What you know can't hurt you (for long): A field experiment on relative performance feedback*

Preliminary and Incomplete. Please do not circulate. Comments are greatly appreciated.

Ghazala Azmat† Manuel Bagues‡ Antonio Cabrales§ Nagore Iriberri¶

September 10, 2015

Abstract

This paper studies the effect of providing feedback to college students on their position in the distribution of grades using a randomized field experiment. This information was updated every six months during a three-year period. Absent the treatment, students underestimate their position in the distribution of grades, and after the three-year treatment, only students exposed to the treatment are correctly informed about their relative position. We find that treated students experience a significant short-term decrease in their educational performance and a short term increase in their self-reported satisfaction. The negative impact on performance fades away after six months.

Keywords: Relative performance feedback, ranking, randomized field experiment, school performance.

*We would like to thank César Alonso, Javier Ruiz-Castillo and participants in presentations at Aalto University, CSEF, Universit de Paris I, Norwegian School of Economics, Universidad Carlos III de Madrid, Nagoya University, Universidad de Navarra, Institute of Fiscal Studies, Warwick University, East Anglia University, Workshop on Experimental Economics at Granada, the Institute for Economic International Studies, Zurich University and Lancaster University for their useful comments. We also acknowledge the support of Isabel Gutiérrez and the administration of Universidad Carlos III for conducting this study. Manuel Bagues acknowledges financial support from Ministerio de Economía y Competición (ECO 2012-31358). Antonio Cabrales acknowledges financial support from Ministerio de Economía y Competición (ECO ECO2012-34581). Nagore Iriberri acknowledges financial support from Ministerio de Economía y Competición (ECO2012-31626), Departamento de Educación, Política Lingüística y Cultura del Gobierno Vasco (IT869-13) and Ministerio de Ciencia e Innovación (ECO2011-25295). All remaining errors are our own.

†Queen Mary, University of London and Centre for Economic Performance, LSE. E-mail: `g.azmat@qmul.ac.uk`.

‡Aalto University and IZA, Helsinki, Finland. Email: `manuel.bagues@aalto.fi`.

§University College London. E-mail: `a.cabrales@ucl.ac.uk`.

¶University of the Basque Country UPV-EHU, IKERBASQUE, Basque Foundation for Science. E-mail: `nagore.iriberri@gmail.com`.

1 Introduction

Feedback information on individual performance is often provided in many professional and educational environments. Students typically receive report cards, teachers learn about their students' satisfaction from teaching evaluations and, in many occupations, firms provide employees information about individual performance, as measured by sales data or revenue.

In addition to this information, individuals may also receive feedback about their performance relative to a relevant reference group. For instance, in some schools students receive information about their relative position within the class (Azmat and Iriberri, 2010). In many universities teachers are informed about their position in the distribution of teaching evaluations relative to other professors teaching in the same Department, Degree or University. Similarly, in the workplace agents learn about their position in the distribution of performance within the company (e.g. Blanes-i-Vidal and Nossol, 2011, and Barankay, 2011). Sometimes rankings are made public and individuals learn not only about their own position in the distribution of performance but also the relative position of their peers, such that status concerns may arise. In this paper we focus on situations where individuals receive feedback on their relative performance privately.

From a theoretical perspective the impact of relative performance feedback on effort is ambiguous. Agents' reaction depends on the agents' prior beliefs about own and others' ability, the new information inferred from the feedback, as well as on the agents' inherent motivations. For example, if ability is complementary to own effort for the purpose of achieving a particular outcome, positive (negative) news about own ability will make individuals work more (less). In addition, agents might care about their relative standing, showing a "competitive" motivation in their preferences, perhaps because corporate recruiters or graduate school admissions officers value relative on top of absolute performance. If that is the case, learning that others' ability is lower (higher) than initially thought could make agents exert a lower (higher) level of effort. We introduce a theoretical model that includes different motivations, and where ability and effort are complements, to help interpret the possible reactions to the provision of relative performance feedback. We show that both the different motivations as well as the informativeness of the feedback relative to agents' prior beliefs, are crucial when predicting a particular direction in the change of effort.

A number of empirical studies have studied the impact of feedback on relative performance. There are now a few lab experiment that test for the effect of relative performance feedback. Studies differ on important aspects such as the underlying incentive scheme (piece-rate or flat-rate), precision of relative performance feedback (one another individual's effort or output, average or rank), and on whether the relative performance feedback is privately or publicly provided. Hannan et al. (2009) find that individuals choose a higher (non-real) effort under piece-rate when the average is provided. Eriksson et al. (2009), on the contrary, find no effect on the choice of real effort under piece-rate when another participant's output is revealed. Charness et al. (2013), and Gerhards and Siemer (2014), find that individuals choose higher effort under flat-rate incentives when they are provided with their rank in the session and when individuals are revealed to be the best performers, respectively, both when this information is privately and publicly provided. Azmat and Iriberri (2014) show that relative performance feedback increases performance but it also increases inequality in satisfaction when performance is related to pay (piece-rate) but not when it is independent of pay (flat-rate). In sum, most studies find that the overall effect is positive when some type of relative performance feedback is provided. Importantly, these studies do not control for individuals' beliefs before the information is provided. Khunen and Tymula (2012), under flat-rate incentives, find that those individuals who rank lower than expected increase effort and those who rank higher than expected reduce effort, where the overall effect is positive.

Other studies have examined the impact in the field. Falk and Ichino (2006), and Mas and Moretti (2009), find that when agents can observe other's work and when other agents can observe one's work, which allows agents to make relative assessment of their effort and output, the overall performance increases even under flat-rate incentives.

Azmat and Iriberri (2010) show that the performance of High School students improved notably when, due to a change in the IT of the school, the report card added information on the average grade obtained by students in the class. Blanes-i-Vidal and Nossol (2011) find that workers increase their effort after they start to receive feedback on their relative performance, when their pay was related to output. More recently, several authors have performed randomized control trials. Tran and Zeckhauser (2012) find that Vietnamese students increase their effort and perform better in an English course when provided with their rank position, thus finding support for an inherit preference for ranking high even though this information was privately provided. Katreniakova (2014) conducted an experiment on the impact of feedback on relative performance in 53 Ugandan schools. The provision of feedback improves students' performance, particularly when financial or reputation rewards are also present. While most studies tend to find a positive impact of feedback on relative performance, Barankay (2011) reports evidence showing that the effect might be sometimes negative. He conducts a three-year randomized control trial with full-time furniture salespeople. Initially, all salespeople receive feedback on relative performance. A random sample of them stops receiving this information. Unlike the previous papers, in the case of Barankay (2011) removing rank feedback actually increases sales performance. Neither of these studies control for individuals' beliefs prior to the provision of information. A possible explanation for these mixed results is that in different context agents may hold different priors about their relative performance or possibly different objective functions.

Given the ambiguous predictions of the theory and the mixed empirical evidence, in this paper we study the impact of providing feedback on relative performance to college students, where we control for students' expectations on their relative ranking. We conducted a field experiment over three years (2010-2013) in a large Spanish university. A cohort of approximately 1,000 students enrolled in several degrees were randomly assigned into treatment and control groups. Students in the control group receive only information on their own performance (as is the norm in this university). Students in the treatment group are provided access to information on their relative performance, decile position in the distribution of performance, in addition to the information on their absolute performance. The treatment starts at the beginning of their 2nd year of study and the information is updated every six months. The intervention lasts for three years and concludes at the end of the fourth and final year. We analyze how the intervention affects students' information, their performance and their (self-reported) effort and satisfaction. Approximately 70% of candidates in the treatment group checked the information that was provided. Most of the students who do not check for the information are from the lower part of the grade distribution. Information from a survey suggests that, prior to the intervention, students were relatively uniformed about their position in the ranking and, in general, they tended to underestimate their ranking. Initially the average student self-reported position is 18 percentiles lower than her true position. The intervention improved significantly the information available to students in the treatment group. At graduation this gap has decreased to 6 percentiles among students in the treatment group, but is still around 11 percentiles in the control group.

The academic performance of students in the treatment and control groups was similar during their first university year. However, after the intervention is introduced, at the beginning of the second year, we observe a significant short-term decrease in the performance of students in the treatment group. During their second year, treated students pass on average 0.4 exams less than students in the control group (10% of a standard deviation). Treated students catch up half of this gap when they resit their exams in the second year, and the rest of the gap fades away afterwards. The performance of the treatment and control groups is statistically indistinguishable at the time of graduation. There is no long-term effect on the likelihood to graduate or on the average AGPA at end of degree. We do not observe any impact either on the type of elective courses chosen by students.

