Fakultät für Wirtschaftswissenschaften
Facoltà di Economia
Faculty of Economics and Management

BEMPS –
Bozen Economics & Management
Paper Series

NO 95/ 2022

Wage Expectations and Access to Healthcare Occupations: Evidence from an Information Experiment

Juliana Bernhofer, Alessandro Fedele, Mirco Tonin

Wage Expectations and Access to Healthcare Occupations: Evidence from an Information Experiment*

Juliana Bernhofer † Alessandro Fedele ‡ Mirco Tonin §

February 2024

Abstract

We analyze how financial incentives affect performance on the admission tests for medical and healthcare schools, a crucial step for aspiring healthcare professionals. To this end, we conducted a randomized information experiment with Italian applicants. We first elicited applicants' expectations about the starting wage of the healthcare job for which they intend to study. We then informed the treatment group about the true starting wages, while providing no information to the control group. Finally, we collected the test scores. Applicants expecting a lower wage tend to perform worse, but correcting wage expectations eliminates this difference; indeed, the treatment enhances the test scores when expectations are lower than the true wage level, while negative effects occur when expectations are higher. Moreover, the treatment does not induce adverse selection of low-altruism applicants.

Keywords: Information Experiment; Wage Expectations; Financial Incentives; Test Scores; Health Occupations

JEL Classification: I1, I23, J3, C9.

*We thank Scott Barkowksi, Max Warner, Daniel Wiesen, the seminar participants at the University of Pisa (2022) and the EuHEA Seminar Series (online, 2021), and the participants at Solomon Lew Conference on Behavioral Economics (Tel Aviv, 2022), AFE (Chicago, 2022), ESA (Bologna, 2022), EHEW (Ispra 2022), BUE-EBEL (Cairo, online, 2022), AEHE (Barcelona, 2022), Euregio Economics Meeting (Bolzano, 2022), BEEN (Lucca, 2022), LESE (Lisboa, online, 2022), Joint Workshop of Applied Macro- and Microeconomics (Bolzano, 2021), CEEL (Trento, 2021), and HELM (online, 2021), for helpful discussions and suggestions. We also acknowledge MUA (South Tyrolean University Movement) for their support in data collection. Finally, Alessandro Fedele and Mirco Tonin gratefully acknowledge financial support from the Free University of Bozen-Bolzano (research project WW202C). Declarations of interest: none.

†Department of Economics, University of Bologna, juliana.bernhofer@unibo.it

‡Faculty of Economics and Management, Free University of Bozen-Bolzano, alessandro.fedele@unibz.it

§Faculty of Economics and Management, Free University of Bozen-Bolzano; Institute for the Evaluation of Public Policies - Bruno Kessler Foundation; IZA; CESIfo, mirco.tonin@unibz.it

1 Introduction

In this experimental paper, we explore the impact of monetary incentives on a crucial step to access healthcare occupations: the admission test for medical and healthcare schools.

The related literature, which we review in the next section, has mostly investigated the effect of financial benefits, such as wages and career opportunities, on the *decision* to apply for a career in the health sector. Findings show that better incentives attract more applicants, enhances their average ability, and have ambiguous effects on their average other-regarding behavior, such as altruism and prosocial motivation, that are crucial determinants of the healthcare services' quality.

An important feature of the health sector is that, in most countries, admission to schools involves passing a selection process. This is usually based on high school grades or other academic records and/or the score on admission tests, such as the International Medical Admissions Test, or, in the US, the Medical College Admission Test; in Italy, where we conducted this study, admission is based exclusively on test scores, and entry is rather competitive.[1] As a result, the decision to apply is no guarantee even to start the path to become a healthcare worker.

On the above basis, in this paper we depart from the extant literature and consider individuals that *have already decided* to apply to healthcare schools (physiotherapy, nursing, and obstetrics, among others) and medical schools, but *have not been admitted yet*; our aim is to study how financial incentives impact their performance on the admission test. This hitherto overlooked step is a crucial one because only those passing the test will have an opportunity to develop a career in the health sector.

We consider students who obtained a high school diploma and, during the summer months, are attending preparatory courses for the admission tests. We designed the following randomized information experiment. We first administered a questionnaire eliciting students' beliefs about the starting wage of the medical or healthcare job for which they intended to study, as well as other information. A few weeks later (approximately four weeks before students took their admission tests), we informed a randomly selected half about the true starting wage (treatment group), while the remaining students received no information (control group). Finally, we collected the scores they obtained on the admission tests.

We find that wage expectations are biased in that more than 70% of applicants in our sample underestimate the true starting wage, while less than 30% overestimate it.[2] It follows that our treatment can indeed convey information, in particular as a *positive shock* to wage expectations for those who underestimate and as a *negative shock* for those who overestimate. Our main results, arising both from non-parametric tests and from regression analysis, are as follows: applicants with lower wage expectations about the starting wage perform worse in the entry exam, but this gap disappears as participants are informed about the correct wage, with those receiving a positive shock, i.e., being informed that the actual wage is higher than what they expected, doing better, while those receiving a negative shock doing worse.

[1] For instance, in 2020, 66,638 students applied for 12,362 seats in Italian medical schools, meaning that more than 80% of applications were rejected (for this and other ranking analyses by UNID Formazione and Testbusters, see https://tinyurl.com/2p8aafbw and https://tinyurl.com/2p9e76c9 respectively; accessed February 23, 2024).

[2] Our sample is composed of (potential) soon-to-be first-year students: the literature eliciting students' expectations about earnings shows, quite intuitively, that the most biased expectations tend to be those of first-year students (e.g., Betts, 1996; Jerrim, 2011).

Our treatment revolves around monetary stimuli; as such, it could trigger adverse selection of low-altruism candidates. Using incentivized dictator games vis-à-vis charities as tools to measure altruism, we show that a positive shock to wage expectations does not improve the test scores of selfish individuals more than those of altruistic ones and that a negative shock discourages mostly selfish individuals. We also show that the treatment effect on exam scores is more relevant for applicants to medical schools, while prospective health professionals appear to react more on the extensive margin, i.e., on the decision to actually take the exam after attending the preparatory course.

In terms of our contribution, we offer experimental evidence that financial incentives are effective on a new margin, the admission test performance. This finding is far from straightforward as our students have already decided to pursue a healthcare career - as their attendance of a preparatory course testifies - and one could therefore expect them to be fully motivated when taking the exam. On top of that, we verify that the treatment does not trigger adverse selection. Overall, we provide novel evidence that monetary stimuli may be effective also in a labor sector, namely healthcare, where non-monetary incentives are generally believed to play a prominent role in attracting and motivating people. In the next section, we provide a review of the relevant literature and further detail the contribution of the paper.

Our information experiment belongs to an increasingly popular category of survey experiments in economics which study the causal link between expectations and behavior: for recent reviews, see Haaland et al. (2023); and Fuster and Zafar (2023). More importantly, our experimental subjects operate in a high-stakes field environment, in which the admission test outcome will determine their future career paths and may delay or deny access to healthcare occupations.

The remainder of the paper is organized as follows. Section 2 reviews the literature. Section 3 outlines the institutional context and the experimental design we have utilized. In Section 4, we present and discuss our results. Section 5 concludes.

2 Literature Review

In this section, we review four streams of literature our paper contributes to.

