without having previously received any relative performance information in these three subjects. We call these subjects "non-incentivized" because students' performance in these subjects is not taken into account in the calculation of the university admission grade in any of the regimes. Students take these exams in both regimes. We use these subjects as the main counterfactual group.

[17] These four subjects differ from the one track to the other. The Tracks are: Classics, Exact Science and Information Technology.

In our analysis, we use the rankings instead of absolute scores for a couple of reasons. First, using the tenth-grade ranking allows us to make comparisons across cohorts and across schools. Notice that we do not observe the different feedback policies in the same year. Thus, we use the within-school rank of a student to compare students who are exposed to different peer groups and teachers. Second, a given twelfth-grade national exam score does not represent the same ability level in different years. However, it is important to make sure that students of the same ability obtain the same rank in different years. The comparison of students' absolute scores across cohorts would be problematic, if the difficulty of the exam changes from one year to another. Additionally, calculating the ranking of a student in each subject takes into account the potential difference in the difficulty of the exam from one subject to another. Thus the ranking allows us to compare a students' performance across different subjects. Also note that any school grade inflation that might occur in the tenth grade does not affect our prior performance measure (tenth grade GPA). Grade inflation would make the teacher more lenient in the overall grading procedure, which implies that the ranking of the students remains unaffected. The national exams in twelfth grade are externally graded. As a result, the teacher in a student's school has no way to affect their national exam final scores. Furthermore, the national exam procedure does not receive any grade curving.

3 Empirical Strategy

This section identifies the effect of relative performance information on students' senior year exam performance. First, we define our measures of rank. Second, we identify if there is an effect. Because we use as an outcome variable the rank in the twelfth grade, the effect is -if anything- of a distributional nature. Finally, we discuss the empirical method to identify the effect of feedback on students' relative final-year performance.

3.1 Calculation of the rank

We use the following normalization in order to calculate our measure of prior performance that allows comparisons of students in tenth grade across schools and cohorts:

$$Rank_{10isc} = \frac{n_{isc}-1}{N_{sc}-1}$$

where n_{isc} is the ordinal ranking of student i within school s in cohort c in tenth grade[18] and is increasing in GPA and N_{sc} is the school cohort size of school s in cohort c. The higher the $Rank_{10isc}$, the higher the ranking of student i in tenth grade in his school s and cohort c. Moreover $Rank_{10isc}$ is bounded between between 0 and 1, with the lowest ranked pupil in each school having $R_{10isc} = 0$. For example, in a school consisting of 100 students ($N_{sc} = 100$), the student with the fifth highest GPA ($n_{isc} = 95$) will have $Rank_{10isc} = 0.95$ while the student with the first lowest GPA will have ($n_{isc} = 5$) so his rank will become $Rank_{10isc} = 0.05$.

The ranks of the student within his school in the twelfth grade and nationwide are calculated using the following normalisations:

$$Rank - school_{12isc} = \frac{k_{isc}-1}{K_{cs}-1}$$

$$Rank - nationwide_{12ic} = \frac{r_{ic}-1}{R_c-1}$$

Where k_{isc} is the ordinal ranking of student i in school s in cohort c in twelfth grade and is increasing in the national exam grade. K_{cs} is the cohort size c in school s. The $Rank-school_{12isc}$ is projected into the [0,1] interval and the lowest ranked pupil in each school cohort has $Rank - school_{12isc}$=0. Notice that there are five subjects, so we first find the ordinal rank of the student based on the average in the five scores, and then we normalise it using the above formula . $Rank - nationwide_{12ic}$ is calculated in a similar way but is independent of the school the student attends. So both $Rank - school_{12isc}$ and $Rank - nationwide_{12ic}$ are calculated based on the twelfth grade national exams in the incentivized subjects, but they measure relative performance in the school and the country respectively. For example, in a cohort with 50,061 students (R_c=50,061), the student with the tenth highest twelfth grade national exam score (r_{ic}=50,051) will have a national rank of $Rank-nationwide_{12ic}$=0.999. If the same student has 78 schoolmates (K_{cs}=79) and he has the second highest score within his school in that cohort (k_{isc}=77), then the school rank of this student becomes $Rank - school_{12isc}$=0.974.

3.2 Identifying the effect

Figure 3 shows the average rank nationwide of each performance group in the twelfth grade exams, conditional on students' prior performance. Cohorts up to 2005 have received the relative performance information. We observe that the lines are parallel in the treatment period (cohorts 2003, 2004 and 2005). This means that the time

[18] Based on the average of the thirteen subjects, ie.the tenth grade GPA.

trends for each quintile of prior performance follow a similar pattern from year to year. Identification is achieved through a difference approach for each prior performance group. The 2006 cohort is the first cohort affected by the abolition of the relative performance information. We observe that from 2005 to 2006 the slopes of the time trends change, meaning that the treatment affected students in all performance groups considerably except the middle quintile, which remained unchanged. In particular, the top quintile achieved a higher average rank nationwide in the twelfth grade when feedback was provided compared to the period after 2006. The opposite applies to the bottom quintile where students end up lower in the distribution of twelfth grade rank when they are aware of their previous relative performance compared to the period after 2006.

Another important observation is that the slopes remain relatively stable after 2006, which is the first affected cohort. So, the change in the slope of the time trends between 2005 and 2006 can be attributed to the abolition of the relative performance information. We produce this figure using students' rank nationwide (Figure 3) and rank within the school (Figure 4). Their measures on rank are derived using the average rank in the general or core subjects.

3.3 Method

We adapt two strategies to quantify the effect of feedback provision on future performance.

First, we use the following specification to estimate the effect of feedback information on students' later rank, conditional on their prior performance.

$$Rank-nationwide_{12ic} = \alpha + \beta_{quintile} Feedback_c * Quintiles_{10isc} + \lambda_{quintile} Quintiles_{10isc}$$
$$+\psi Feedback_c + X'\gamma + \psi_c + \phi_s + \epsilon_{ic} \text{ (1a)}$$

$$Rank-school_{12isc} = \mu + \delta_{quintile} Feedback_c * Quintiles_{10isc} + \kappa_{quintile} Quintiles_{10isc}$$
$$+\xi Feedback_c + X'\zeta + \theta_c + \omega_{isc} \text{ (1b)}$$

where $Quintiles_{10isc}$ is a dummy variable that takes the value of one if the student is in the corresponding quintile based on his tenth grade performance in his school. Moreover, $Feedback_c$ is a dummy variable equal to one if the student takes the eleventh grade national exam ie. if the graduation year is smaller than 2006 (feedback regime). The parameter of interest β (δ) measures the effect of feedback on student's rank nationwide (within his school) in the subsequent year, conditional on tenth grade performance. In some specifications, we control for unobserved time and school invariant factors that may affect final year's rank using time and school fixed effects. We also control for students' characteristics (X) like the age and the gender of the student. Specification (1b) exploits within school variation, thus we use (1a) without the school fixed effects when we are interested in exploiting across schools time invariant variation.

In addition to the first strategy, we now use the following difference specification to find the effect of feedback on each decile of students' twelfth grade performance. We run the following specifications for each decile of tenth grade performance $\theta \in [0, 1]$:

$$Rank - nationwide_{12ic\theta} = \delta_\theta + \alpha_\theta X_i + \beta_\theta D_{\mathrm{c}} + \psi_c + \epsilon_{ic\theta} \quad (2a)$$

$$Rank - school_{12isc\theta} = \omega_\theta + \alpha_\theta X_i + \gamma_\theta D_{\mathrm{c}} + \theta_c + u_{isc\theta} \quad (2b)$$

where δ_θ captures a performance group-specific fixed effect. D_{c} is a feedback dummy that takes the value one in the feedback regime and it takes the value zero in the non-feedback regime. The parameter of interest β is estimated separately for each one of the ten deciles, including clusters at the school level. A similar regression across all decile groups gives the pooled OLS estimator of β_θ which is exactly zero, because as we explained before, the provision of feedback has a zero average effect. A negative coefficient of β_θ orγ_θ implies that feedback induces a deterioration in the rank nationwide or within his school for students at this decile.

4 Main Results

4.1 Effect on performance

Main OLS results are reported in Table 4. The first column in Table 4 corresponds to the basic specification (1a) without school and year fixed effects. The dummy for the third tenth grade quintile is omitted as a point of comparison. This shows that when feedback is provided, a student in the top quintile in his school has a 0.042 percentile rank gain in his twelfth grade national exam performance compared to a student who is in the median quintile in his school, ceteris paribus. Similarly, a student who receives feedback and is in the bottom quintile in his school has a 0.088 percentile rank loss in his twelfth grade national performance compared to a student in the median quintile in his school. In columns 2 and 3, we see that the results of column 1 are robust when controlling for unobserved heterogeneity across schools and years respectively. Adding school and year fixed effects slightly change the coefficients estimates, which remain statistically significant at a 1 % significance level. In all specifications, we control for a set of pupil characteristics and we cluster the standard errors at the school level.

In specification (1b) we exploit the within school variation and results are in Table 5. The effect of feedback on students' within school performance in the incentivized subjects is reported in columns (1) and (2) and in the non-incentivized subjects in columns (3) and (4). In the first column, we show that students in the top quintiles, 5 and 4, based on the tenth grade performance, benefit from feedback. This gain is associated with 0.045 and 0.040 school percentile ranks respectively, compared to the third quintile. Similarly, quintiles 2 and 1 (bottom ones) experience a loss of 0.038 and 0.079 school percentile ranks when feedback is provided. In column 2 we control for unobserved heterogeneity across years and as we expect, results are similar to Table 4 column 3 when we controlled for unobserved heterogeneity across years and schools in the national analysis.

We replicate the analysis using the school rank in the non-incentivized subjects as the outcome variable. In columns 3 and 4 (Table 5), we find that the coefficients are not statistically significant and there is no evidence that the provision of feedback affects students' performance in these subjects. What is important here is that students do not receive any social comparison information regarding the non-incentivized subjects neither in the feedback regime nor in the non-feedback regime. These results are also meaningful because they provide evidence that there are no spill-over effects from the feedback towards the non-incentivized subjects. In other words, students do not react to feedback by studying more or less for the school exams rather than the national exams. These findings support that the effects on students' final-year performance are generated by the relative performance information that is provided in the eleventh grade.

We then run specification (2a) and in Figure 5 we plot the β_θ coefficients of the rank nationwide and the associated 95 % confidence interval. We observe that receiving information about someone's relative performance has a negative effect for students below the 45th percentile and a positive effect for students above it. At the highest two deciles, the curve is slightly decreasing implying that there is a ceiling effect. In other words, there is some upper bound on how much improvement feedback induces for the highest performing students. Thus, receiving relative performance information the year before the university admission, high stake exams improves the final-year rank nationwide of the high achieving students by up to 5 percentiles. By contrast, when the relative performance information is provided then the final-year rank nationwide of the low achieving students drops by up to 8 percentiles. In Figure 6, we report γ_θ coefficients and the associated 95 % confidence interval, which shows the effect of feedback on the final-year rank within the school (and not national rank as before). The estimated treatment effects on the final-year rank within the school (in Figure 6) are very similar to the ones found before, when the national rank was considered (in Figure 5). For schools above or below the average quality school a student's rank within the school differs from his rank nationally. However, on average (across all schools) the school rank for each decile might not dramatically differ from the national rank given that the school sample is a representative one in terms of many observed characteristics.

Figure 7 plots the treatment effect coefficients for the non-incentivized subjects that we use as the main non-treated subjects and we explained previously in this section. In line with Table 5, we find no evidence that providing feedback affects students' performance in these subjects.

We then standardize the twelfth-grade scores in each year and school to give a zero mean and a standard deviation of one. Then we run a specification similar to (2b), but the outcome variable is the twelfth-grade standardized score of student i in school s in cohort c in each decile θ. We run this regression for each decile of tenth-grade performance, and we plot the coefficient of the feedback dummy D_c. The treatment effects line for each decile of prior performance is presented in Figure 8. There, the gain for students above the 40th percentile is up to 0.15 standard deviations while the performance of students who are below the 40th percentile drops by up to 0.3 standard deviations.

4.2 Gender

Next, we turn to the gender analysis. As literature on evaluating social programs has shown, individuals respond differently to the same policy (**?**). To test whether boys and girls react differently to the provision of feedback, we estimate the following regression:

$$Rank - nationwide_{ic} = \delta + \beta Feedback_c * Female_i + \kappa Feedback_c$$
$$+ \lambda Female_i + \alpha X_i + \mu_t + \epsilon_{ic} \text{ (3a)}$$

$$Rank - school_{isc} = \delta + \beta Feedback_c * Female_i + \kappa Feedback_c$$
$$+ \lambda Female_i + \alpha X_i + \mu_t + \epsilon_{isc} \text{ (3b)}$$

where X_i includes the tenth grade GPA performance, a dummy for early enrollment in school and dummies for the speciality chosen in the twelfth grade. OLS results are shown in Table 7. Although girls outperform boys, girls end up in a lower rank on average when feedback is provided. This is the case for both; the rank nationwide and the rank within their school.[19] Running specification (2b)[20] for boys and girls separately produces Figure 10, which presents the treatment lines for boys (on the left) and girls (on the right).

For both genders, the effect of feedback is positive for high-achieving students and negative for low-achieving students. We make two important points here: first, the average effect of feedback is positive for boys' final-year rank and negative for girls' final-year rank, as shown by the horizontal line, which is generated by a regression across all deciles (Figure 9). Second, the effects of feedback are more pronounced for girls. As indicated by the steeper treatment line in Figure 9, girls exhibit greater sensitivity to rankings.

Our evidence is consistent with the literature supporting a differential gender effect of feedback, with females responding more to additional information. In an experimental context, **?** shows that women and men may react differently in the absence of feedback information because of different levels of self-confidence. Also using an experimental context, **?** argue that women never have the same level of self-confidence as men because women expect less of themselves than men do. Our gender differential negative feedback effects are consistent with the existing literature on gender specific perceptions regarding competition. (**?**, **?**). Among women and men of the same prior performance, women are less effective than men when they take the competitive national exams.

4.3 Long term outcomes

In this section we examine the effect of feedback provision on students' long term outcomes. Students who have not been admitted to their chosen university department may re-apply a year (or more) after graduation using their school grades and re-taking

[19]In Table 7, if we include school fixed effects in columns (3) and (4), we account for heterogeneity across schools and the coefficient estimates become the same as in columns(1) and (2).

[20](2a) gives results almost identical to (2b) for both genders.

national exams in all subjects. Those students usually do not attend any school/college, or pursue any job, or do military service after graduation and before the next admission period.

We use binary response models to examine whether the provision of feedback affects the decision to retake the exam. In Table 8, we observe that a significant percentage out of the cohort population repeats the exams one year after graduation from senior high school[21].

We define as "misplacement" the difference between the tenth grade rank each student achieves within the school and the rank nationwide in the twelfth grade. Thus, the misplacement variable is bounded between minus one and one. Students with larger differences between the tenth- and twelfth- grade ranks would have a large change in their relative performance. The misplacement variable takes the value zero for students whose twelfth-grade rank happens to correspond exactly to the tenth-grade rank. But it can also take positive (negative) values if the student achieves a better (worse) performance in the tenth grade relative to the twelfth.

To examine if feedback provision affects someone's decision to retake the national exams through the misplacement effect we run the following specification:

$$Retake_{i,t+1,s,d} = a + X'_{itsd}\gamma + \delta Misplacement_{itsd} Feedback_t + \beta Feedback_t$$

$$+\omega Misplacement_{itsd} + \zeta Z_{td} + \xi_s + \omega_t + \epsilon_{itsd}$$

The decision to retake the national exam one year after graduation depends also on the opportunity cost for the student. Thus, we control for the unemployment rate in each year t and district d of student's residence.