The evidence suggests that providing information on relative performance to college students does not necessarily improve their performance and might even have a negative impact. More precisely, the impact of the treatment might depend crucially on students' priors and on their preferences. In the case of the college students analyzed here, learning that they were doing better than expected had on average a negative impact on their performance.

This behavior is easily interpreted in light of our model. Two assumptions are necessary. The first assumption requires that initial knowledge of own ability is more precise than knowledge of others' ability, something natural in college where most peers are new. The second assumption requires that concerns for relative standing are stronger than the desire to reach an absolute goal in terms of grades, a reasonable assumption in a university setting, where important rewards, such as internships are awarded based on relative performance. Under these two assumptions our theoretical framework predicts agents should decrease effort. First, given own effort and others' ability are complements in the "competitive" motivation of preferences, when learning that others' ability is lower than the expected, which our data clearly shows it is the case, agents should decrease their effort levels. Second, given the strategic complementarity of efforts, again, in the "competitive" motivation of preferences, all decision makers' equilibrium efforts should decrease.

The paper is organized as follows. Section 2 presents the theoretical framework. Section 3 describes the institutional background and the design of the experiment, as well as the additional surveys we carried out in the field. Section 4 presents data. Section 5 shows the empirical analysis and finally, Section 6 concludes.

2 Theoretical Framework

Let the utility of an individual depend on her output, F, where output is a function of individual's effort x_i and ability θ_i in a complementary fashion, and $0 < \delta < 1$ is a constant.

$$F\left(x_i, \theta_i\right) = \left(\theta_i x_i\right)^{\delta}$$

Given the complementarity between x_i and θ_i, the marginal output of effort x_i is increasing in ability θ_i

$$\frac{\partial F\left(x_i,\theta_i\right)}{\partial x_i \partial \theta_i}=\delta^2\left(\theta_i x_i\right)^{\delta-1}>0 \tag{1}$$

Assume further that individuals have a "competitive" motivation in their preferences, so that their utility also depends on the relative standing in the group. For example, the individuals are competing for a prize and the probability that individual i wins the price is given by the expression

$$G\left(x_i,\theta_i,x_{-i},\theta_{-i}\right)=\left(1-e^{-\left(\theta_i x_i-\theta_{-i}x_{-i}\right)}\right) \tag{2}$$

where clearly a higher talent θ_i or effort x_i of individual i makes it more likely that she wins the prize, while a higher talent θ_{-i} or effort x_{-i} of opponents makes it less likely.

Note that own effort and others' effort are strategic complements in $G\left(.\right)$ since

$$\frac{\partial G\left(x_i,\theta_i,x_{-i},\theta_{-i}\right)}{\partial x_i \partial x_{-i}}=\theta_i\theta_{-i}e^{-\left(\theta_i x_i-\theta_{-i}x_{-i}\right)}>0 \tag{3}$$

and that marginal product of own effort x_i in the competitive motivation function is increasing in the ability of others θ_{-i}

$$\frac{\partial G\left(x_i,\theta_i,x_{-i},\theta_{-i}\right)}{\partial x_i \partial \theta_{-i}}=\theta_i x_{-i}e^{-\left(\theta_i x_i-\theta_{-i}x_{-i}\right)}>0 \tag{4}$$

but in terms of the competitive motivation, own effort x_i and own ability θ_i may be complements or substitutes, since the sign of the derivative

$$\frac{\partial G\left(x_i,\theta_i,x_{-i},\theta_{-i}\right)}{\partial x_i \partial \theta_i}=(1-\theta_i x_i)e^{-\left(\theta_i x_i-\theta_{-i}x_{-i}\right)} \tag{5}$$

depends on whether $\theta_i x_i$ is smaller or bigger than 1.

Total utility is given by:

$$\alpha F\left(x_i,\theta_i\right)+\beta G\left(x_i,\theta_i,x_{-i},\theta_{-i}\right)-C\left(x_i\right)$$

Relative performance feedback can be informative about own ability as well as about others' ability.

Assume first that relative performance feedback informs the decision maker that others' ability, θ_{-i} is lower than expected, and thus that they were underestimating their relative position. Then, the reaction function for effort of agent $x_i(\theta_i, x_{-i}, \theta_{-i})$ will shift down from the effect on the competitive motivation (eq. 4). And if everyone lowers their estimate of the ability of opponents, given the strategic complementarity between own effort and others efforts (from eq. 4), then the equilibrium effort x_i^* will go down for everyone.

Assume next that relative performance feedback reveals that own ability θ_i is higher than initially thought. Then the effect is more complicated. On the one hand, from the complementarity of own effort and ability in $F(.)$, the reaction function for effort should shift up (see eq. 1), but since the relationship between own ability and effort in the competitive motivation $G(.)$ could be one of substitutability, the reaction function for effort could shift down (if $\theta_i x_i > 1$, see eq. 5). Then, if the shift of the reaction function is the same for everyone (up or down depending on the relative sizes and signs of effects on $F(.)$ or $G(.)$), the strategic complementarity of own and others' efforts should shift the equilibrium choice of effort for everyone in the same direction, up or down, as the individual reaction functions. People with a high relative desire for maximizing their own output versus having a high standing within the cohort ($\alpha >> \beta$) could increase effort after learning their relative position is better than expected, whereas people with a high relative desire for standing within the cohort ($\beta >> \alpha$), and a value for $\theta_i x_i > 1$ (so that own effort x_i and own ability θ_i are substitutes in $G(.)$) could decrease effort after learning their relative position is better than expected.

The final effect therefore depends on the prior knowledge of own ability θ_i, versus the knowledge of others ability θ_{-i}. If information about θ_{-i} is the only novelty, the effect would be an unambiguous decrease in effort, provided $\beta > 0$. If information about θ_i is the novelty, the effect would be ambiguous.

This theoretical framework shows that different motivations in the utility, the expectations individuals have prior to the provision of information, and whether feedback informs about own ability or about others' ability, are important determinants of effort choices, which can lead to different reactions in effort. Note however, that the fact that the framework allows for different responses does not mean that the model does not provide guidance as to what effects we should find. Particular directions for the effect depend on particular types of information. For example, it is natural to expect that knowledge of own ability θ_i, is more precise than knowledge of others ability θ_{-i}, particularly in a university, where all peers are relatively new for most students. Therefore, the feedback will make individuals update their knowledge of others' ability than their knowledge of their own ability. Also, the grades in a university serve as a signal of ability to potential employers and to graduate school admissions officers. This means that although some students will have an intrinsic motivation to have better grades, it is likely that many of them will have an even stronger desire to do well with respect to others. If this is the case, the dominant force will be the one that shifts effort up or down in the presence of a negative or positive surprise.

3 Institutional framework

The Randomized Control Trial was conducted at University Carlos III, Madrid. The university offers several four-year and six-year degrees in three different campuses. The majority of students do their degree in Spanish but a small minority do it in English. Our study involves students enrolled in the Spanish track of four of these four-year degrees - Business, Economics, Finance, Law - and one six-year degree - Business and Law.[1] Two of these degrees, Business and Business and Law, are held simultaneously in two different locations, the Northern and the Southern campuses. The study therefore involves students in seven different degree-locations. Below we explain the most relevant features of this university and the design of the experiment.

[1]The choice of degrees and campuses was based on data availability and size. We did not consider degrees where there is only one lecture group.

3.1 Educational Institution

In Spain access to university degrees is based on applicants' *entry grade*, which is calculated as a weighted average of their High School accumulated GPA (60%) and the grade obtained in a standardized exam known in Spanish as *Selectividad* (40%). University Carlos III offers the most selective degrees in the region according to the required minimum entry grade. Out of the seven degree-locations included in our study, the degree in Business and Law in the Southern campus is the most selective one, while the least selective degrees are Economics and Finance. [2]

Each year includes two 14-week terms. The first term takes place from September to December, with exams taken in January. The second term takes place from February to April, with exams taken in May. Students that fail to pass an exam on either of the two terms have the chance to resit that exam in June.

Each week students attend each one lecture and one tutorial. The assignment of students to lecture and tutorial groups is based on their surname initial.[3] As an illustration, Figure 1 shows how students enrolled in 2010 in the 1st year of the Business degree in the Southern campus were distributed across groups. For instance, students whose surname initial was "A" or "B" were assigned to tutorial group number 74 and lecture group "74-75-76" (which combines tutorial groups 74, 75 and 76). As we show below, in the Spanish context surname order is uncorrelated with socio-economic status or academic performance and, as a result, performance across groups tends to be balanced.

All courses in the 1st and 2nd year of the degree are compulsory. Courses in the 3rd and 4th year of the degree tend to be optional. In each course the final grade is usually a weighted average of the grade obtained in the end of term exams (60%), midterm evaluations (20%) and group presentations/assignment (20%). The end of term exam is usually the same in different groups of the same subject.