First, our paper is related to a growing literature that examines how financial incentives affect the selection into the health sector and/or the performance of health professionals. A positive impact of expected earnings on the number of applicants to nursing schools is found by Schweri and Hartog (2017), who consider healthcare trainees in Switzerland, and Kugler (2022), who uses the data of German 14- to 15-year-olds. Deserranno (2019) runs a field experiment with applicants for a health-promoter position in Uganda. Communicating different values of the existing promoters' actual income distribution randomly, the position is advertised as high, medium, or low paying. Using this experimental variation in expected earnings, the author finds that higher financial incentives attract more applicants, but crowd out prosocially motivated applicants, who stay longer on the job and demonstrate better job performance than less motivated agents. This adverse selection outcome is in line with the theoretical predictions provided by Heyes (2005) in the nursing sector. A different result, albeit not limited to the health sector, is obtained by Dal Bó et al. (2013). The authors rely on a recruitment drive for community development positions in the Mexican public sector, where two different salaries are announced randomly across recruitment sites; they observe that a higher wage attracts agents who are more able and more motivated to provide public services. Such an advantageous selection outcome is in line with Fedele (2018), who derives theoretical conditions under which a higher pay in the nursing sector attracts more able and more motivated individuals. Ashraf et al. (2020) run a field experiment in the context of a recruitment drive for healthcare positions in Zambia. The recruitment ads either highlight career opportunities, or mention helping the community as the main benefit, so that the former is perceived as yielding a higher present value of future earnings. The authors show that the treatment attracts high-skilled individuals at the expense of prosocial individuals; this adverse selection, however, vanishes in the pool of successful applicants, who are chosen among the most skilled ones; as a result, agents in the career opportunities group provide better work outcomes.[3] Unlike the above contributions, our information experiment involves individuals that have already decided to apply to medical and healthcare schools - indeed, they are attending a preparatory course for the entry exam - and shows that money also matters for performance on the admission tests, a crucial and hitherto overlooked step for aspiring healthcare professionals.

Second, our paper contributes to recent literature on information experiments using students' expectations about earnings. Wiswall and Zafar (2015a,b) collect New York University undergraduate students' beliefs about the population distribution of earnings and elicit the students' probability of choosing various majors. The authors then inform a randomly selected subset of students about the true population earnings. Students' beliefs are shown to be biased in that the average earnings of workers with no college degree are underestimated, while those of graduates in economics and business are overestimated. This outcome enables the authors to observe that students in the treatment group correct their own earning expectations (i.e., underestimators revise them upward, while overestimators revise downward) and increase the probability of selecting higher-earning majors. A similar information experiment is conducted by Conlon (2021) on a sample of Ohio State University first-year students. The author finds that students underestimate salaries in almost every field and that correcting their expectations affects their actual choice of majors. Our contribution is different in that we consider students who have already chosen but have not yet been admitted to their course of study; this is because we are interested in the information treatment effect on the admission test outcomes rather than on the students' choice of major.

Third, there is related experimental literature investigating how financial incentives impact student performance in high-stakes exams. Angrist and Lavy (2009) study the impact of sizable monetary rewards (up to $2,400) for passing the Bagrut tests, a formal pre-requisite for university admission in Israel, and find effects on pass rates for female students. Kremer et al. (2009) consider a scholarship program in which Kenyan girls doing well on the secondary school admissions exam have school fees paid and receive a grant; the authors observe a positive effect on exam scores in one of the two school districts where the program is implemented. Burgess et al. (2021) depart from the existing works by estimating the impact of incentives - money or prizes like trips to theme parks - on various dimensions of effort (e.g., attendance, homework) of UK students over a school year: they find little average impact of incentives on GCSE test scores, which serve as

[3] A similar result concerning the performance of health professionals is obtained by Propper and Van Reenen (2010). Relying on the fact that nursing wages are standardised across the UK, the authors find that regions where the current relative pay of nursing staff is higher - because outside wages are lower - show increased hospital quality when measured by hospital deaths for emergency heart attacks.

the main gatekeeper to progress to university, in the overall population, while students with lower performance at baseline are responsive. The above papers consider the general population of students, while our focus is on aspiring health professionals, a category that is of particular interest and has peculiarities, e.g., regarding the importance of altruistic motives.

Finally, our experimental analysis of the link between financial considerations and the test performance of prospective medical and health students complements health economics literature that focuses on non-financial factors and current medical students. For instance, Hennig-Schmidt and Wiesen (2014) find that medical students in Germany are more patient-regarding and more willing to sacrifice their own profit compared to non-medical students. Li et al. (2017) instead show that medical students are less altruistic and more efficiency-focused than the average individual in the US.[4] Attema et al. (2023) observe that the degree of German medical students' altruistic preferences change during the course of their medical training: altruism is highest for freshmen, declines over the course of medical studies, and increases again for last year students; the authors also find that altruistic students have lower income expectations.

To the best of our knowledge, this is the first paper looking at the impact of financial incentives on test outcomes in the entry exam for medical and healthcare schools.

3 Data

Institutional Background. In Italy, applicants to medical and healthcare schools are selected on the sole basis of their score on two different tests, one for medicine and one for the other healthcare professions. This is the only selection that takes place until, in the case of doctors, they choose their specialty six years later; instead, healthcare professionals already choose their specialty at this stage.[5] The test for medicine is standard across all Italian universities; the test for health professions is instead administered at the individual university or healthcare school level. However, the two tests share the following characteristics: (i) they are based on 60 multiple-choice questions, each with 5 choices and only one correct answer; (ii) they contain 5 topics, namely general knowledge, logical reasoning, biology, chemistry, and physics plus mathematics; (iii) the duration is 100 minutes; (iv) 1.5 is the mark for correct answers, -0.4 for wrong ones, and 0 for no answers; therefore, the maximum score is 90 and the minimum score is -24. Moreover, both tests are held once a year in the first two weeks of September, generally on two distinct dates. Finally, students applying for medical schools may select one or more preferred universities before taking the test; there is a single ranking at the national level for medicine and a higher score increases both the likelihood of acceptance to the preferred option(s) and of acceptance more generally. By contrast, students applying for other health professions take the test at the specific university which they have previously

[4] Using a similar experimental framework, Li et al. (2022) compare actual US physicians to a US representative sample, a subsample who hold a graduate degree and have an annual household income over $100,000, and a nationwide sample of medical students. They compare these populations in terms of altruism and preferences regarding efficiency. The authors find that physicians are remarkably more altruistic than all the other samples and equally efficiency-focused.

[5] Healthcare specialties include Physical Therapy, Nursing, Obstetrics, Speech Therapy, Nutritional Therapy, Occupational Therapy, Prevention Techniques, Biomedical Technology, Dental Hygiene, and Medical Radiation Technology.

selected and compete only with other applicants to the same university and profession.[6]

Students and Questionnaire. In the summer of 2018, we launched the first wave of a paper-based questionnaire to applicants to Italian medical and healthcare schools; two additional waves followed in 2019 and 2020 for a total of 408 participants.[7] Our respondents were high school graduates attending three admission test preparatory courses organized by *Movimento Universitario Altoatesino* (MUA - South Tyrolean University Movement), a student organization located in Bolzano-Bozen, South Tyrol, Italy. The organization offers a course for prospective physicians and two courses for prospective health professionals. Though the majority of applicants in the second group take the very same test at the local healthcare school *Claudiana*, MUA organizes one course in Italian and one in German, as both languages are officially spoken in South Tyrol. Accordingly, each year we administered the questionnaire to three distinct groups of applicants: prospective physicians and health professionals attending courses in Italian, to whom we administered a questionnaire in Italian, and prospective health professionals attending the course in German, for whom the questionnaire was in German.

