Correcting Beliefs about Job Opportunities and Wages: A Field Experiment on Education Choices

Bart K. de Koning[1], Robert Dur[2,4,5], and Didier Fouarge[3,4,6]

[1] *Cornell University*
[2] *Erasmus University Rotterdam*
[3] *Maastricht University*
[4] *IZA*
[5] *CESifo*
[6] *Netspar*

March 20, 2025

Preliminary version

Click here for the most recent version of the paper

Abstract

We run a field experiment in which we provide information to students about job opportunities and hourly wages of occupations they are interested in. The experiment takes place within a widely-used career guidance counseling program in the Netherlands, and involves 28,267 pre-vocational secondary education students in 243 schools over two years. The information improves the accuracy of students' beliefs, both in the short run (for job opportunities and hourly wages) and up to seven months later (for job opportunities only). Administrative data that covers up to four years after the experiment show that students choose (and remain in) post-secondary education programs with better labor market prospects as a result of the information treatment.

Keywords: Education choice, labor market information, field experiment.
JEL codes: C93, D83, I26, J24
AEA Social Science Registry Entry: AEARCTR-0003220

Corresponding author: Bart K. de Koning (dekoning@cornell.edu). This project received funding from the Dutch Ministry of Education, Culture, and Science. Fouarge further acknowledges the financial support of NRO, UWV, Randstad, S-BB (grant: 405-17-900). This study is registered in the AEA RCT Registry and the digital object identifier (DOI) is: https://doi.org/10.1257/rct.3220. Ex-post IRB approval was obtained from the Cornell University IRB: Protocol #0148756.

1 Introduction

Each year, millions of teenagers around the world face a choice that has far-reaching consequences, both for themselves and for society: the choice of post-secondary education program. This choice is important for themselves, as the program from which they earn a degree is an important determinant of future labor market outcomes (see e.g., Bleemer and Mehta, 2022; Ketel et al., 2016; Kirkeboen et al., 2016). It is also important for society, as it affects future shortages and excess supply of labor in important occupations. Despite its huge importance, students often decide on their field of study without having accurate information about the labor market prospects of different programs (Baker et al., 2018; Hastings et al., 2016) and careers (Arcidiacono et al., 2012; Betts, 1996). As a result, many teenagers end up choosing programs that have a bleak outlook, both in terms of job opportunities and wages.

Motivated by this issue, a number of experimental studies have focused on the impact of providing information about the earnings prospects of different majors on enrollment decisions. The results of these experiments tend to be sobering. Even though students' choices move in the direction of education programs with better labor market prospects, the size of these effects tends to be limited, not seldomly statistically indistinguishable from zero (see e.g., Bonilla-Mejía et al., 2019; Conlon, 2019; Hastings et al., 2015; Pekkala Kerr et al., 2015). An explanation for this puzzling finding may be that a large part of the differences in earnings prospects by field of study can be explained by occupation choice (Altonji et al., 2012), which students seem to be well aware of (Arcidiacono et al., 2012). Information about average earnings by field of study may therefore not be of particular interest, as it does not accurately reflect the earnings opportunities conditional on the students' occupations of interest. In this paper, we therefore ask whether *occupation-specific* information about job opportunities and hourly wages affects beliefs, preferences over occupations, and educational choices of students.

We conduct a natural field experiment in which we provide a random selection of students with personally targeted information about the labor market prospects of a small set of occupations they are interested in. To our knowledge, we are the first to do so. We study whether the information leads students to correct their beliefs about the labor market prospects of these occupations and shifts students' preferences over occupations immediately after the intervention, and up to one and a half year later through a post-experimental survey. Using linked administrative data, we also study whether it influences their actual education choices up to four years after the experiment. Our multi-year field experiment involves 28,267 students at 243 different schools for pre-vocational education in the Netherlands. The students take part in the experiment in grades 8 to 10[1] and are between 13 and 16 years old.

The experiment takes place as part of a widely-used career guidance counseling program. In this computer-based program, students perform numerous assignments that help them discover what they are interested in, what they are good at and, ultimately, which occupations would be a good fit for them. As part of one of these assignments, students take an extensive 'skills and personality' test that results in a short-list of twenty (out of 353) occupations that fit their

[1]The second to fourth year of pre-vocational secondary education in the Netherlands.

abilities and interests best according to the answers they provided. Our experiment starts right after completing this test.

Our experiment proceeds as follows. First, we ask students in which secondary-school specializations (i.e., set of subjects; called "profiles") they are most interested. Next, we show students their shortlist of twenty occupations and ask them to select the five that they like the most. We then ask them to state their beliefs about the job opportunities and hourly wages for these five occupations, and to rank the occupations based on how much they would like to work in them. Subsequently, we provide students of randomly selected schools with information about the job opportunities and, for a random subset of these schools, supplemented by hourly wages of the selected occupations.[2] Students at the remaining schools do not receive any information and form our control group. In a sub-randomization arm that we explain in more detail later, we alter the 'sender' of the information to be either a labor market research institute, or a specific researcher (male or female and experienced or less experienced). The identity of the 'sender' is randomized within the treatment group. Next, students in both the treatment and control group watch an information video about the Dutch economy that does not convey any information about specific occupations. This gives students the opportunity to reflect and, if they so wish, update their stated beliefs. Finally, we ask students to again state their beliefs of wages and job prospects, and to re-rank occupations according to preferences. These answers are our first set of outcome measures. In addition to these data, we obtain (i) post-experimental survey data (up to one and a half year later) on the students' beliefs and preferences, and (ii) up to four years of administrative data on their education choices during and after secondary education.

In line with the earlier studies (Baker et al., 2018; Betts, 1996; Hastings et al., 2016), our results show that students have highly inaccurate beliefs about the job opportunities[3] and hourly wages of the occupations that they like. They tend to overestimate both, particularly for those occupations they like the most. Our information intervention is effective in correcting beliefs. Immediately following the intervention, treated students overestimate the job opportunities and hourly wages to a smaller degree, make smaller absolute errors, and are more likely to hold correct beliefs. The improved accuracy is mostly driven by students correcting overestimations. Our post-experimental survey data show that these effects partly persist: those who received the information in their final school year have more accurate expectations about the job opportunities up to seven months later.

We also find evidence that the treatment increases the likelihood that students change their favorite occupation. If students do so, they tend to substitute the initial occupation for one with better job opportunities or hourly wages. We do not find evidence that this ranking persists in the survey fielded after the experiment. However, this may be driven by selection into the survey. We find, unexpectedly, that the sample of surveyed students differed from the full

[2]Our motivation for restricting information provision to job opportunities in the first treatment is possible information overload. Processing information about both job opportunities and hourly wages for five occupations may be too demanding, rendering the treatment less effective.

[3]We elicit beliefs and provide information about the job opportunities in a qualitative manner, so inaccuracies may stem from students having beliefs that deviate from the truth, or not agreeing on the terms we assign.

sample in the experiment in that the former was less likely to change their favorite occupation for one with better prospects during the experiment than the latter. Using administrative data containing four years of student enrollment decisions, we do find that our information intervention impacts students' educational decisions. Students who receive positive news about the job opportunities of a profile (i.e., set of subjects) are more likely to choose this profile for their last two years of secondary education. Similarly, we provide evidence that treated students enroll (and remain) in study programs associated with occupations that provide better labor market prospects.

The identity of the sender of the information that is mentioned in the intervention – the labor market research institute or a researcher from this institute, either senior or junior, either female or male – appears inconsequential for the subsequent beliefs and preference ranking of occupations.

Our study contributes to a growing body of literature on the role of labor market expectations in education choices. As stated, previous studies have invariably found that students have highly noisy beliefs about the labor market returns of different study programs (Baker et al., 2018; Betts, 1996; Hastings et al., 2016). There are, however, some notable heterogeneities. Students who are more concerned with the labor market prospects of programs are less likely to overestimate these prospects (Hastings et al., 2016). The differences in concerns about these prospects are large between men and women (Wiswall and Zafar, 2017). Men tend to care more about pecuniary outcomes, whereas women care more about job security and flexibility. Similarly, we find in our data that male students select occupations with better job opportunities and higher hourly wages. However, they are also more likely to overestimate these and make larger absolute errors. A number of studies further document that students from low socioeconomic status backgrounds have less accurate expectations (Baker et al., 2018; Hastings et al., 2015, 2016). This could be explained by a lack of access to reliable information from, e.g., parents (Bleemer and Zafar, 2018; Lergetporer et al., 2018). We indeed confirm that students from higher socioeconomic status neighborhoods make smaller absolute errors and are more likely to be correct about the hourly wages of the occupations they select, but this does not hold for the job opportunities.