[2] Information on minimum entry grades is available at `http://portal.uc3m.es/portal/page/portal/acceso_universidad/notas_corte_pc/notas_corte_09_10/notasmadrids09.pdf`, retrieved on April 30 2015.

[3] The only exception are second year students in the English track. That is why we do not consider these students in our analysis and restrict the analysis to students in the Spanish track.

Students' permanence in the university is subject to certain requirements. During their first year at Carlos III, students must pass at least two courses. By the end of their second year, they must have passed every first year course. Finally, they cannot fail the same exam more than three times. If any of these conditions is not satisfied, students cannot pursue their studies.[4]

Students receive regularly information on the grades that they have obtained in each subject. The university summarizes this information through an official measure of Accumulated Grade Point Average (AGPA), which students can also access at any point in time in the intranet of the university.[5] Students do not receive information about their position in the distribution of AGPAs, relative to other students, or about the AGPA of any other students.

Students are not explicitly rewarded for their relative performance, except for a prize given to the best student in the cohort.[6] Nonetheless, relative performance might be relevant. For instance, many students enroll in the Erasmus exchange program, typically during their third or fourth year. Whether students are admitted to the program or not is based on their performance in a language exam and their position in the distribution of grades. The relative position of students in the distribution of grades might also play a role when students apply for an internship, typically during the last year of the degree, or later after graduation, when they enter the labor market.

3.2 Field Experiment: Design and Procedure

The intervention was restricted to students who had entered the university in Fall 2009 and who were registered in at least one second-year course in Fall 2010. This condition excludes approximately 10% of the 2009 cohort, in general students who were expelled

[4]More detailed information is available at the webpage of the university `http://portal.uc3m.es/portal/page/portal/conocenos/nuestros_estudios/normativa_09/Permanencia`), retrieved on February 11 2015.

[5]The university calculates the accumulated grade point average adding up the grades obtained by the student, modified with a penalty for the number of times the exam is taken, and dividing this sum by the total number of courses taken. There is no penalty if the exam for the course is taken only once. If the student failed once the course grade is multiplied by 0.95, twice by 0.90 and so on.

[6]This prize, known as *premio extraordinario* is awarded by the ministry of education upon graduation.

because they did not manage to satisfy the permanence requirement: passing at least two courses during the first year.

Students' were assigned randomly to the treatment or to the control group based on the lecture group in which they were enrolled in.[7] We selected randomly one of the 432 different possible assignments. The set of possible assignments was subject to the constraint that there is one treated group per degree-location. As a result of the random draw, 623 students were assigned to the treatment group and 354 to the control group. Figure 6 shows the distribution of students to the control and the treatment group by degree.

The intervention starts in early December of 2010 and it concludes three years later, at the end of the fourth academic year. During this period students in the treatment group were granted access to feedback on their relative performance every six months. More precisely, treated students received every six months an email message from a corporate account saying:

> This email is part of a pilot project of academic assessment management. If you want to see your average grade, and your relative position in terms of average grade among the students that started the degree the same year you did, you can do it by clicking **here**

After logging in with their university login and password, students get access to a screen where they can observe their own GPA and also their position in the distribution of grades, measured in deciles (Figure 2).

Note that students receive information about their position in the ranking in terms of their AGPA. Given that by construction the influence of each additional course on their ranking decreases overtime, it is not surprising that students' position in the ranking varies increasingly less over time. As shown in Figure 3, while 45% of students experienced a variation in their ranking at the beginning of their 2nd year, at the end of the 4th year only 25% of students experience any such variation.

[7] A few students were enrolled in several groups. They were assigned to the group where they attended the majority of the courses.

3.3 Survey data

We also collected information from three different surveys: (i) teaching evaluations filled by students (ii) a survey about students' knowledge of their relative position in the distribution of grades, to a sample of 2nd year students, who were not affected by the intervention, (iii) a similar survey to a sample of graduating students belonging both to the treatment and the control group.

3.3.1 Teaching evaluations

Teaching evaluations are collected by the University administration. They are anonymous, so we cannot match the teaching evaluations to the students in our sample. But given we know the tutorial group the teaching evaluations belong to, we can assign teaching evaluations to the treatment and the control group based on the tutorial group during the academic year 2010-2011, when students are registered in compulsory 2nd year courses. Unfortunately we cannot match the information during the third and the fourth academic years, when most courses are elective. Students complete teaching evaluations twice a year.

During the academic year 2010-2011, students completed their 1st term teaching evaluations before the intervention took place and they completed their 2nd term teaching evaluations after they had received feedback on their relative performance and also the results of exams conducted in January. We were provided with teaching evaluations for a random sub-sample of 347 students, for 142 students before the treatment and 165 students after the treatment.

3.3.2 Survey: Students' information on relative performance at the beginning of the 2nd year

In order to obtain information on Carlos III students' prior information about their position in the distribution of grades, we conducted a survey among a group of students of the 2010 cohort at the beginning of their second year (November 2011). The survey was administered during the lecture of a compulsory course and in total 57 Economics students participated.[8] We decided not to conduct this survey among students belonging to the treated cohort (2009 cohort) in order to avoid the introduction of any confounding effects that might perhaps affect their performance later on.

Students were asked to answer privately the following question:

When you enrolled one year ago in this degree your cohort included N students. If we were to rank all students in this cohort by their Accumulated Grade Point Average (AGPA), such that number 1 is the student with the highest AGPA and number N is the student with the lowest AGPA. In which position do you think you would be?

Note: N was equal to 300, which corresponds to the number of students who enrolled in 2010 in the Economics degree offered by Universidad Carlos III in its Southern Campus

3.3.3 Survey: Students' information on relative performance at the time of graduation

We also surveyed a sample of students from the treatment and the control groups three years after the intervention about their relative ranking. The survey was conducted at end of the undergraduate thesis presentation, which is the last requirement that students satisfy before graduation.[9] The sample includes 97 students from Economics, Business and Finance degrees. Four students did not reply to the survey. By construction the sample of students who was surveyed is not a random sample of all students. Students in the upper part of the grade distribution are over-represented. We discuss the results of this survey in the empirical analysis section.

[8] More precisely, we surveyed students enrolled in Game Theory, Degree in Economics, groups 63, 64, 68, 69. 21 people did not attend the lecture the day of the survey. All attending students except one participated in the survey.

[9] To prevent (treated) students from having access to the information provided, they were not allowed to access internet during the survey.

4 Data

4.1 Individual characteristics

Table 1 provides information on the predetermined individual characteristics of the 977 students who participated in the intervention. A little over half of the students are women, and practically all of them are Spanish. In general they attended previously a High School, only 5% have a vocational training background. Around two thirds of the students come from the Madrid region and, within this region, most of them come from the center of Madrid (31%). Approximately 22% come from municipalities located in the Southern part of the metropolitan area. This group has probably a relatively lower socio-economic status.

The average entry grade into the university is 7.24 (out of 10). Interestingly, students' grades are significantly lower at the university than their entry grade during their first year at university. As shown in Figure 4, grades shift down along the whole distribution. The average grade decreases from 7.24 to 6.02, roughly about one standard deviation. Information on academic performance is also available for the second, third and fourth year, after the treatment was implemented. This information includes the elective courses chosen, the number of exams taken in each course, the grade obtained in each exam and students' AGPA.

We test formally whether the predetermined characteristics of individuals assigned to the treatment and the control group are statistically significant using the following regression:

$$X_{s,d,g} = \alpha + \beta Treatment_{d,g} + \gamma Z_d + \epsilon_{s,d,g} \tag{6}$$

where $X_{s,d,g}$ refers to a given predetermined characteristic of student s, enrolled in degree d and tutorial group g. $Treatment_{d,g}$ takes value one if the student is exposed to the treatment and the equation also includes a set of degree fixed effects (Z_d). As one can see in Table 1, columns 2-3, the two groups are very similar. Out of eleven observable characteristics, in no dimension the difference is significant at the 5% and in two dimensions the difference is significant at the 10% (Table 1, columns 4). An

F-test confirms that one cannot statistically reject that the assignment was random.

Table 2 provides information on students' academic performance during the intervention. During the regular exam season of their second year students take on average eleven exams and they pass approximately eight. Students have the chance in June to resit exams that they had failed. During the second year resit season students on average take around three exams and pass one of them. The number of exams taken and passed during the third and the fourth year is slightly lower. By September of their fourth year approximately half of the students in our sample have managed to graduate and 15% had dropped out, typically during their second year.[10]

4.2 Teaching evaluations and survey data

First, according to the information provided by teaching evaluations, students were relatively satisfied with the quality of the courses they receive before the intervention took place. In a scale from 1 (not at all) to 5 (very satisfied), students' average assessment is equal to 3.8 (Table 3, column 1). They are slightly less satisfied with the fairness of grading. Again using a scale from 1 and 5, the average answer is 3.6. Students devote each week roughly between 4 and 7 study hours to each subject.[11] Taking into account that there are typically 5 or 6 courses per term, this implies that on average students spend each week approximately 32 hours studying.