The questionnaire contains questions on wage expectations and family background, as well as measures of cognitive skills and altruism. We elicited wage expectations through the following question (English translation):

> In your opinion, what is the monthly net starting wage of South Tyrolean health authority employees that practice the profession you are preparing for? Provide a single value: ______ Euro.[8]

Treatment. Our treatment consisted in sending an e-mail with information on the true starting wage provided to us by the South Tyrolean health authority: the monthly net starting salary of physicians was 3650 Euro in 2018, which increased to 3850 Euro in 2019, no matter the medical specialty; by contrast, the initial wage of all the other healthcare workers, 1600 Euro, remained fixed over the period under analysis (source: Ufficio Pensioni, Azienda Sanitaria dell'Alto Adige, personal communication). As an example, we report the English translation of the e-mail sent to prospective health professionals:[9]

> Dear student (female), dear student (male),
> Thank you for completing our questionnaire during the MUA course! Concerning one of the questions, we wish to inform you that the monthly net starting wage of South Tyrolean health authority employees that practice the profession you are preparing for is **1600 euro**.
> Kind regards,

Importantly, we elicited expectations and provided information about the exact same wage. This wage is the one most relevant for students who plan to work in the region, as the South Tyrolean health authority is the dominant regional employer in the healthcare sector and, for some specialties, a local monopsonist. Even if prospective students are

[6] Sources: Italian ministerial decrees no. 385, 2018 14-05; 542, 2019 18-06; 218, 2020 16-06.

[7] Due to the pandemic, in 2020 we switched from the paper-based version to a digital version that was designed using the software Qualtrics.

[8] Note that the monthly net wage is the standard popular method for discussing wages in Italy: see, e.g., the annual report on Italian graduate employment status provided by Almalaurea, a consortium of Italian Universities (https://www.almalaurea.it/en/node/27992, accessed February 24, 2024).

[9] The email sent to prospective physicians is identical, with the exception of the salary level.

unsure about their future location, our treatment is still relevant as it conveys information about the level of wages that can be expected in the sector.

The recipients of our treatment were identified by matched-pair randomization, which took place in four steps. First, each year the subjects were divided into 3 different sets: those attending (i) the course for physicians, (ii) the course for health professionals in Italian, and (iii) the course for health professionals in German. Second, within each set, subjects were ranked in descending order from the individual expecting the largest wage to the individual expecting the lowest wage. Third, within each set, subjects were grouped into pairs, with the first pair consisting of those expecting the two highest wages, the second pair of those expecting the third and fourth highest wages, etc. Finally, an automated randomization procedure was carried out within each pair to assign one subject to treatment (i.e., receiving the email) and the other to control (i.e., not receiving the email).

To minimize the probability that subjects in the treatment group shared information with those in the control group, the e-mails were sent in August 2018-2020, after the MUA courses had already ended during the last week of July 2018-2020. At the same time, to give subjects enough time to (possibly) react to the treatment, the e-mails were sent around four weeks before the admission tests took place in the first two weeks of September 2018-2020. In case some information exchange might have occurred between subjects in the treatment and control groups during the time between the treatment and the dates of admission tests, our results would be interpreted as a lower bound of the true effect of receiving information on starting wages.

Test Scores. After the admission tests took place, the questionnaire data of each subject were matched with her/his test score. The test outcomes of prospective health professionals were collected with the support of the local healthcare school Claudiana. Conversely, scores of medical school applicants were not accessible through institutional sources; therefore, we contacted the applicants directly through e-mail, WhatsApp, and telephone.[10]

For better comparability across the six different admission tests (physicians and health professionals in 2018, 2019, and 2020), we standardized the test scores on a 0 to 100 scale using the min-max procedure; that is, the standardized score of subject i is

$$Score_i\,(std.) = \frac{score_{i,k} - minimum\ score_k}{maximum\ score_k - minimum\ score_k} \times 100, \tag{1}$$

where $k = 1, ..., 6$ denotes the combination of test and year to which subject i belongs.[11]

Descriptive Statistics. From the full data set, we excluded: (i) 2 subjects who had already participated in previous years; 17 subjects who answered a control question that should not have been answered, therefore showing a propensity to provide random an-

[10] Because prospective physicians' exam results could not be retrieved from administrative sources, one can be afraid that the scores were over-reported. Though we are not able to directly check whether this is the case, we can compare the self-reported scores of our sample to the official scores achieved by the actual Italian population of applicants to medical schools: no statistically significant difference is found (more details can be read at the end of this section). In addition, even if over-reporting was present, there is little reason to think that it would be correlated with treatment and, therefore, that our estimated treatment effect would be affected.

[11] While in the randomization process we have separately considered Italian-speaking and German-speaking applicants for health professions because they attended different preparatory courses, here we joined them because they took the same bilingual admission test.

swers;[12] (iii) 27 subjects who indicated the correct starting wage; our treatment does not provide any additional information to these respondents, who are therefore expected not to react to it and are too few to be analyzed separately.

The final sample contains 362 subjects, while the test scores are available for 296 subjects. In the results section we explore in details the determinants of score availability. In Table 1, we provide the descriptive statistics of our variables of interest for both the full sample and the sample with available score. We also distinguish between applicants whose expected wage is under -or over- the correct one, and in the remainder of the paper, we refer to them as *underestimators* and *overestimators*, respectively.

Table 1: Summary Statistics

	Full Sample						Available-Score Sample					
	All		Underest.		Overest.		All		Underest.		Overest.	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Treatment	0.51	0.50	0.51	0.50	0.54	0.50	0.53	0.50	0.53	0.50	0.54	0.50
Score (std.)							57.4	25.4	55.36	25.64	62.88	23.96
Female	0.78	0.42	0.79	0.41	0.74	0.44	0.78	0.41	0.80	0.40	0.74	0.44
Age	19.83	2.68	19.57	2.19	20.59	3.69	19.84	2.78	19.56	2.30	20.59	3.70
Cognitive Skills	7.64	2.81	7.86	2.81	6.99	2.71	7.84	2.78	8.12	2.76	7.11	2.70
Charitable Giving	69.80	26.92	71.34	26.32	65.06	28.32	70.16	26.59	71.73	26.04	65.78	27.77
Family Network	0.49	0.50	0.51	0.50	0.43	0.50	0.50	0.50	0.52	0.50	0.43	0.50
Physician	0.40	0.49	0.50	0.50	0.11	0.31	0.39	0.49	0.49	0.50	0.11	0.32
Observations	362		271		91		296		216		80	

Notes: *Score (std.) stands for standardized scores.*

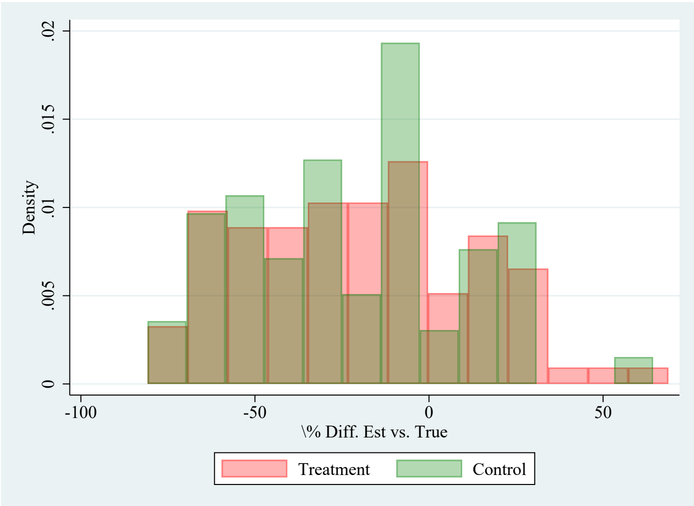
We first observe that the wage expectations of 75% of the subjects (271/362) are lower than the true starting wage in the full sample of 362 individuals; the percentage is similar, 73% (216/296) in the available-score sample of 296 individuals. *Treatment* identifies the proportion of subjects who received the email with wage information: 51% of the full sample and 53% of the available-score sample; the slight imbalance is due to uneven subject numbers in some of the six groups and post-intervention drops. The mean of the standardized score is approximately 57 points and lower for underestimators vis-à-vis overestimators. 78% of the applicants are females, who seem to be more present among underestimators. *Age* shows that our applicants are on average less than 20 years old (the median age is 19). *Cognitive skills* indicates the number of correct answers in the 12-item Raven's Standard Progressive Matrices Test (Set E);[13] the mean value is around 8 and underestimators tend to perform better. *Charitable Giving* stems from two incentivized dictator games, in which respondents decide to donate an amount between 0 and 100 EUR to the World Wide Fund for Nature (WWF) and Médecins Sans Frontières (MSF) and two participants per group (6 in total each year) are randomly selected to actually receive the money; this variable measures the applicants' average donation to WWF and MSF. The mean value is around 70 EUR and underestimators tend to donate more. *Family Network* is a dummy that equals 1 if at least one physician (health professional) is present in the family network (i.e., parents, siblings, grandparents, aunts, and uncles)

[12] The control question was part of a 10-item set of 5-point Likert scale questions and was formulated as follows: "This is a control question: please do not answer."