Using Linear Probability (LPM), Probit and Logit models we find that when feedback is provided, students with higher misplacement are more likely to repeat the national exams one year after graduation. In Table 10, we interact dummies that capture the magnitude of misplacement with the feedback dummy and we observe that students in the top misplacement quintile (5) are more likely to retake the national exams when feedback is provided. The Top Misplacement Quintile (5) is the most positive one and contains students who get a better rank in the tenth grade compared to the twelfth. In the feedback years, these are the low achieving students. In other words, low-achieving students are more likely to resit the national exams when feedback is provided. By

[21] The number of students retaking the exam is calculated using the Ministry of Education dataset. The data about the labor force capacity are collected from the National Statistical Authority.

contrast, high-achieving students are less likely to retake the exams when feedback is provided.

Having a particular placement in university admission affects an individual's employment and earnings prospects. We examine whether feedback influences the matching of students to university departments. We first rank all programs [22] according to their average cut-offs over the seven years period. Each program's cut-off reflects the demand for this particular university department, with highly demanded programs exhibiting high cut-offs. Students apply 671 to programs based on preferences, social status and expected earnings. There are programs in total. We estimate the effect of feedback on the difference in the selectiveness position and rank of the program admitted conditional on tenth grade performance. Figure 10 presents the treatment effect line for the selectiveness position (on the left) and rank of the admitting program (on the right). The provision of feedback has a positive effect on the selectiveness position and rank of the admitting program in the upper half of the prior performance distribution, and a negative effect on the lower half. In particular, high-achieving students move up the university selectiveness ladder by 30 positions, which is 0.15 of a standard deviation. When feedback is provided, low achieving students move down the program selectiveness ladder by 35 positions which is 0.18 of a standard deviation. Different placements in university admission induce different gains related to the returns to college.

Enrolling into a specific university department may affect students' career paths and their lifetime earnings. Using Labor Force Survey data, we match salaries for each occupation to each university department. In particular, we use the 2003 Labour Force Survey to map each college major into the most related occupation and then into the expected annual earnings after graduation (in Euros).[23] We then use these figures as the expected earnings of current students after graduation from the particular program. In Figure 11, we present the effect of feedback on the expected annual earnings, conditional on the tenth grade performance. For students above the 50th percentile, annual expected earnings increase by 250 Euros per year, which is equivalent to 0.17 of a standard deviation. For students below the 50th percentile, the decline in their expected annual earnings corresponds to 0.20 of a standard deviation.

[22]By program we mean each combination of university department.

[23]Mean:12,758 with 1,473 standard deviation.

4.4 Social Mobility

In this section, we examine if the provision of feedback changes the relationship between parental income (proxied by neighborhood income) and post-secondary opportunities (indicated by the program the student enrols in). A priori, we might expect that students coming from more advantaged neighborhoods would have better chances of embarking on a better and more-selective program with higher expected returns than students coming from less-advantaged neighborhoods. Could the provision of feedback affect this flow of students from high-income families to high-expected-income programs? Providing relative performance information might have a different effect on students whose parents have varying levels of income; the difference in the role feedback plays may be related to other family resources (financial support or social networks) or students from different income backgrounds might value the ranking information differently.

To investigate whether feedback has a differential effect on students from different income backgrounds, we create quintiles based on the neighborhood income and the selectiveness of the program admitted to. In Table 13 we report for each quintile of neighborhood income, the percentage of students who enroll into each quintile of programs by selectiveness, in the feedback and the non-feedback regime. We then calculate the difference between the feedback and the non-feedback percentage. In the last row of Table 13, we vertically add the percentages of students who enroll in any program for each quintile of neighborhood income, to find the total difference of enrolled students between the feedback and the non-feedback period. In the last column of Table 13, we horizontally add the percentages of students who enroll in each quintile of programs. We do that to examine if feedback provision affects the total percentage of students who enroll in higher education. We find that 2.2 % more students (83.7% Vs 81.5%) enroll in a program in the feedback regime.

In Table 13, we find descriptive evidence that more students coming from the lowest-income neighborhoods (Quintile1) enroll in any program when feedback is provided (2.2% more students). A possible explanation is that low achieving students discover that if they do not exert more effort they will not be admitted to any program in tertiary education. Or they might discover that they are not worse than the low achieving students from high income neighborhoods and that they still have a chance to enroll in university. So, they might decide to exert more effort. This may show that feedback benefits students from low-income neighborhoods by reducing social inequalities and possibly future income inequalities. On the other hand, high achieving students from low-income neighborhoods discover in the eleventh grade that they are highly ranked on a national scale and they might react by exerting more effort.

We also find descriptive evidence in Table 13, that feedback provision alters the parental income (proxied by the neighborhood income) composition of students who are admitted into the top-ranked programs (Quintile 5). More students from low income neighborhoods are admitted to the most-selective programs that provide students with the highest expected earnings after graduation (such as engineering and law), when feedback information is provided (2.9% Vs 2.6 %). This implies that providing relative performance information encourages social elevation and improves economic opportunity for these students.

It is crucial from a policy perspective to understand if providing feedback is beneficial for the society as a whole. On one hand, high performing students are usually the ones responsible for innovation and technological breakthroughs. The technological diffusion is beneficial for the society as a whole, because technological innovation is one of the driving forces behind a country's economic prosperity and productivity advance (**?**). On the other hand, our study shows that providing relative performance information improves the performance of high-achieving students whereas low-achieving students perform even worse. This widening of the performance gap caused by feedback may be translated into a wage gap later. We find evidence to this direction using students expected wages. This might be detrimental especially for low achieving students. An economist may be fond of the efficiency achieved though information provision as high achieving students end up higher in the society and the spillover effects of the technological advances to the whole society. Nevertheless, at the end of the day its up to the society to decide whether efficiency can be traded for equality.

Additionally, our descriptive statistics evidence show that providing the relative performance information may encourage students from low income families to enroll in university and especially to more selective programs. From this perspective, providing the relative performance information encourages social elevation for students coming from low income neighborhoods. Thus, feedback decreases the performance or income inequality between students coming from low and high income neighborhoods.

5 Mechanisms

In this section, we discuss the most likely mechanisms that could explain our findings. Although it is impossible to distinguish between students' and parents' reaction to the social comparison information, we are able to rule out the possibility that teachers are driving the results. This is because the national exams are externally marked and teachers have no way to affect these grades. Neither students nor parents can select which school to attend, because this allocation is centrally managed and is based on geographical criteria.[24]Additionally, teachers cannot allocate students to classes in a way that facilitates sorting because it is prohibited by the law. Students are allocated to classes based on alphabetical order and they cannot switch classes within the school. Teachers have neither direct nor indirect financial incentives to react to the social comparison information. Teachers' compensation is not linked to teachers' quality nor is it a function of school-based performance: it is based on years of teaching experience and level of education attained. Also, schools' financing is not affected by their students' success in the national exams.

5.1 Mechanism 1: Priors- Positive Vs Negative Surprise

In this section, we examine whether students respond to the specific type of feedback that they get. Students might not only compare themselves with their class, school or cohort but they may also compare their own relative performance in different periods in time. We exploit within-school variation in the 134 senior high schools and we restrict this part of the analysis to the feedback years.

A recent paper by **?** highlights the importance of students' priors, when evaluating the effects of feedback. The authors provide relative performance information to university students, and that decreases their short-term performance. After conducting a survey, the authors find that students tend to be under-confident and the provision of feedback only increases their self-reported satisfaction.

Although we do not observe students' exact priors, we assume that their priors will be a function of their tenth-grade performance. In the feedback regime, students update their priors using the eleventh-grade performance. If a student receives information that he is in a higher decile in the eleventh grade than in the tenth grade, then the student receives a positive shock, that can be translated into a "positive surprise". On the other

[24]Parents have to submit an application to the local authority with proof of address

hand, if the student discovers that he is in a lower decile in the eleventh grade than in the tenth grade, then this student receives a “negative surprise”. Intuitively, students who receive a positive surprise in the eleventh grade might increase their expectations of themselves and exert more effort in the twelfth grade, whereas they might reduce their effort if they receive a negative surprise. In Figure 12, we graph the effect on the twelfth-grade rank for each combination of deciles in tenth and eleventh grade.

The horizontal axis represents the eleventh-grade rank of students, and the vertical axis represents the tenth-grade rank. Different colours express different magnitudes of the treatment effects on the twelfth-grade rank. The diagonal starting from zero towards the right upper edge of the box, represents the case of ”no value feedback”: those students whose eleventh-grade percentile rank equals their tenth-grade percentile rank. The treatment effect is positive for most students experiencing a positive surprise. These are students who are on the right of the diagonal of ”no value feedback”. On the left of the diagonal, feedback effects are mainly negative, meaning that students' twelfth grade rank declines when they receive a negative surprise.

A concern here is that students might not be aware of their tenth grade percentile rank, especially if they attend a school with more than one class. However, the analysis here uses deciles of performance and not percentiles, allowing students to have priors that do not accurately express their exact tenth grade rank.

5.2 Mechanism 2: School quality revelation

An alternative mechanism could be that students use the information obtained by the publication of their scores to infer the quality of their senior high school.[25]Students who take the eleventh grade national exams discover their school rank and their national rank, and the comparison of the two ranks reveals information about the quality of the school. If a student discovers that his national rank is greater than the school rank then his school is of good quality. Conversely the school is of lower quality if the national rank is lower than the school rank. The revelation of the school quality in the eleventh grade might affect students' choice of effort in the twelfth grade. Thus, we exploit the across schools variation in quality to identify the effect of feedback on students' rank

[25]We measure school quality based on the schools' average national exam performance in the twelfth grade from 2003 to 2009. Then we construct a rank measure for school quality that varies from zero to one. The average quality of the schools in our sample is 0.52 (sd 0.21, min, 0.018 and max. 0.985) which means that our school sample is of a representative quality.

nationwide.

In Figure 13, we produce the treatment lines separately for students who discover that the school they attend is worse (on the left) and better (on the right) than the average quality school. In Figure 14, we repeat the same exercise and we produce these figures using the standardised national exam score.[26] The average effect for students who realise that they attend a worse-than-average quality school is negative, whereas it is positive for those who realise that they attend a better-than-average quality school.

Starting with the bottom of the prior performance distribution, we observe that low achieving students in good schools do better that those in lower quality schools. Surprisingly, there is a huge increase in the national rank for the top students in the worse schools and this increase even offsets the increase in the national rank of the top students in good schools. There are two possible explanations: first, high performing students in the low quality schools take the eleventh grade national exams and when they receive feedback, they realise that they are actually exceptional on a national scale. Thus they might decide to exert more effort in the next time period, so feedback acts as a motivation boost for these students.

Second, the realisation of their national rank acts as a rude awakening for these students who might initially have a wrong perception about the national competition and about their school's quality. These students might be the top students in their class or school but they now learn that they are behind. In the next time period, they exert more effort to catch up with the national standards.

If students realise the quality of the high school through the eleventh grade national exams, then the response to the feedback would be more consistent across the school group. For example, if students from all parts of the performance distribution in school X discover that their school is of low quality and they are concerned with university admission, they might all exert more effort to catch up with national standards.

5.3 Mechanism 3: Parental investment

Another possible mechanism is that parents decide to invest more or less in students based on the eleventh grade results. Parents may start devoting more time helping the child with the homework or they may invest in external support (such as supplementary material/books, private tutors etc). It is true that there might be variation in family

[26] Standardised within each year with zero mean and a standard deviation of one.

income within the school and the neighborhood income represents the average income in each region. We observe considerable differences in neighborhood income.[27]

In Figure 15, we draw the treatment lines for the bottom and the top quintile of neighbordhood income. We run specification (2b) separately for the top and the bottom quintiles of neighborhood income. This may not fully reflect the family income but we examine the effects of feedback across regions by average reported income. A wealthy family may have the financial resources to invest in the child and thus the student may improve his performance in the subsequent year exam. On the other hand, families from low income neighborhoods may not be able to pay enough to further support the student. In Figure 16, we observe that disclosing rankings increases the average subsequent national rank for students coming from the highest-income neighborhoods. The average effect on the subsequent national rank for students from the lowest-income neighborhoods is negative. In high-income neighborhoods the positive effects of feedback hold for students above the 40th percentile while only students above the 60th percentile from low income neighborhoods benefit from feedback. If parental investment was the only driver of the findings, then we would expect students from highest-income neighborhoods to improve at all parts of the prior-performance distribution. That implies that there might be, to some extent, differential parental investment in students by family income (proxied by neighborhood income) but that cannot fully generate our results.

5.4 Mechanism 4: Practice

It could be argued that students can accurately place themselves within their class, even if they are not explicitly informed about their rank. This is likely to occur due to repeated interactions among classmates throughout high school. However, here students receive new information that is broader. Consider the within school rank: students receive information about how well they did within their school. In Figure 16, we report the treatment lines for students in schools of different capacity in the eleventh grade. We make four broad categorisations. First, we consider schools with only one class where it is likely that students already know their relative standing and the social comparison information has no extra value (Panel A). Nevertheless, in a school with only two classes students might know their relative performance in their class but not

[27] mean: 23,517 standard deviation: 8,609 min.:13,005 max.: 66,521

in the whole school. Thus, we see that there is a small positive feedback effect on students who are above the 40th percentile and a small negative effect on those below it (Panel B). Additionally, the treatment lines become steeper when we consider schools with three classes (Panel C). In this case, the information is much broader than that which students can collect from interaction with their classmates. This is even more pronounced when we look at students in schools with more than three classes (Panel D). Summary statistics about the capacity of schools in our sample are presented in Table 6. Figure 16 shows that the effect of feedback depends on whether the additional information is actually informative about someone's relative performance.

That could allay the concern that the eleventh grade national exam might provide students with experience or training instead of information about their relative performance. School exams in the eleventh grade have the same format as national exams in the eleventh grade and the past papers are available in both regimes. Students practice on past questions and are aware of the structure and the types of questions in both cases. If students were experienced from sitting the eleventh grade national exams, then the experience or training effect would not vary by the size of the school. In other words, if that was the mechanism then students in small schools would have no reason to react differently than students in regular schools.

5.5 Mechanism 5: Learning about own ability

Another possibility is that students have imperfect information about their own ability and they compare their own absolute score with the average school/cohort score to infer their own ability. We adapt a theoretical model proposed by **?**.[28] In the non-feedback regime, students in the eleventh grade sit school exams and they receive information about their own absolute performance only. In the feedback regime, they receive information about their own performance, but also about the school and cohort average performances.

Students take exams in two time periods; the eleventh and the twelfth grade.

Period 1: This is the learning stage. The eleventh grade own performance provides students with some information about their ability (and the easiness of the exam). This

[28] **?** presents a principal-multiple agents model where agents have imperfect information about their abilities under multiple types of contracts. The model is also used by **?**. The natural experiment they study gives students information about the average class grade, while here the social comparison information refers to the average school and cohort grade.

performance acts as a private signal s_i for the student. In the feedback regime, students also observe the average score in the school or nationwide $\bar{s} = \{\bar{s}_{school}, \bar{s}_{cohort}\}$ which is the average signal in the school and the country respectively. In the feedback regime student i may compare his own signal with the average signal (in the school and/or in the cohort) and that may affect student's perceived belief about his own ability. That could, in turn, determine the amount of effort he decides to exert in the second period. The amount of effort students decide to exert in the twelfth grade affects their final year's scores.