A number of field-experimental studies have tested the effects of interventions aimed at improving students' knowledge about the returns to – and costs of – education. Evidence from the Dominican Republic shows that providing students with information about the returns to attending secondary school increases enrollment (Jensen, 2010). For the general secondary education student population in industrialized countries, providing information about the returns to further education does not seem to influence actual enrollment (Pekkala Kerr et al., 2015; Bonilla-Mejía et al., 2019). There is some evidence that it does increase intended enrollment, particularly for students from low socioeconomic status backgrounds (Oreopoulos and Dunn, 2013; McGuigan et al., 2016; Peter and Zambre, 2017).

Most closely related to our paper are a number of studies that focus on providing information about the returns to specific study programs or institutions. These generally find some impact

on enrollment decisions, though the exact margin of impact differs. Hastings et al. (2015) show that, after being provided with such information, students are more likely to enroll in higher-return study programs. Similarly, Ballarino et al. (2022) show that informed students were *less* likely to enroll in fields of study with *weak* labor market prospects, though this seems to be driven by selection out of higher education for students who would otherwise enroll in these fields. Bonilla-Mejía et al. (2019) show no such impact, but find that students are more likely to enroll in prestigious institutions. Lastly, Conlon (2019) shows that students are more likely to enroll in a study program about which the student receives information, regardless of the content of this information. All-in-all, the results of information interventions targeting field of study are mixed.

Our study further draws on, and contributes to, recent work on role models. Porter and Serra (2020) show that female students are more likely to enroll in economics classes when they get to listen to a female role model talk about her experiences in university, as well as her career path and achievements. Moreover, Del Carpio and Guadalupe (2021) ran an experiment studying female enrollment in a 5-month software coding program. They show that removing a 'success story' of a female participant from the information page decreases enrollment by four percentage points. Our inclusion of the different 'information senders' provides a further look into how the characteristics of a person providing information affects the degree to which it is used.

The main contribution of our intervention is that, to the best of our knowledge, we are the first to present students with information on the labor market prospects of occupations rather than specific study programs. Our setting provides a unique opportunity to do so, as vocational education programs are strongly tied to occupations. Furthermore, the information we provide is highly detailed (based on 113 occupational groups), and tailored to the student's self-declared interests. Lastly, with the exception of Hastings et al. (2015), all field-experimental studies we know of required students to attend a presentation, take a survey, or visit a website they otherwise would not have. Our intervention is designed within an established career guidance program actually used as part of students' curriculum in school, and thus more closely matches the definition of a natural field experiment (Harrison and List, 2004). The setup of our experiment further allows us to provide a detailed picture of how students use the information and how it affects their beliefs. Using the time-stamped experimental data, we can show how much time students spend to process the information we provide them with, and whether this is sufficient for knowledge to persist in the long-term. Lastly, administrative data allows us to follow students for up to four years, and to estimate the impact of the treatment not only on initial decisions but to verify that these decisions are not reversed in subsequent years. Beyond the details of the intervention, we do not just treat students who are close to post-secondary education, but also those who still have to decide on their specialization in secondary education. This allows us to analyze what the impact of our information treatment is at different stages of students' educational careers. The intervention is low-cost, as it relies on readily available information, and easy to replicate in similar career guidance programs. Our partner company

is currently investigating the optimal way to implement occupational information provision at different levels of secondary education in its online career guidance platform.

The rest of this paper is structured as follows. Section 2 explains the institutional context: the Dutch education system and career guidance practice. Section 3 shows how we recruited schools and randomized them into treatment groups. Section 4 describes the experimental design. Section 5 lays out the data and Section 6 presents the results. Section 7 concludes.

2 Institutional Context

In this experiment, we focus on students enrolled in pre-vocational secondary education in the Netherlands. Pre-vocational secondary education is one of the three main tracks of Dutch secondary education[4]. As the name suggests, it is vocationally-oriented and offers a broad range of subjects. It is also the largest track in terms of student numbers: in the 2017/2018 school year, about 53% of Dutch children in secondary school attended pre-vocational secondary education (Dutch Inspectorate of Education, 2020).

The pre-vocational secondary education program takes four years to complete (Nuffic, 2019). At the end of the second year, students choose a 'learning pathway' (i.e., level of theoretical rigor). Pre-vocational secondary education is divided into four 'learning pathways': the basic vocational program, advanced vocational program, combined vocational-theoretical program, and theoretical program (Nuffic, 2019). In the theoretical program, students mostly take general subjects. The combined program drops one general subject in favor of four hours of vocational training, but is otherwise the same. In the basic and advanced vocational programs, students receive approximately 12 hours of vocational training instead of general subjects. General subjects are taught at a lower level compared to the combined and theoretical programs, with the level at the advanced vocational program being above that of the basic vocational program. Within the learning pathways, students also choose a 'profile', which determines the subjects they are taught (Government of the Netherlands, n.d.a).[5] Both the learning pathway and profile a student chooses have important consequences for the opportunities for further education at the time the student graduates, on which we expand below.

At the end of the fourth year, students decide how to continue their education. Notably, Dutch law dictates that students cannot leave education until they are either eighteen years of age or have a 'starting qualification' (i.e., an intermediate vocational education or senior general secondary education degree). The vast majority of students can therefore not leave education after graduating from their pre-vocational secondary education program. This leaves them with essentially two options: move on to post-secondary intermediate vocational education or

[4]Pre-vocational secondary education is known as 'vmbo' in Dutch. The two other tracks are higher general secondary education (havo) and pre-university education (vwo).

[5]For the basic vocational, advanced vocational, and mixed program there are ten available profiles: 1. Building, housing and interiors, 2. Engineering, fitting out and energy, 3. Transport and mobility, 4. Media, design and IT, 5. Maritime and technology, 6. Care and welfare, 7. Business and commerce, 8. Catering, baking and leisure, 9. Animals, plants and land and 10. Services and products. For the theoretical program, there are four options: 1. Care and welfare, 2. Engineering and technology, 3. Business and 4. Agriculture.

enroll in a different (sub)track of secondary education. Graduates from all learning pathways are eligible to enroll in intermediate vocational education. Programs in intermediate vocational education generally train students for a specific occupation. The exact level at which graduates can enroll depends on the chosen learning pathway. Graduates from the basic vocational program can enroll in qualification level 2 of intermediate vocational education only. Graduates from the other three programs can enroll in levels 2, 3 and 4 (Government of the Netherlands, n.d.b). In recent years, over 95% of students in the basic vocational, advanced vocational, and combined vocational-theoretical program directly enroll in intermediate vocational education. Of the remaining 5%, approximately half enrolls in intermediate vocational education at some later point in time. Among graduates from the theoretical program, about 80% directly enrolls in intermediate vocational education, with an additional 5% enrolling at a later point in time. The remaining 15% most commonly enrolls in *secondary general education*, a track that prepares students for higher vocational education (Dutch Ministry of Education, Culture and Science, n.d.). The occupations in the Qompas system focus on those attainable with an intermediate vocational education degree.

Schools are required by law to provide career guidance counseling to students. To structure their career guidance counseling efforts, schools often make use of programs offered by private companies. For this experiment, we partner with a company called Qompas, a widely-used career guidance counseling program developer in the Netherlands. As part of the program, students complete a number of assignments aimed at helping them learn more about themselves and the choices they will have to make. While students can access the program at any time, schools generally use Qompas during their career guidance counseling classes at set times during the school week. Usually, students individually go through the assignments in a classroom setting. All assignments the students complete are saved and stored in their personal file, which they are supposed to review periodically. We implement the experiment described in this paper within one of Qompas's assignments: the occupation assignment. While the Qompas system has a suggested order for doing the different assignments, schools decide in which year students actually do it. Schools usually let students do the occupation assignment in the second, third or fourth year of education. We expand on this in Section 4.