[13] Raven's standard progressive matrices test is a visual task of abstract reasoning used to measure cognitive skills. The test requires examinees to infer a rule to generate the next items in a series, or to determine whether a presented design is consistent with the rule (Leavitt, 2011). Raven test's Set E is the most difficult and it was selected after validation with first-year undergraduate students in the bachelor's program in Economics and Management, Free University of Bozen-Bolzano; the other (simpler) sets provided insufficient variation across students in the number of correct answers.

of the prospective physician (health professional); this occurs to half of our samples. Finally, *Physicians* show that 40% of the subjects attended the course to enter medicine and 60% took the course to enter another health profession; prospective physicians are more likely than prospective health professionals to underestimate the starting wage; this comes as no surprise because the physicians' wage is significantly larger.

Figure 1: Percentage Deviation of Expected Wages from True Wages



Notes: *The figure compares the distributions of the % deviations from true wages across treatment and control groups.*

Balance between Treatment and Control Groups. To show that subjects randomly assigned to treatment and control groups have similar wage expectations, in Figure 1 we consider the full sample of 362 individuals and display the distribution of percentage deviations of wage expectations from true wages across the two groups. The distributions are indeed very similar (two-sample Kolmogorov-Smirnov test, $D = 0.0463$, p-value$= 0.998$) confirming that treatment and control groups are balanced along this crucial variable. In Table A1 of the Appendix, we run a logistic regression with the likelihood of being assigned to either the treatment or control group as the dependent variable and our main observables as regressors. The absence of any statistically significant effects and the Wald test failing to reject the null hypothesis of joint non-significance of the model predictors support the claim that subjects in the treatment and control groups are balanced. This claim holds true also when considering the available-score sample of 296 individuals.

General Population of Applicants. A possible issue concerning our sample is that it includes subjects who are attending a preparatory course. Preparatory courses are not expensive (the course for medicine costs 160 euro and that for healthcare professions 90 euro) and attendance is quite common due to the selectivity of the admission tests, but clearly not universal: students attending preparatory courses may be positively selected if, for instance, they are particularly motivated to succeed, or negatively selected if, for example, they use the course to acquire the necessary knowledge they lack. We have two ways to assess how our sample compares to the general population of applicants.

First, for some statistics we can compare our available-score sample to the actual population sitting the tests. For prospective physicians, we can exploit individual-level data on scores, which are available at the national level for the year 2018. We find no statistically significant difference between raw scores in our sample and nationwide raw scores (33.35 vs. 35.69; $t(40420) = -1.52$, $p = .129$). We repeated a similar exercise with the 2018 data on prospective health professionals, for whom we had data on the test scores, high school grades and gender for the entire population taking the exam at the local healthcare school. Subjects in our sample achieve higher raw scores (50.78 vs. 43,16; $t(268) = -3.52$, $p = .0005$), but there are no significant differences in high school grades (77.24 vs. 78.35; $t(261) = -0.78$, $p = .44$) or division by gender (85% vs. 80,77%; $t(266) = -0.745$, $p = .457$).[14]

Second, in 2020 our questionnaire was administered online to a representative population of 18-19 years old living in Northern Italy (including South Tyrol). Participants were selected by the Italian survey company SWG, for a total of 349 individuals. Among them, 35 want to pursue a career in the health sector. Compared to our full sample of 362 applicants, we find no difference in terms of professional peer family network and gender, while our applicants do better in the Raven test.[15] Results are robust when considering the available-score sample.

Overall, the above tests suggest no major differences between our sample and the general population of applicants. The fact that our sample may contain individuals with relatively high cognitive skills makes our analysis even more relevant: with selective entrance exams, these individuals have a higher chance to enter the profession.

4 Results

This section provides the results of our analysis, which is divided into two parts. In Subsection 4.1, we consider only the 296 individuals for whom test scores are available; our aim is to measure how the information treatment affects their exam performance (i.e., the intensive margin). In Subsection 4.2, we focus on the full sample of 362 subjects, including those without test scores; doing so, we aim at investigating whether students' participation to the test (i.e., the extensive margin) is affected by the treatment.

4.1 Intensive Margin

Underestimators in the treatment group experience a positive shock to their wage expectations because they learn that actual wages are higher than they had assumed; accordingly, in case of a reaction, they can be expected to feel more motivation and to perform better on the exam. Conversely, overestimators in the treatment group experience a negative shock and can be expected to perform worse.

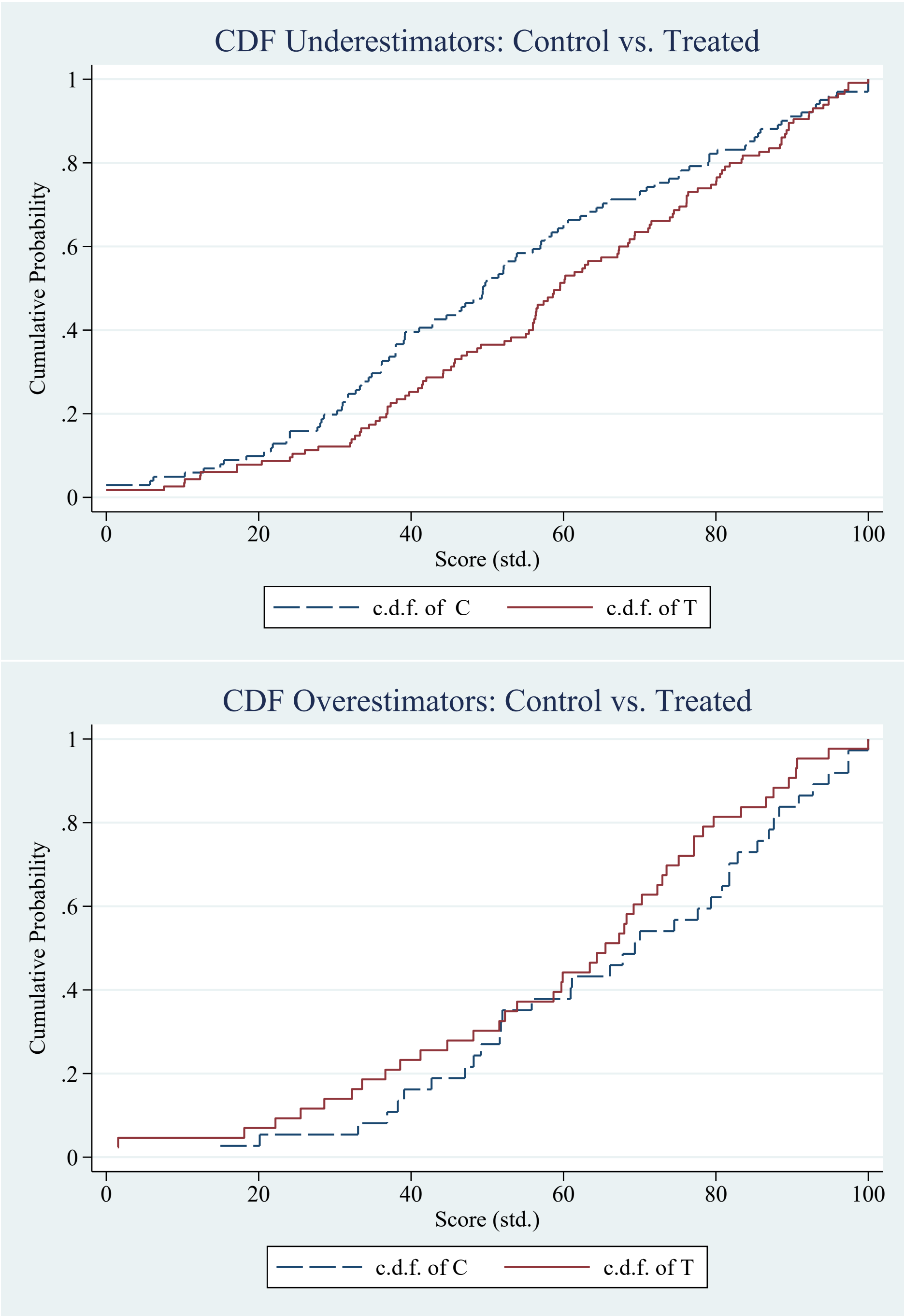
In Figure 2, we plot the cumulative distributions of the standardized test scores for underestimators (upper panel) and for overestimators (lower panel); in each panel, we compare the treatment group (solid line) to the control group (dashed line). If we look at the underestimators, the distribution for the treatment group is clearly shifted towards