Period 2: Following the realisation of the signals, in the second period students choose the effort to exert (e_i). Students' objective is to maximize the second period performance (q_i) after choosing the effort to exert. Assuming that the performance production is a linear function in effort and that effort and ability are complements[29] in the performance production function[30] it follows that: $q_i = e_i \alpha_i$. There is also a cost associated with the effort exerted that is $c(e_i)$ and is increasing in effort and convex. [31] In the absence of the social comparison information students receive only the private signal and they maximise:

$$u^{NF} = E[p_i(\alpha_i, e_i) - c(e_i)|s_i] = E[\alpha_i|s_i]e_i - c(e_i)$$

and the F.O.C simplifies to $E[\alpha_i|s_i] - c'(e_i^{NF*}) = 0$ (1)

In the feedback regime where social comparison information is provided the student observes the average signal (which could be either the school average signal or the cohort average signal) and maximises:

$$u^F = E[p_i(\alpha_i, e_i) - c(e_i)|s_i, \bar{s}] = E[\alpha_i|s_i, \bar{s}]e_i - c(e_i)$$

and the F.O.C simplifies to $E[\alpha_i|s_i, \bar{s}] - c'(e_i^{F*}) = 0$ (2)

The proof can be found in the web appendix and it comes from the comparison of equations (1) and (2). We also assume that the private signal is given by $s_i =$

[29] Notice here that there is no pass-fail scheme and students do not try to achieve a performance threshold. University cut-offs are determined endogenously based on demand and pre-specified supply of seats. In other words, the model makes these predictions based on the fact that ability and effort are complements in the production function. In a different setting where university cut-offs are pre-determined, effort and ability could be substitutes in the production function. In that case, a student who is above average in the eleventh grade may choose to exert less effort in the twelfth grade in order to achieve a specific performance threshold.

[30] $\frac{dqi}{d\alpha_i de_i} > 0$

[31] $(c'(e_i) > 0, c''(e_i) > 0, c'(0) = c''(0) = 0)$

$\alpha_i + c for i = 1, 2, ...$ and it depends on student's i ability level (α_i) and a shock that is common to all students[32] ie. the easiness of the exam (c). Let us summarize now the main hypothesis about the effect of the eleventh grade social comparison information on the twelfth grade performance.

Null Hypothesis: Students do not react to the social comparison information

That would suggest that students are not uncertain about their ability or that students have already figured out their relative performance information and the explicit addition of it is redundant or that the private signal that students get in the feedback regime equals the average signal.

Alternative Hypothesis: Positive effect for high ability students and negative effect for low ability students

That would suggest that students will react differently to feedback. Based on the model, high ability students will perform better when the social comparison information is provided. On the other hand, low ability students will perform worse when the social comparison information is provided. Our findings support the alternative hypothesis, implying that students whose eleventh grade absolute score is above the school/cohort average score, might be encouraged by their relative performance and that makes them exert a higher amount of effort in the twelfth grade. On the other hand, students who realise that they score below the school/cohort average might be discouraged by that and exert a lower amount of effort in the twelfth grade. This channel highlights the importance of non-cognitive skills on educational outcomes and especially self-perception about own ability and confidence. The importance of non-cognitive skills is well established in the literature (**?**, **?**, **?**)

6 Threats to identification

Attrition

In our attempt to evaluate the impact of feedback on different performance groups, the problem of attrition cannot be ignored. If attrition is random and affects different performance groups in a similar way in both regimes, then the estimates remain unbiased.

[32] In the school or the cohort depending on the feedback or non-feedback regime.

Differential attrition here could arise because students from the lowest percentiles are more likely to drop out from school in comparison to students from the highest percentiles, when they realise their relative performance. What could bias our estimates, is if differential attrition follows the abolition of feedback.[33] In Figure 17, we observe that attrition rates differ for each quintile of prior performance but the attrition rates following the abolition of feedback do not change dramatically compared to previous years.

Notice here that students drop from our sample either because they drop out from school or they move to a different school. The unique student code that identifies students across grades within a school changes if the student switches to another school. We cannot follow students who move to a different school.

Exploiting within school variation, we use the following specification to check for differential attrition that changes with feedback:

$$Drop - out_{12-10isc} = \alpha + \beta_{quintile} Feedback_c * Quintiles_{10isc} + \lambda_{quintile} Quintiles_{10isc}$$

$$+\psi Feedback_c + X'\gamma + \theta_c + \varphi_s + \epsilon_{isc}$$

Table 11 reports OLS results. The attrition rate is larger for the lowest quintile than any other, compared to the third quintile, when feedback is provided. But most importantly, none of the coefficients of interest are statistically significant. This implies that there is differential attrition, but it does not vary with feedback policies.

Robustness checks

In this section, we construct robustness exercises to complement our main analysis.

One concern might that the change in the variation of performance over time is caused by time trends and not the provision of ranks. Exploiting the within school variation [34], we run specification (2b) but without pooling feedback and non-feedback years together. Instead, we just compare every pair of consecutive years in the sample. We present the placebo treatment lines in Figure 18. Panel A compares the cohort 2003 to the 2004, as if feedback was abolished in 2004. We find similar cohort behavior

[33] The first affected cohort for which feedback is abolished is the cohort that was in the twelfth grade in 2006. Thus, this cohort was in the tenth grade in 2004. This is the first cohort that did not sit national exams in the eleventh grade.

[34] very similar results if we exploit the across schools variation.

from 2003 to 2005 as the treatment lines are flat for these pairs of years. The only pair of years that we expect to find a differential response of cohorts is 2005-2006 (the year of the reform). Panel C corresponds to the actual reform and we observe that the treatment effects are negative for all percentiles below the 50th percentile and positive above it. For every other pair of years, we expect to find similar cohort behaviour. We find no evidence that other time specific factors could generate our results or drive ours results. Regarding any policy anticipation effects, the reform was announced in December of 2003-2004. We find very small treatment effects in Panel D, which is the first non-treated cohort. Students in the first non-feedback cohort might observe how last year's peers of similar tenth grade performance did and use this information to slightly adjust their behavior. Again after 2007, the curve is almost flat throughout the ability distribution implying a similar cohort behavior.

We conduct some other placebo exercises to verify that the effect does not depend on the numbers of subjects examined. In Figure 19, we draw the treatment lines for each subject separately. Before 2005, students take national exams on five core or general subjects (Modern Greek, Mathematics, History, Biology and Physics). From 2005 onwards, two core subjects are examined at a national level whereas the other three core subjects are examined at a school level. Students take national exams on Modern Greek which is the only compulsory core subject in both regimes. In the non-feedback regime, students choose the second core subject on which they take national exams among the options of: Mathematics, History, Biology and Physics and they sit school exams on the other three remaining subjects[35]. Panel D presents the feedback line in Modern Greek, which is compulsory examined at a national level in both regimes. We observe that the treatment lines follow a similar pattern for all subjects indicating that the number of subjects examined does not drive our results.

In Table 1, we calculate the twelfth grade rank based on different subjects. In column (1) we find the effect of feedback on the final year rank that takes into account the Electives or Track subjects on top of the core subjects[36] and the results are very similar to those reported so far. In column (2) we take into account the effect of feedback on students' performance in Modern Greek which is a common subject in both regimes

[35]Around 80% of students choose to take national exams on Mathematics together with the compulsory Modern Greek subject. These students sit school exams on: History, Biology and Physics.

[36]Students sit national exams in four Elective subjects. So the overall rank in calculated based on nine subjects.

and takes a special weight in the calculation of the university admission grade. Notice, that in the non-feedback regime two subjects are examined nationally and three within the school. In column (3) we calculate the twelfth year's rank based on the scores in the national exams: five subjects in the feedback regime and the two subjects in the non-feedback regime. Results remain very similar. Feedback effects remain positive for the top quintiles of prior performance, whereas it is negative for the bottom quintiles of prior performance.

7 Conclusion

In this paper, we examined the effects of providing information on a student's rank on the student's short- and long-term outcomes. Knowing one's rank may affect investment decisions and, thus, later productivity. Following an unexpected policy change that took place in Greece, we carried out a large-scale, primary data-collection process. Using unique, detailed data on students' performance throughout senior high school and school quality data, we examine the effects of receiving information about the relative performance of students within their own school and across the nation.

We find that disclosing information on rankings has a positive effect on high-achieving students' short- and long- term outcomes. In particular, we find the following results for high-achieving students: Feedback information improves their subsequent performance by 0.15 of a standard deviation; they enroll into more selective university departments by 0.15 of a standard deviation, and their expected annual earnings increase by 0.17 of a standard deviation. The effects on low achieving students are negative: Their subsequent performance drops by 0.3 standard deviations; they are admitted to university departments which are less selective by 0.18 of a standard deviation, and their expected annual earnings decrease by 0.20 of a standard deviation. We also find that the results are more pronounced for females, indicating greater sensitivity to feedback. Our results show that, absent feedback on rankings, high-achieving students are more likely to re-take the exams. The resulting delay of the most-able students into university and/or labor market is an important loss of human capital for society.

We also find suggestive evidence that feedback encourages students from low-income neighborhoods to enroll in university and to study in more selective programs. This may, in the long run, reduce income inequality.

We outline several potential mechanisms that may explain why students react to the provision of feedback: 1) priors of students, 2) school quality revelation, 3) parental investment, 4) practice, and 5) learning about own ability. The last mechanism highlights the importance of non-cognitive skills such as self-perception about one's own ability and confidence.

Our findings have important policy implications. Providing rankings is a low-cost instrument that has the potential to affect not only students' high school performance but also labour market outcomes. The relative nature of the above mentioned results restricts the broad implementation of providing feedback, but makes it very important in a competitive process. If the social information is provided and parents can choose the best school for their child, then the relative position of the student among his school peers cannot be ignored. A crucial question concerns the social information transparency and future research is needed to understand which mechanism drives the effects. Our analysis highlights the importance of knowing one's rank on high-stake exams in influencing scholastic and labour market outcomes, and we believe that the rank could be a new factor in the education production function.

