Labour Economics 93 (2025) 102688

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Wage expectations and access to healthcare occupations: Evidence from an information experiment[☆]

Juliana Bernhofer[a], Alessandro Fedele[b,*], Mirco Tonin[b,c,d,e]

[a] Department of Economics, Ca' Foscari University of Venice, Italy
[b] Faculty of Economics and Management, Free University of Bozen-Bolzano, Italy
[c] Institute for the Evaluation of Public Policies - Bruno Kessler Foundation, Italy
[d] IZA, Germany
[e] CESIfo, Germany

ARTICLE INFO

JEL classification:
I1
I23
J3
C9

Keywords:
Information experiment
Wage expectations
Test scores
Health occupations

ABSTRACT

We investigate how correcting students' wage expectations affects their performance on admission tests for medical and healthcare schools, a critical step for aspiring healthcare professionals. Using a randomized information experiment with Italian applicants, we first elicited their expectations about the starting wage of the healthcare profession they intended to pursue. The treatment group was then informed of the actual starting wages, while the control group received no such information. Finally, we collected and analyzed their test scores. Our findings reveal that applicants with lower wage expectations tend to perform worse on the test. However, correcting these expectations eliminates the performance gap: providing accurate wage information enhances test scores for applicants who initially underestimated wages, while it negatively impacts those who overestimated them.

1. Introduction

Wage expectations and beliefs are powerful motivators in career choice. However, pessimistic wage beliefs are common and often undermine potential applicants' motivation to pursue certain professions (e.g., Conlon, 2021). Correcting inaccurate wage beliefs is therefore particularly important for professions and sectors experiencing severe worker shortages.

This experimental study focuses on one such sector: healthcare. We examine how correcting students' wage expectations impacts their performance on a crucial step toward entering healthcare occupations - the admission test for medical and healthcare schools. A growing body of literature, reviewed in the next section, explores the effect of financial stimuli - such as expected wages, actual wages, or career opportunities - on the *decision* to apply for a career in the health sector. These studies find that better financial incentives attract more applicants, improve their average ability, and have ambiguous effects on their average other-regarding behaviors like altruism and prosocial motivation, which are key determinants of healthcare quality.

An essential feature of the healthcare sector is the selective nature of admission to training programs. In many countries selection relies on high school grades or other academic records, and/or standardized admission tests (e.g., the International Medical Admissions Test or the Medical College Admission Test in the US). In Italy, where we conducted this study, admission is solely based on test scores, and entry is rather competitive.[1] Consequently, deciding to apply does not guarantee entry into the field, as only those who pass the admission test can pursue a healthcare career.

☆ 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) and Claudiana 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).

* Corresponding author.
E-mail addresses: juliana.bernhofer@ec.europa.eu (J. Bernhofer), alessandro.fedele@unibz.it (A. Fedele), mirco.tonin@unibz.it (M. Tonin).

[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).

https://doi.org/10.1016/j.labeco.2025.102688
Received 28 February 2024; Received in revised form 16 January 2025; Accepted 1 February 2025
Available online 22 February 2025

Building on this context, our study departs from existing literature by focusing on individuals who *have already decided* to apply to medical and healthcare schools (e.g., physiotherapy, nursing, and obstetrics), but *have not yet been admitted*. We investigate whether correcting their wage beliefs influences their performance on the admission test - a crucial, yet previously overlooked, step in gaining access to a healthcare career.

We consider students who obtained a high school diploma and are attending preparatory courses during the summer months for the admission tests. We designed the following randomized information experiment. First, we administered a questionnaire to elicit students' beliefs about the starting wage for the medical or healthcare profession they intended to pursue, along with other relevant information. A few weeks later, approximately four weeks before the admission tests, we informed a randomly selected half of the students (the treatment group) about the true starting wage. The remaining students (the control group) received no such information. Finally, we collected their admission test scores. Remarkably, information experiments in Western healthcare contexts remain rare, as highlighted by Haaland et al. (2023).

Our findings reveal that wage expectations among applicants are systematically biased: over 70% of students in our sample underestimate the true starting wage, while fewer than 30% overestimate it.[2] It follows that our treatment can indeed convey information: a *positive shock* for those who underestimated wages and a *negative shock* for those who overestimated them. Our main results, derived from both non-parametric tests and regression analysis, are as follows: applicants with lower wage expectations about the starting wage perform worse on the admission test. However, this performance gap disappears when participants are informed of the correct wage. Specifically, those who receive a positive shock (learning that the actual wage is higher than expected) perform better, while those who receive a negative shock perform worse.