[14] The difference in degrees of freedom compared to scores is due to missing data concerning high school grades and gender-neutral names in the general population.

[15] 45.71% of the general-population respondents provided more than 6 correct answers, compared to 67.40% in our sample of applicants, Fisher's exact test $p = .0015$. However, such performance gap might be overestimated because respondents in our sample completed it during their preparatory lectures in a relatively controlled environment, while questionnaires given to the general population were administered via CAWI web surveys, where more opportunities for distraction may be present.

the right, especially in the central part; as for the overestimators, the opposite is true as the distribution for the treatment group is shifted towards the left, even if the difference appears less pronounced and the two distributions actually overlap in the central part. This is already an indication that being informed about the correct wage encourages underestimators to do better, while somehow discouraging overestimators.

Figure 2: Cumulative Distributions of Scores



The difference between treatment and control for underestimators is confirmed by the non-parametric tests reported in Table 2. The p-value of the Kolmogorov-Smirnov (KS)

Table 2: Non-parametric Test Results

	Underestimators n = 216		Overestimators n = 80	
	C	T	C	T
Mean Score	54.5	58.8	66.4	59.9
Kolmogorov-Smirnov p-value	0.025		0.405	
Mann-Whitney p-value	0.027		0.240	

test is 0.025, while using the Mann-Whitney (MW) test the p-value is 0.027. Thus, we can reject the null that the two distributions are equal. In line with the more limited difference already evident in Figure 2, the tests fail to reject equality of distributions for overestimators. Non-parametric tests thus show that our information treatment affects the underestimators' test scores, pushing them to do better; instead, the impact on the overestimators' scores goes in the opposite direction, but it is not significant.

It is also interesting to compare the distributions of test scores for underestimators and overestimators in the control groups, as well as for the ones in the treatment groups: see Figure A1 in the appendix for a graphical comparison. From the mean scores reported in Table 2, it appears that the two groups differ in absence of treatment, with overestimators doing better (66.4 vs. 54.5). Indeed, both the KS and the MW tests reject the equality of the distributions of test scores for under- and over-estimators in the control group (p-value 0.025 and 0.003, respectively). After aligning their expectations by informing them about the correct wage, however, this difference disappears. Indeed, mean scores are very similar (58.8 vs. 59.9), and both the KS test (p-value 0.73) and the MW test (p-value 0.76) fail to reject the null hypothesis that the distributions of test scores is the same between under- and over-estimators in the treatment group. Aligning wage expectations thus also aligns performance.

Next, we explore the same issues using regression analysis, estimating the following OLS model,

$$Score_i\,(std.) = \beta_0 + \beta_1 U_i + \beta_2 T_i \times (1 - U_i) + \beta_3 T_i \times U_i + \beta_4 \sum_{j=1}^{5} YP + \epsilon_i \qquad (2)$$

where: *Score$_i$(std.)* denotes subject i's standardized score as given by (1); U_i takes value 1 if subject i is an underestimator and 0 otherwise; T_i equals 1 if subject i is in the treatment group and 0 otherwise. Accordingly, the coefficient β_1 captures any difference between underestimators and overestimators unrelated to the treatment; β_2 captures the treatment effect on overestimators; β_3 shows the treatment effect on underestimators. The second to last term on the right-hand side of equation (2) describes fixed effects by professional category (prospective physicians or health professionals) and year (2018-2020), capturing, for instance, differences in the difficulty of the exam or any other factor that is common for the year/professional category combination.

Table 3 presents our OLS regression results. Model (1) implements equation (2), while in model (2) we control for some additional variables. In particular, *High Cognitive* is a dummy variable equal to 1 when the correct answers provided by an applicant to medical (healthcare) schools are weakly above the median value computed over the applicants to medical (healthcare) schools; it equals 0 otherwise.[16] Similarly, *High Altruism* is a

[16] The median values, computed on the full sample of 362 individuals, are 9 correct answers out of 12

dummy variable equal to 1 when an applicant to medical (healthcare) schools donates weakly more than the median values of the average donation to WWF and MSF.[17] In models (3) and (4), we further divide the sample of underestimators into stronger and weaker underestimators: the former (latter) are defined as underestimators whose wage expectations are weakly below (strictly above) the median value for underestimators by professional category. In this way, we can assess whether the intensity of treatment matters. Indeed, strong underestimators receive an even more positive news than weak underestimators in terms of the difference between their expected wage and the actual one. We perform this additional exercise only for underestimators as they are much more numerous than overestimators.

Table 3: Treatment Effects on Scores

	(1) Score (std.)	(2) Score (std.)	(3) Score (std.)	(4) Score (std.)
β_1: Wage underestimation	-10.6**	-11.5**		
	(4.6)	(4.5)		
β_{1_w}: Weak Wage Underestimation			-6.6	-8.3*
			(5.1)	(5.0)
β_{1_s}: Strong Wage Underestimation			-15.6***	-15.9***
			(5.2)	(5.3)
β_2: Treatm. x Wage overest.	-7.0	-8.0*	-7.0	-8.1*
	(4.5)	(4.7)	(4.5)	(4.7)
β_3: Treatm. x Wage underest.	6.7**	6.0*		
	(3.3)	(3.2)		
β_{3_w}: Treatm. x Weak Wage Underestimation			1.7	2.4
			(4.5)	(4.4)
β_{3_s}: Treatm. x Strong Wage Underestimation			12.7***	10.5**
			(4.5)	(4.6)
Female		1.6		2.2
		(3.2)		(3.2)
Age		0.8*		0.8*
		(0.5)		(0.4)
High Cognitive		10.8***		10.5***
		(2.9)		(2.9)
High Altruism		1.0		1.1
		(2.9)		(2.9)
Family Network		1.8		1.3
		(2.7)		(2.7)
Constant	64.9***	39.2***	65.1***	40.8***
	(4.3)	(11.6)	(4.3)	(11.4)
Observations	296	296	296	296
R-squared	0.1	0.2	0.2	0.2
Additional Controls:				
*Year*Professional Category*	Yes	Yes	Yes	Yes

Notes: *Estimation method: OLS. Dependent Variable: Standardized Test Score, as defined in (1). Robust standard errors are clustered at the unit-of-randomization level. *** p<0.01, ** p<0.05, * p<0.1*

Models (1) and (2) provide three findings. First, the point estimate of parameter β_1,

for prospective physicians and 7 for health professionals. Our results are robust when we consider the medians defined over the available-score sample, in which case the value for health professionals rises to 8, while it remains unchanged for prospective physicians.