References

Figure 1: Map of schools in the sample

Figure 2: Timing

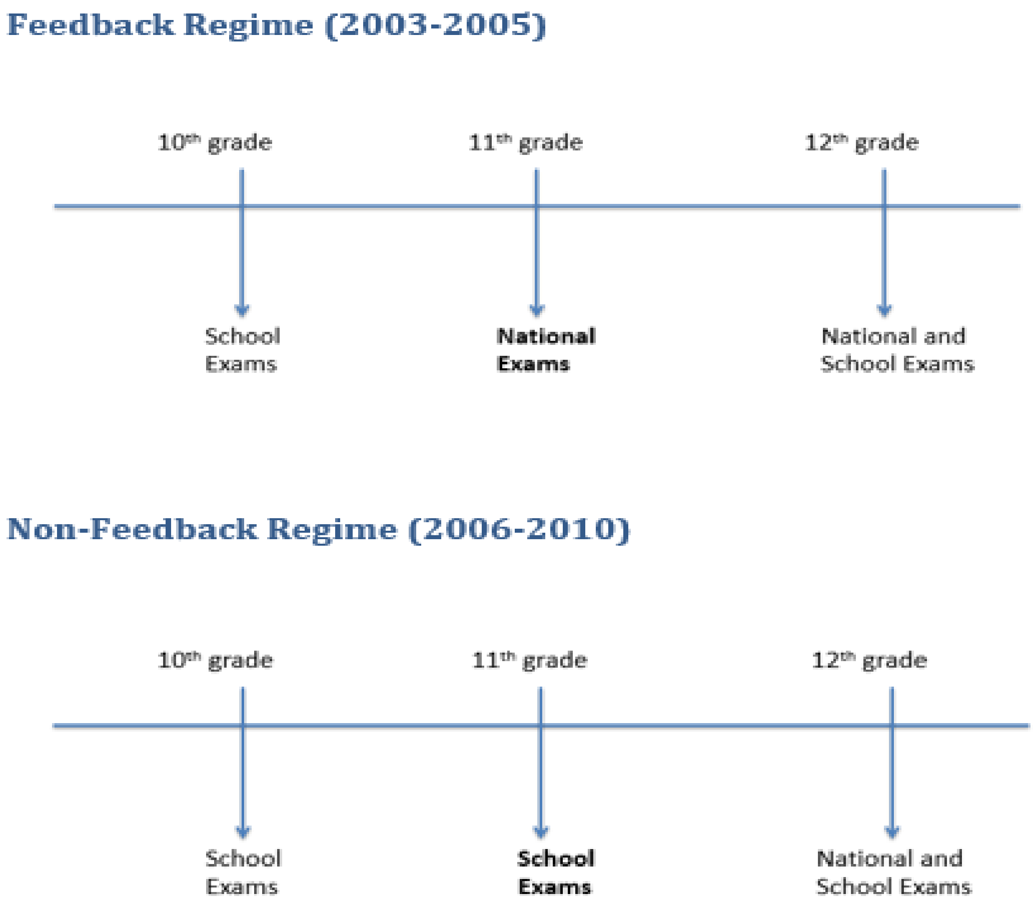


Figure 3: Announcement of results

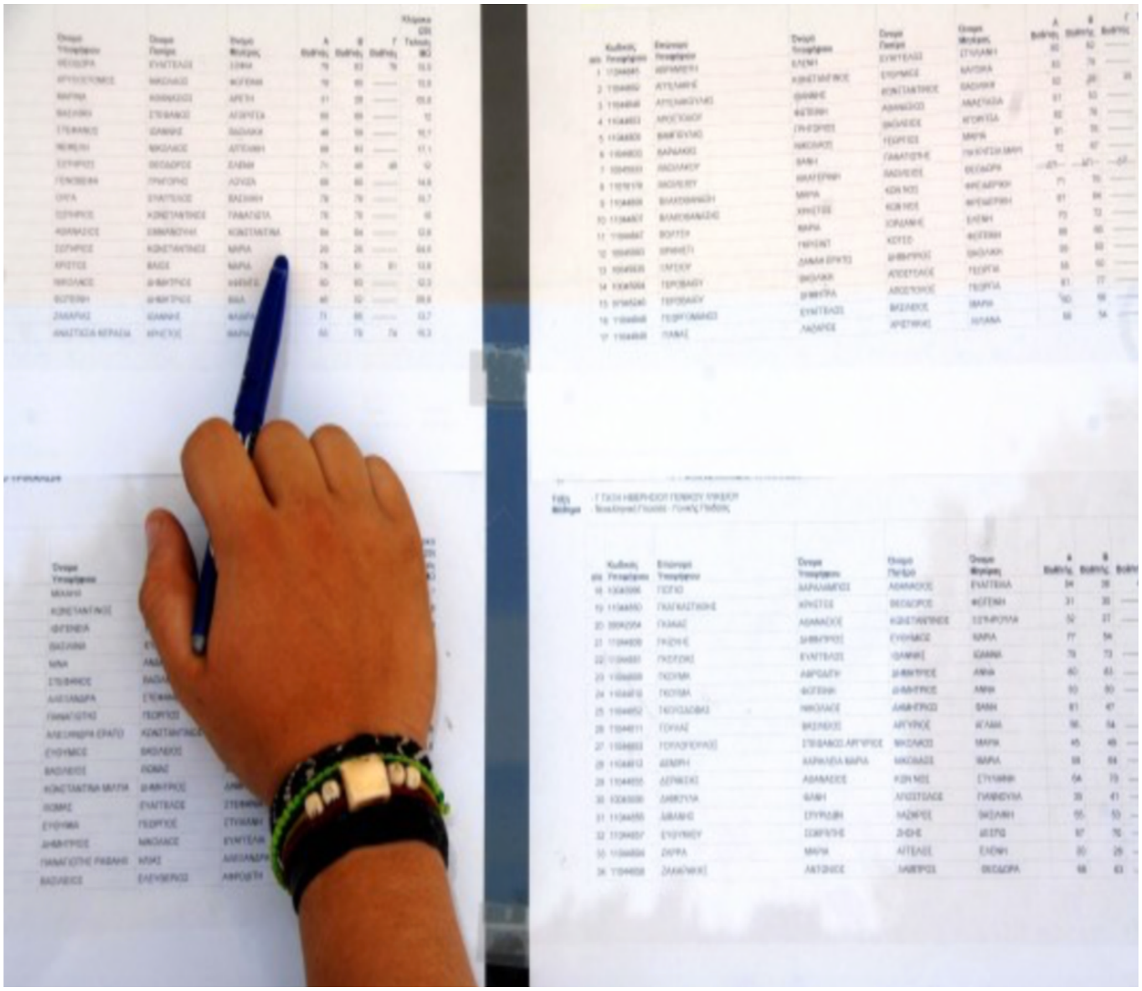
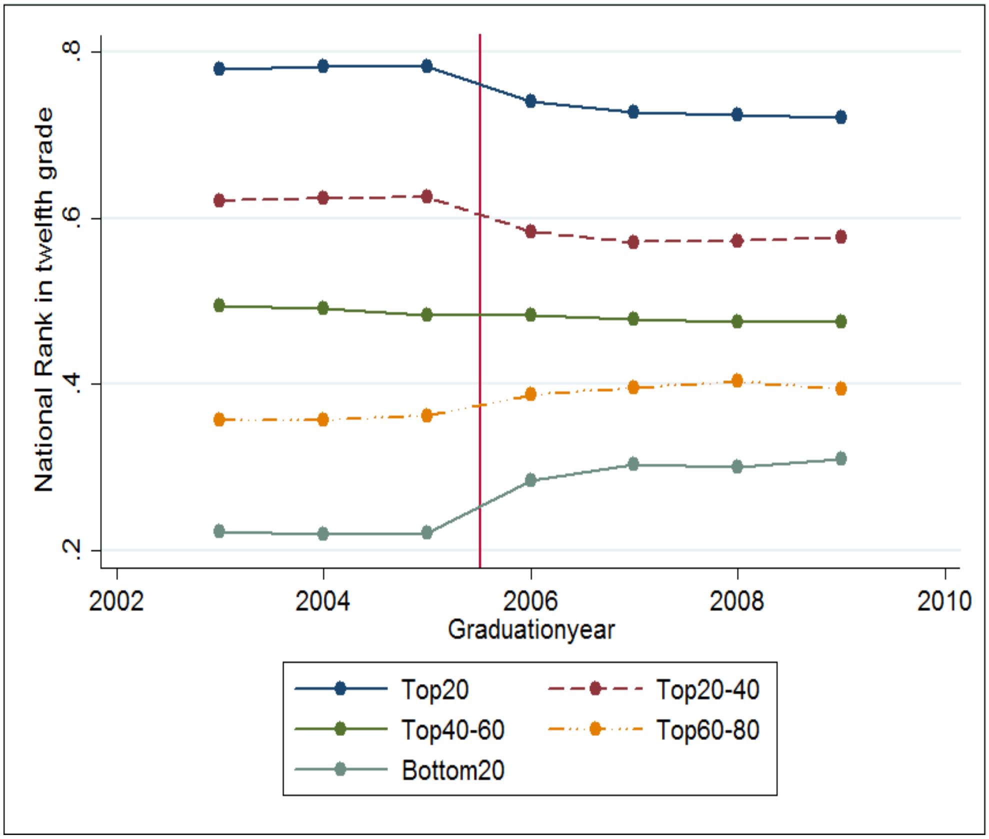


Figure 4: Announcement of school results-Zoom in

Α/Α	Κωδικός υποψηφίου	Επώνυμο	Όνομα	Όνομα πατρός	Α Βαθμός	Β Βαθμός	Γ Βαθμός	Τελικός Βαθμός
2	15030628	ΑΛΑΒΕΡΑ	ΜΑΡΙΑ ΕΛΕΝΗ	ΙΩΑΝΝΗΣ	84	79		16,3
4	15030629	ΒΑΡΚΑ	ΙΩΑΝΝΑ	ΘΕΟΔΟΣΙΟΣ	93	84		17,7
8	15030630	ΔΕΜΑΓΚΟΣ	ΚΩΝΣΤΑΝΤΙΝΟΣ	ΙΩΑΝΝΗΣ	27	38		6,5
14	15030631	ΕΞΑΡΧΟΠΟΥΛΟΥ	ΠΑΝΩΡΑΙΑ	ΔΗΜΗΤΡΙΟΣ	80	89		16,9
17	15030632	ΖΕΓΙΑΝΝΗ	ΜΑΡΙΑ ΠΑΝΑΓΙΩΤΑ	ΜΙΧΑΗΛ ΣΤΑΜΑΤΙΟΣ	51	45		9,6
19	15030633	ΚΑΠΠΑ	ΙΩΑΝΝΑ	ΠΕΤΡΟΣ	73	74		14,7
21	15030634	ΚΑΡΑΚΑΤΣΙΩΝΗ	ΒΑΣΙΛΙΚΗ	ΚΩΝΣΤΑΝΤΙΝΟΣ	33	22		5,5
22	15030635	ΚΑΡΠΑΘΑΚΗΣ	ΚΩΝΣΤΑΝΤΙΝΟΣ	ΙΩΑΝΝΗΣ	77	57	70	14,7
23	15030636	ΚΑΡΠΑΘΙΟΥ	ΝΟΜΙΚΗ ΗΛΙΑΝΑ	ΜΗΝΑΣ	40	34		7,4
24	15030637	ΚΑΤΑΚΑΛΕΑΣ	ΑΝΤΩΝΙΟΣ	ΠΑΝΤΕΛΗΣ	82	90		17,2
25	15030638	ΚΑΤΣΑΜΑΚΗ	ΧΑΡΑ ΣΕΒΑΣΤΙΑΝΑ	ΕΥΠΡΕΠΙΟΣ	52	48		10
29	15030639	ΚΟΥΜΠΟΓΙΑΝΝΗΣ	ΑΝΑΣΤΑΣΙΟΣ	ΑΘΑΝΑΣΙΟΣ	90	93		18,3
31	15030640	ΚΡΙΜΙΖΗ	ΧΡΥΣΟΒΑΛΑΝΤΟΥ ΔΟΜΝΙΚΗ	ΜΙΧΑΗΛ	28	24		5,2
32	15030641	ΚΥΔΩΝΑΚΗ	ΣΟΦΙΑ	ΑΝΤΩΝΙΟΣ	80	91		17,1
34	15030646	ΜΑΥΡΙΔΟΥ	ΑΙΚΑΤΕΡΙΝΗ	ΑΛΕΞΑΝΔΡΟΣ	14	15		2,9
43	15030648	ΝΤΙΝΩΡΗ	ΕΛΕΥΘΕΡΙΑ ΣΑΒΒΟΥΛΑ	ΠΑΝΑΓΙΩΤΗΣ	91	95		18,6
45	15030649	ΟΙΚΟΝΟΜΟΥ	ΑΝΝΑ ΦΙΛΙΑ	ΝΙΚΟΛΑΟΣ	96	94		19
51	15030650	ΠΑΡΒΕΡΗ	ΣΕΒΑΣΤΗ ΕΥΑΓΓΕΛΙΑ	ΑΝΑΣΤΑΣΙΟΣ	87	88		17,5
57	15030652	ΣΜΑΡΑΓΔΑΚΗ	ΕΥΑΓΓΕΛΙΑ	ΓΕΩΡΓΙΟΣ	100	99		19,9
58	15030656	ΣΩΤΗΡΙΔΗΣ	ΔΗΜΗΤΡΙΟΣ	ΓΕΡΑΝΤΙΟΣ	77	77		15,4
60	15030659	ΤΡΙΑΝΤΑΦΥΛΛΟΥ	ΔΗΜΗΤΡΑ	ΘΕΜΕΛΗΣ	67	53	68	13,5

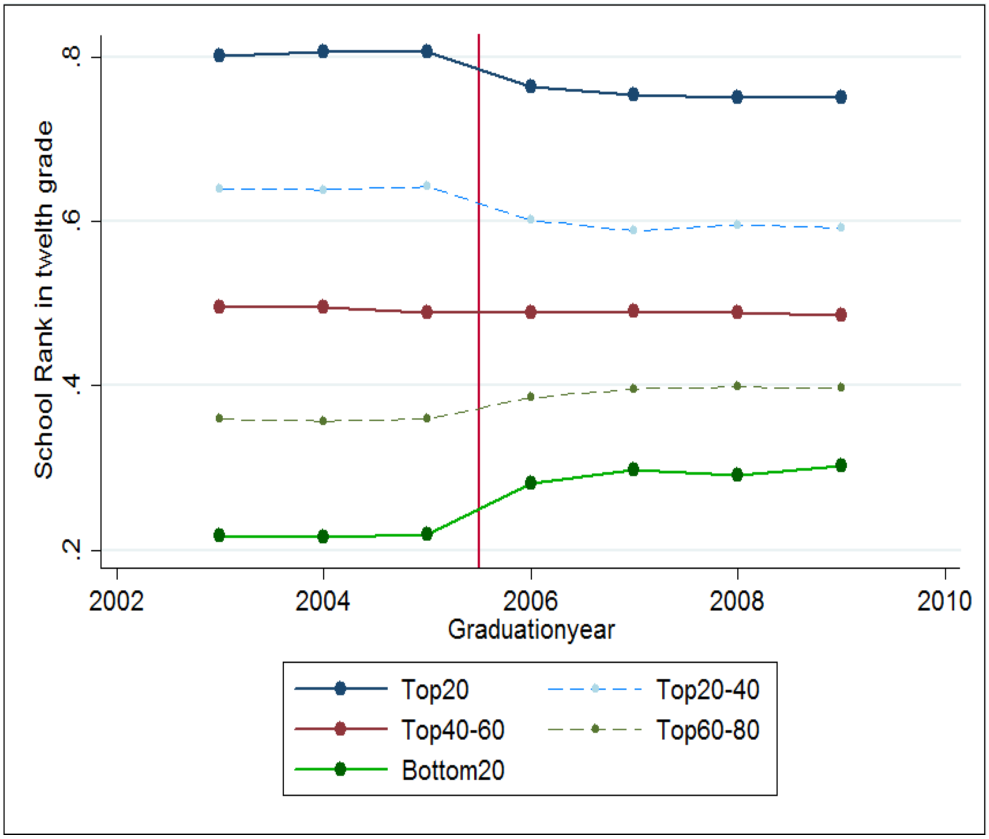
Note: This is the format of the publicly announced results. In particular, publicly announced are the following: a random code, a unique code, a student's surname, a student's first name, the father's name, the score given by the first examiner, the score given by the second examiner, the score given by the third examiner-if necessary- and the final score. The final score is the average of the score given by the first and second examiners. Students are sorted by alphabetical order. If the difference between the score given by the first and the second examiner is greater than or equal to 13, then a third examiner is required. The scores given by the first, second and third-if necessary- examiners are from 0 to 100 while the final score ranges from 0 to 20. If a third examiner is required, then the final score is the average between the two highest scores given by any examiners.

Figure 5: Time trends for twelfth grade rank nationwide



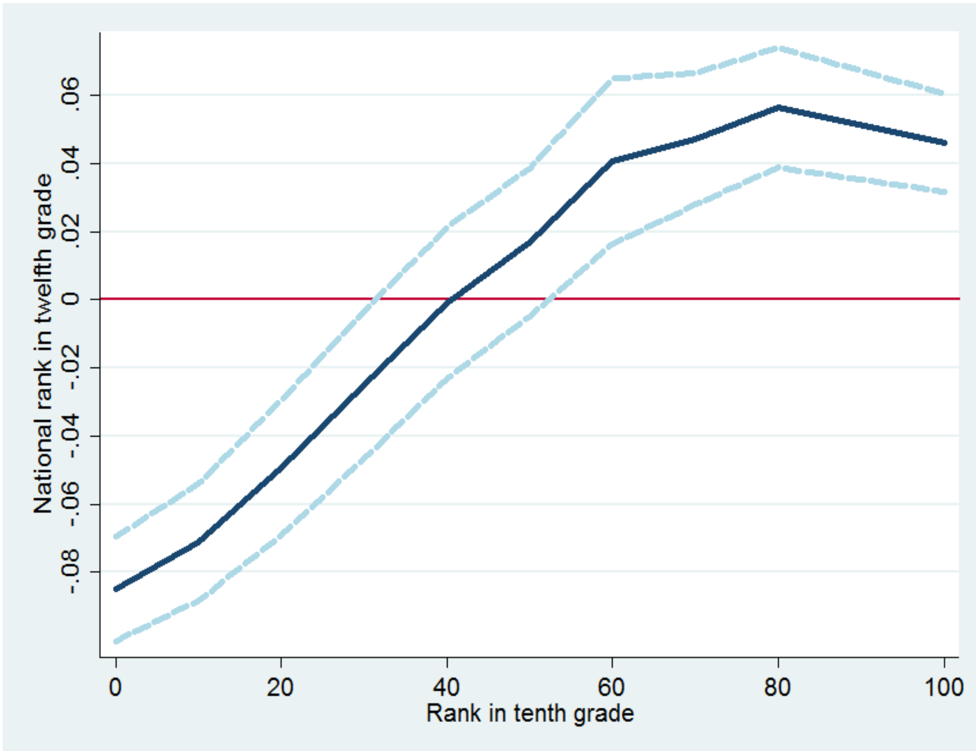
Note: Feedback provision for cohorts 2003-2005.The 2006 cohort is the first one for which feedback is abolished. Outcome variable: The national rank in twelfth grade. The trends correspond to different performance groups based on the tenth grade performance. The vertical line is between the last tenth grade cohort that receives feedback in the twelfth grade and the first cohort that does not receive feedback in the twelfth grade

Figure 6: Time trends for twelfth grade rank within the school



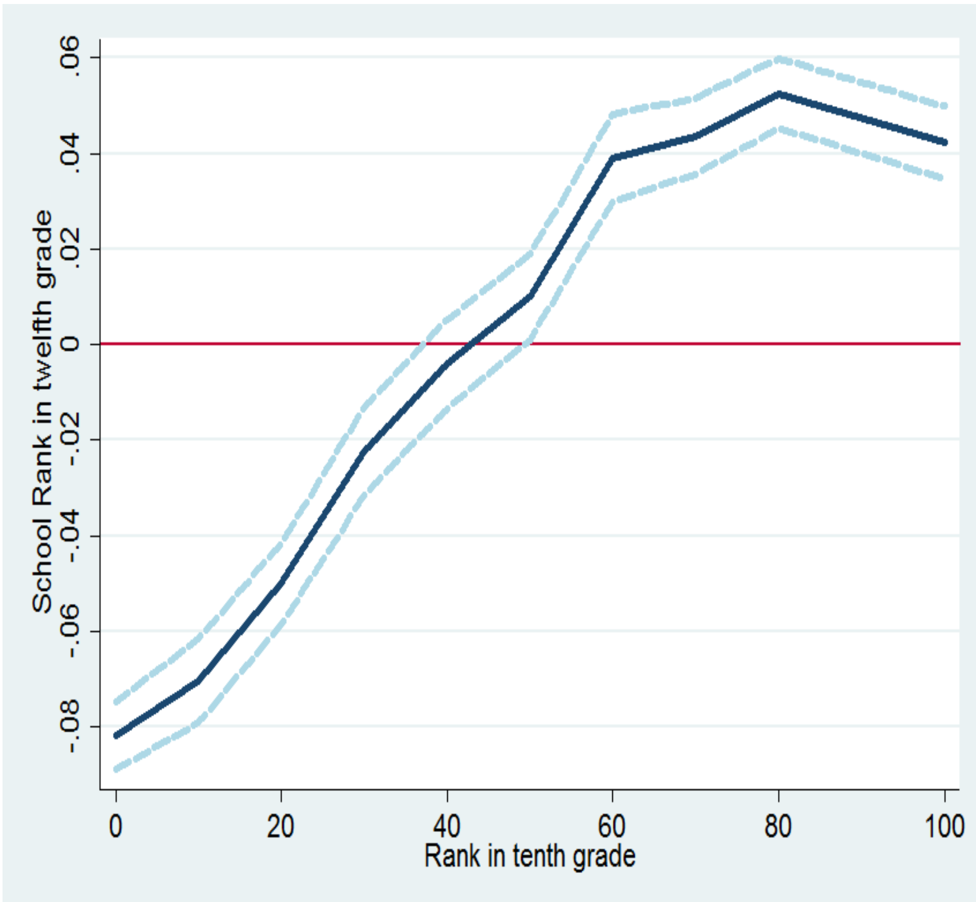
Note: Feedback provision for cohorts 2003-2005.The 2006 cohort is the first one for which feedback is abolished. Outcome variable: The rank in twelfth grade within the school. The trends correspond to different performance groups based on the tenth grade performance. The vertical line is between the last tenth grade cohort that receives feedback in the twelfth grade and the first cohort that does not receive feedback in the twelfth grade