Our treatment focuses on monetary considerations, raising the question of whether it could disproportionately stimulate low-altruism candidates. Using incentivized dictator games vis-à-vis charities to measure altruism, we find that a positive shock to wage expectations does not improve the test scores of selfish individuals more than those of altruistic ones. Moreover, a negative shock discourages mostly selfish individuals. We also show that the treatment effect on exam scores is more pronounced among applicants to medical schools. In contrast, prospective health professionals are more likely to respond on the extensive margin - namely, deciding to take the exam after attending a preparatory course.

In terms of our contribution, we offer experimental evidence that financial incentives are effective on a new margin, the admission test performance and participation. This finding is far from straightforward, given that our subjects had already demonstrated commitment to a healthcare career, as evidenced by their enrollment in preparatory course testifies. It might be expected that such individuals would already be fully motivated when taking the exam. 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.

Our information experiment contributes to the growing field of survey experiments in economics that examine the causal link between expectations and behavior. For recent reviews, see Haaland et al. (2023) and Fuster and Zafar (2023). Importantly, our experimental subjects operate in a high-stakes field environment, where admission test outcomes 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. Section 4 presents and discusses the results. Section 5 concludes.

2. Literature review

In this section, we briefly review five related streams of literature and detail the contribution of the paper. In a nutshell, this is the first paper looking at the impact of financial incentives, in the form of expected wages, on test outcomes in the entry exam for medical and healthcare schools.

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) and Kugler (2022). Deserranno (2019) runs a field experiment with applicants for a health-promoter position in Uganda: she finds that higher financial incentives attract more applicants, but crowd out prosocially motivated applicants; 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 observe that a higher wage for community development positions in the Mexican public sector attracts agents who are more able and more motivated to provide services; such an advantageous selection outcome is in line with Fedele (2018). Ashraf et al. (2020) implement a field experiment with applicants for healthcare positions in Zambia: a higher present value of future earnings is shown to attract high-skilled individuals at the expense of prosocial individuals; this adverse selection, however, vanishes in the pool of successful applicants. 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 a recent literature on information experiments using undergraduate students' expectations about earnings. Findings show that students tend to underestimate salaries and that correcting their expectations affects their actual choice of majors (e.g., Wiswall and Zafar, 2015a,b; Conlon, 2021). 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 showing that financial incentives positively impact student performance in high-stakes exams (e.g., Angrist and Lavy, 2009, Kremer et al., 2009). This stream considers the general population of students, while our focus is on aspiring health professionals, a category that is of particular interest and has peculiarities, for instance regarding the importance of altruistic motives.

Fourth, the presence of biased beliefs in our sample enables us to provide both negative and (especially) positive wage shocks to students. This connects our paper to the experimental literature on gift exchange that studies how workers respond to unexpected changes in wages (e.g., Gneezy and List, 2006; Bellemare and Shearer, 2009; Ockenfels et al., 2015). Since the seminal theory paper by Diamond (1971), the role of biased consumer beliefs for search market outcomes has been widely investigated. In labor economics, the focus is generally on workers' beliefs about wages. For instance, the recent paper by Jäger et al. (2024) observes that workers wrongly anchor their beliefs about

[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).

outside options on their current wage. Our paper considers prospective students rather than workers and, in line with the gift exchange literature, finds that unexpected increases (decreases) in wages trigger responses in terms of better (worse) performance.

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 (e.g., Hennig-Schmidt and Wiesen, 2014; Li et al., 2017; Attema et al., 2023).

3. Data

In this section, we first describe the institutional background and the admission test. We then provide details about the students who are part of the sample and about the questionnaire, the treatment and the outcome variable, i.e., the test score. Finally, we provide some descriptive statistics, discuss the balance between treatment and control groups, and compare our sample with the general population of applicants.

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

Test. 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 min; (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 selected and compete only with other applicants to the same university and profession.[3] For medicine, it is necessary to reach a minimum score of 20 to be included in the national ranking. Given that demand is higher than supply, however, the minimum score to actually be assigned a place in the period under consideration is higher and, in particular, it was 43.2 in 2018, 41.7 in 2019, 39.5 in 2020.[4] For health professions, the minimum score fixed by law to be included in the local ranking is 0. Unfortunately, we do not have information on the minimum score needed to actually be assigned a place, score that varies year-by-year among the ten different specialties listed in the previous subsection.