According to survey information provided by teachers, the attendance rate to lectures is above 80%.[12] Each course includes four hours of weekly lectures, which implies that a student enrolled in 5.5 courses who attended 80% of lectures, would spend 18 hours weekly sitting in class. Adding up study time and lectures, the average student devotes around 50 hours a week to college related work (according to the self-reported survey information).

[10] This calculation excludes 200 students who were enrolled in the Business and Law degree, which has a six-years length.

[11] This information is only available at the group level. Hours of study takes value 1 if the individual studied less than an hour per week; 2, between one and four hours; 3, four to seven hours; 4, seven to ten hours and 5 more than ten hours.

[12] Information available at `https://portal.uc3m.es/portal/page/portal/calidad/Resultados_encuestas_a_alumnos_y_profesores/00_Informe_1_cuatrimestre_2012_2013.pdf`, retrieved on April 30 2015.

Second, we also measure students' knowledge about their position in the distribution of grades, absent of any intervention. The answers are reported in Figure 5. The x-axis reports the actual position of the student in the ranking, normalized between 0 (lowest grade) and 1 (highest grade) among students who enrolled in Economics in Fall 2009. The y-axis provides information on their self-reported relative performance, normalized in a similar way. Most observations lie far below the diagonal, reflecting that students tend to uninformed. In particular, students underestimate their position in the distribution of grades. The average student makes an error in his prediction of 22 percentiles and she tends to underestimate her relative ranking by 18 percentiles.

5 Empirical analysis

5.1 Do students take the treatment?

Students in the treatment group receive an email with a link to a personalized webpage where they can find feedback on their relative performance. 72% of students checked this information at least once. The average student checked four times the ranking during the duration of the treatment. As shown in Figure 7, the probability to check is strongly correlated with the position in the ranking. In the top quartile almost 90% of students accessed the information, in the bottom quartile less than half did. Female students are also slightly more likely to check, but the difference only marginally significant once we take ranking into account (Table 4).

Unfortunately we cannot disentangle why some students do not check the information. Some individuals might not read emails from corporate accounts, some others perhaps read the email but prefer not to find out about their position in the ranking. One third of students that did not check their ranking were expelled from the university at the end of their second year due to the unfullfilment of the permanence requirements. It is possible that these students were not active students at the time of the intervention.

5.2 Was the treatment informative?

An important element of the intervention is whether it managed to have a differential impact on the information available to students in the treatment and the control groups. The intervention was designed to minimize information spillovers, but it is still possible that students from the control group received some information from treated students. Students in both groups might also increase over time their knowledge about their position in the distribution, independently of the intervention.

We study this issue using the information provided by the survey conducted at graduation. The information displayed in Figure 8 reveals two interesting patterns. First, compared to students at the beginning of their 2nd year, students at the end of their 4th year have more accurate information about their relative performance. The average error has decreased from 22 percentiles to 12 percentiles. Second, students in the treatment group are better informed than students in the control group. The average error is equal to 9 percentiles among students in the treatment group and equal to 15 percentiles among students in the control group. This difference is statistically significant at the 5% level (Table 5).

For students in the control group, this improvement might potentially reflect learning over time or potential information spillovers. Unfortunately we cannot disentangle these two hypothesis. Note also that students in the treatment group do not perfectly predict their position in the ranking. This might be due to several factors. First, students were asked about their exact position in the ranking, while the intervention provided access only to their position in terms of decile. Second, the survey was conducted after the final exams but before students could access information on their final ranking, the last update of the ranking information took place shortly after we conducted the survey. Third, a few students in this group (less than 10%) had never checked the information provided. Last, some students may have forgotten their position in the ranking.

5.3 Feedback effect on academic performance

We analyze the impact of the treatment on academic performance. More precisely we estimate the following regression:

$$Y_{s,d,g,t+i} = \alpha + \beta Treatment_{d,g} + \boldsymbol{X_{s,d,g,t}\gamma} + \lambda Z_d + \epsilon_{s,d,g,t+i} \tag{7}$$

where $Y_{s,d,g,t+i}$ stands for the performance of student s, enrolled in degree d and tutorial group g, in the academic term $t+i$, and t refers to the time of the intervention. $Treatment_{d,g}$ takes value one if the student is exposed to the treatment, and the equation also includes a set of degree fixed effects (Z_d). In some specifications we also include the set of predetermined individual controls $X_{s,d,g,t}$ listed in subsection 4.1. To account for potential existence of common shocks, we report standard errors clustered at the tutorial group level (45 groups).

As expected, the performance of the treatment group and the control group was similar before the intervention took place (Table 6, 1st row). The intervention took place in Fall of the second year. We do not observe any impact on the number of exams taken by students during the regular exam period. However, the performance of the treatment group is significantly worse. On average, students in the treatment group passed 0.41 (10% of a standard deviation) fewer exams during the regular exam period. Columns 3 and 4 provide information about resits, which are scheduled in June. Students in the treatment group take 0.38 more resits, reflecting their higher failure rate during the year, and they manage to recover half of the gap. During the third and the fourth years there are no significant differences in performance between the treatment and the control group. If anything the performance of the treatment group is slightly better. By the end of the fourth year there are no significant difference between students in the treatment or the control group in terms of the number of exams passed, the dropout rate, time to graduation or the accumulated grade point average (Table 7). In sum, the treatment group experiences a short-term negative impact on performance but in the longer term the gap disappears.

5.4 Heterogeneity analysis

Are all students equally affected by the provision of information on relative performance? We consider several sources of heterogeneity. First, we try to infer which students are receiving 'good news', students learn they are ranked higher than expected as they were underestimating their rank, or 'bad news', students learn they are ranked lower than expected as they were overestimating their rank. We do not have direct information on the priors of students that participated in the intervention, but we can exploit the information provided by the survey that was conducted during the second year among the students who ere not affected by the treatment. Using this information we can predict the expected news that students are expected to receive when they get access to the ranking information based on their observable characteristics. We estimate the following equation:

$$Y_s = \alpha + \boldsymbol{X_s}\beta + \epsilon_s \tag{8}$$

where Y_s refers to the difference between the self-reported and the actual relative ranking. The dependent variable takes positive values when students overestimate their own ranking and negative otherwise. The set of controls $\boldsymbol{X_s}$ includes gender, entry grade, and performance during the 1st year. As shown in Table 8, underconfidence is relatively stronger among students in the upper part of the ranking and it tends to be lower among students that used to obtain high grades in High School. Women underestimate their position by 7 percentiles, but this gap is not statistically significant. These observable characteristics explain 50% of the variation.

We use $\hat{Y_s}$ to predict whether a given student was positively or negatively surprised by the feedback on relative performance. More precisely, we infer a positive surprise whenever the actual ranking is larger than the predicted self-reported ranking and a negative surprise otherwise. Using this methodology we infer that 729 students were positively surprised and 248 were negatively surprised. We regress equation 7 separately for these two groups of students, using as dependent variable the number of exams passed during the second year in the regular exam period. According to our estimates, receiving positive news reduces by 0.45 the number of exams passed during the second year and it has virtually no effect on students who are expected to receive a positive (Table 9, columns 2 and 3). These estimates are imprecise and it is not possible to reject at the 5% level that they are equal, but overall they are consistent with the hypothesis that receiving good news affects performance negatively.

Another possible way to infer which students were positively surprised by the feedback on relative performance is to examine how their grades have evolved from High School to University. Most students experience a drop in their average grade and a few they obtain a slightly higher grade at university. As shown in Table 9, columns 4 and 5, only the later group seems to be affected by the treatment.

We also examine whether the impact of the treatment differs by gender. The point estimate is slightly larger in the case of women, but the magnitude is statistically similar (columns 6 and 7). We do not find any differential effect either according to whether individuals are above or below the mean (columns 8 and 9).

5.5 Feedback effect on teaching evaluations

We analyze the impact of the treatment on teaching evaluations using the following regression:

$$Y_{c,g,d} = \alpha + \beta Treatment_{c,g,d} + \boldsymbol{X_c}\boldsymbol{\gamma} + \boldsymbol{Z_d}\boldsymbol{\lambda} + \epsilon_{c,g,d} \tag{9}$$

where $Y_{c,g,d}$ stands for some average self-reported measure in course c (e.g. Econometrics I), tutorial group g (e.g. group 72) and degree d (e.g. Business in the Southern Campus). The regression includes a set of course fixed effects ($\boldsymbol{X_c}$) and degree fixed effects ($\boldsymbol{Z_c}$). Before the intervention, students in the treatment and the control group report very similar values in terms of their overall satisfaction with courses, the fairness of the grading and the hours of study (Table 10, columns 1-3). During the second term both groups report statistically similar values in term of hours of study and fairness of grading but the satisfaction of the treated group is significantly larger (approximately one third of a standard deviation).