Figure 7: Treatment effects on the rank nationwide conditional on prior performance



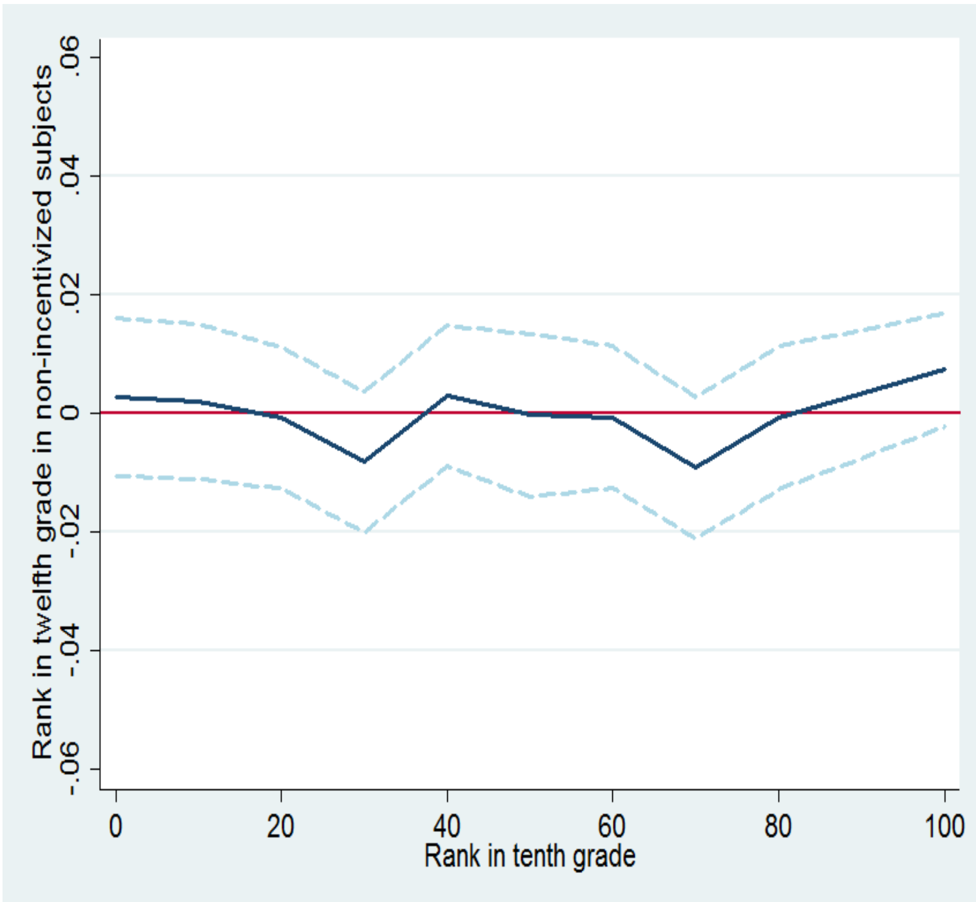
Note: The estimated effect of feedback on the national rank in the twelfth grade at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The national rank is calculated based on the five core-education subjects (incentivized). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 8: Treatment effects on the rank within the school in incentivized subjects conditional on prior performance



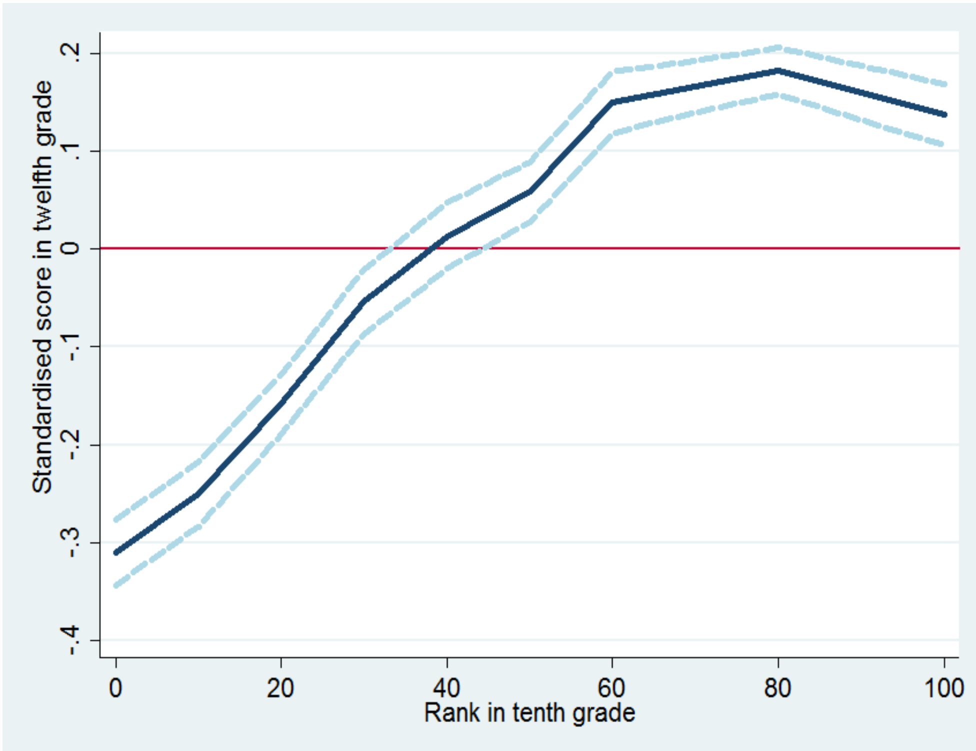
Note: The estimated effect of feedback on the school rank in the twelfth grade at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The school rank is calculated based on the five core-education subjects that students take in the twelfth grade and determine the university admission grade (incentivized subjects). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 9: Treatment effects on the rank within the school in non-incentivized subjects conditional on prior performance



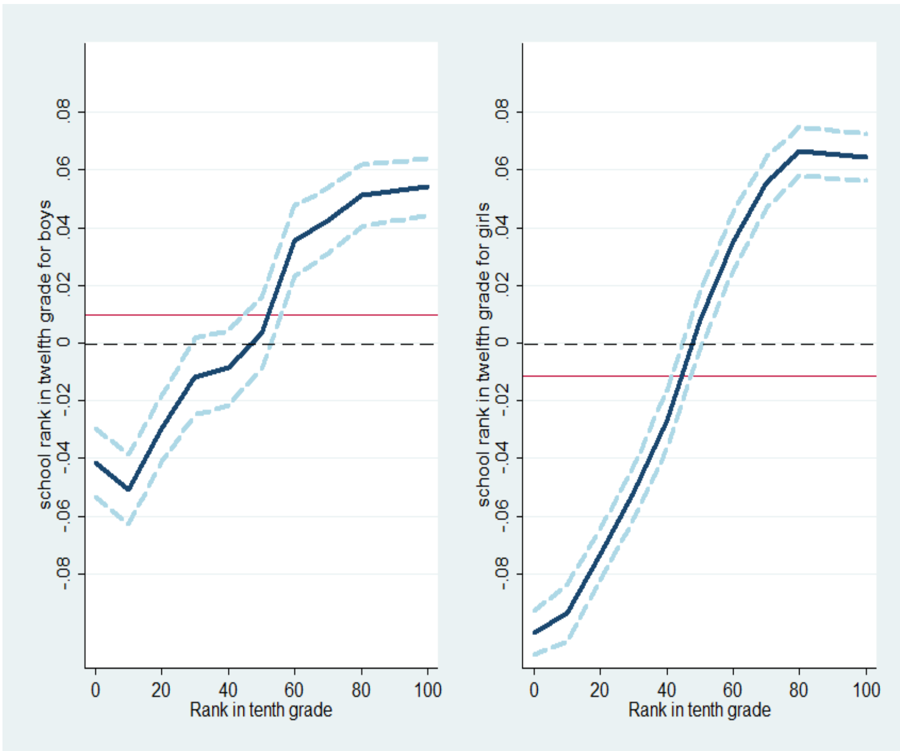
Note: The estimated effect of feedback on the school rank in the school exams at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The school rank in the school exams is calculated based on the three non-incentivised subjects that all students take in the twelfth grade and these subjects are not taken into account in the calculation of the university admission grade. Students never receive social comparison information in these subjects. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 10: Treatment effects on the standardised score conditional on prior performance



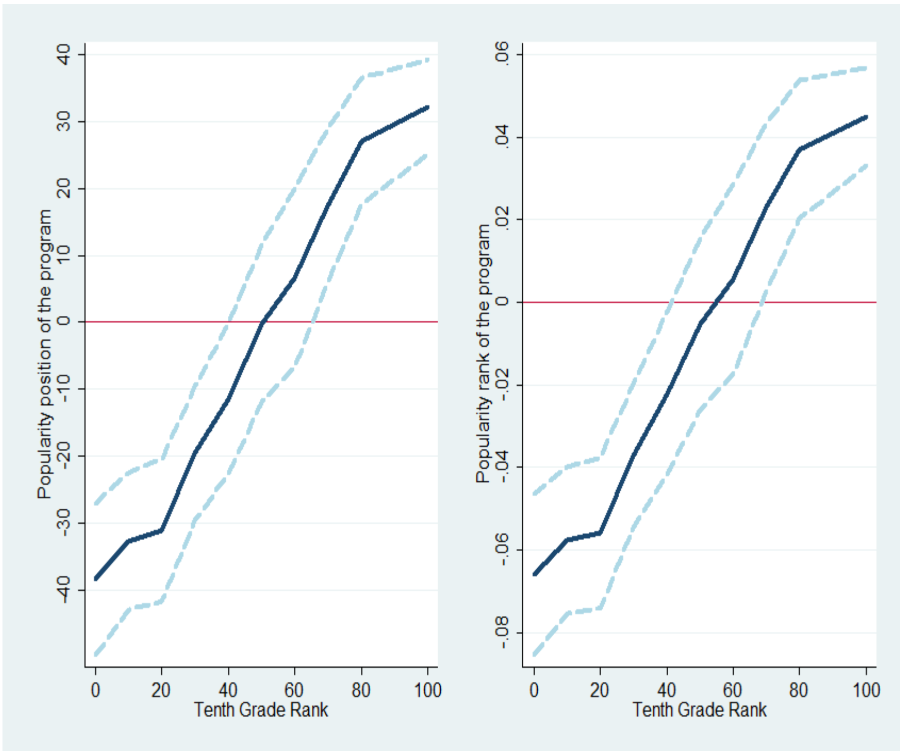
Note: The estimated effect of feedback on the standardised score in the twelfth grade at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The standardised score is calculated based on the five core-education subjects (incentivized). The standardised score has a mean of zero and a standard deviation of one in each year. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 11: Treatment effects on the rank within the school by gender conditional on prior performance



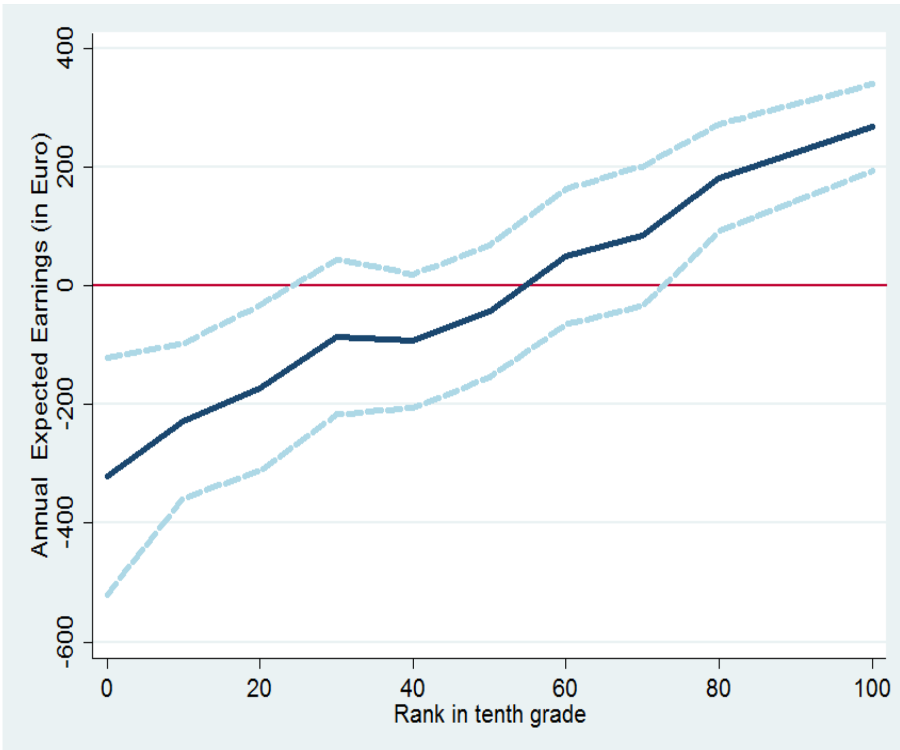
Note: The estimated effect of feedback on the school rank in the twelfth grade by gender at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. Males are depicted on the left and Females on the right. The school rank is calculated based on the five core-education subjects (incentivized). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 12: Treatment effects on the selectiveness/prestigiousness position and rank of the program admitted conditional on prior performance



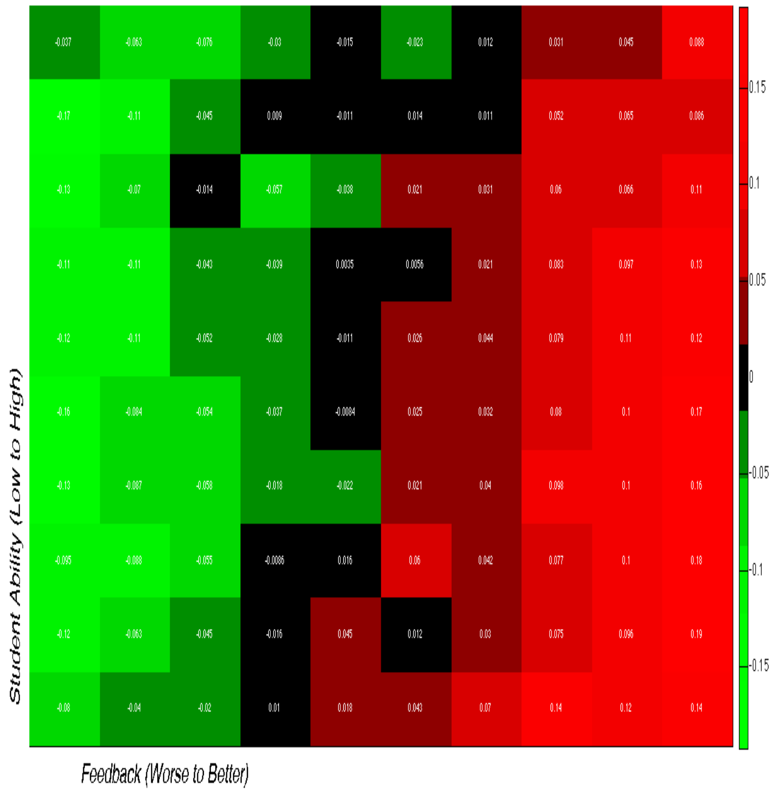
Note: The estimated effect of feedback on the popularity/prestigiousness position (on the left) and rank(on the right) of the program admitted and the associated 95 % confidence interval. There are 672 programs in total. Popularity position and rank measured by the average university department cut-off score over seven years. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 13: Treatment effects on the annual expected earnings conditional on prior performance