Students and Questionnaire. In the summer of 2018, we launched the first wave of a paper-based questionnaire to applicants to medical and healthcare schools; two additional waves followed in 2019 and 2020 for a total of 408 participants. This sample size is determined by the availability of subjects in the preparatory courses over the three years of data collection. Due to the pandemic, in 2020 we switched from the paper-based version to a digital version designed using the software Qualtrics. 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: see Appendix C for the full text. 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.[5]

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). Below, we report the English translation of the e-mail sent to prospective health professionals. The email sent to prospective physicians is identical, with the exception of the salary level.

> 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 most relevant one 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 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,

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

[4] See https://www.unidformazione.com/risultati-punteggio-minimo-test-medicina-2024-sessione-luglio/.

[5] 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).

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. Moreover, 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. Subjects in the control group did not receive any email before the admission test. If information exchange occurred between subjects in the treatment and control groups during the period between the treatment and the admission test dates, our results should be interpreted as a lower bound of the true effect of receiving information about 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.

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

Given that prospective physicians' individual exam results could not be retrieved from administrative sources, there is concern that self-reported scores might be inflated. 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 and we find no statistically significant difference (we provide more details at the end of this section). In addition, even if over-reporting was present, it would be problematic only if it was correlated with treatment. Moreover, if one expects that over-reporting is more likely among low-performance respondents, the actual treatment effect would be bigger because - due to the overall positive effect of the treatment - such respondents are more numerous in the control group.

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 answers[7]; (iii) 27 subjects who indicated the correct starting wage, as 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. In Section 4.1, however, we report a robustness check where we include these subjects.

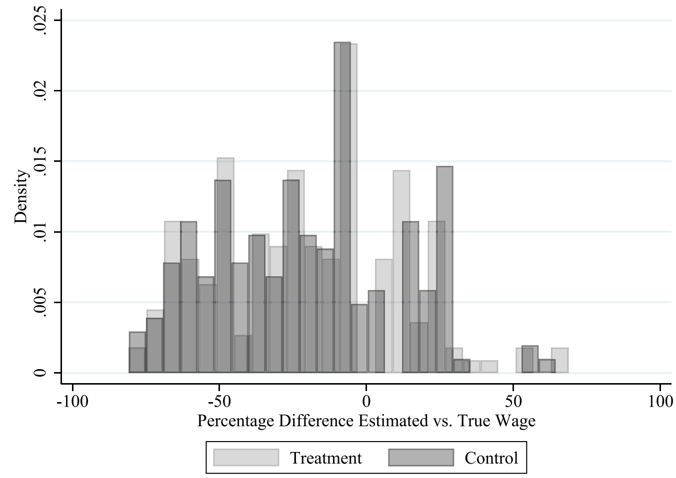

Fig. 1. Percentage deviation of expected wages from true wages.
Notes: The figure compares the distributions of the % deviations from true wages across treatment (186 obs.) and control (176 obs.) groups. The sample includes 362 subjects, regardless of score availability but with the exclusion of subjects who indicated the correct starting wage. A two-sample Kolmogorov–Smirnov test (D = 0.0308, p-value = 1.000) confirms treatment and control groups are balanced..

The final sample contains 362 subjects, while test scores are available for 296 subjects. In the results section, we explore in detail the determinants of score availability. In Fig. 1, we display the distribution of percentage deviations of wage expectations from true wages, considering the full sample of 362 individuals and distinguishing between the treatment and control groups. Figure A1 in Appendix A presents the same figure separately for (prospective) physicians and health professionals. The two figures show that most subjects underestimate the true wage, with some overestimating it. In the remainder of the paper, we refer to the former as *underestimators* and to the latter as *overestimators*. In Table 1, we provide summary statistics for our variables of interest for both the full sample and the sample with available score.

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. This, of course, implies that we have less power to detect the impact of correcting overly optimistic expectations, and should be taken into consideration when interpreting the results. *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)[8]; the mean value is around 8 and underestimators tend to perform better. *High Altruism* stems from two incentivized dictator games, in which respondents decide to donate an amount between 0 and 100 euro

[6] 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.