5.6 Robustness checks

5.6.1 Instrumental variables

Not all students in the treatment group accessed the information. As shown in Table 4, tend to be relatively better students and they are slightly more likely to be female. We also conduct the analysis using an instrumental variable (IV) strategy where we use the (random) assignment to the treatment group as an instrument for accessing the information. The point estimates from the IV exercise are slightly larger but overall the results are statistically similar (Table A1).

5.6.2 Placebos

To be done

5.6.3 Randomization inference

To be done

6 Conclusions

Students at Carlos III University appear to be uninformed about their relative performance and they tend to underestimate their position in the distribution of grades. A random sample of students were given access to learn about their relative position in the distribution of grades. The treatment was effective in informing students about their rank, while students in the control group remained relatively uninformed and underestimate their rank. We find that providing feedback on students' relative performance has a negative impact on their performance during a few months but it has no long-term impact on performance. We also observe a positive effect on self-reported student satisfaction with the quality of the courses, perhaps associated to the positive surprise about their own performance.

The evidence is partly consistent with a context where most students receive positive news, learning they are ranked better than expected, which decreases performance and increases their satisfaction. These results are consistent with individuals who initially underestimate their position in the distribution of performance and learn they are doing better than expected, which leads them to decrease effort. Our results also suggest that the impact of relative performance feedback may depend crucially on individuals prior information and their preferences. In the absence of this information, more field experiments might be very useful to estimate the impact of this policy.

References

[1] Azmat, Ghazala and Nagore Iriberri (2010), "The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment Using High School Students," *Journal of Public Economics*, Vol. 94(7-8), pp. 435–452.

[2] Azmat, Ghazala and Nagore Iriberri (2014), "The Provision of Relative Performance Feedback: An Analysis of Performance and Satisfaction," *Journal of Economics and Management Strategy*, forthcoming.

[3] Barankay, Iwan (2012), "Rank Incentives: Evidence from a Randomized Workplace Experiment", mimeo.

[4] Blanes i Vidal, Jordi and Mareike Nossol (2011), "Tournaments without Prizes: Evidence from Personnel Records," *Management Science* Vol. 57, pp. 1721–1736.

[5] Charness, G., Masclet, D., and M.C. Villeval. (2013), "The Dark Side of Competition for Status," *Management Science* 60(1), 38–55.

[6] Eriksson, T., Poulsen, A., Villeval, M., (2009), "Feedback and incentives: experimental evidence," *Labour Economics* 16, 679–688.

[7] Falk, A., Ichino, A., (2006), "Clean evidence on peer pressure," *Journal of Labor Economics* 24 (1), 39–57.

[8] Gerhards, L. and N. Siemery, (2014), "Private versus public feedback: The incentive effects of symbolic awards," mimeo.

[9] Hannan, R.L., Krishnan, R., Newman, D., (2008), "The effects of disseminating relative performance feedback in tournament versus individual performance compensation plans," *The Accounting Review* pp.83–4.

[10] Katreniakova, Dagmara (2014), "Social Comparison, Rewards and Incentives to Learn: A randomized control trial in Uganda," CERGE-EI, mimeo.

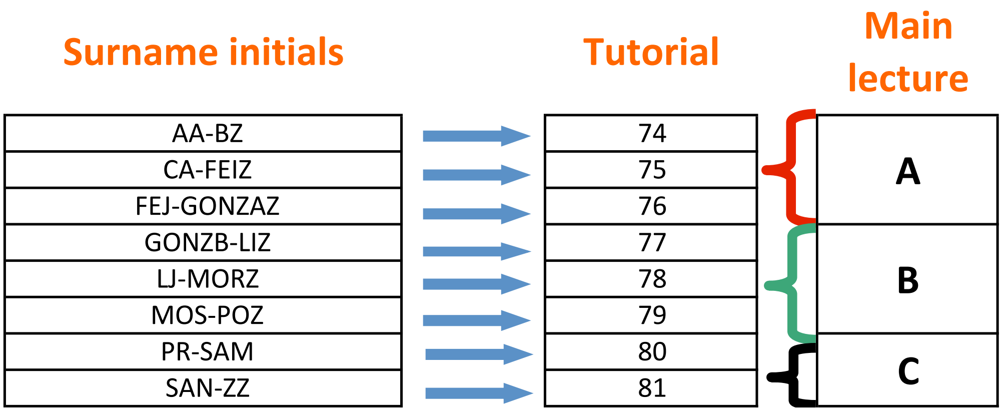
[11] Khunen, Camelia N., and Tymula Agnieszka (2012), "Feedback, Self-Esteem, and Performance in Organizations," *Management Science*, Vol. 58(1), pp. 94–113.

[12] Kuziemko, I., R. Buell, T. Reich, and M. Norton (2013), "Last-place Aversion: Evidence and Redistributive Implications," *Quarterly Journal of Economics* (forthcoming)

[13] Mas, A., Moretti, E., (2009), "Peers at work," *American Economic Review* 99 (1), 112–145.

[14] Tran, Anh, and Richard Zeckhauser (2012), "Rank as an inherent incentive: Evidence from a field experiment," *Journal of Public Economics*, Vol 96, pp. 645–650.

Figure 1: Assignment to Tutorial and Lecture Groups



Note: This assignment corresponds to 1st year students, Business Administration, Getafe, Spanish track, 2010.

Figure 2: Feedback on Relative Performance

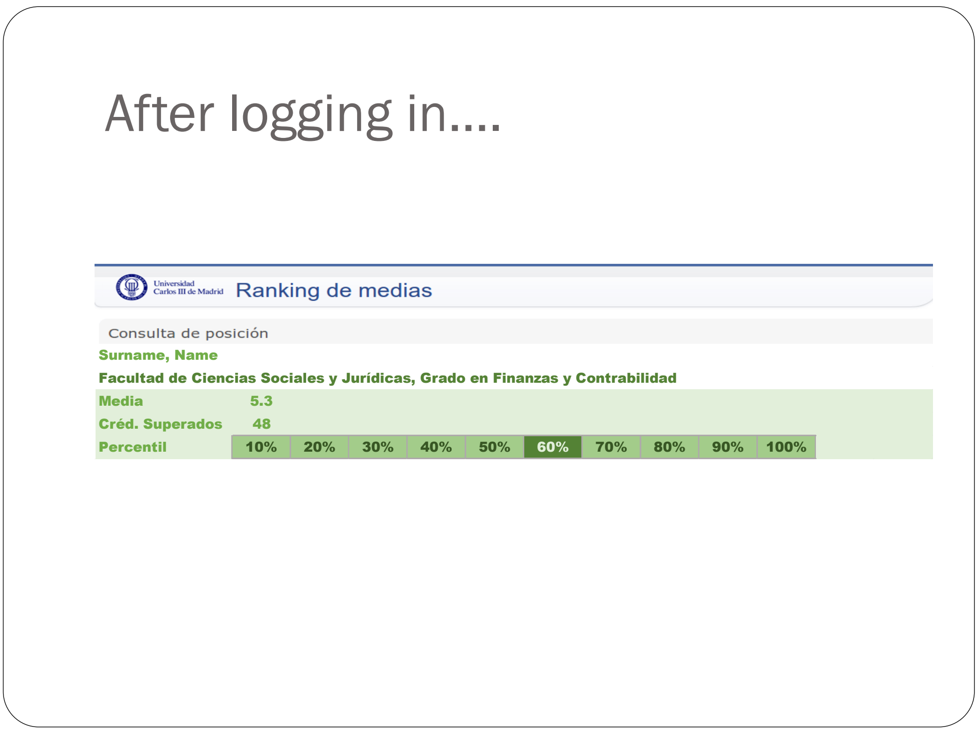
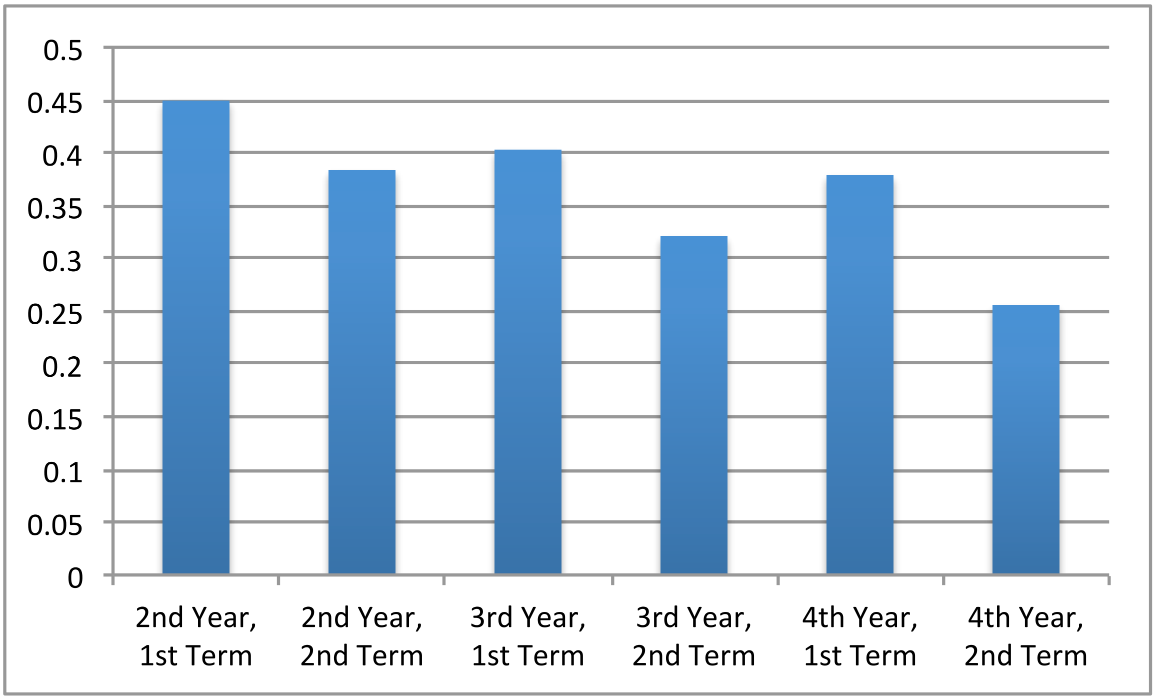