3 Recruitment and Randomization

Qompas recruited schools to participate in the experiment. At the time of recruitment, 300 schools for pre-vocational secondary education were registered as users in the Qompas system, which comprises about a third of all schools of this type in the Netherlands. Of these schools, thirteen were not eligible to participate in the experiment because of missing information. The 287 remaining schools were informed through a system message as well as an email that an experiment would take place. Qompas informed schools that they, together with a research institute of Maastricht University, were asked by the Ministry of Education, Culture, and Science to do research into the effects of labor market information on the choices of pre-vocational

secondary education students. Qompas further explained to schools that the research would be conducted by way of an experiment within the Qompas's career guidance counseling platform. Schools also received contact details of the person responsible for the experiment at Qompas in case they had any questions or complaints. Appendix A provides the original version as well as an English translation of the message. Only a single school indicated that it did not want to be a part of the experiment. This left us with 286 schools.

To randomize schools, we employed a stratified procedure at the school level. The reason for randomizing at the school level instead of at the student level is twofold. First, it reduces the risk of spillovers. Second, we expected that schools would be less willing to participate if some of their students were to be provided with information, whereas others were not.

We randomized schools into three main groups of approximately equal size: a control group, a treatment group that receives information about just job opportunities, and a treatment group that receives information about both job opportunities and hourly wages. The latter two groups were randomly assigned to receive information from either a research institute or a specific researcher from this institute. Columns 2 and 3 of Table 1 display the exact division of schools assigned to the different groups. We discuss the transition from assignment to actual participation in detail in Section 4.3

[Table 1 about here.]

We stratified schools on the basis of three characteristics: the number of broad profiles offered in the school, the number of students who completed the occupation test in the year before the experiment, and the quality of life indicator of neighborhoods the students come from. For the available profiles, we relied on data from Qompas. Qompas also registered the number of students who completed the occupation test in the previous year. However, data was not available for all schools. If no data was available, we predicted the number using the number of newly registered students in the Qompas system and the total number of students in the school itself[6]. If data on one of the two was not available, we predicted the number using just the available measure. For the quality of life in neighborhoods students came from, we relied on the quality of life indicator developed by the Dutch Ministry of the Interior and Kingdom Relations.[7] All neighborhoods (defined by their 4-digit postal code) in the Netherlands have a score, ranging from 1 (very low quality of life) to 9 (very high quality of life). For every school, we calculated the weighted average quality of life indicator score of the neighborhoods the school's student body came from.[8] If no data on the residential location of students was available, we predicted the average quality of life indicator score using the score of the school's neighborhood.

We used a block design to randomize. Because the profile choice is one of our outcome variables and largely determines the variety of occupations the students are likely to be inter-

[6]Data on the number of students in the school itself is provided as open data by the Dutch education executive agency; `https://duo.nl/open_onderwijsdata/databestanden/vo/leerlingen/leerlingen-vo-2.jsp`; Retrieved: 22-06-2018

[7]`https://data.overheid.nl/dataset/leefbaarometer-meting-2018`; Retrieved: 22-06-2018

[8]This information is available in the data set referred to in footnote 6.

ested in, we first sought balance on this dimension. We divided the schools into three groups: predetermined choice (only one theoretical profile available), limited choice (between one and three theoretical profiles available), and all four profiles available. Within these groups, we subsequently ranked schools based on the number of students who completed the occupation test last year. We split these groups into three more equal groups based on this dimension. As schools vary a lot in size, we hoped to improve balance in terms of sample size in this way. Lastly, within each of the now nine groups, we ranked schools on the basis of the weighted average of the quality of life indicator score. We then further split these groups into two. Increased balance on this dimension is important as we estimate heterogeneous effects based on the indicator. In the end, we were left with eighteen strata.

Within each stratum, schools were randomly assigned to the different treatment groups according to the division specified in Table 1. As not every stratum contained a perfect multitude of six schools, not all schools could be assigned in one go. We dealt with the unassigned schools by recreating strata as mentioned above, omitting the division in two based on the weighted average of the quality of life indicator score. Within each of the now nine strata, schools were again randomly assigned. For unassigned schools arising from this procedure, we repeated the procedure once more, now stratifying only based on the freedom of profile choice. The last ten remaining unassigned schools were sorted based on the freedom of profile choice and then assigned based on a randomly ordered list of the control and treatment groups. Figure B1 in Appendix B provides a visual representation of the procedures.

4 Experimental Design

In this section, we describe each stage of the experiment in detail in chronological order. The accompanying Online Appendix shows screen captures of the screens students in each of the treatment and control groups see in the experiment.

4.1 Occupation test

Our intervention is preceded by (and makes use of information collected in) an extensive occupation test. The test was designed by Qompas and had already been used for a number of years before our involvement. During the test, Qompas asks students to answer 90 questions about themselves and their attitudes towards a number of salient occupations (e.g., waiter/waitress, mason, mechanic). The aim of this test is to predict what sort of occupations the student might be interested in. Based on the answers, Qompas calculates a score for each of the 353 occupations in their system. This score represents how well the various occupations fit the student's preferences and abilities. Qompas subsequently uses the results of this test in a reflective assignment that students must do, which contains our intervention.

4.2 Elicitation of baseline information

During the reflective assignment, we collect baseline data on students' preferences and beliefs. We ask students about their intended profile choice, which second-year students still have to make at this point. They can pick multiple options in case they are not sure yet. We subsequently show students the twenty occupations that fit them best according to the test and ask them to select the five occupations they are most interested in. Students then receive information on the day-to-day activities in these occupations. After they read the information, we ask the students to rank the occupations in order of how much they would like to work in them later in life. Lastly, using a slider, we ask students to state their beliefs about the job opportunities and gross hourly wages of the five occupations they selected.[9] The options for job opportunities are "very poor", "poor", "reasonable", "good", and "very good". The options for the hourly wage range between €10.- and €26.-, with €1.- intervals.

During the first year of the experiment (the 2018/2019 school year), the sliders had a default option: "reasonable" for the job opportunities and €18.- for the hourly wages. Qompas removed this default option for the 2019/2020 school year. Moreover, in the 2018/2019 school year, students were able to alter their prior beliefs later on in the experiment by returning to them after receiving the information. Qompas corrected this error for the 2019/2020 school year. Because of these issues, we only consider the students who went through the experiment in the 2019/2020 school year whenever prior beliefs are relevant.

4.3 Information provision

After eliciting the baseline preference ranking and beliefs about the labor market prospects, we present students who have been assigned to the treatment information about the labor market prospects of the occupations they selected. Control group students do not receive any labor market information. For treatment groups 1 and 2, we provide information about the forecasted job opportunities. In treatment groups 3 and 4 we add information about the occupations' median hourly wage levels. Maastricht University's Research Center for Education and the Labor Market (ROA)[10] provided us with the information. As part of one of its research programs, ROA develops labor market forecasts for job opportunities of 113 different occupational groups in the next six years.[11] ROA also calculated the median hourly wage of intermediate vocational education graduates for these 113 occupations. To this end, they used data from the Dutch Labor Force survey, matched to administrative records. To construct the information set, we matched the Qompas occupations to these occupational groups.

[9] We ask for gross hourly wage because many youngsters in the Netherlands have a side job, e.g., in a supermarket, and are likely to have a good understanding of what they earn per hour with this job. The Dutch income tax system features a quite sizeable tax-free sum. Consequently, for most youngsters, gross earnings equal net earnings.

[10] www.roa.nl

[11] For information on methods, validity, and the governance of this project, see https://roa.maastrichtuniversity.nl/research/research-projects/project-onderwijs-arbeidsmarkt-poa. These forecasts are used by the national unemployment agency and for the accreditation of new study programs.

In treatments 1 and 3, we tell students that the information is sent by a researcher affiliated with ROA. This sender is randomized at the individual level within these treatments. The sender takes four different identities: inexperienced male researchers, experienced male researchers, inexperienced female researchers, and experienced female researchers.[12]. For each identity, one of four possible names is chosen. We show the name and experience on the screen.[13] We do not explicitly mention gender, but the names of all senders are indicative of their gender and the Dutch word for 'researcher' is different for men and women. We do not show pictures of the senders, so as to avoid bias caused by appearance unrelated to status or gender.