[7] 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".

[8] 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.

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
Altruism	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: The table reports summary statistics for our main variables of interest. *Score (std.)* is the dependent variable in most regression analyses and represents the rescaled version (minmax normalized within the matrix of test year and professional category) of the admission test scores. *Treatment = 1* if the questionnaire respondent received our email containing information on wages. *Female = 1* if the subject is female. *Age* is expressed in years. *Cognitive Skills* is the number of correct answers to the Raven's Standard Progressive Matrices Test (Set E). *Altruism* is expressed in euro and is the average donation out of 100 euro in two dictator games. *Family Network = 1* if prospective doctors (health professionals) have a doctor (health professional) in their close family. *Physicians = 1* if the applicant attended the preparatory course for medical school (as opposed to health school).

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 euro 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) of the prospective physician (health professional); this is the case for half of our sample. 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.

Balance between Treatment and Control Groups. To show that subjects randomly assigned to treatment and control groups have similar wage expectations, we test whether the distributions presented in Fig. 1 are significantly different. The distributions are indeed very similar (two-sample Kolmogorov–Smirnov test, $D = 0.0308$, p-value $= 1.000$), 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 have 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. 81%; $t(266) = -0.745$, $p = .457$).[9]

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.[10] 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 Section 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 Section 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

[9] The small 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.

[10] 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.

Fig. 2. Cumulative distributions of scores.
Notes: The figures compare cumulative distribution functions of standardized test scores of treated (solid line) and untreated (dashed line) subjects for both underestimators (top) and overestimators (bottom). A shift to the right of the functions indicates higher standardized scores. See Table Table 2 for non-parametric tests of equal distributions.

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 Fig. 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). Among the underestimators, the distribution for the treatment group is clearly shifted to the right, particularly in the central region. In contrast, for the overestimators, the distribution for the treatment group shifts to the left. However, this shift is less pronounced, with noticeable overlap between the two distributions in the central region. This is already an indication that being informed about the correct wage encourages underestimators to do better, while somehow discouraging overestimators.

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) 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 Fig. 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 A2 in the appendix for a graphical comparison. From the mean scores reported in Table 2, it appears that the two groups differ in the 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. The positive link between test scores and deviation of expected wages from true wages is confirmed by Fig. 3, where we present a scatter plot with the percentage difference to actual monthly salary on the horizontal axis and the standardized test score on the vertical axis, with data presented separately for medical school and healthcare school applicants. In both cases, the relationship is positive, though the slope is steeper for healthcare school applicants.

Next, we move to the regression analysis and estimate the following OLS model,

$$Score_i\,(std.) = \beta_0 + \beta_1 I_i + \beta_2 T_i \times \left(1 - I_i\right) + \beta_3 T_i \times I_i + \sum_{j=1}^{5} \beta_{4j} Y P_j + \epsilon_i \quad (2)$$

where: $Score_i(std.)$ denotes subject i's standardized score as given by (1); I_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. Finally, $\sum_{j=1}^{5} \beta_{4j} Y P_j$ represents 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 Eq. (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.[11] Similarly, *High Altruism* is a dummy variable equal to 1 when an applicant to medical (healthcare) schools donates weakly more than the median values of the average between individual donations to the WWF and MSF.[12] 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.

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

[12] 75 euro is the median value for both professions and in both samples.

Table 2
Non-Parametric test results.

	Underestimators n = 216			Overestimators n = 80		
	C		T	C		T
Mean score	51.45		58.79	66.38		59.87
Standard deviation	26.02		24.91	22.90		24.69
Combined Kolmogorov–Smirnov p-value		0.025			0.405	
Mann–Whitney p-value		0.027			0.240	

Notes: The table reports mean scores of control (C) and treatment (T) groups for under- and overestimators and the p-values of two two-sided non-parametric tests (Kolmogorov–Smirnov for differences in the distribution of scores and Mann–Whitney for differences in average ranks).