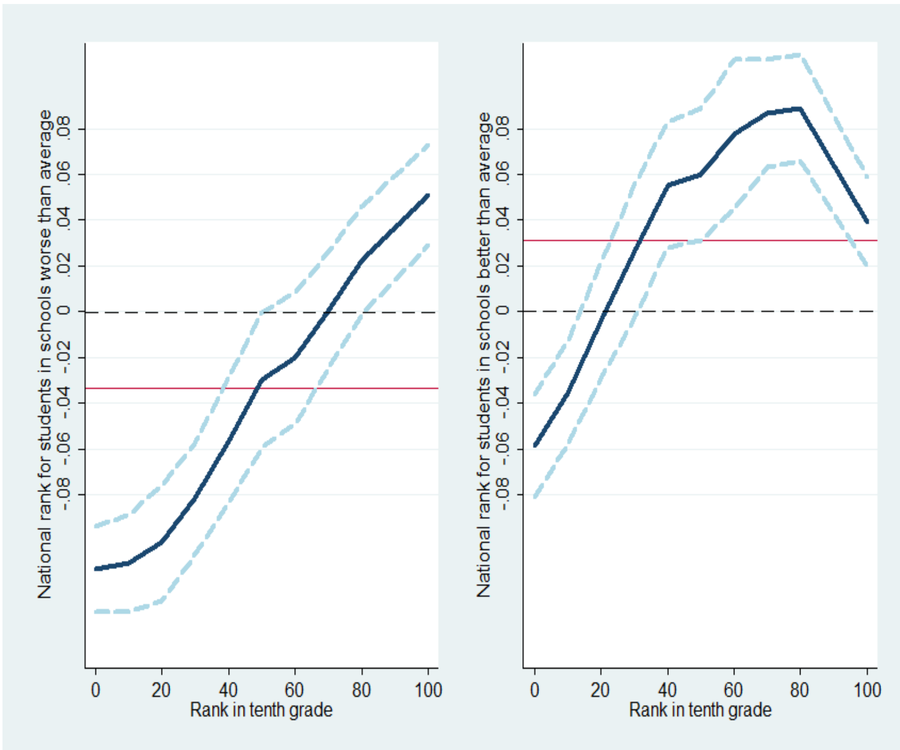
Note: The estimated effect of feedback on the expected annual wage at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The annual expected earnings are calculated based on the actual annual earnings of older graduates who studied the same college field. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 14: Positive and Negative Surprise



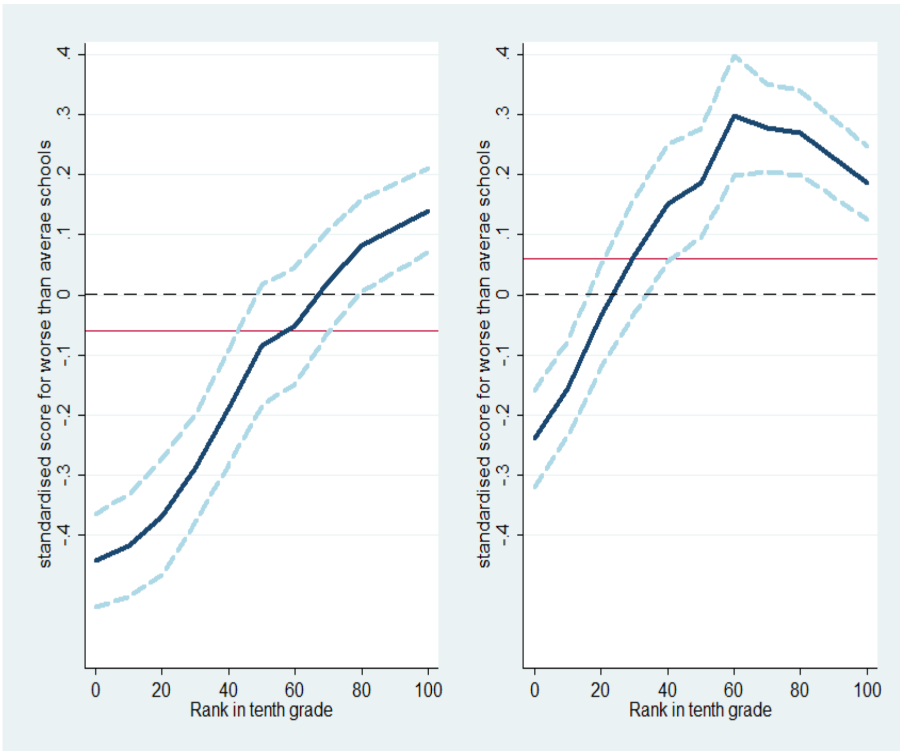
Note: Treatment effect for students with positive or negative surprise. Student performance (in deciles) in tenth grade on the vertical axis and student performance in eleventh grade (in deciles) on the horizontal axis.

Figure 15: Treatment effects on the rank nationwide by school quality conditional on prior performance



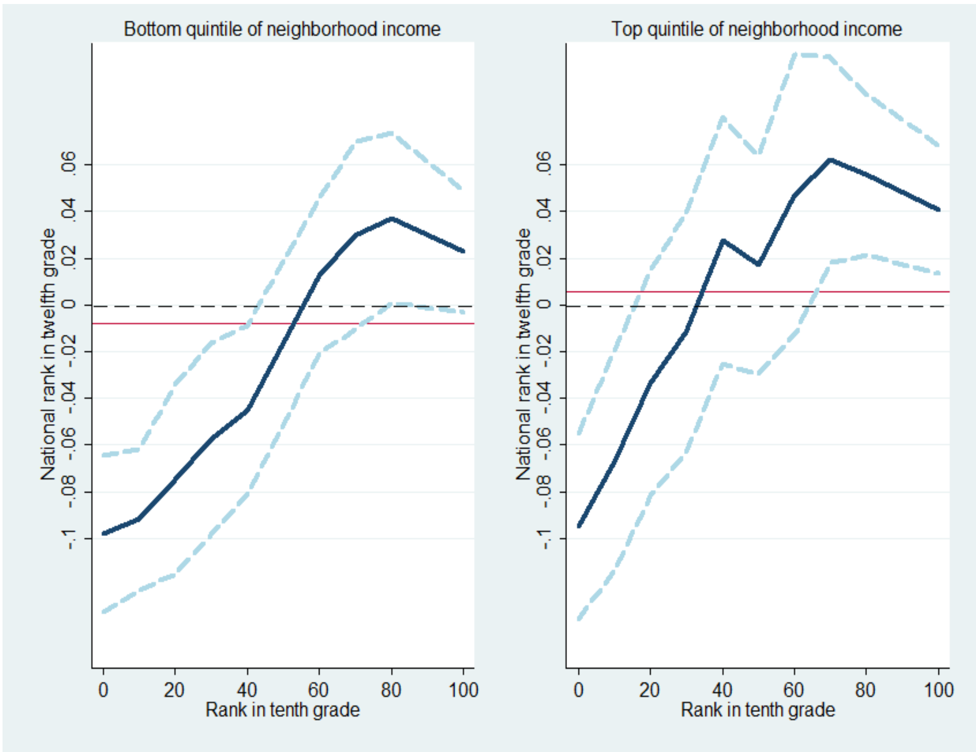
Note: The estimated effect of feedback on the national rank in the twelfth grade by quality of school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The effect of feedback on students' national rank when they discover they attend a school that is worse than the average quality school (on the left) and better than the average quality schools (on the right). The national rank is calculated based on the five core-education subjects (incentivized). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 16: Treatment effects on the standardised score by school quality conditional on prior performance



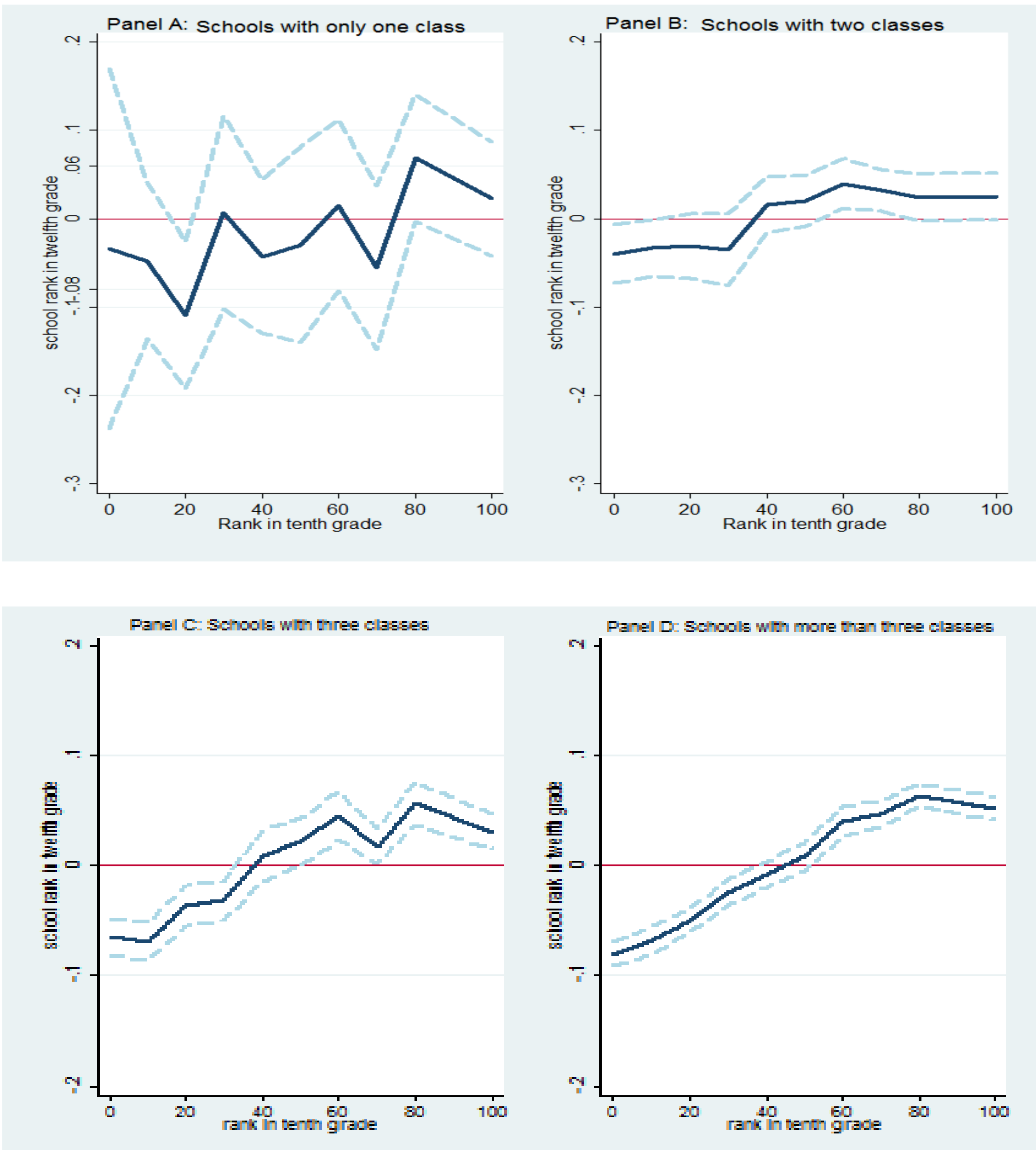
Note: The estimated effect of feedback on the standardised score in the twelfth grade by quality of school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The effect of feedback on students' standardised score when they discover they attend a school that is worse than the average quality school (on the left) and better than the average quality schools (on the right). The standardised score is calculated based on the five core-education subjects (incentivized). The standardised score has a mean of zero and a standard deviation of one in each year. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 17: Treatment effects on twelfth grade national rank for the bottom and top quintiles of neighborhood income



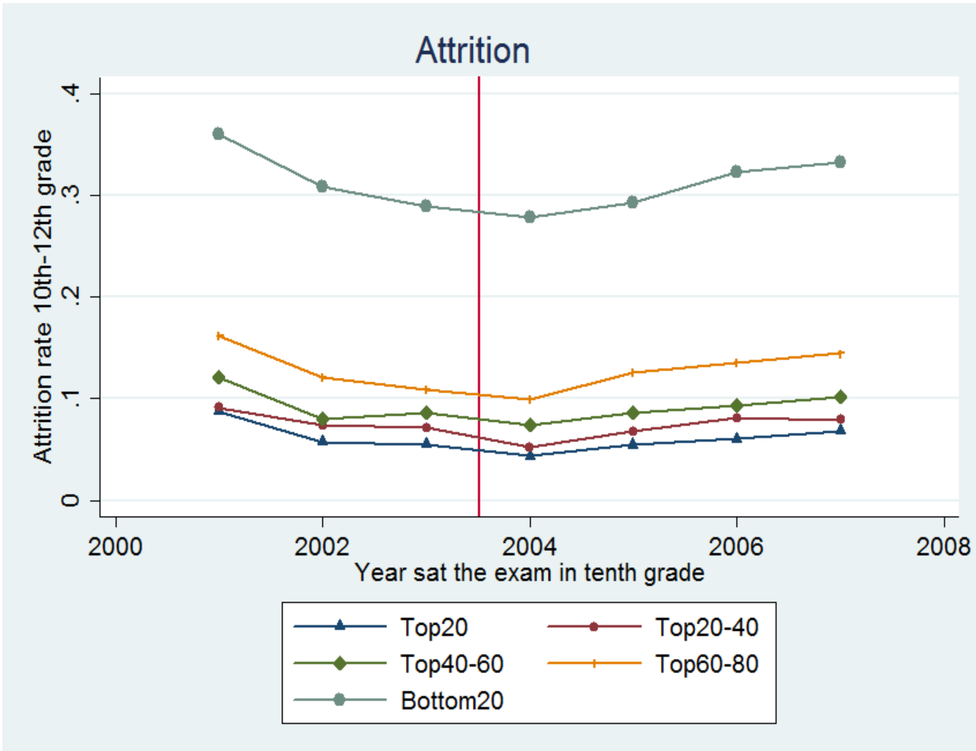
Note: The estimated effect of feedback on the national rank in the twelfth grade for the bottom (on the left) and top (on the right) quintiles of neighborhood income at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The national rank is calculated based on the five core-education subjects (incentivized). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 18: Treatment effects on the rank within the school conditional on prior performance for schools of different capacity