In treatment groups 2 and 4, we do not specify a human information sender. Instead, we tell students that the Research Center for Education and the Labor Market will provide them with the information. As we do not provide students in the control group with any information, we do not show them a sender either.

4.4 Video

Next, we show students in all groups a short video about work in general.[14] The video does not mention any particular occupations or the importance of job opportunities and wages. The main reason to show the video is to create some time between the first and second elicitation of beliefs for the control group. Without the video, students in the control group would be asked to state their beliefs a second time right after the first.

4.5 Elicitation of posterior beliefs and ranking

To estimate the initial effect of the treatment on beliefs and preferences, we elicit the students' ranking and beliefs a second time after the video. We show students their initial ranking and beliefs and ask them if they want to change anything.

4.6 Alternative occupations

37.7% of students select only occupations of which the job opportunities are forecasted to be "very bad", "bad" or "reasonable". We suggest to those students (in both the control and treatment groups) a set of alternative occupations with better labor market prospects and provide information on the day-to-day activities of these occupations. To treated students, we state that the labor market prospects for their chosen occupations are not very good, and that the proposed alternatives have better prospects. We do not tell control group students why we offer them alternatives. If students get to see the alternative occupations, they get the opportunity to include these occupations in their ranking. Initially, we place these alternative occupations at the bottom of the ranking in a randomized order.

[12] In Dutch: 'beginnend onderzoek(st)er' and 'ervaren onderzoek(st)er'.

[13] With their consent, we use the actual names of employees at the Research Center of Education and the Labor Market.

[14] https://www.youtube.com/watch?v=YJ78VDQrO3c

As just described, information about the labor market prospects of the alternative occupations was supposed to only be provided to students in the treatment groups. However, due to a programming error, control group students also received information about the job opportunities of the alternative occupations as well as about their initial set of occupations. Because of this error, we remove all students who were suggested alternatives from our post-intervention analyses. That also means we do not consider the alternative occupations in our analysis at all.

4.7 Elicitation of posterior intended profile choice

At the end of the experiment, we once again ask students what profile they intend to choose. We show them their initial selection and allow them to alter it.

5 Data

5.1 Sample

We collected data between September of 2018 and July of 2020, covering the 2018/2019 and 2019/2020 school years. 249 schools actually participated in the experiment, for a total of 40,176 individuals. At the other 37 schools, the part of the platform that included our experiment appears not to have been used by any student. As schools could not know their treatment assignment before going through the experiment, this forms no threat to our internal validity.

A small fraction of the individuals who went through the experiment were either first-year students (1,855) or school administrators involved in study guidance (48). We exclude them from the data, as specified in our pre-analysis plan. Of the remaining group of students, 1,082 did not make a first ranking of their selected occupations. As these students bring no data worth analyzing, we also exclude them. Qompas also records the creation and mutation date of each variable. 8,924 students changed their initial preference ranking on a different day than on the day they went through the experiment. This could be because these students went through the experiment multiple times, making the belief and ranking measures unreliable. We therefore remove these students from the sample as well. None of these sample restrictions are related to treatment status. After imposing our restrictions, we are left with 28,267 individuals from 243 schools. Columns (4) to (7) of Table 1 show how these numbers relate to the number of assigned schools. Table 2 shows that covariates are balanced between the control and treatment groups.

[Table 2 about here.]

5.2 Survey data

In addition to the experimental data collected through the career guidance counseling program, we conducted a survey among graduating students in the 2019/2020 school year. For this

purpose, we invited all students who took part in our experiment in their third year in the 2018/2019 school year and students who took part in their fourth year in the 2019/2020 school year. The survey was fielded between the 15^{th} of April and the 20^{th} of May, 2020, around the time at which they would need to finalize their enrollment decisions for next school year. To incentivize responses, we announced that we would raffle off 20 €25.- vouchers for a large Dutch e-tailer among survey respondents. The survey was sent to 9,510 students of which 1,061 responded, implying a response rate of 11%.

Again, we impose a number of sample restrictions. In our analysis, we only consider students who went through the experiment, did not change their prior ranking on a different day than they created it, and did not see the alternative occupations. After we impose our sample restrictions, we are left with 4,389 survey invitees, and 405 respondents. In the survey, we once again ask students to state their beliefs about the labor market prospects of the occupations they selected as well as to rank the occupations based on how much they would like to carry them out later in life. For the analysis of the survey respondents' beliefs and preferences, all above mentioned sample restrictions apply as well. In Table C1 in the Appendix, we present the results of a linear probability model of answering the survey depending on treatment status and a number of demographic characteristics. We show that the survey response rate is not affected by treatment status. We do observe that older students and male students are less likely to respond to the survey.

5.3 Administrative data

To study long-term outcomes of our intervention, we match our experimental data to administrative records at the Dutch Executive Education Agency. The Dutch Executive Education Agency is an agency of the government of the Netherlands responsible for all administrative and informational matters related to Dutch education. As such, they manage all registrations in official education programs; including for secondary and post-secondary education. This means that we are able to track students' educational status at any point in the education system. We matched students to their administrative records using their name and zip code from the Qompas system. To ensure the resulting matches from this method are accurate, we impose a number of additional sample restrictions. First, we remove 32 experimental observations for which we find a duplicate match. Second, we drop students for whom the grade they are registered in differs between the administrative and experimental data. Last, we impose our 'no mutation' and 'no alternative occupation' restrictions. Ultimately, we match 53% of the eligible students from the analysis sample. Table C2 in the Appendix shows that treatment status is unrelated to any observable characteristics for the matched sample.

Through the administrative data, we are able to observe students' educational status from October of the academic year they went through the experiment (i.e., 2018 or 2019) until 2022. This means we have four years of data for students who went through the experiment in the 2018/2019 academic year and three years for those who went through the experiment in the 2019/2020 academic year. We are interested in a number of outcomes. First, for the students

who went through the experiment in their second year of pre-vocational secondary education, we are interested in their profile choice the following year. We observe this for all such students in our administrative sample.

The second outcome we are interested in is their enrollment decision upon graduation. Our pre-registered outcome of interest is the likelihood of choosing a study program associated with one of the occupations in their choice set. Each occupation in the Qompas system is associated with a study program identifier. While useful, it turned out that these identifiers are in most cases deprecated. To match these identifiers to current program identifiers, we use a crosswalk provided by the Vocational Education and Industry Partnership.[15] This crosswalk provides a list of all current and deprecated program identifiers, allowing us to match the program identifiers in the Qompas system to the current program identifiers we obtain as part of the survey and administrative data. Unfortunately, not all matches are 1-to-1 meaning that a program identifier from the Qompas system can be matched to multiple current program identifiers and vice versa. We are able to match 467 of 496 program identifiers we observe in the administrative data to at least one program identifier in the Qompas system. A student is coded as having chosen a program associated with an occupation if the program identifier of the study program the student enrolled in had a crosswalk connection to the program identifier in the Qompas system. Note that we do not require this match to be unique. Students not enrolled in intermediate vocational education (e.g., because they went on to general secondary education) are coded to not have enrolled in a study program associated with any of the occupations in their choice set.

Another outcome of interest – though not explicitly pre-registered – are the job opportunities and hourly wages of the chosen study programs. One could imagine that our intervention did not only affect students' choices between the occupations in their choice set, but made them more broadly aware of the labor market consequences of their choices. To calculate the labor market prospects of each program, we take the frequency-weighted[16] average of the labor market prospects of all occupations associated with the current program identifier.

6 Results

6.1 Descriptive statistics

6.1.1 Selected occupations

Figure 1 shows the job opportunities and hourly wages of the occupations students selected for their top five before the intervention. Most selected occupations have job opportunities that are either poor (category 2), reasonable (3) or good (4). Hourly wages generally range between €12.- and €18.-. The Figure also shows that before the interventions there is no difference between the control and treatment groups in terms of job opportunities and hourly wages for

[15] `https://kwalificatie-mijn.s-bb.nl/Lijsten/Output/46604`; Retrieved: 25-05-2024.