Figure 3: Information over time



Note: Each bar reflects the proportion of people who experienced mobility from one term to the next in terms of their decile in relative distribution. For instance, approximately 45% of individuals were placed in a different decile at the end of the 1st term of their 2nd year relative to their position at the end of the 1st year.

Figure 4: Entry grade and 1st year grades at college

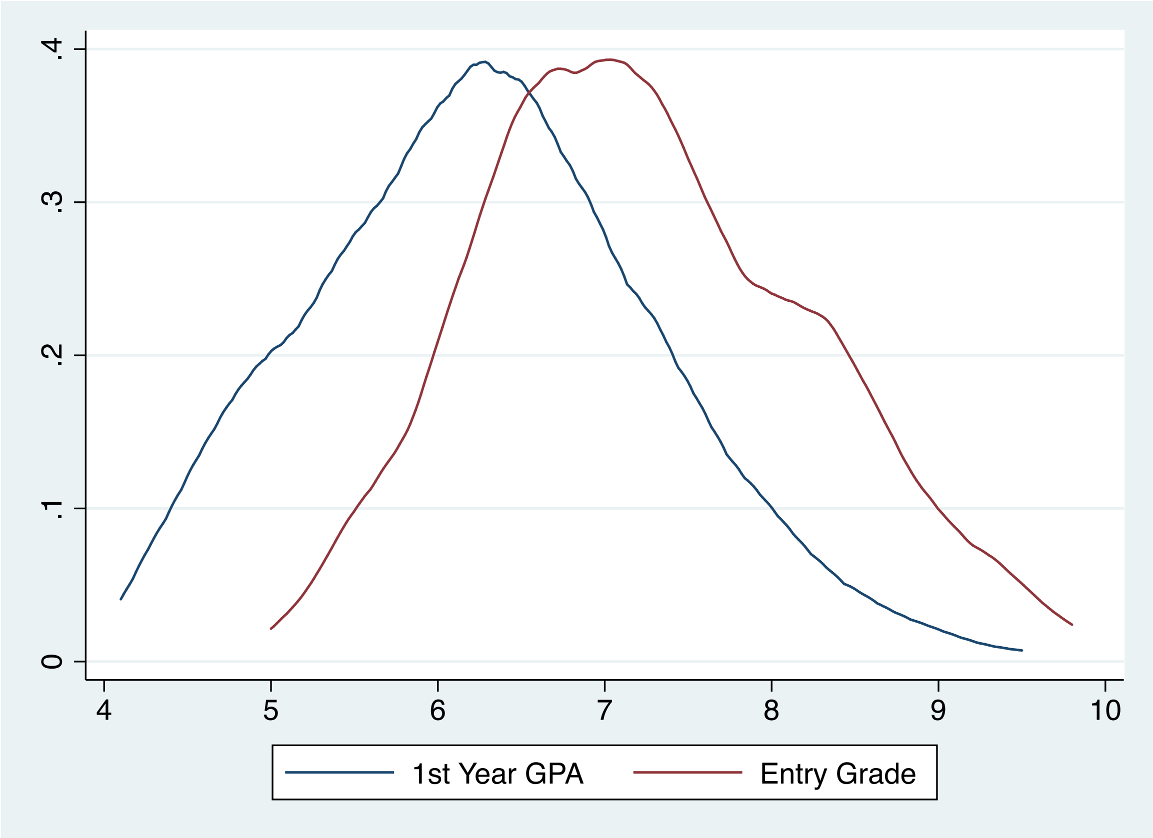
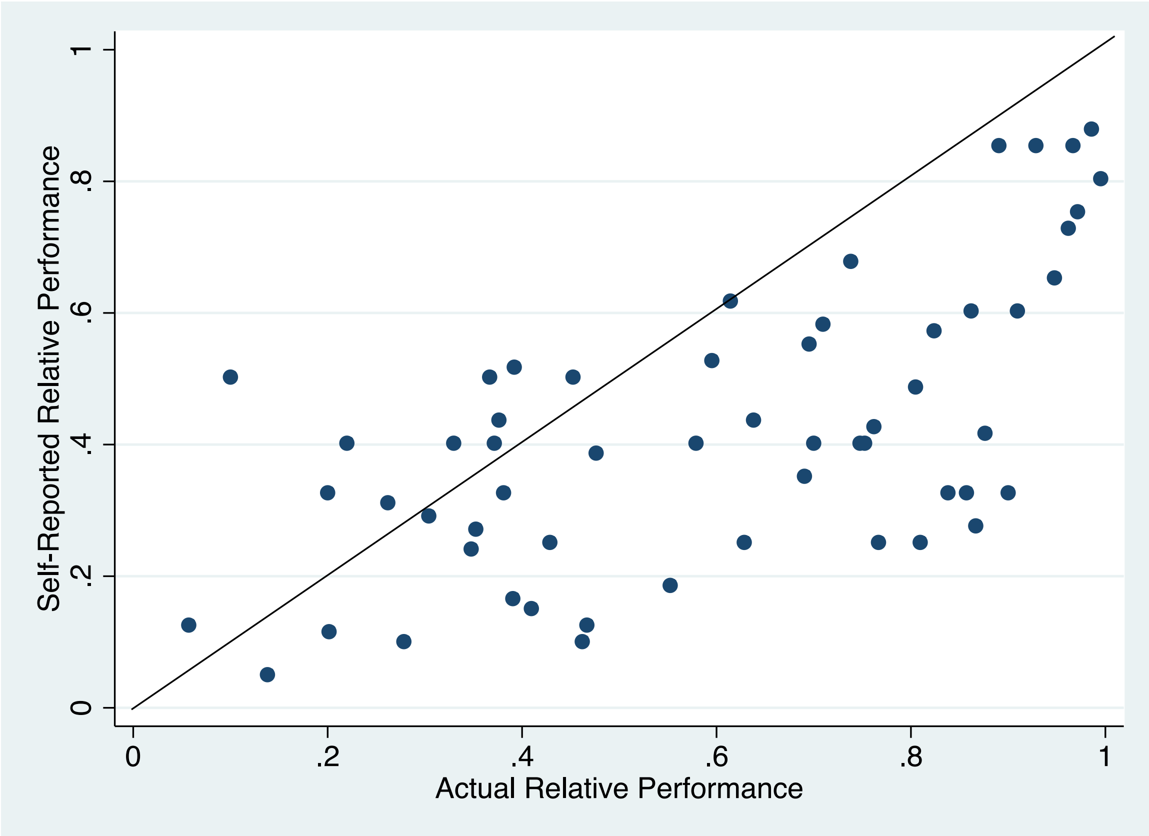


Figure 5: Relative performance at the beginning of the 2nd year



Note: The figure includes information from 57 second year Economics students, class of 2014, who were surveyed in November 2011. The x-axis reports the actual position in the ranking, normalized between 0 (lowest grade) and 1 (highest grade) among students who enrolled in the same degree in Fall 2009. The y-axis provides information on the self-reported relative performance, normalized in a similar way.