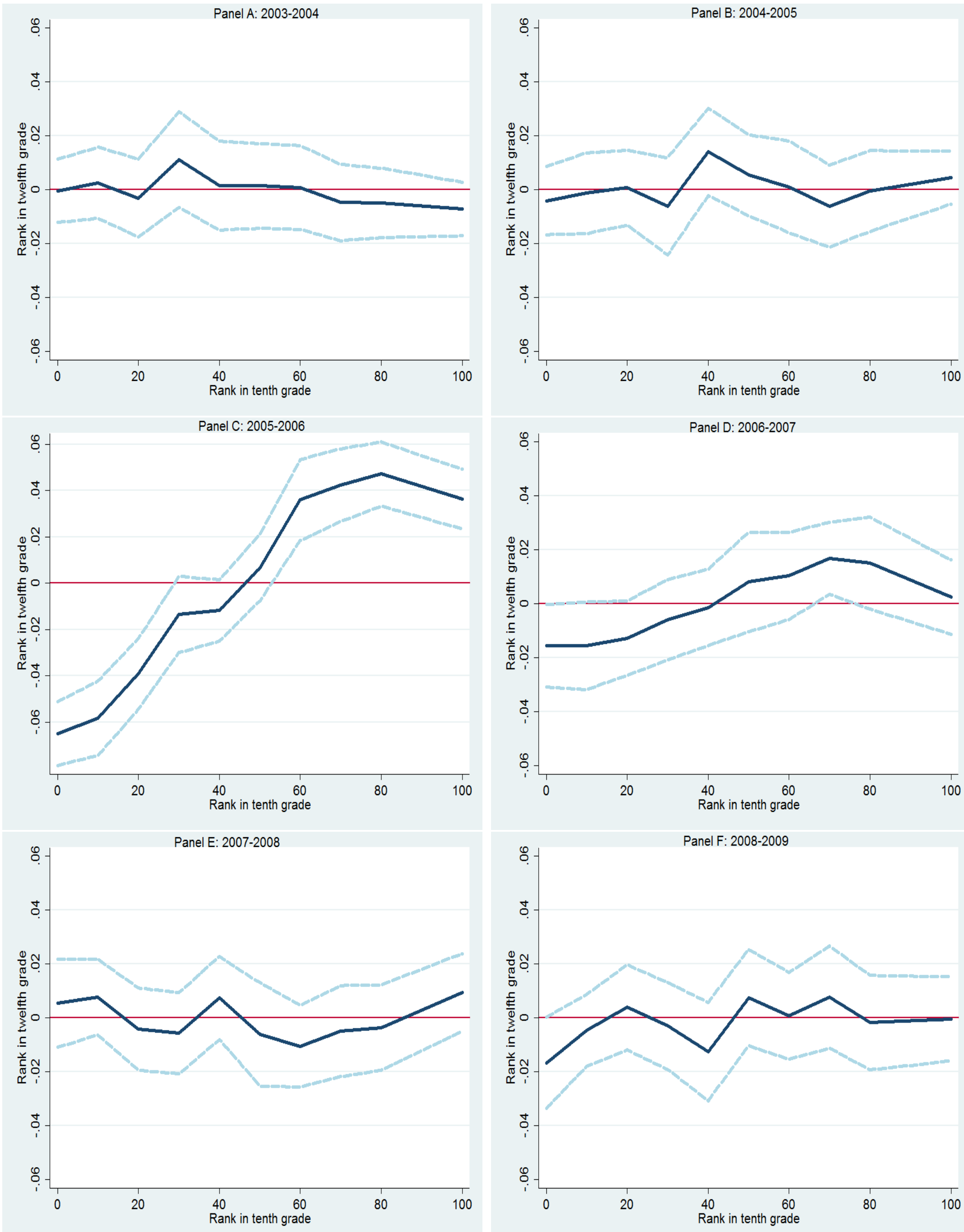
Note: The estimated effect of feedback on the school rank in the twelfth grade by capacity of school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 19: Drop out rates for each quintile of students' prior performance



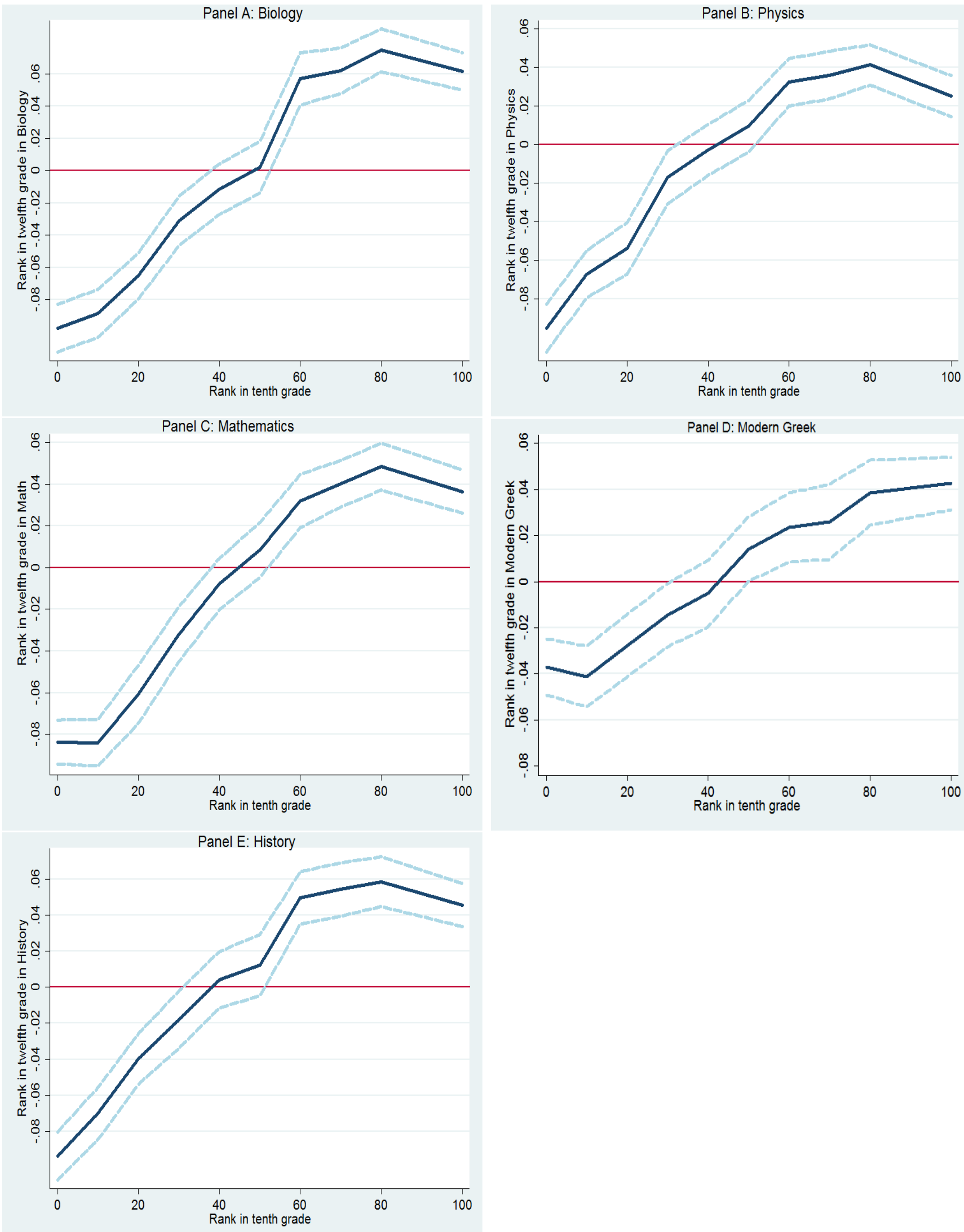
Note: Attrition rates between the tenth and the twelfth grade for each quintile of students' GPA performance in the tenth grade. Cohorts that are in tenth grade from 2001 to 2003 sit national exams in eleventh grade. Cohorts that are in the tenth grade after 2004 they do not sit national exams in the eleventh grade. The vertical line is between the last tenth grade cohort that receives feedback in the twelfth grade and the first tenth grade cohort that does not receive feedback in the twelfth grade.

Figure 20: Placebo Tests



Note: As if feedback was abolished in 2004 (Panel A), 2005 (Panel B), 2006 (Panel C), 2007 (Panel D), 2008 (Panel E) and 2009 (Panel F).

Figure 21: Feedback effects on twelfth grade rank for each subject separately



Note: The estimated effect of feedback on the twelfth grade rank within the school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The regressions are conditional on the following students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Table 1: Descriptive Statistics in twelfth grade

Variable	Mean	Std. Dev.	Min.	Max.
Student Characteristics				
Age	17.875	0.466	17	27
Early enrollment	0.167	0.373	0	1
Female	0.566	0.496	0	1
School cohort size	78.518	31.17	10	170
School GPA	85.930	10.186	49.44	100
National exam grade	62.843	19.362	7.550	98.857
Cohort size	63,186	8,710	50,061	71,796
Retake the national exam	0.115	0.319	0	1
Specialty Characteristics				
Specialty:Classics	0.359	0.48	0	1
Specialty:Exact Sciences	0.164	0.371	0	1
Specialty:Information Technology	0.477	0.499	0	1
School Characteristics				
Private School	0.039	0.193	0	1
Experimental School	0.061	0.24	0	1
Public School	0.9	0.3	0	1
Urban	0.973	0.161	0	1
logIncome(in 2009 Euro)	9.999	0.270	9.473	11.105
University Admission				
Admitted	0.823	0.381	0	1
College district different from school district	0.677	0.468	0	1
Number of university departments	8.293	10.543	1	242
Rank of admitted college in preference list	24.699	21.618	1	254
Places in tertiary education	60,960	6,268	52,450	68,136

Note: 45,746 obs. 7 cohorts. The variable "places in tertiary education" is calculated as the average across admitted students.

Table 2: Sample and Population

	Sample (134schools)	**Population** (1189schools)	**Difference**
Variable	**Mean**	**Mean**	**(b/s.e.)**
Student Characteristics			
Age	17.875	17.892	-0.017***
			(0.003)
Early enrollment	0.167	0.167	-0.0004
			(0.002)
Female	0.566	0.565	0.002
			(0.003)
School cohort size	78.518	75.358	3.160***
			(0.197)
Specialty: Classics	0.359	0.366	-0.007
			(0.004)
Specialty: Exact Sciences	0.164	0.159	0.005
			(0.002)*
Specialty: Information Techno-logy	0.477	0.475	0.002
			(0.003)
School Characteristics			
Private school	0.039	0.080	-0.041***
			(0.001)
Public schools	0.900	0.901	-0.001
			(0.002)
Experimental school	0.061	0.019	0.042***
			(0.001)
Urban	0.973	0.892	0.082***
			(0.002)
logIncome (in 2009Euro, annual)	9.999	9.938	0.060***
			(0.001)

Note: 45,746 obs. in sample and 431,469 obs. in population. There are in total 1,323 senior high schools in operation. Evening schools are excluded from the sample and the population.

Table 3: Treatment and Control Group

Variable	Feedback Mean	No Feedback Mean	Difference (b/s.e.)
Student Characteristics			
Age	17.835	17.909	0.074***
			(0.004)
Early enrollment	0.209	0.129	-0.080***
			(0.004)
Female	0.553	0.579	0.026***
			(0.005)
School cohort size	88.083	70.030	18.053***
			(0.288)
Specialty: Classics	0.344	0.377	0.033***
			(0.004)
Specialty: Exact Sciences	0.176	0.154	-0.022***
			(0.004)
Specialty: Information Technology	0.480	0.469	-0.011**
			(0.005)
School Characteristics			
Private school	0.037	0.037	0.0003
			(0.002)
Public schools	0.905	0.897	-0.008
			(0.005)
Experimental school	0.058	0.066	0.007
			(0.005)
Urban	0.972	0.974	0.002
			(0.002)
logIncome (in 2009Euro,annual)	9.988	10.005	0.017***
			(0.003)

Note: 21,965 obs. in treatment group and 23,781 obs. in control group. The feedback period is the pooled period from 2003 to 2005 while the non-feedback period consists of the pooled period from 2006 to 2009.

Table 4: Estimation results: Rank nationwide

Dependent Variable: Rank nationwide in incentivized subjects			
	Specifications		
Variable	(1)	(2)	(3)
Feedback*quintile5	0.042***	0.043***	0.045***
	(0.004)	(0.006)	(0.004)
Feedback*quintile4	0.036***	0.037***	0.040***
	(0.004)	(0.005)	(0.004)
Feedback*quintile2	-0.045***	-0.045***	-0.038***
	(0.004)	(0.004)	(0.005)
Feedback*quintile1	-0.088***	-0.088***	-0.079***
	(0.004)	(0.004)	(0.004)
Feedback	0.009***	0.009	-0.001
	(0.003)	(0.009)	(0.003)
quintile5	0.234***	0.235***	0.251***
	(0.003)	(0.004)	(0.004)
quintile4	0.094***	0.094***	0.102***
	(0.003)	(0.003)	(0.003)
quintile2	-0.081***	-0.083***	-0.093***
	(0.003)	(0.003)	(0.003)
quintile1	-0.176***	-0.177***	-0.192***
	(0.003)	(0.003)	(0.003)
Female	-0.008***	-0.008***	-0.011***
	(0.001)	(0.002)	(0.002)
Age	-0.010***	-0.010***	-0.010***
	(0.003)	(0.003)	(0.003)
Early enrollment	-0.006**	-0.007**	-0.006*
	(0.002)	(0.003)	(0.003)
Specialty: Science	0.042***	0.041***	0.041***
	(0.002)	(0.002)	(0.002)
Specialty: Classics	-0.022***	-0.021***	-0.021***
	(0.002)	(0.002)	(0.002)
Log Income	0.055***		
	(0.003)		
Experimental school	0.029***		
	(0.004)		
Private school	0.145***		
	(0.004)		
Urban	0.021***		
	(0.004)		
Year FE.	no	no	yes
School FE.	no	yes	yes
Observations	45,746	45,746	45,746
R squared	0.635	0.666	0.675
No of schools	134	134	134

Note: A constant is also included. Clusters at school level. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 5: Rank within the school in incentivized and non-incentivized subject

Dependent Variable: School Rank in incentivized and non-incentivized subjects				
	Incentivized subjects		**Non-Incentiv. subjects**	
Variable	(1)	(2)	(3)	(4)
Feedback*quintile5	0.045***	0.045***	0.005	0.005
	(0.004)	(0.004)	(0.006)	(0.006)
Feedback*quintile4	0.040***	0.040***	-0.005	-0.005
	(0.004)	(0.004)	(0.006)	(0.006)
Feedback*quintile2	-0.038***	-0.038***	-0.004	-0.003
	(0.004)	(0.005)	(0.006)	(0.006)
Feedback*quintile1	-0.079***	-0.079***	0.005	0.005
	(0.004)	(0.004)	(0.005)	(0.006)
Feedback	0.001	-0.001	0.003	0.001
	(0.003)	(0.003)	(0.004)	(0.004)
quintile5	0.251***	0.251***	0.256***	0.256***
	(0.004)	(0.004)	(0.005)	(0.004)
quintile4	0.102***	0.102***	0.103***	0.105***
	(0.003)	(0.003)	(0.003)	(0.004)
quintile2	-0.093***	-0.093***	-0.094***	-0.095***
	(0.003)	(0.003)	(0.004)	(0.005)
quintile1	-0.193***	-0.192***	-0.200***	-0.200***
	(0.003)	(0.003)	(0.004)	(0.006)
Female	-0.009***	-0.011***	0.054***	0.054***
	(0.001)	(0.002)	(0.003)	(0.002)
Age	-0.010***	-0.010***	0.002	0.002
	(0.003)	(0.003)	(0.002)	(0.002)
Early enrollment	-0.006*	-0.006*	0.006	0.007
	(0.003)	(0.003)	(0.005)	(0.005)
Specialty: Science	0.047***	0.048***	0.033***	0.034***
	(0.001)	(0.002)	(0.002)	(0.003)
Specialty: Classics	-0.019***	-0.021***	0.097***	0.097***
	(0.001)	(0.002)	(0.002)	(0.003)
Log Income	0.051***	0.049***	0.007	0.007
	(0.0004)	(0.001)	(0.004)	(0.004)
Experimental school	-0.041***	-0.038***	-0.003	-0.004
	(0.002)	(0.004)	(0.004)	(0.003)
Private school	-0.003	-0.004	0.030	0.032
	(0.003)	(0.004)	(0.016)	(0.018)
Urban	-0.017***	-0.016***	-0.003	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)
Year FE.	no	yes	no	yes
Observations	45,746	45,746	45,746	45,746
R squared	0.674	0.675	0.542	0.543
No of schools	134	134	134	134