[17] 75 EUR is the median value for both professions and in both samples.

corresponding to the dummy for being an underestimator, is negative and significant: underestimators achieve lower average scores than overestimators. Second, the estimated parameter β_2 - the treatment effect on overestimators - is negative, and, when controlling for additional covariates, marginally significant. Third, parameter β_3 - the treatment effect on underestimators - is positive and significant: the underestimators' test scores improve by 6.7 standardized points (or, as an alternative metric, 6.7/25.4=0.26 standard deviations) due to being informed about the correct wage, an economically significant magnitude. It is worth noticing how the treatment effect on over- and under-estimators is symmetrical. Finally, looking at the covariates in model (2), we see that, as expected, test scores positively correlate with applicants' cognitive skills, while there is no significant coefficient associated with gender, charitable giving or family network.

We shift our attention to models (3) and (4), where we distinguish between weak and strong underestimators. We see from coefficient β_{1_s} and β_{3_s} that strong underestimators do particularly bad in absence of treatment, but they are highly reactive when informed about the correct wage. Coefficients β_{1_w} and β_{3_w} show instead that weaker underestimators do slightly worse than overestimators and that their response to treatment, albeit positive, is much smaller compared to strong underestimators and non-significant.

Overall, the regression analysis confirms the non-parametric tests reported at the beginning of this section: the treatment has a beneficial effect on the performance of subjects whose wage expectations are upwardly corrected, while subjects whose wage expectations are corrected downward tend to perform worse. Another interesting finding is that, while in the absence of treatment the underestimators perform worse than the overestimators, this gap is eliminated by our information treatment. Indeed, looking for instance at model (2) and comparing the differential treatment effect between under- and overestimators to the differential level irrespective of treatment, i.e., (β_3 - β_2) + β_1, gives a non-significant coefficient of 2.5 ($p = 0.6$).

Take together, the results from the non-parametric tests and from the regression analysis are consistent with a role played by deferred financial incentives - in the form of future wages - in the effort to pass the entry exam to medical and healthcare schools. Individuals who expect low wages seem to be less motivated compared to those who expect high wages; once expectations are corrected, the gap in terms of test scores disappears.

Given the positive relationship between cognitive skills and test scores shown in Table 3, one might alternatively think that underestimators perform worse because of lower skills. However, this is unlikely to be the case because the performance indeed converges after wrong expectations are corrected. To check that cognitive skills do not play a role, in Table A2 of the Appendix, we look into the determinants of being an underestimator. We use logit regressions, in which the dependent variable takes value 1 if the subject underestimated wages and 0 otherwise, and find that indeed cognitive skills have no effect on the likelihood that applicants are underestimators. Table A2 also shows that female students are more prone to be underestimators than male students, in line with recent evidence,[18] as is an applicant having someone in the family active in the same profession.[19]

[18] Briel et al. (2020), for instance, conduct a survey at Saarland University, Germany, and observe that women expect a lower average starting salary than men. The same finding is obtained by Favara et al. (2021), who survey a sample of 14- to 15-year-old students in Peru.

[19] This could be explained, for instance, by people having a tendency to complain about their wage within the family environment, which is taken by youngsters as indicative of the wage being rather

In the remainder of this section, we are interested in further understanding the intensive margin effect of our information treatment. To this aim, we investigate heterogeneity along different dimensions. We first split the sample along the cognitive skills and charitable giving dimensions to investigate adverse selection. We then analyze separately applicants to medical and healthcare schools. Conducting these analysis is rather demanding with a sample that is not particularly large. As such, the results should be taken only as indicative, although we believe them to be informative.

Cognitive skills and altruism. Since our treatment revolves around financial incentives, one might wonder whether it gives rise to adverse selection by favoring less altruistic (or even less skilled) applicants to do well in the entry exam. To verify whether this is the case, we explore whether a shock to wage expectations has a differential effect depending on the level of skills and altruism.

Starting with cognitive skills, as measured through the Raven test, we rely on the median-split dummy *High Cognitive* defined above: a total of 184 (112) subjects turned out to provide a number of correct answers that is weakly above (below) the median values by professional category.[20] In Table 4, models (1) and (2), we perform the analysis separately for these two groups; models (3) and (4) include the additional covariates. We see that both the encouragement arising from correcting pessimistic expectations about the wage and the discouragement arising from correcting overly optimistic expectations are particularly strong for the subset of high-skilled subjects, even if the large standard errors imply that the differences in the treatment effects between high and low cognitive samples are not significant.

Moving to altruism, measured by the two aforementioned incentivized dictator games, we use the median-split dummy *High Altruism*. Therefore, in models (5) and (6), we conduct the analysis separately for those who gave at or above the median (158 subjects) and below the median (138 subjects), adding controls in models (7) and (8).

Here, we can notice three interesting patterns. (i) The baseline difference in test scores between underestimators and overestimators, as captured by the coefficient β_1, is particularly relevant for subjects with low altruism. This is compatible with financial incentives being particularly relevant for low-altruism individuals. (ii) Correcting overly optimistic wage expectations through our treatment reduces the performance only among subjects with low levels of altruism: this can be seen from the strong negative coefficients for β_2 in models (6) and (8) and the almost-zero coefficients in models (5) and (7). Accordingly, selfish students are particularly discouraged when they find out the actual wage is lower than expected. (iii) This is instead not the case when correcting low wage expectations. The coefficient β_3 is indeed stronger for subjects with high altruism - models (5) and (7) - than for subject with low altruism - models (6) and (7). Again, these differences are not statistically significant, but, overall, we can safely exclude that our treatment would give rise to adverse selection within the pool of applicants along the dimension of other-regarding behavior, given the combination of discouragement for selfish overestimators and some possible encouragement of altruistic underestimators. In terms of cognitive abilities, the treatment appears not to trigger adverse selection either.

low in the profession of the complainer.

[20] As mentioned, results are robust when we consider the median values defined over the available-score sample, rather than those computed on the full sample of individuals.

Table 4: Heterogeneous Reactions to Treatment: Cognitive Skills and Altruism

	(1) High Cog.	(2) Low Cog.	(3) High Cog.	(4) Low Cog.	(5) High Altr.	(6) Low Altr.	(7) High Altr.	(8) Low Altr.
β_1: Wage underestimation	-12.1*	-9.6	-11.1*	-7.5	-5.2	-15.9**	-7.7	-13.9**
	(6.1)	(6.9)	(6.2)	(6.7)	(6.4)	(6.4)	(6.7)	(6.4)
β_2: Treatm. x Wage overest.	-11.4*	1.5	-9.3	-0.06	0.1	-14.4**	-1.2	-15.2**
	(6.5)	(6.5)	(6.7)	(6.6)	(6.9)	(6.5)	(7.2)	(6.7)
β_3: Treatm. x Wage underest.	8.4**	4.8	8.3**	2.7	8.3*	4.2	8.5**	2.4
	(3.6)	(6.1)	(3.6)	(5.8)	(4.2)	(5.4)	(4.2)	(5.0)
High Altruism			-1.5	7.4				
			(3.9)	(5.0)				
Female			-1.1	9.5*			5.3	3.3
			(4.1)	(5.1)			(4.7)	(4.3)
Age			0.2	1.2			0.7	0.9
			(0.7)	(0.8)			(0.7)	(0.6)
Family Network			7.5**	-6.2			0.9	0.8
			(3.6)	(4.2)			(3.7)	(4.3)
High Cognitive							4.9	16.5***
							(4.3)	(4.9)
Constant	68.6***	56.6***	61.3***	21.5	55.5***	75.2***	35.7*	39.8**
	(5.7)	(5.9)	(16.0)	(17.4)	(5.9)	(5.7)	(18.4)	(16.3)
Observations	184	112	184	112	158	138	158	138
R-squared	0.09	0.3	0.1	0.4	0.2	0.1	0.2	0.2
Additional Controls:								
*Year*Professional Category*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *Estimation method: OLS. Dependent Variable: Standardized Test Score. Robust standard errors are clustered at the unit-of-randomization level.*
**** p<0.01, ** p<0.05, * p<0.1*

Profession. In Table 5, we divide subjects into applicants to medical and healthcare schools, considering that these two career tracks, albeit both related to the health sector, clearly differ. Models (1) and (3) show that prospective physicians react more strongly to the information treatment compared to prospective health professionals, even if the differences in the treatment effects are not statistically significant. It is of note that the health professionals' starting wage is less than half than the physicians' one; accordingly, there is less room to underestimate it. Indeed, among underestimators the average percentage difference between expected and true wage is 48% when considering prospective physicians and only 19% when focusing on applicants for health professions and, as we have seen in Table 3, strong underestimators react more to the treatment than weak ones. Again, it is worth noting that for both prospective physicians and prospective health professionals (β_3 - β_2) + β_1 is close to zero and statistically insignificant, meaning that the difference in performance in the entry exam between under- and over- estimators disappears when expectations are aligned through our information treatment.