Figure 6: Assignment to Treatment and Control

	South Madrid		North Madrid	
	Treatment	Control	Treatment	Control
Finance and Accounting	1	1		
Economics	1	2		
Business	1	2	1	1
Law	1	2		
Law and Business	1	1	1	1

Figure 7: Share of individuals who checks the ranking, by quartile

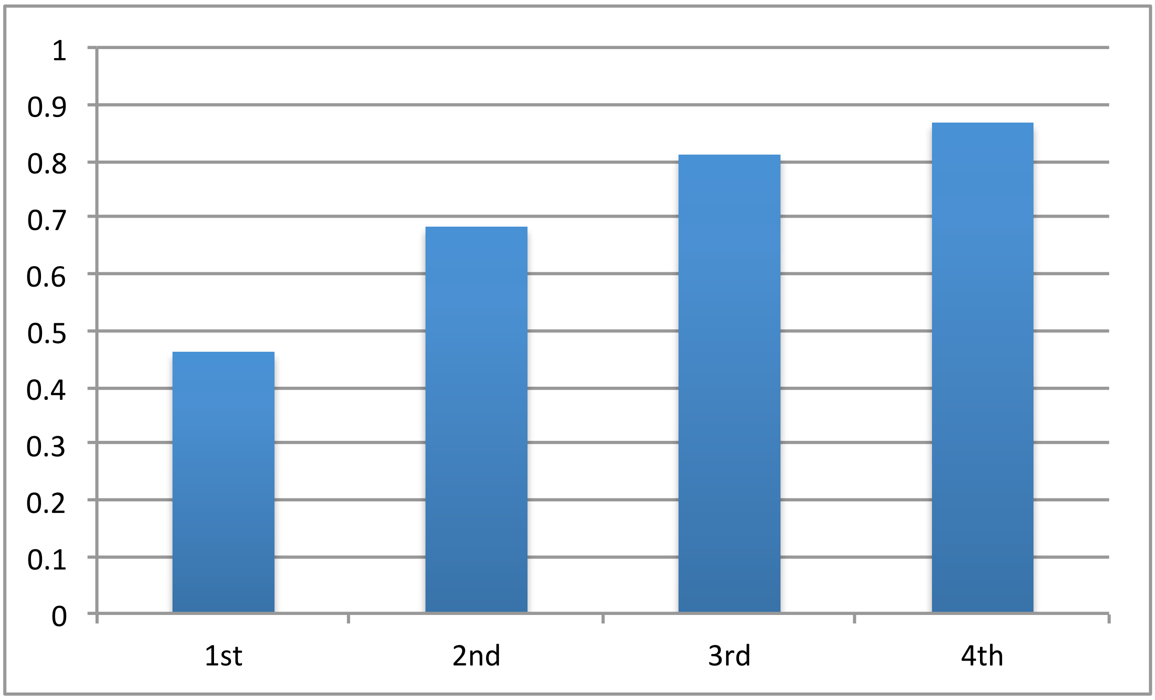
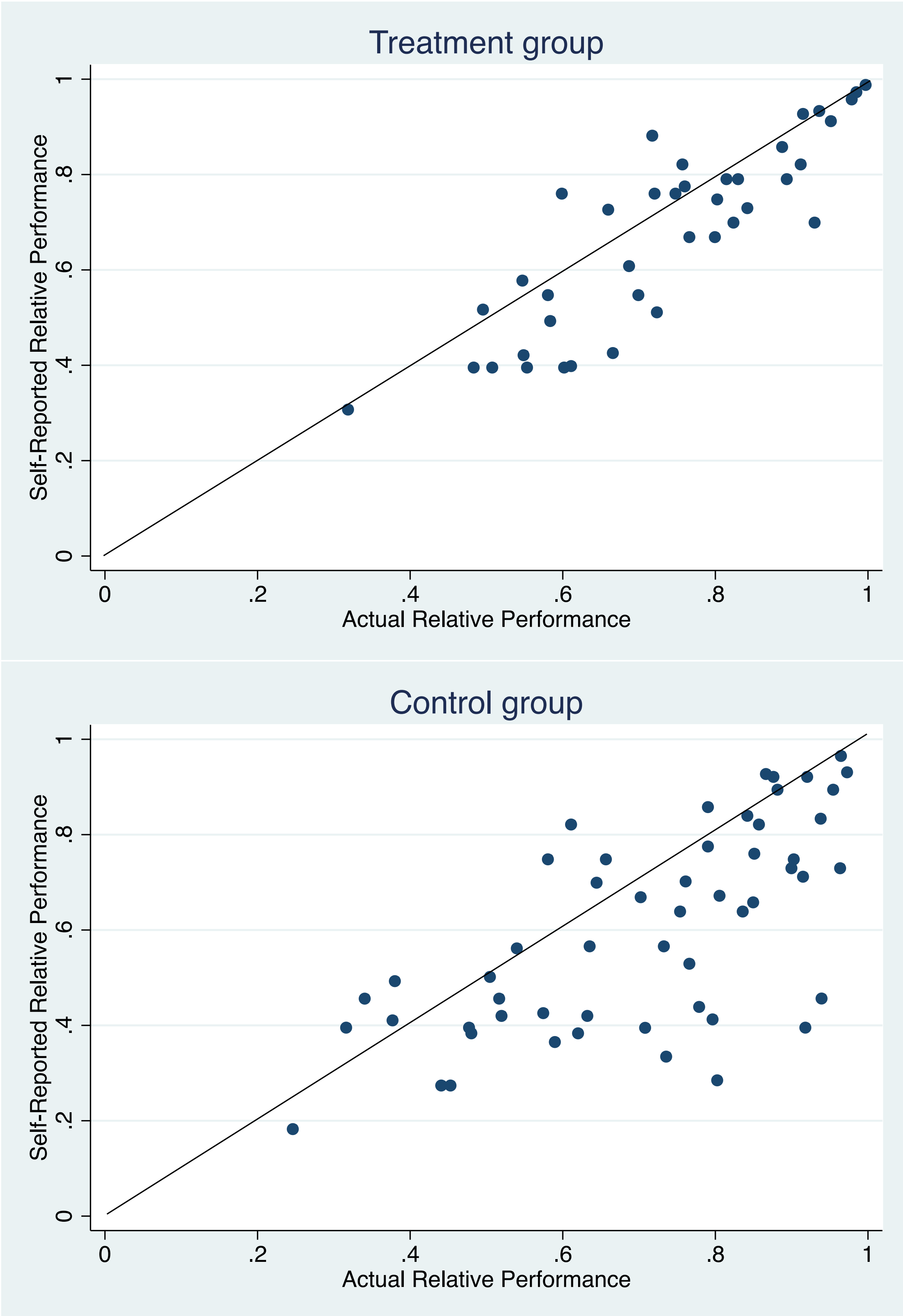


Figure 8: Relative performance at graduation, treatment group



Note: The figure includes information from 93 students in Economics and Business who were surveyed in the summer of 2013, at the time of graduation. The upper (lower) panel includes students in the treatment (control) group. The x-axis reports the actual position in the ranking, normalized between 0 (lowest grade) and 1 (highest grade), relative to students from the same cohort. The y-axis provides information on the self-reported relative performance.

Table 1: Descriptive statistics, individual level

	1	2	3
	All	**Treated-Control**	
	Mean	Difference	p-value
Female	0.54	0.03	0.43
Foreigner	0.03	-0.00	0.71
High School	0.95	-0.02	0.17
Central Madrid	0.31	-0.01	0.81
Western Madrid	0.11	0.01	0.72
Southern Madrid	0.22	0.05	0.07
Other regions	0.30	-0.04	0.19
Entry Grade	7.24	-0.10	0.07
Grade	6.02	-0.05	0.45
Percentile	0.54	-0.02	0.40
Exams taken	4.89	-0.06	0.23
Exams passed	3.70	-0.09	0.64

Note: The table includes information on 977 students that took part in the intervention. Column (1) provides information on the mean value of each characteristic. Column (2) reports the difference between the treatment and the control group, conditional on degree. Column (3) reports the p-value of this difference.