[16] We base the weight on the number of choice sets the occupation was included in.

the occupations the students selected for their top five. Tables C3 and C4 in Appendix C confirm this. Although Table C4 does show that students in the first treatment group select occupations with about 1% lower hourly wages, the joint significance tests do not allow us to reject that the selection process between the treatment and control groups is the same. These Tables also show that there is no difference in the labor market prospects of the selected occupations between the first and second year of the experiment.

There are some interesting patterns in the selection of the occupations. Tables C5 and C6 in Appendix C show that male students generally select occupations with better job opportunities and higher hourly wages than female students do. This is in line with the finding of Wiswall and Zafar (2017) that male students care more about remuneration than female students. The latter finding is interesting in particular, since the literature (see e.g., Bleemer and Zafar, 2018; Lergetporer et al., 2021) shows that these students are generally less informed about earnings.

[Figure 1 about here.]

6.1.2 Prior beliefs

Figure 2 shows the prior belief accuracy of the control and treatment groups in the two years of the experiment. We denote the prior beliefs of individual i about the job opportunities of occupation j by $O_{i,j}^{Prior}$, and the actual job opportunities for that occupation by O_j^{Actual}. We apply the same logic to the hourly wages, which we denote as W. To measure belief accuracy, we first consider the difference between individual i's belief about the prospects of occupation j and its actual prospects: $O_{i,j}^{Prior} - O_j^{Actual}$ and $W_{i,j}^{Prior} - W_j^{Actual}$. These differences, which we report in Figure 2, allow us to analyze the degree of over- and underestimation of job opportunities and hourly wages. In the 2018/2019 school year, treated students show significantly more accurate expectations about the job opportunities and hourly wages than do control group students. This must be due to the fact that, due to a programming error, they could correct their initial beliefs, as discussed in Section 4.3. As stated there, we will not include students from the 2018/2019 school year in our analyses whenever prior beliefs are relevant. In the 2019/2020 school year, when the programming error was fixed, there is no difference between the beliefs of control and treatment group students, as shown in the panels on the right-hand side of Figure 2.

[Figure 2 about here.]

The figure shows left-skewed distributions, which indicates that students tend to overestimate the labor market prospects of their preferred occupations. This observation should be interpreted with caution, however. First, job opportunities are measured using ordinal categories. Differences between students' beliefs and the provided information could thus be caused by misperceptions about the true job opportunities, or differences in the interpretation of what sort of job opportunities are (very) poor, reasonable and (very) good. We have no way of disentangling the two. Second, the distribution of 'true' values is right-skewed. That is, occupations

with very poor and poor job opportunties and hourly wages between €10.- and €17.- are more common than occupations with good and very good job opportunities and hourly wages €19.- to €26.-. Since the midpoint of the slider was set at 'reasonable' and €18 for job opportunities and hourly wages respectively, this may have induced students to select nearby values, leading to excessive overestimations.

[Figure 3 about here.]

That said, Figure 3 does lend some credibility to the idea that students may be overestimating the labor market prospects of the occupations they are most interested in. The graph shows that the overestimation of both job opportunities and hourly wages is most pronounced for the initially highest-ranked occupation, and gradually decreases with the occupation's rank. There are two likely explanations for this phenomenon: (i) students may attempt to justify their occupational preferences by assigning better labor market prospects to those they have ranked higher, or (ii) students may be more interested in occupations with better labor market prospects, which would put those for which they draw higher beliefs at the top of their ranking.

When using central tendency measures, errors in beliefs that have opposite directions may cancel each other out. We therefore consider two additional metrics to assess the accuracy of students' beliefs and how these differ by a number of characteristics. First, we analyze the absolute values of the belief errors: $|O_{i,j}^{Prior} - O_j^{Actual}|$ and $|W_{i,j}^{Prior} - W_j^{Actual}|$. Second, we analyze how often beliefs are exactly correct (i.e., $O_{i,j}^{Prior} - O_j^{Actual} = 0$ and $W_{i,j}^{Prior} - W_j^{Actual} = 0$).

Table C7 in Appendix C presents regression results showing that male students tend to overestimate both job opportunities and hourly wages to a larger degree. They also make larger absolute errors and are less likely to be correct. Third and fourth year students, who are closer to making a decision than second year students don't do much better when it comes to the job opportunities, but make smaller absolute errors for the hourly wages. Students in schools where more profiles are available seem to make smaller absolute errors and are somewhat more likely to be correct about both job opportunities and hourly wages. As shown in Figure 3, higher ranked occupations are overestimated to a much larger degree – also when accounting for occupation fixed effects to avoid a mechanical correlation between overestimation and the occupation's labor market prospects. The difference between the number one and number five ranked occupation is almost an entire category for the job opportunities and €1.50 for the hourly wages.

6.1.3 Time spent processing

A natural question to ask prior to assessing the impact of the treatment is whether students in the treatment group actually paid attention to the provided information. A sensible proxy for this is the time it took students to go from the initial beliefs elicitation to the posterior ranking (i.e., the time spent on studying the information screens and watching the video about work). To process the information provided, treated students should take more time to get to the posterior rankings than control students, who see no information screens. Beyond seeing

whether treated students actually paid attention to the information we provided, this exercise is also useful to see how the treatments providing information about just the job opportunities differ from those where information about both the job opportunities and hourly wages are provided. If students spend equal amounts of time processing each piece of information[17], we expect the treatment groups that received information about both the job opportunities and hourly wages to spend up to twice as much time on processing the information as the groups that just received information about the job opportunities.

[Figure 4 about here.]

Figure 4 shows how the time spent differs between the groups. The control group takes about 70 seconds from stating their initial beliefs to making the posterior ranking. Most of this time is likely spent on watching the video about work. [18] Students in the job opportunities treatment take about 14 more seconds to get to the posterior ranking. Since the five pieces of information is the only difference between this group and the control group, this indicates they spend a little under 3 seconds on each piece of information. Those students who additionally receive information about the hourly wages spend a little over 24 seconds longer on the information than students in the control group. F-tests show that we can reject the null hypothesis that students in the job opportunities treatments spend as much time as those in the job opportunities and hourly wages treatment ($p<0.01$), indicating that students exert additional time to learn about the hourly wages. In fact, we cannot reject that students in the job opportunities and hourly wages treatment spend *twice as much* time as those in the job opportunities treatment. This indicates that students spend approximately an equal amount of time on each piece of individual information.

6.2 Short-term and medium-term treatment impact

6.2.1 Posterior beliefs

The analyses of time spent clearly show that students spend time to process the information we provide. That still leaves the possibility that they do not understand or believe the information. It is therefore useful to study how beliefs change immediately after receiving the information as a manipulation check. Figure 5 shows the posterior belief accuracy for the control group and relevant treatment groups. We denote the posterior beliefs of individual i about the job opportunities of occupation j by $O_{i,j}^{Post}$ and that of the hourly wages by $W_{i,j}^{Post}$. The graphs show that in both years, students in the treatment groups are much more likely to be correct about the job opportunities and the hourly wages of their selected occupations. This is largely driven by the correction of overestimations. Students who initially underestimated the labor market

[17] We define a piece of information to be an information about one characteristic of one occupation (e.g., the job opportunities of the second-ranked occupation).

[18] The video lasts 1 minute and 49 seconds, see: `https://www.youtube.com/watch?v=YJ78VDQrO3c`. Since students in the control group on average take 70 seconds to arrive at the posterior ranking, it follows that not all students watched the video completely.

prospects of their occupations react much less strongly than those who initially overestimated them. Tables C8 and C9 in Appendix C confirm this for the 2019/2020 cohort, where we can use students' prior beliefs in the analysis.

[Figure 5 about here.]

Table C10 shows that the treatment is equally effective when we state that a researcher provides the information, compared to an institute. Zooming in on the specific researcher, Table C11 shows that neither whether a male or a female researcher provides the information, nor whether the sender was an 'experienced' or 'beginning' researcher matters for the degree to which beliefs are updated.

Table C12 shows that when it comes to job opportunities, third and fourth year students react more strongly to the treatment than second year students. The same holds for female versus male students. Table C13 shows that treated fourth year students are more often correct than earlier year students, although the change is smaller than for the job opportunities and only marginally significant.