Table 5: Heterogeneous Reactions to Treatment: Profession

	(1) Doc	(2) HP	(3) Doc	(4) HP
β_1: Wage underestimation	-27.4**	-6.7	-31.4***	-5.9
	(10.4)	(5.0)	(11.2)	(5.0)
β_2: Treatm. x Wage overest.	-18.2	-5.7	-16.8	-6.3
	(15.0)	(4.8)	(13.7)	(5.1)
β_3: Treatm. x Wage underest.	9.1*	4.3	12.0***	2.9
	(5.0)	(4.4)	(4.5)	(4.4)
High Altruism			5.6	-1.7
			(4.0)	(3.6)
High Cognitive			18.7***	7.1*
			(4.4)	(3.9)
Female			6.0	-0.3
			(3.9)	(4.5)
Age			-2.5	0.9**
			(1.5)	(0.4)
Family Network			10.7**	-0.8
			(4.2)	(3.7)
Constant	70.6***	62.9***	99.4***	41.0***
	(9.5)	(4.4)	(31.1)	(10.4)
Observations	115	181	115	181
R-squared	0.1	0.09	0.3	0.1
Additional Controls:				
Year	Yes	Yes	Yes	Yes

Notes: *Estimation method: OLS. Dependent Variable: Standardized Test Score. Robust standard errors are clustered at the unit-of-randomization level. *** p<0.01, ** p<0.05, * p<0.1*

All in all, the results from the non-parametric and the regressions analyses show that financial incentives matter for the test scores in the entry exam and that adverse selection, an issue that is particularly relevant for the health sector, does not take place.

4.2 Extensive Margin

So far, we have investigated the treatment effect on the intensive margin, that is, on the test score when taking the exam. An additional margin of interest is the participation to the exam. During the summer months, people in our sample are taking a course preparing them for the admission exam into medical or healthcare schools, but may end up not taking it for a variety of reasons.

In Table 6, we study whether wage expectations and the treatment correcting them matters for this extensive margin: we consider the full sample of applicants to healthcare and medical schools, including those for whom we have no information on test scores, and use the score availability as the dependent variable. As mentioned, the type of data available to us differs depending on whether participants apply to medical or healthcare schools and, therefore, we conduct the analysis separately for the two groups.

In particular, for prospective health professionals, the test outcomes come directly from administrative sources. This means that we do have the score of all those who did take the exam at the local healthcare school. The lack of score availability may then happen due to three possible reasons: (i) participants may have decided not to enter healthcare school at all, changing their career; (ii) participants may have decided to enter a healthcare school in another province; (iii) other reasons (e.g., illness on the day of the exam). In terms of our experimental treatment, it is indeed plausible to assume that positive (negative) information about wages can encourage (discourage) participants to continue pursuing their career in the health sector; moreover, given that we provide information about the wage paid by the local health authority, a positive (negative) information about local wages could indeed encourage (discourage) people to pursue their education locally. Between the two possible reasons, we consider the change in career to be the most likely reason for people to drop out, given that wages offered by the public health sector in South Tyrol are generally better than in the rest of Italy. Overall, for prospective health professionals, lack of scores means a failure to attend the exam they were preparing for, i.e., the exam at the local healthcare school.

Scores of medical school applicants were instead not accessible through institutional sources; therefore, we contacted the applicants directly, first through e-mail, then, if they did not reply, through WhatsApp; finally, if they still did not reply and provided their phone number in the survey, by telephone. We may therefore have missing scores also for participants who did take the exam.[21] For instance, for some participants test scores could be unavailable due to problems with contact tools or details (e.g., emails going into spam, wrong or missing telephone number) or, alternatively, to lack of willingness to reply. This latter option is of course very plausible when using emails or messaging apps, less so when contacted by phone, as participants cannot plausibly screen out our number and, therefore, do not know who is calling when deciding whether or not to accept the call. In any case, lack of scores for prospective physicians cannot strictly be interpreted as evidence that participants did not take the exam they were preparing for.

In Table 6, model (1) and (2), we conduct the analysis for prospective physicians and we find no significant effect. For applicants to healthcare schools, models (3) and (4), we see that a positive shock to wage expectations increases the likelihood of score availability (i.e., of showing up for the exam in the local school), with a marginally significant coefficient. For overestimators, the treatment results instead in a negative coefficient, thus indicating a lower likelihood of taking the exam, but it is not significant. Also in this case, (β_3 - β_2) + β_1 is close to zero and insignificant, meaning that preexisting differences between under- and over-estimators disappear once expectations are aligned through our information treatment.

[21] As mentioned, for medical schools the exam is national, and, anyhow, there is no medical school in the province. Therefore, what mentioned above for healthcare schools about the relevance of the local institution does not apply for medical students.

The analysis above suggests that prospective health professionals show some tendency to react along the extensive margin, i.e., attendance to the exam. On the contrary, this is not the case for prospective physicians.

Table 6: Likelihood of taking the test

	Doc	Doc	HP	HP
β_1: Wage underestimation	-0.050	-0.012	-0.129	-0.145*
	(0.191)	(0.207)	(0.082)	(0.085)
β_2: Treatm. x Wage overest.	0.000	0.000	-0.037	-0.056
	(.)	(.)	(0.096)	(0.097)
β_3: Treatm. x Wage underest.	0.052	0.055	0.106*	0.107*
	(0.076)	(0.072)	(0.057)	(0.057)
Female		-0.028		0.084
		(0.066)		(0.061)
Age		-0.025		0.004
		(0.021)		(0.008)
High Cognitive		0.195***		0.065
		(0.062)		(0.049)
High Altruism		0.031		0.047
		(0.075)		(0.046)
Family Network		0.125*		-0.044
		(0.068)		(0.050)
Observations	141	141	216	216
Wald Test of Joint Significance	0.69	11.81	6.31	12.58
Model p-value	0.957	0.225	0.277	0.248
Pseudo R-squared	0.005	0.091	0.030	0.055
Additional Controls:				
Year	Yes	Yes	Yes	Yes

Notes: *Estimation method: Logit. Dependent Variable: Available Score, a dummy equal to 1 (0) if the individual test score is (is not) available. In models (1) and (2), five observations perfectly predict success. Coefficients are reported as marginal effects. Robust standard errors are clustered at the unit-of-randomization level. *** p<0.01, ** p<0.05, * p<0.1*

5 Concluding Remarks

In this section, we bring the results together and provide some concluding remarks.

We have shown through an information experiment that financial incentives have an impact on the recruitment process in the health sector. Indeed, applicants with different wage expectations about initial wages have different performances in the entry exam. However, this difference disappears as they are informed about the correct wage, with those receiving a positive shock, i.e., being informed that the actual wage is higher than what they expected, doing better, while those receiving a negative shock doing worse.[22] This margin appears to be particularly effective for applicants to medical schools. By contrast, applicants to healthcare schools appear to react more along the extensive

[22] The relatively high share of applicants who underestimate wages implies that the estimated effects are more precise for a positive shock compared to a negative one.

margin, as they are more likely to actually show up for the exam once informed about a prospective higher wage.