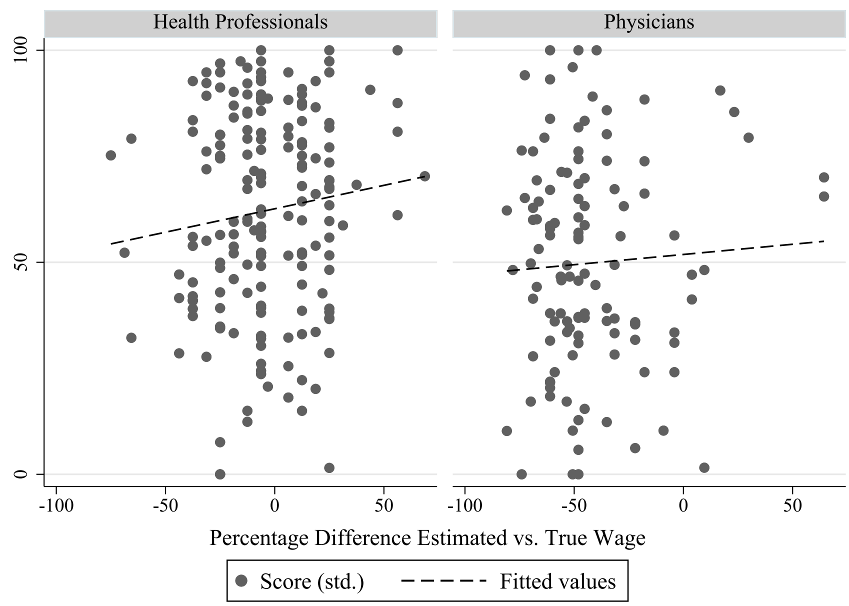

Fig. 3. Standardized scores and percentage deviation of expected wages from true wages - by profession.
Notes: The figure reports % deviations from true wages across treatment (x-axis) and standardized scores (y-axis) by professional category. The sample includes all 296 subjects with available scores.

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

We shift our attention to models (3) and (4), where we distinguish between weak and strong underestimators. We see from coefficients 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 downwards tend to perform worse, albeit the impact for the latter group is measured with less precision. As mentioned in Section 3, the fact that only a quarter of subjects are overestimators implies that we have less power to detect an impact on them. 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., $(\beta_3 - \beta_2) + \beta_1$, gives a non-significant coefficient of 2.5 (p-value 0.6).

As mentioned in Section 3, we have excluded 27 subjects who indicated the correct starting wage. Test scores are available for 24 of these. In a robustness check, we introduce a new group "Correct wage estimation", where we include these subjects plus 60 more whose wage expectations were within a range of 10% of the correct figure. Overall, this new sample contains 296 + 24 individuals. We cannot use smaller ranges because no additional subject is within a $\pm 1\%$ range and only 8 would be within a $\pm 5\%$ range. The split among the three groups using a range of $\pm 10\%$ is that 170 subjects are underestimators, 84 are (almost) correct, and 66 are overestimators. Table B1 in Appendix B replicates Table 3 using this alternative specification. The coefficient of the treatment dummy for those guessing correctly within $\pm 10\%$ is small and insignificant, while the treatment effect for the underestimators remains positive and significant. The outcome for underestimators remains qualitatively unchanged by the inclusion of the group of those guessing almost correctly also when distinguishing between weak and strong underestimators. The treatment coefficient for overestimators is negative, but not significant.

Taken 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

Table 3
Treatment effects on scores.

	(1)	(2)	(3)	(4)
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 Overestimation	−7.0	−8.0*	−7.0	−8.1*
	(4.5)	(4.7)	(4.5)	(4.7)
β_3: Treatm. x Wage Underestimation	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	0.140	0.184	0.151	0.191
Additional controls:				
*Year*Professional Category*	Yes	Yes	Yes	Yes

Notes: Estimation Method: OLS. Dependent Variable: Standardized Test Score, which expresses the admission test score on a minmax scale from 0 (minimum score achieved within the year*professional category of reference) to 100 (maximum score achieved within the year*professional category of reference). Strong Wage Underestimation, High Cognitive and High Altruism are median splits by professional category measuring strong (vs. weak) wage underestimation, high (vs. low) cognitive abilities and high (vs. low) attitudes towards charitable giving. Family Network = 1 with at least one physician (health professional) within the close family network. Robust standard errors are clustered at the unit-of-randomization level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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,[13] as is an applicant having someone in the family active in the same profession.[14]

In the remainder of this section, we further explore the intensive margin effect of our information treatment by investigating heterogeneity across different dimensions. First, we split the sample along the cognitive skills and charitable giving dimensions. Next, we analyze applicants to medical and healthcare schools separately. Given the limited sample size, these additional analyses should be taken with caution, though we believe they provide valuable insights.