Next, we study how persistent the effects on posterior beliefs are. Table C14 shows that beliefs about the job opportunities remain more accurate at the time of the survey for students treated in the 2019/2020 school year (that is, up to seven months after treatment). This does not hold for those treated in the 2018/2019 school year (who completed the survey over a year after the treatment). However, we cannot ascribe the difference to time since treatment alone. The reason is that information on job opportunities and hourly wages may become more important as students get closer to their post-secondary education decision. As we survey graduating students, the students who received the information most recently were also much closer to the end of their secondary school career. As such, the reason these students better recall the information may be that they paid more attention to it, not that that they received it more recently. With our data, we cannot distinguish between these two mechanisms. For the hourly wages, we find that treated students do not have more accurate beliefs than the control group for both years of the experiment.

6.2.2 Rankings

Table 3 shows how the treatment affects the likelihood of students changing their favorite occupation between the first and second elicitation; a pre-registered outcome. We observe that students in the treatment group indeed change their favorite occupation significantly more often than those in the control group. The effect size is fairly small, however. In the control group, approximately 5.53% of students change their favorite occupation. In the treatment groups, this fraction is 0.88 to 2.16 percentage points higher.

The fact that students in the treatment group change their favorite occupation (slightly) more often does not tell the whole story, however. Table 3 also shows whether students in the treatment group switch towards occupations with better labor market prospects. ΔO_j^{Actual} and ΔW_j^{Actual}, respectively, denote the difference in the job opportunities and hourly wages

between the number one ranked occupation at first elicitation and the number one ranked occupation at second elicitation. If a student does not change his or her favorite occupation between the first and second elicitation, $\Delta O_j^{Actual} = \Delta W_j^{Actual} = 0$. Columns 2 and 4 show the effect unconditional on actually changing the number one ranked occupation. The job opportunities in the treatment groups rise by anywhere from 0.0190 to 0.0305 categories. For the wage treatments, the hourly wages rise by €0.09. Columns 3 and 5 show the change for students who did change their favorite occupation. For students in the treatment groups, the job opportunities move up by 0.285 to 0.447 categories and hourly wages by €1.12 to €1.20. It is important to note that in both cases, the job opportunities and hourly wages do not move at all for control group students.

[Table 3 about here.]

Table C15 in Appendix C shows that treated students in schools with four profiles available are not more likely to change their favorite occupation compared to schools with fewer profiles available, but do switch to occupations with better job opportunities when they do. This may be driven by the fact that these students have a larger set of options to choose from. This Table, together with Table C16 shows that we find no further evidence for heterogeneous treatment effects. Table C17 shows there is no effect of the identity of the information sender either.

We do not find evidence that treated students still prefer occupations with better prospects in the survey. However, Columns (1) and (3) of Table C18 show that the treated students in the survey did not switch to occupations with better prospects directly after the intervention either. Hence, the lack of an effect at time of the survey is likely due to selection into survey participation, a problem we do not face when using the administrative data on actual education choices, to which we turn now.

6.3 Impact on educational decisions

We analyze the treatment impact on the profile choice in secondary school, and post-graduation outcomes. As we have not found any evidence that the identity of the sender impacts the response to the treatment, we collapse the four treatments into two:

1. Job Opportunities Information Treatment
2. Wage and Job Opportunities Information Treatment

This makes the interpretation of the results easier, and increases power of the individual estimates as the sample size for each collapsed treatment is larger.

6.3.1 Profile choice

The outcome of interest for second-year students is whether we observe the student as enrolled in profile P in the next school year. We estimate a conditional logit model McFadden (1974), which models the probability of choosing profile P as a function of its attributes. In our

case, those attributes are the (i) news and (ii) average information provided about the job opportunities and hourly wages for the control and treatment groups. The advantage of using the conditional logit model is that it is essentially a fixed effects model, taking care of any individual-level variation and increasing precision. For this analysis, we restrict the sample to students that we first observed in their second year (i.e., before they made their profile choice).

We define the information that students received about each profile in two different ways. First, we calculate the 'news' about each profile P as $\mathcal{O}_P = \sum_{j \in P} \left(O_j^{\text{Actual}} - O_{i,j}^{\text{Prior}}\right)$ and $\mathcal{W}_P = \sum_{j \in P} \left(W_j^{\text{Actual}} - W_{i,j}^{\text{Prior}}\right)$. These values increase for every occupation j belonging to profile P for which the student received good news (i.e., they underestimated the occupation's labor market prospects). This variable therefore contains information on the amount of news the student received, as well as the quality of that news.[19] The main advantage of this operationalization is that we can set this value to 0 for profiles not associated with any occupations the student selected. This means it is defined for all possible profiles the student can choose. Using $\mathcal{O}_P$ and $\mathcal{W}_P$ also comes with downsides. While this method captures the 'surprise' element of the treatment, measurement error in the prior beliefs of the labor market prospects of the occupations may lead to attenuation bias. Additionally, we have to restrict the sample to the 2019/2020 treatment year, as prior beliefs are essential for this analysis and a programming error led to bias in our measure of prior beliefs in the 2018/2019 school year, as discussed before. As a robustness check, we run another analysis in which we use the average information received about the job opportunities and hourly wages. Specifically, we calculate $O_P = \frac{\sum_{j \in P}\left(O_j^{\text{Actual}}\right)}{N_P}$ and $W_P = \frac{\sum_{j \in P}\left(W_j^{\text{Actual}}\right)}{N_P}$. These variables take care of the downsides that come with the news values, but do have their own drawbacks. Beyond the fact that it does not take prior beliefs into account, it is also problematic that these covariates are only defined for profiles associated with occupations that the student picked. Together, we believe they provide a comprehevensive analysis of the profile choice.

[Table 4 about here.]

Table 4 shows the results of this analysis, with the coefficients expressed as odds ratios. The first three columns show the impact of the news, and the last three columns show the impact of the average presented information. Within each set of columns, the first column shows the overall impact. The second column shows the impact on students on the theoretical pathway (who can choose among four profiles), and the third column shows the impact for those on the practical pathways (who can choose among ten profiles). To increase precision of our estimates, we include a dummy to indicate whether the profile was part of the student's intended profile choice and a set of dummies for the number of occupations the student selected associated with the profile.

Column 1 shows that for students who received information about just the job opportunities compared to the control group, the odds of choosing profile P increase by 0.09 if the

[19] This aligns with our preregistration, in which we state: "*[we] expect that the more and the better the news one receives about occupations associated with a particular study profile compared to the news about occupations associated with the prior intended profile choice, the more likely it is that the profile is chosen*".

job opportunities were 'underestimated' by one category. Looking at columns 2 and 3, the magnitude of the coefficients are not substantially different between the two groups, although only column 3 shows (marginal) significance. The coefficient for students in the wage and job opportunities treatment is considerably smaller and insignificant. Since Sections 6.2.1 and 6.2.2 do not show strong differential impacts between the two treatments, we do not think one should put too much weight on the differences between the two treatments here. Reassuringly, the χ^2 test shows that the impact of the news about the job opportunities is jointly significant among the two treatments. The negative coefficient on the job opportunity news for the control group is worth discussing as well. A higher value means that the job opportunities in profile P were initially underestimated to a larger degree, which is associated with decreased odds of choosing the profile. The treatment does not fully make up for this initial overestimation but as mentioned in Section 6.2.1, beliefs may be correlated with the degree to which students like occupations and their associated profiles. We do not find any evidence that wage news impacts the profile choice. Looking at columns 2 and 3, the magnitude of the coefficients are not substantially different between the two groups, although only column 3 shows (marginal) significance.

When we move to the average information provided in Columns 4 to 6, the sample size decreases despite including students from the 2018/2019 school year. This is because the average information provided is only defined for profiles for which the student has selected occupations associated with it, and all other profiles available to the student therefore drop out of the analysis. Moreover, if a student selected five occupations that are all associated with the same profile, they fall out of the sample as well as covariates need to be defined for at least two profiles. The results in these columns are mixed. While treated students select profiles for which they observe higher average job opportunities, the opposite is true for the hourly wages. Notably, the positive impact of a one standard deviation increase in the average observed job opportunities is almost exactly negated by a one standard deviation increase in the average observed hourly wages. These contrasting findings are hard to rationalize. All-in-all, we conclude that the treatment had at most a moderate impact on the profile choice of students through news about the job opportunities.