Passing the entry exam into medical or healthcare school is a crucial step to start a career in the health sector. Our novel experimental evidence shows that monetary incentives may not only affect the decision to apply to health professions, as pointed out by previous literature, but also the likelihood to pass admission tests. This crucial step has been hitherto neglected by the literature, maybe because it was assumed that students who had already decided to study for the admission test would try to pass it to the best of their ability, leaving no room for incentives. The evidence we present suggests this is not necessarily the case.

References

Angrist, J. and Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *The American Economic Review*, 99(4):1384–1414.

Ashraf, N., Bandiera, O., Davenport, E., and Lee, S. S. (2020). Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services. *The American Economic Review*, 110(5):1355–1394.

Attema, A. E., Galizzi, M. M., Groß, M., Hennig-Schmidt, H., Karay, Y., L'haridon, O., and Wiesen, D. (2023). The formation of physician altruism. *Journal of Health Economics*, 87:102716.

Betts, J. R. (1996). What Do Students Know about Wages? Evidence from a Survey of Undergraduates. *The Journal of Human Resources*, 31(1):27.

Briel, S., Osikominu, A., Pfeifer, G., Reutter, M., and Satlukal, S. (2020). Overconfidence and Gender Differences in Wage Expectations. *IZA Discussion Paper Series*, (13517).

Burgess, S., Metcalfe, R., and Sadoff, S. (2021). Understanding the response to financial and non-financial incentives in education: Field experimental evidence using high-stakes assessments. *Economics of Education Review*, 85:102195.

Conlon, J. J. (2021). Major Malfunction. *Journal of Human Resources*, 56(3):922–939.

Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *The Quarterly Journal of Economics*, 128(3):1169–1218.

Deserranno, E. (2019). Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda. *American Economic Journal: Applied Economics*, 11(1):277–317.

Favara, M., Glewwe, P., Porter, C., and Sanchez, A. (2021). Expecting Better? How Young People Form Their Earnings Expectations. *IZA Discussion Paper Series*, (14289).

Fedele, A. (2018). Well-paid nurses are good nurses. *Health Economics*, 27(4):663–674.

Fuster, A. and Zafar, B. (2023). Survey experiments on economic expectations. In *Handbook of Economic Expectations*, pages 107–130.

Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, 61(1):3–40.

Hennig-Schmidt, H. and Wiesen, D. (2014). Other-regarding behavior and motivation in health care provision: An experiment with medical and non-medical students. *Social Science & Medicine*, 108:156–165.

Heyes, A. (2005). The economics of vocation or 'why is a badly paid nurse a good nurse'? *Journal of Health Economics*, 24(3):561–569.

Jerrim, J. (2011). Do UK higher education students overestimate their starting salary? *Fiscal Studies*, 32(4):483–509.

Kremer, M., Miguel, E., and Thornton, R. (2009). Incentives to learn. *The Review of Economics and Statistics*, 91(3):437–456.

Kugler, P. (2022). The role of wage beliefs in the decision to become a nurse. *Health Economics*, 31(1):94–111.

Leavitt, V. M. (2011). Raven Progressive Matrices. In *Encyclopedia of Clinical Neuropsychology*, pages 2114–2115. Springer New York, New York, NY.

Li, J., Casalino, L. P., Fisman, R., Kariv, S., and Markovits, D. (2022). Experimental evidence of physician social preferences. *Proceedings of the National Academy of Sciences*, 119(28):e2112726119.

Li, J., Dow, W. H., and Kariv, S. (2017). Social preferences of future physicians. *Proceedings of the National Academy of Sciences*, 114(48):E10291–E10300.

Propper, C. and Van Reenen, J. (2010). Can pay regulation kill? Panel data evidence on the effect of labor markets on hospital performance. *Journal of Political Economy*, 118(2):222–273.

Schweri, J. and Hartog, J. (2017). Do wage expectations predict college enrollment? Evidence from healthcare. *Journal of Economic Behavior and Organization*, 141:135–150.

Wiswall, M. and Zafar, B. (2015a). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824.

Wiswall, M. and Zafar, B. (2015b). How do college students respond to public information about earnings? *Journal of Human Capital*, 9(2):117–169.

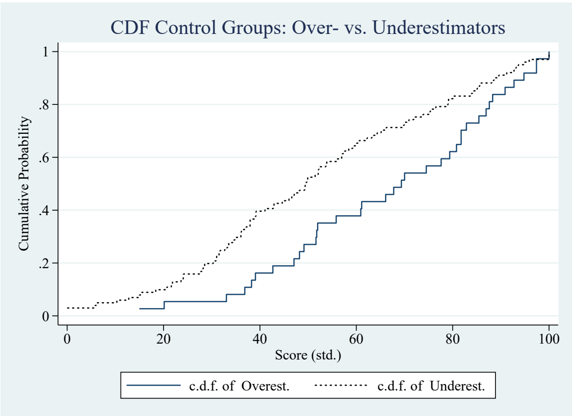
A Appendix: Additional Tables and Figures

Table A1: Sample Balance: Likelihood of being assigned to Treatment

	(1)		(2)	
Wage underestimation	0.875	(0.213)	0.884	(0.234)
Survey Year	0.986	(0.125)	0.996	(0.127)
Female			1.195	(0.327)
Age			0.986	(0.040)
Physicians			0.937	(0.231)
High Altruism			0.861	(0.189)
High Cognitive			1.381	(0.305)
Observations	362		362	
Wald Test of Joint Significance	0.308		3.669	
Model p-value	0.857		0.817	
Additional Controls:				
Year	Yes		Yes	

Notes: *Estimation method: Logit. Dependent Variable: Treatment (d.), where Treatment (d.)=1 if the individual received true wage information by email. Coefficients are reported as odds ratios. Robust standard errors are clustered at the unit-of-randomization level. *** p<0.01, ** p<0.05, * p<0.1*

Figure A1: Cumulative Distributions of Scores - Over- vs Under-estimators



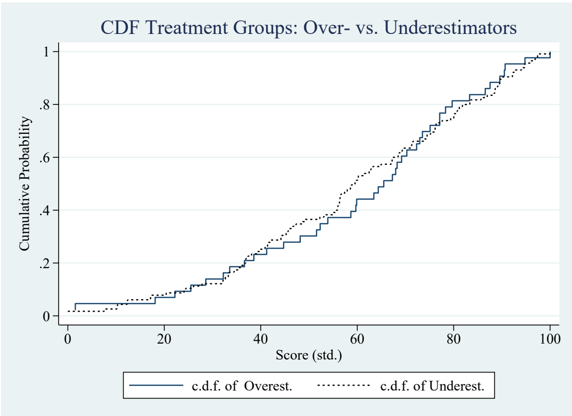


Table A2: Determinants of Wage Underestimation

	Full Sample	Available-score Sample
Female	0.131**	0.162***
	(0.052)	(0.057)
Age	-0.006	-0.005
	(0.007)	(0.008)
High Cognitive	0.007	0.020
	(0.045)	(0.049)
Family Network	0.101***	0.111***
	(0.038)	(0.043)
High Altruism	0.060	0.066
	(0.039)	(0.044)
Observations	362	296
Wald Test of Joint Significance	46.54	43.50
Model p-value	0.000	0.000
Pseudo R-squared	0.1851	0.1837
Additional Controls:		
*Year*Professional Category*	Yes	Yes

Notes: *Estimation method: Logit. Dependent Variable: Wage Underestimator (d.), which takes value 1 if the individual underestimated prospective wages. Robust standard errors are clustered at the unit-of-randomization level. *** p<0.01, ** p<0.05, * p<0.1*