Cognitive Skills and Altruism. Since our treatment revolves around financial incentives, one might wonder whether it favors 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 (strictly below) the median values by professional category.[15] 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).

[13] 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.

[14] 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 low in the profession of the complainer.

[15] 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. × Wage Overestimation	−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. × Wage Underestimation	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	0.090	0.321	0.114	0.382	0.209	0.141	0.224	0.226
Additional controls:								
*Year*Professional Category*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *Estimation Method: OLS. Dependent Variable: Standardized Test Score, which expresses the admission test score on a minmax scale of 0 (minimum score achieved within the year*professional category of reference) to 100 (maximum score achieved within the year*professional category of reference). High Altruism and High Cognitive are median splits by professional category measuring high (vs. low) cognitive abilities and high (vs. low) attitudes towards charitable giving. Family Network = 1 with at least one physician (health professional) within the close family network. Models (1) - (4) compare the subpopulation with high cognitive skills (models (1) and (3)) to the subjects with low cognitive skills (models (2) and (4)). Models (5) - (8) compare the population with high altruistic tendencies (models (5) and (7)) to those with low altruism (models (6) and (8)). Robust standard errors are clustered at the unit-of-randomization level.* *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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 subjects with low altruism - models (6) and (7). Again, these differences are not statistically significant, but, overall, we can safely exclude that our treatment would select selfish people within the pool of applicants, given the combination of discouragement for selfish overestimators and some possible encouragement of altruistic underestimators. The same conclusion can be drawn in terms of cognitive abilities.

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 $(\beta_3 - \beta_2) + \beta_1$ 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.

Given that, as reported in Section 3, we know the threshold test scores to access medical schools, as a robustness check we assess whether informing students affects entry into medical schools. To this aim, we create a binary outcome variable that indicates whether applicants scored above the year-specific threshold. The results, reported in Table B2, Appendix B, show that, indeed, de-biasing positively affects the likelihood of being above the threshold among underestimators, while the effect on overestimators has the opposite sign.

4.2. Extensive margin

So far, we have analyzed the treatment effect on the intensive margin, focusing on test scores. Another key margin of interest is exam participation. While all individuals in our sample attended a preparatory course for the medical or healthcare school entrance exam during the summer, some ultimately chose not to take the exam for various reasons.

In Table 6, we examine whether wage expectations and the treatment correcting these expectations influence this extensive margin. The analysis includes the full sample of applicants to healthcare and medical schools, even those with missing test scores. Since the type of data available differs depending on whether participants applied to medical or healthcare schools, we conduct the analysis separately for each group.

For prospective healthcare professionals, test outcomes are obtained directly from administrative sources. This allows us to identify the scores of all individuals who took the local healthcare school exam. Missing scores in this group may arise from three main reasons: (i) participants opted out of pursuing a healthcare career entirely, changing their career path; (ii) participants enrolled in a healthcare program in a different province; or (iii) other factors, such as illness on the exam day. Our experimental treatment could plausibly affect participation and positive (negative) information about wages might encourage (discourage) individuals to continue pursuing a healthcare career. Furthermore,

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. × Wage Overestimation	−18.2	−5.7	−16.8	−6.3
	(15.0)	(4.8)	(13.7)	(5.1)
β_3: Treatm. × Wage Underestimation	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	0.103	0.092	0.275	0.119

Notes: Estimation Method: OLS. Dependent Variable: Standardized Test Score, which expresses the admission test score on a minmax scale of 0 (minimum score achieved within the year*professional category of reference) to 100 (maximum score achieved within the year*professional category of reference). High Altruism and High Cognitive are median splits by professional category measuring high (vs. low) cognitive abilities and high (vs. low) attitudes towards charitable giving. Family Network = 1 with at least one physician (health professional) within the close family network. We also control for the year in which the test was taken. In the above table we compare the subpopulation of prospective physicians (models (1) and (3)) to prospective health professionals (models (2) and (4)). Robust standard errors are clustered at the unit-of-randomization level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

since the treatment provides information specific to local wages, positive (negative) wage information could also influence the decision to take the exam locally. Given that public healthcare wages in South Tyrol are generally higher than the national average, however, we consider career changes to be the most likely reason for non-attendance at the local healthcare school exam. Overall, for prospective healthcare professionals, missing scores typically indicate non-attendance at the local exam.