6.3.2 Post-graduation outcomes

The administrative records allow us to follow students for up to four years from the year in which they took part in our experiment. The first set of outcomes we are interested in is whether students make different educational decisions at the extensive margin upon graduation (i.e., are they more or less likely to enrol in intermediate vocational education or general secondary education). We also study whether treated students perform differently in secondary school as compared to control students. Table 5 shows that this is not the case, conditional on a set of baseline covariates. There is no significant impact of the treatment on ever being retained (Column 1), starting general secondary education (Column 2), or starting post-secondary vocational education (Column 3). However, there are some notable differences in who starts

post-secondary vocational education, which is relevant for the sample of students for whom we observe the program choice. Students who took part in 2019/2020, men, and students in schools with more profiles available are less likely to enroll in post-secondary vocational education. Students from richer neighborhoods, those who took part in the experiment in a later grade, and those on the practical pathway are more likely to do so. The latter in particular is relevant: approximately 80% of students on the theoretical pathway enroll in post-secondary vocational education, whereas 95% of those on the practical pathway do so; in line with the national average as described in Section 2. With these sample considerations in mind, Table 6 provides the answer to the most policy-relevant question of the paper: do students who receive information enroll in programs associated with occupations that provide better labor market prospects?

[Table 5 about here.]

We consider two programs in this analysis: (i) the first programs that students enrolled in [first choice], and (ii) the last program in which we observe the student enrolled [final choice]. Including both provides us with a sense of the impact of the treatment on the initial choice, but ensures we are not simply picking up that students initially pick programs with better prospects but then leave them for the programs they initially intended to enroll in anyway. Beyond the covariates listed in the regression table, all regressions contain fixed effects for each of the occupations the students selected to be in their top five to account for initial preferences.

[Table 6 about here.]

The results in Column 1 of Table 6 show that treated students initially enroll in programs associated with occupations that provide better job opportunities; around 0.05 to 0.06 categories on average, or a 2% increase over the control group mean. The impact of the two treatment types is jointly significant at the 10%-level. The same holds true for the hourly wages of the first choice in Column 2, which increase by €0.13.- to €0.19.-; a 1% increase of the control group mean. It may seem surprising that the coefficient on the job opportunities treatment is positive as well but the correlation between job opportunities and hourly wages is around 0.5. It is therefore not surprising that those who receive information about the job opportunities and respond to it also choose programs with higher hourly wages. Columns 3 and 4 show even stronger results than Columns 1 and 2, indicating that the treatment's impact did not exclusively impact the initial choice upon graduation but have a persistent impact. The two treatments are jointly significant at the 10-% and 5-% level in Columns 3 and 4, respectively.

It is worth comparing the treatment effect in Table 6 to that on the immediate change in job opportunities and hourly wages of the number one ranked occupation in Table 3. One might expect that, similar to the treatment impact on belief accuracy, the immediate impact would be large but decrease over time. This is not what we find. In fact, the results in Table 6 show a much larger impact of the treatment on the job opportunities and hourly wages of the chosen study programs than on the number one ranked occupation. The most likely explanation for this

is that students were restricted in their ranking by their initial choice set. These restrictions do not apply for their actual program choice[20]. The results in Table 6 indicate that the treatment does more than just help students decide between occupations in their initial set of interest. The results suggest that, in response to the treatment, students gathered information about labor market prospects on a broader range of occupations and, consequently, made choices associated with better prospects.

While we consider the results in Table 6 to be the most policy relevant, they do not speak to the mechanism through which the treatment operates. In our pre-analysis plan, we describe a likely mechanism: that students are more likely to pick a study program associated with an occupation about which they receive good news.[21] To this end, we estimate a conditional logit model. The outcome of interest is whether the student selected a study program associated with occupation j, and the news and information provided about the occupation, interacted with the treatment are the main covariates of interest. Note that this restricts our analysis to the 25% of students who picked a program associated with one of their initial occupations of interest. Table 7 shows the results of this analysis.

[Table 7 about here.]

Column 1 shows the impact of news about the job opportunities and wages on the first choice upon enrolling in post-secondary vocational education. While all odds ratios of the interactions between the treatment and news values are above 1 (indicating a positive impact on the probability of choosing a study program associated with the occupation), none are significantly different from zero. Joint significance tests of both the impact of job opportunities news in both treatments, and the impact of job opportunities and wage news in the job opportunities and wage information treatment fail to reject the null hypothesis that there is no impact. The impact on the final choice in Column 3 looks similar. Column 2, which considers only the information – which allows us to include more individuals in the sample, and avoids measurement error related to inaccurate belief elicitation – shows larger impacts. We observe that information about the job opportunities increases the odds of picking a program associated with the occupation by almost 0.1, and an impact of about 0.03 for the wage information in the treatment that receives information about it. In this case, tests of joint significance of the two pieces of information in the job opportunities and wage information treatment do allow us to reject the null hypothesis that information in this treatment does not have any impact. Similar to before, Column 4 looks similar, although the impact of the wage information appears to be more muted.

All-in-all, Tables 6 and 7 show that the treament has an impact on the labor market prospects of programs the students choose. Some of this impact looks to be driven by switches

[20]In fact, we observe that only 25% of students end up picking a program associated with an occupation that appeared in their initial five, of which 55% pick a program associated with their initial favorite.

[21]Specifically, we state "*[we] expect that the better the news about an occupation's prospects compared to the news about the other occupations, the more likely it is that the student decides to pursue the study program associated with this occupation.*".

within the initial choice set, but it is likely that the treatment also made students more broadly aware about labor market prospects.

7 Conclusion

In this paper, we presented a field experiment aimed at improving the accuracy of Dutch pre-vocational education students' beliefs about the job opportunities and hourly wages of occupations they are interested in. In line with the literature, we find that students' prior beliefs are highly inaccurate. In our sample, both job opportunities and hourly wages are strongly overestimated, particularly for students' favorite occupations. This could be innocuous, and simply the results of students rationalizing their choices. However, since our results indicate that students do indeed attach some value to the labor market prospects of occupations when making educational decisions, another explanation is likely. If students gather noisy information and tend to gravitate towards the occupations for which they learn the labor market prospects are best, these will often be the occupations for which the information was least accurate in a winner's curse fashion. This underlines the importance of providing students with accurate information.

Our results show that providing such information is effective in correcting belief errors in the short term. However, survey data shows that these beliefs stick for at most a couple of months and only for the job opportunities. Students who receive information are more likely to change their favorite occupation between the first and second elicitation of the ranking and, if they do so, switch towards occupations with better labor market prospects. We are unable to confirm whether this change in preferences holds in the long term, however. Despite belief accuracy decaying and a lack of evidence about persistent changes in preferences, our treatment does affect educational decisions. News about the job opportunities affects students' choices for secondary school specializations and students who receive labor market information enroll in programs associated with occupations that provide better labor market prospects.

References

Altonji, Joseph G, Erica Blom, and Costas Meghir, "Heterogeneity in human capital investments: High school curriculum, college major, and careers," *Annu. Rev. Econ.*, 2012, *4* (1), 185–223.

Arcidiacono, Peter, V Joseph Hotz, and Songman Kang, "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 2012, *166* (1), 3–16.

Baker, Rachel, Eric Bettinger, Brian Jacob, and Ioana Marinescu, "The effect of labor market information on community college students' major choice," *Economics of Education Review*, 2018, *65*, 18–30.

Ballarino, Gabriele, Antonio Filippin, Giovanni Abbiati, Gianluca Argentin, Carlo Barone, and Antonio Schizzerotto, "The effects of an information campaign beyond university enrolment: A large-scale field experiment on the choices of high school students," *Economics of Education Review*, 2022, *91*, 102308.

Betts, Julian R, "What do students know about wages? Evidence from a survey of undergraduates," *Journal of human resources*, 1996, pp. 27–56.