Note: Standard errors are clustered at the school level. A constant is also included. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 6: Estimation results : Differential Response by Gender

Dependent Variable: Rank in twelfth grade				
	Rank within the school		**Rank nationwide**	
Variable	(1)	(2)	(3)	(4)
Female*Feedback	-0.028***	-0.028***	-0.027***	-0.027***
	(0.005)	(0.005)	(0.005)	(0.005)
Female	0.054***	0.054***	0.052***	0.052***
	(0.003)	(0.003)	(0.003)	(0.003)
Feedback	0.009***	0.002	0.009	0.008
	(0.003)	(0.004)	(0.008)	(0.009)
Speciality in Science	0.198***	0.199***	0.198***	0.196***
	(0.004)	(0.004)	(0.004)	(0.004)
Speciality in Classics	-0.040***	-0.039***	-0.039***	-0.040***
	(0.003)	(0.003)	(0.005)	(0.003)
Age	-0.051***	-0.051***	-0.060***	-0.060***
	(0.005)	(0.005)	(0.005)	(0.005)
Early enrollment	-0.047***	-0.048***	-0.058***	-0.058***
	(0.005)	(0.005)	(0.005)	(0.005)
Income	-0.0001	-0.0001	0.0002***	0.0002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Private	-0.015*	-0.015*	0.134***	0.134***
	(0.008)	(0.008)	(0.016)	(0.017)
Experimental	-0.015**	-0.015**	0.017	0.017
	(0.006)	(0.006)	(0.018)	(0.018)
Urban	-0.029***	-0.029***	0.007	0.007
	(0.007)	(0.007)	(0.015)	(0.015)
R^2	0.14	0.14	0.16	0.16
N	45,746	45,746	45,746	45,746
Year FE		✓		✓
No of schools	134	134	134	134

Note: Standard errors are clustered at the school level. A constant is also included. *,**,*** denotes significance at the 10%,5% and 1% level respectively. The rank in the twelfth grade here takes into account only the incentivized subjects. It is calculated within the school for columns (1) and (2) and across schools in columns (3) and (4)

Table 7: Capacity of schools

Variable	Mean	Std. Dev.	Min.	Max.
Schools with one class				
Public	0.899	0.302	0	1
Private	0.101	0.301	0	1
Experimental	0	0	0	0
Urban	0.378	0.485	0	1
Class size	18.130	5.717	10	29
No of schools	14			
No of students	522			
Schools with two classes				
Public	0.932	0.252	0	1
Private	0	0	0	0
Experimental	0.068	0.252	0	1
Urban	0.832	0.375	0	1
Class size	16.000	4.739	10	27
No of schools	38			
No of students	3,709			
Schools with three classes				
Public	0.941	0.235	0	1
Private	0.053	0.223	0	1
Experimental	0.006	0.077	0	1
Urban	0.986	0.115	0	1
Class size	18.211	4.998	10	32
No of schools	63			
No of students	9,959			
Schools with three classes				
Public	0.881	0.324	0	1
Private	0.035	0.184	0	1
Experimental	0.084	0.277	0	1
Urban	1	0	0	1
Class size	20.072	6.973	10	33
No of schools	74			
No of students	26,354			

Note: 111 senior high schools provided us with the eleventh and twelve grade classroom information. The number of classes in a school may not be stable across years. Some schools may expand and some others may shrink in some years.

Table 8: Loss of human capital in terms of labour force participants

Year	Students Retaking	Potential Impact on Labour Market
2003	7925	0.167%
2004	7223	0.150%
2005	6387	0.131%
2006	10421	0.213%
2007	6642	0.135%
2008	5730	0.116%
2009	4576	0.092%
2010	7680	0.153%

Table 9: Decision to Retake and Feedback

Dependent Variable: Repeat the national exams				
	LPM		**Probit**	**Logit**
Variable	(1)	(2)	(3)	(4)
Feedback* Misplacement	0.058	0.059	0.345	0.602
	(0.016)***	(0.016)***	(0.092)***	(0.181)***
Feedback	0.012	0.019	0.070	0.131
	(0.006)*	(0.007)**	(0.036)*	(0.074)*
Misplacement	-0.014	-0.015	-0.071	-0.099
	(0.014)	(0.015)	(0.077)	(0.142)
Age	-0.014	-0.019	-0.076	-0.157
	(0.003)***	(0.006)***	(0.039)*	(0.062)**
Early Enrolled	-0.005	-0.006	-0.011	-0.033
	(0.008)	(0.008)	(0.022)	(0.082)
Female	-0.007	-0.007	-0.044	-0.073
	(0.003)*	(0.004)*	(0.020)*	(0.038)*
Specialization in Classics	-0.020	-0.018	-0.113	-0.200
	(0.004)***	(0.007)*	(0.024)***	(0.046)***
Specialization in Science	0.013	0.016	0.090	0.169
	(0.005)**	(0.004)***	(0.026)***	(0.049)***
District Unemployment	0.005	0.002	0.025	0.046
	(0.002)**	(0.002)	(0.012)*	(0.019)**
If admitted in first place	-0.212	-0.218	-1.041	-1.964
	(0.008)***	(0.008)***	(0.035)***	(0.070)***
Internal Migration	0.064	0.072	0.445	0.889
	(0.005)***	(0.005)***	(0.037)***	(0.077)***
logIncome	-0.009			
	(0.011)			
Urban	0.024			
	(0.013)*			
Private	-0.056			
	(0.007)**			
Public	-0.039			
	(0.009)***			
R^2 or pseudo-R^2	0.05	0.06	0.07	0.07
Log likelihood			-13,432	-13,439
School FE		✓	✓	✓
Year FE	✓	✓	✓	✓
N	45,746	45,746	45,746	45,746

Note: A constant is also included. Standard errors are clustered at the school level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 10: Decision to Retake, Feedback and Misplacement

Dependent Variable: Repeat the national exams				
	LPM		**Probit**	**Logit**
Variable	(1)	(2)	(3)	(4)
Feedback	-0.031	-0.002	-0.007	-0.017
	(0.007)***	(0.008)	(0.047)	(0.090)
Feedback* Misplacement Quintile 5	0.045	0.040	0.219	0.412
	(0.010)***	(0.009)***	(0.050)***	(0.095)***
Feedback* Misplacement Quintile 4	0.023	0.023	0.120	0.231
	(0.010)**	(0.009)**	(0.049)**	(0.095)**
Feedback* Misplacement Quintile 2	0.004	0.007	0.049	0.103
	(0.010)	(0.011)	(0.054)	(0.101)
Feedback* Misplacement Quintile 1	-0.034	-0.031	-0.151	-0.274
	(0.010)***	(0.010)***	(0.052)***	(0.098)***
Misplacement Quintile 5	-0.017	-0.018	-0.103	-0.184
	(0.007)**	(0.007)**	(0.038)***	(0.073)**
Misplacement Quintile 4	-0.025	-0.025	-0.139	-0.262
	(0.007)***	(0.007)***	(0.038)***	(0.072)***
Misplacement Quintile 2	0.017	0.016	0.076	0.143
	(0.007)**	(0.008)**	(0.039)*	(0.073)**
Misplacement Quintile 1	0.030	0.031	0.148	0.273
	(0.007)***	(0.009)***	(0.043)***	(0.080)***
Female	-0.010	-0.010	-0.056	-0.105
	(0.004)***	(0.004)***	(0.020)***	(0.037)***
Age	0.002	-0.001	-0.001	-0.002
	(0.007)	(0.007)	(0.035)	(0.067)
Early Enrolled	0.011	0.009	0.047	0.087
	(0.008)	(0.008)	(0.041)	(0.078)
Unemployment	0.005	0.002	0.010	0.020
	(0.001)***	(0.002)	(0.011)	(0.021)
Internal migration	-0.024	-0.022	-0.109	-0.211
	(0.007)***	(0.008)***	(0.038)***	(0.075)***
Specialization in Science	-0.007	-0.004	-0.018	-0.036
	(0.005)	(0.005)	(0.025)	(0.048)
Specialization in Classics	-0.018	-0.017	-0.093	-0.175
	(0.004)***	(0.004)***	(0.024)***	(0.045)***
Private	-0.087			
	(0.011)***			
Public	-0.040			
	(0.009)***			
LogIncome	-0.033			
	(0.008)***			
Urban	0.006			
	(0.010)			
R^2 or pseudo-R squared	0.03	0.04	0.06	0.06
Log likelihood			-14,062	-14,063
School FE		✓	✓	✓
Year FE	✓	✓	✓	✓
N	45,746	45,746	45,746	45,746

Note: A constant is also included. Standard errors are clustered at the school level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 11: Estimation results : Drop out

Dependent Variable: Dummy for drop out		
	Specifications	
Variable	(1)	(2)
Feedback*quintile5	0.009	0.010
	(0.007)	(0.007)
Feedback*quintile4	0.007	0.007
	(0.007)	(0.007)
Feedback*quintile2	0.009	0.010
	(0.008)	(0.008)
Feedback*quintile1	0.013	0.014
	(0.015)	(0.016)
Feedback	0.017	0.041
	(0.019)	(0.033)
quintile5	0.000	-0.001
	(0.004)	(0.004)
quintile4	-0.006	-0.006
	(0.005)	(0.005)
quintile2	0.025***	0.025***
	(0.006)	(0.006)
quintile1	0.153***	0.153***
	(0.003)	(0.014)
Female	-0.011***	-0.011***
	(0.003)	(0.004)
Absences10	0.001***	0.002***
	(0.0001)	(0.0001)
Year FE.	no	yes
Observations	56,041	56,041
R squared	0.130	0.203
No of schools	134	134

Note: A constant is also included. Clusters at school level. *,**,*** denotes significance at the 10%,5% and 1% level respectively. Quintiles are constructed based on the school performance in tenth grade used.

Table 12: Drop out rate and Transfers

	Drop out 10 11	**Transfers 10 11**	**Drop out 11 12**	**Transfers 11 12**
	%	%	%	%
2000-2001	8.89			
2001-2002	12.31	8.89		
2002-2003	11.07	8.4	6.07	7.02
2003-2004	8.87	7.93	5.14	9.47
2004-2005	10.71	7.46	6.67	6.45
2005-2006	9.41	6.35	5.67	6.58
2006-2007	12.59	9.46	6.15	8.60
2007-2008	11.71	5.85	8.63	6.80
2008-2009	13.71	7.52	6.13	6.58
2009-2010	10.56	6.45	6.01	8.38
2010-2011	9.92	5.76	6.19	8.61

Note: The first column shown the percentage of students who drop out from school between the tenth and eleventh grade. The second column shown the percentage of students who transfer to a school in the eleventh grade. The third column shown the percentage of students who drop out from school between the eleventh and twelfth grade. The fourth column shown the percentage of students who transfer to a school in the twelfth grade. Data from 134 schools are used.

Table 13: Estimation results: Different outcome variables

Dependent Variable: Rank in twelfth grade			
	Specifications		
Variable	(1)	(2)	(3)
Feedback*quintile5	0.026***	0.030***	0.050***
	(0.004)	(0.007)	(0.005)
Feedback*quintile4	0.022***	0.015***	0.032***
	(0.004)	(0.007)	(0.005)
Feedback*quintile2	-0.029***	-0.032***	-0.042***
	(0.004)	(0.007)	(0.005)
Feedback*quintile1	-0.052***	-0.045***	-0.066***
	(0.004)	(0.006)	(0.005)
Feedback	0.002	0.008	-0.0004
	(0.003)	(0.005)	(0.003)
quintile5	0.257***	0.247***	0.245***
	(0.003)	(0.005)	(0.003)
quintile4	0.109***	0.110***	0.107***
	(0.003)	(0.005)	(0.003)
quintile2	-0.097***	-0.100***	-0.091***
	(0.003)	(0.005)	(0.003)
quintile1	-0.207***	-0.231***	-0.210***
	(0.003)	(0.005)	(0.003)
Female	-0.019***	0.030***	-0.014 ***
	(0.001)	(0.002)	(0.001)
Early Enrollment	0.010***	0.009***	0.011 ***
	(0.002)	(0.002)	(0.002)
Specialty: Science	0.006***	0.019*** *	0.023***
	(0.002)	(0.004)	(0.002)
Specialty: Classics	0.010***	0.098***	-0.059
	(0.002)	(0.003)	(0.002)
Observations	45,746	45,746	45,746
R squared	0.661	0.674	0.625
No of schools	134	134	134

Note: A constant is also included. The outcome in the first column is the rank calculated based on the five core subjects and the four Track subjects. The outcome in the second column is the rank in Modern Greek. The outcome variable in the third column is calculated based on five subjects in the feedback regime and two subjects in the non-feedback regime. Standard errors clustered at the school level. Year fixed effects included. Clusters at school level. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 14: Descriptive Evidence of Social Mobility

Quintiles of program' popularity	**Quintiles of Neighborhood Income**											
	Quintile1		**Quintile2**		**Quintile3**		**Quintile4**		**Quintile5**		**Total**	
	Feed.	No Feed.	Feed.	No Feed.	Feed.	No Feed.	Feed.	No Feed.	Feed.	No Feed.	Feed.	No Feed.
Quintile1	4.6	3.4	4.2	3.2	4.5	2.9	2.7	2.1	3.1	2.7	19.1	14.3
	+1.2		+1		+1.6		+0.6		+0.4		+4.8	
Quintile2	3.9	3.5	3.5	3.1	4.1	3.5	2.5	2.7	3.1	3.2	17.1	16
	+0.4		+0.5		+0.6		-0.2		-0.3		+1.1	
Quintile3	3.6	3.4	3.2	3.2	3.8	3.5	2.7	3	3.2	3.4	16.5	16.5
	+0.2		0		+0.3		-0.3		-0.3		0	
Quintile4	3.1	3.0	2.8	3.2	3.5	4.3	2.5	3.2	3.4	3.8	15.3	17.5
	+0.1		-0.4		-0.8		-0.7		-0.4		-2.2	
Quintile5	2.9	2.6	2.6	3.0	3.9	4.1	2.8	3.2	3.6	4.2	15.8	17.1
	+0.3		-0.42		-0.2		-0.4		-0.6		-1.3	
Total	18.1	15.9	16.3	15.7	19.8	18.3	13.2	14.2	16.3	17.4	**83.7**	**81.5**
	2.2		+0.6		+1.5		-1		-1.1		**+2.2**	

Note: "Quintile 1" represents the bottom quintile of program's popularity/ prestigiousness and neighborhood income. "Quintile 5" denotes the top quintile of the program's popularity/ prestigiousness and neighborhood income. For each quintile of neighborhood income two percentages are reported: the first one represents the percentage of students who enroll in university in the feedback period and the second one the corresponding percentage in the non-feedback period. The differences between the percentage in the feedback period and the non-feedback period for each quintile of program's popularity prestigiousness are also reported.