Conversely, for medical school applicants, exam scores were not accessible through institutional sources. To gather this information, we contacted applicants initially via email, then through WhatsApp if they did not respond and had provided a number in the survey, and finally by phone. Missing scores in this group may result from various issues, such as technical problems (e.g., emails filtered as spam, incorrect or missing phone numbers) or unwillingness to share scores. While unwillingness to share is a plausible explanation for non-responses via email or messaging apps, it is less likely in phone interactions, as participants would not typically know who is calling before answering. Therefore, for prospective medical students, missing scores cannot be strictly interpreted as evidence of non-participation in the exam.[16]

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, $(\beta_3 - \beta_2) + \beta_1$ is close to zero and insignificant, meaning that preexisting differences between under- and over-estimators disappear once expectations are aligned through our information treatment.

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, who instead react along the intensive margin, as shown in Section 4.1. As a robustness check, in Table B3, Appendix B, we report for the full sample the results of a Cragg hurdle regression, jointly modeling the intensive and extensive margins. This analysis confirms the positive impact of the treatment on underestimators, while the effect on overestimators is not statistically significant.

Overall, our findings that underestimators tend to react positively to the treatment may be interpreted as a positive reaction to an unexpected increase in their future wage. This interpretation is in line with the gift exchange literature findings cited in Section 2. For overestimators the impact has the opposite sign, but is less precisely measured, as they are only a quarter of the sample.

5. Concluding remarks

We have shown through an information experiment that financial incentives influence the recruitment process in the health sector. Specifically, applicants with differing expectations about initial wages exhibit varying performances on the entry exam. However, this performance gap closes when applicants are informed of the correct wage. Those receiving a positive shock, learning that the actual wage is higher than expected, perform better, while those receiving a negative shock perform worse. As mentioned, the relatively high proportion of applicants underestimating wages results in more precise estimates for the effects of positive shocks compared to negative ones. This effect is particularly pronounced among medical school applicants, where the intensive margin (exam performance) is more significantly impacted. In contrast, healthcare school applicants primarily respond along the extensive margin, as they are more likely to attend the exam when informed of a higher prospective wage.

Passing the entry exam for medical or healthcare school is a pivotal step toward starting a career in the health sector. Our novel experimental evidence demonstrates that monetary incentives not only influence the decision to pursue health professions, as highlighted by previous research, but also affect the likelihood of passing admission tests. This

[16] Since the medical school entrance exam is national and there were no medical schools in the province, the significance of the local institution, which is relevant for healthcare applicants, does not apply to medical school applicants.

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. × Wage Overestimation	0.000	0.000	−0.037	−0.056
	(.)	(.)	(0.096)	(0.097)
β_3: Treatm. × Wage Underestimation	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

Notes: Estimation Method: Logit. Dependent Variable: Available Score, a dummy equal to 1 if the individual test score is available and 0 otherwise. In models (1) and (2), five observations perfectly predict success. High Altruism and High Cognitive are median splits by professional category measuring high (vs. low) cognitive abilities and high (vs. low) attitudes towards charitable giving. Family Network = 1 with at least one physician (health professional) within the close family network. We also control for the year in which the test was taken. Coefficients are reported as marginal effects. Robust standard errors are clustered at the unit-of-randomization level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

critical step has been largely overlooked in the literature, likely due to the assumption that students who prepare for admission tests will already be fully motivated to perform to the best of their ability, leaving no room for external incentives. However, our findings suggest this assumption is not entirely accurate.

In terms of policy implications, educational guidance counselors for secondary school students, as well as those responsible for recruitment and admissions in medical and healthcare schools, may benefit from providing applicants with more comprehensive job-related information, including, but not limited to, details about wages. Our findings reveal that applicants' expectations can be biased and correcting these biases enables more informed career decisions. Importantly, our results indicate that wage information does matter and suggest that concerns about potential negative selection caused by financial disclosures in professions where non-monetary motivations are essential may be less warranted than commonly assumed.

That said, it is crucial to exercise caution when drawing general policy conclusions from a single study involving a relatively small sample from a specific region. Additionally, our research does not investigate the mechanisms through which the treatment impacts performance, such as increased study time before the test or greater focus during the exam. Exploring these channels would be a valuable avenue for future research.

Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.labeco.2025.102688.

Data availability

Data will be made available on request.

References

Angrist, J., Lavy, V., 2009. The effects of high stakes high school achievement awards: Evidence from a randomized trial. Am. Econ. Rev. 99 (4), 1384–1414.
Ashraf, N., Bandiera, O., Davenport, E., Lee, S.S., 2020. Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services. Am. Econ. Rev. 110 (5), 1355–1394.
Attema, A.E., Galizzi, M.M., Groß, M., Hennig-Schmidt, H., Karay, Y., L'haridon, O., Wiesen, D., 2023. The formation of physician altruism. J. Heal. Econ. 87, 102716.
Bellemare, C., Shearer, B., 2009. Gift giving and worker productivity: Evidence from a firm-level experiment. Games Econom. Behav. 67 (1), 233–244.
Betts, J.R., 1996. What do students know about wages? Evidence from a survey of undergraduates. J. Hum. Resour. 31 (1), 27.
Briel, S., Osikominu, A., Pfeifer, G., Reutter, M., Satlukal, S., 2020. Overconfidence and gender differences in wage expectations. IZA Discuss. Pap. Ser. (13517).
Conlon, J.J., 2021. Major malfunction. J. Hum. Resour. 56 (3), 922–939.
Dal Bó, E., Finan, F., Rossi, M.A., 2013. Strengthening state capabilities: The role of financial incentives in the call to public service. Q. J. Econ. 128 (3), 1169–1218.
Deserranno, E., 2019. Financial incentives as signals: Experimental evidence from the recruitment of village promoters in Uganda. Am. Econ. J.: Appl. Econ. 11 (1), 277–317.
Diamond, P.A., 1971. A model of price adjustment. J. Econom. Theory 3 (2), 156–168.
Favara, M., Glewwe, P., Porter, C., Sanchez, A., 2021. Expecting better? How young people form their earnings expectations. IZA Discuss. Pap. Ser. (14289).
Fedele, A., 2018. Well-paid nurses are good nurses. Heal. Econ. 27 (4), 663–674.
Fuster, A., Zafar, B., 2023. Survey experiments on economic expectations. In: Handbook of Economic Expectations. pp. 107–130.
Gneezy, U., List, J.A., 2006. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. Econometrica 74 (5), 1365–1384.
Haaland, I., Roth, C., Wohlfart, J., 2023. Designing information provision experiments. J. Econ. Lit. 61 (1), 3–40.
Hennig-Schmidt, H., Wiesen, D., 2014. Other-regarding behavior and motivation in health care provision: An experiment with medical and non-medical students. Soc. Sci. Med. 108, 156–165.
Heyes, A., 2005. The economics of vocation or 'why is a badly paid nurse a good nurse'? J. Heal. Econ. 24 (3), 561–569.
Jäger, S., Roth, C., Roussille, N., Schoefer, B., 2024. Worker beliefs about outside options. Q. J. Econ. qjae001.
Jerrim, J., 2011. Do UK higher education students overestimate their starting salary? Fisc. Stud. 32 (4), 483–509.
Kremer, M., Miguel, E., Thornton, R., 2009. Incentives to learn. Rev. Econ. Stat. 91 (3), 437–456.
Kugler, P., 2022. The role of wage beliefs in the decision to become a nurse. Heal. Econ. 31 (1), 94–111.
Leavitt, V.M., 2011. Raven progressive matrices. In: Encyclopedia of Clinical Neuropsychology. Springer New York, New York, NY, pp. 2114–2115.
Li, J., Dow, W.H., Kariv, S., 2017. Social preferences of future physicians. Proc. Natl. Acad. Sci. 114 (48), E10291–E10300.
Ockenfels, A., Sliwka, D., Werner, P., 2015. Timing of kindness–Evidence from a field experiment. J. Econ. Behav. Organ. 111, 79–87.
Schweri, J., Hartog, J., 2017. Do wage expectations predict college enrollment? Evidence from healthcare. J. Econ. Behav. Organ. 141, 135–150.
Wiswall, M., Zafar, B., 2015a. Determinants of college major choice: Identification using an information experiment. Rev. Econ. Stud. 82 (2), 791–824.
Wiswall, M., Zafar, B., 2015b. How do college students respond to public information about earnings? J. Hum. Cap. 9 (2), 117–169.