Table 2: Descriptive statistics, performance

	1	2	3	4
	Mean	St. Dev	Min	Max
		Second year		
Exams taken	10.69	3.19	0	19
Exams passed	7.75	3.83	0	16
Retakes taken	2.91	2.94	0	17
Retakes passed	1.12	1.25	0	8
		Third year		
Exams taken	10.26	4.52	0	20
Exams passed	8.07	4.06	0	19
Retakes taken	2.15	2.65	0	18
Retakes passed	0.98	1.28	0	7
		Fourth year		
Exams taken	8.59	4.68	0	23
Exams passed	6.69	4.41	0	18
Retakes taken	1.22	2.09	0	15
Retakes passed	0.68	1.11	0	7
		Overall		
All exams taken	36.46	13.91	0	87
All exams passed	25.82	11.54	0	46
Dropout	0.15	0.36	0	1
Graduation in 4 years	0.51	0.5	0	1

Note: The table includes information on 977 students that took part in the intervention, except for the variable *graduation rate*, which excludes 200 students enrolled in the six-years degree in Business and Law. The variables *Exams taken* and *Exams passed* refer respectively to the number of exams taken or passed during the regular exam season (January and May). Variables *Retakes taken* and *Retakes passed* refer exams taken and passed during the retake season (June). The lower panel provides information measured in September 2013, at the end of the fourth academic year. *AGPA* refers to the Accumulated Grade Point Average.

Table 3: Descriptive information, teaching evaluations

	1	2	3	4	5	6
	First semester			**Second semester**		
	N	Mean	St. Dev	N	Mean	St. Dev
Satisfaction	182	3.87	0.76	165	3.63	0.85
Hours of study	182	2.92	0.45	165	3	0.48
Grading	182	3.56	0.67	165	3.15	0.82

Note: The table includes information from 347 tutorial groups from the second year, academic year 2010-2011. The first row provides information on students' self-reported satisfaction with the overall quality of each course, coded in a scale from 1 (not at all) to 5 (very satisfied). The second row reports the average satisfaction with the grading, also coded in a scale from 1 (not at all) to 5 (very satisfied). The third row provides information on the number of hours studied weekly. Hours of study takes value 1 if the individual studied less than an hour per week; 2, between one and four hours; 3, four to seven hours; 4, seven to ten hours and 5 more than ten hours.

Table 4: Who checks the information?

	1	2
Female	0.106**	0.079*
	[0.047]	(0.045)
True rank		0.585***
		(0.097)
Entry grade		-0.047
		(0.034)
Constant	0.665***	0.708***
	[0.034]	(0.229)
Observations	354	347
R-squared	0.084	0.161

Note: The regression includes information from 354 students who were assigned to the treatment group. The dependent variable is a dummy that takes value one if the students checked at least once the information. Robust standard errors in parenthesis. *: $p < 0.10$, ** $p < 0.05$,*** $p < 0.01$

Table 5: Available information at graduation

	1	2
Treatment	-0.050**	-0.048*
	(0.025)	(0.025)
Female		0.048**
		(0.022)
True rank		0.023
		(0.070)
Entry grade		-0.004
		(0.023)
Constant	0.143***	0.129
	(0.019)	(0.142)
Adj. R-squared	0.055	0.073
N	93	93

Note: The regression includes information from 93 students who were surveyed at graduation. The dependent variable is the difference between the self-reported position in the ranking and the actual one, normalized between 0 and 1. Robust standard errors in parenthesis. *: p < 0.10, ** p < 0.05,*** p < 0.01

Table 6: Impact on academic performance

	1	2	3	4
	Regular exams		**Retakes**	
	Taken	Passed	Taken	Passed
First year	-0.061	-0.044	0.037	-0.022
	(0.046)	(0.099)	(0.172)	(0.056)
Second year	-0.070	-0.408**	0.372*	0.211*
	(0.165)	(0.179)	(0.220)	(0.112)
Third year	0.201	0.073	0.075	0.065
	(0.307)	(0.269)	(0.164)	(0.084)
Fourth year	0.119	0.298	-0.181	0.012
	(0.360)	(0.344)	(0.170)	(0.064)

Note: Each cell reports the main estimate from a different regression on the sample of 977 students that took part in the intervention. The independent variable is a dummy that takes value one if the student was part of the treatment group. The first row provides information for the 1st academic year, the second row for the 2nd academic year, and so on. The first two columns report information from exams taken during the regular period (January and May). Columns (3) and (4) provide information from retakes (June). The dependent variable in columns (1) and (3) is the number of exams taken. The dependent variable in columns (2) and (4) is the number of exams passes. All regressions include a control for academic performance during the first year. Standard errors clustered at the tutorial level in parenthesis. *: $p < 0.10$, ** $p < 0.05$,*** $p < 0.01$.

Table 7: **Long term impact on academic performance**

	1	2	3	4
	Exams passed	Dropout	Graduation	True rank
Treatment	0.271	-0.006	0.021	0.008
	(0.782)	(0.022)	(0.031)	(0.012)
Adj. R-squared	0.437	0.260	0.265	0.640

Each column reports the result of a different regression on the sample of 977 students that took part in the intervention. In the first column the dependent variable is the overall number of exams passed between enrollment in 2009 and the summer of 2013. In columns (2) and (3) the dependent variable is respectively a dummy that takes value one if the student had dropped out or had graduated by September 2013, four years after enrollment. In the last column the dependent variable is the position in the ranking in September 2013. All regressions include a control for academic performance during the first year. Standard errors clustered at the tutorial level in parenthesis. *: p < 0.10, ** p < 0.05,*** p < 0.01

Table 8: Available information at the beginning of the 2nd year

	1	2	3	4
	Absolute error		**Average error**	
Female		0.120***		-0.066
		(0.034)		(0.041)
True rank		0.327***		-0.657***
		(0.105)		(0.112)
Entry grade		-0.075**		0.105***
		(0.031)		(0.034)
Constant	0.223***	0.484**	-0.177***	-0.493**
	(0.021)	(0.186)	(0.028)	(0.209)
Adj. R-squared	0.000	0.347	0.000	0.501
N	57	52	57	52

Note: The regression includes information from 57 students who were surveyed at the beginning of their 2nd year in college. In columns (1) and (2) the dependent variable is the difference between the self-reported position in the ranking and the actual one in absolute value, normalized between 0 and 1. In columns (3) and (4) the dependent variable is the difference between the self-reported position in the ranking and the actual one, normalized between 0 and 1. Robust standard errors in parenthesis. *: p < 0.10, ** p < 0.05,*** p < 0.01

Table 9: Heterogeneity analysis

	1	2	3	4	5	6	7	8	9
		News		**Grades**		**Gender**		**Median**	
Sample:	All	Positive	Negative	Higher	Lower	Female	Male	Above	Below
Treatment	-0.408**	-0.448*	-0.020	-1.799***	-0.300	-0.465*	-0.324	-0.363	-0.308
	(0.179)	(0.224)	(0.369)	(0.433)	(0.181)	(0.263)	(0.259)	(0.266)	(0.235)
Adj. R-squared	0.595	0.630	0.562	0.547	0.593	0.584	0.585	0.495	0.652
N	977	729	248	96	878	528	449	442	535

Note: The dependent variable is the number of exams passed during the regular exam period of the 2nd year. All regressions include a control for academic performance during the first year. Standard errors are clustered at the tutorial level.

Table 10: Teaching evaluations

	1	2	3	4	5	6
	Pre-treatment semester			**Post-treatment semester**		
	Satisfaction	Hours of study	Course Easiness	Satisfaction	Hours of study	Course Easiness
Treatment	0.013	0.126	-0.000	0.296***	0.153	0.117
	(0.095)	(0.080)	(0.084)	(0.109)	(0.099)	(0.121)
Adj. R-squared	0.250	0.388	0.461	0.210	0.223	0.368
N	182	182	182	165	165	165

Appendix A: Tables

Table A1: Impact on academic performance - IV estimates

	1	2	3	4
	Regular exams		**Retakes**	
	Taken	Passed	Taken	Passed
First year	-0.083	-0.061	0.050	-0.030
	(0.064)	(0.135)	(0.235)	(0.077)
Second year	-0.095	-0.555**	0.506	0.287*
	(0.224)	(0.247)	(0.303)	(0.154)
Third year	0.274	0.099	0.102	0.089
	(0.419)	(0.367)	(0.224)	(0.114)
Fourth year	0.162	0.405	-0.246	0.016
	(0.490)	(0.469)	(0.231)	(0.087)

Note: Each cell reports the result of a different IV regression on the sample of 977 students that took part in the intervention. The independent variable is a dummy variable that takes value one if the student accessed the information on relative performance, instrumented by being assigned to the treatment. The first row provides information for the 1st academic year, the second row for the 2nd academic year, and so on. The first two columns report information from exams taken during the regular period (January and May). Columns (3) and (4) provide information from retakes (June). The dependent variable in columns (1) and (3) is the number of exams taken. The dependent variable in columns (2) and (4) is the number of exams passes. All regressions include a control for academic performance during the first year and degree fixed effects. Standard errors clustered at the tutorial level in parenthesis. *: $p < 0.10$, ** $p < 0.05$,*** $p < 0.01$.