Bleemer, Zachary and Aashish Mehta, "Will studying economics make you rich? A regression discontinuity analysis of the returns to college major," *American Economic Journal: Applied Economics*, 2022, *14* (2), 1–22.

_ **and Basit Zafar**, "Intended college attendance: Evidence from an experiment on college returns and costs," *Journal of Public Economics*, 2018, *157*, 184–211.

Bonilla-Mejía, Leonardo, Nicolas L Bottan, and Andrés Ham, "Information policies and higher education choices experimental evidence from Colombia," *Journal of Behavioral and Experimental Economics*, 2019, *83*, 101468.

Carpio, Lucia Del and Maria Guadalupe, "More women in tech? Evidence from a field experiment addressing social identity," *Management Science*, 2021.

Conlon, John J, "Major Malfunction: A Field Experiment Correcting Undergraduates' Beliefs about Salaries," *Journal of Human Resources*, 2019, pp. 922–939.

Dutch Inspectorate of Education, "De staat van het onderwijs 2020," 2020. Retrieved from `https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2020/04/22/staat-van-het-onderwijs-2020/Staat+van+het+Onderwijs+2020.pdf` on 2024-05-22.

Dutch Ministry of Education, Culture and Science, "Doorstroom vanuit vmbo naar het mbo," n.d. Retrieved from `https://www.ocwincijfers.nl/begroten-en-verantwoorden/sectoren/studiefinanciering/iedereen-gelijke-kansen/doorstroom-vanuit-vmbo-naar-het-mbo` on 2024-11-17.

Government of the Netherlands, "Pre-vocational secondary education (VMBO)," n.d. Retrieved from `https://www.government.nl/topics/secondary-education/pre-vocational-secondary-education-vmbo` on 2019-08-06.

_ , "Wat zijn de vooropleidingseisen voor het middelbaar beroepsonderwijs (mbo)?," n.d. Retrieved from https://www.rijksoverheid.nl/onderwerpen/middelbaar-beroepsonderwijs/vraag-en-antwoord/wat-zijn-de-vooropleidingseisen-voor-het-middelbaar-beroepsonderwijs-mbo on 2019-08-11.

Harrison, Glenn W and John A List, "Field experiments," *Journal of Economic literature*, 2004, *42* (4), 1009–1055.

Hastings, Justine, Christopher A. Neilson, and Seth D. Zimmerman, "The Effects of Earnings Disclosure on College Enrollment Decisions," Working Paper 21300, National Bureau of Economic Researc 2015.

_ , Christopher A Neilson, Anely Ramirez, and Seth D Zimmerman, "(Un) informed college and major choice: Evidence from linked survey and administrative data," *Economics of Education Review*, 2016, *51*, 136–151.

Jensen, Robert, "The (perceived) returns to education and the demand for schooling," *The Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Kerr, Sari Pekkala, Tuomas Pekkarinen, Matti Sarvimäki, and Roope Uusitalo, "Post-Secondary Education and Information on Labor Market Prospects: A Randomized Field Experiment," IZA Discussion Papers 9372, Institute for the Study of Labor (IZA) 2015.

Ketel, Nadine, Edwin Leuven, Hessel Oosterbeek, and Bas van der Klaauw, "The returns to medical school: Evidence from admission lotteries," *American Economic Journal: Applied Economics*, 2016, *8* (2), 225–54.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad, "Field of study, earnings, and self-selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

Lergetporer, Philipp, Katharina Werner, and Ludger Woessmann, "Does Ignorance of Economic Returns and Costs Explain the Educational Aspiration Gap? Evidence from Representative Survey Experiments," IZA Discussion Papers 11453, Institute for the Study of Labor (IZA) 2018.

_ , _ , and _ , "Does ignorance of economic returns and costs explain the educational aspiration gap? Representative evidence from adults and adolescents," *Economica*, 2021, *88* (351), 624–670.

McFadden, D., "Conditional logit analysis of qualitative choice behavior," Technical Report 105-142, Frontiers in Econometric, New York 1974.

McGuigan, Martin, Sandra McNally, and Gill Wyness, "Student awareness of costs and benefits of educational decisions: effects of an information campaign," *Journal of Human Capital*, 2016, *10* (4), 482–519.

Nuffic, "The Education System of the Netherlands," 2019. Retrieved from https://www.nuffic.nl/documents/459/education-system-the-netherlands.pdf on 2019-08-04.

Oreopoulos, Philip and Ryan Dunn, "Information and college access: Evidence from a randomized field experiment," *The Scandinavian Journal of Economics*, 2013, *115* (1), 3–26.

Peter, Frauke H and Vaishali Zambre, "Intended college enrollment and educational inequality: Do students lack information?," *Economics of Education Review*, 2017, *60*, 125–141.

Porter, Catherine and Danila Serra, "Gender differences in the choice of major: The importance of female role models," *American Economic Journal: Applied Economics*, 2020, *12* (3), 226–54.

Wiswall, Matthew and Basit Zafar, "Preference for the workplace, investment in human capital, and gender," *The Quarterly Journal of Economics*, 2017, *133* (1), 457–507.

Tables

Table 1: Treatment assignment, participation and analysis sample overview

Treatment Group	**Frac. of Schools**	**Assigned Schools**	**Participating Schools**	**Participating Students**	**Schools in Analysis**	**Students in Analysis**
Control Group	1/3	96	83	12,544	81	9,275
Job Opp. Info by Researcher (Treatment 1)	1/6	47	42	6,917	42	5,117
Job Opp. Info by Research Institute (Treatment 2)	1/6	47	40	6,470	40	5,151
Job Opp. & Wage Info by Researcher (Treatment 3)	1/6	48	38	5,580	38	4,254
Job Opp. & Wage Info by Research Institute (Treatment 4)	1/6	48	43	5,680	42	4,470
Total	1	286	246	37,191	243	28,267

Table 2: Balance table

	Control		Treatment 1		Treatment 2		Treatment 3		Treatment 4		
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	P-value joint sign.
Nmbr of Students	164.14	134.24	175.55	147.33	178.00	121.59	158.21	118.49	150.10	110.95	0.82
Nmbr of Profiles Available	3.37	0.99	2.93	1.30	3.30	1.14	3.21	1.07	3.14	1.20	0.38
Age	14.04	1.01	14.02	0.99	14.16	0.99	14.06	0.99	14.05	1.04	0.82
Male	0.52	0.50	0.53	0.50	0.56	0.50	0.52	0.50	0.50	0.50	0.43
Grade	2.47	0.64	2.45	0.64	2.56	0.68	2.43	0.61	2.48	0.67	0.84
QOL Score	6.57	1.35	6.55	1.31	6.77	1.23	6.65	1.42	6.46	1.46	0.59

Note: No. of students and no. of profiles available are school level variables. Age, male, grade and QOL score are individual level variables. QOL score refers to the quality of life indicator score of the neighborhood the student lives in (see Section 3 in the paper). The last column of the Table indicates whether a joint significance tests shows a significant difference between the treatment groups and the control group for each of the variables considered.

Table 3: Treatment effect on likelihood changing favorite occupation and change in prospects

	(1) Pr(Fav. Change)	(2) ΔO_j^{Actual}	(3) ΔO_j^{Actual} (Changed)	(4) ΔW_j^{Actual}	(5) ΔW_j^{Actual} (Changed)
Sender: researcher					
Info: job opportunities	0.00760*	0.0190***	0.295***	0.0130	0.192
	(0.00454)	(0.00622)	(0.101)	(0.0174)	(0.292)
Sender: institute					
Info: job opportunities	0.0120**	0.0305***	0.447***	0.0303**	0.430*
	(0.00534)	(0.00650)	(0.102)	(0.0149)	(0.253)
Sender: researcher					
Info: job opp. & wages	0.0201***	0.0278***	0.358***	0.0876***	1.115***
	(0.00547)	(0.00577)	(0.0880)	(0.0215)	(0.295)
Sender: institute					
Info: job opp. & wages	0.0180***	0.0214***	0.285***	0.0904***	1.197***
	(0.00638)	(0.00640)	(0.0944)	(0.0199)	(0.278)
Constant	0.0565***	0.000666	0.0120	0.00466	0.0843
	(0.00318)	(0.00325)	(0.0590)	(0.0102)	(0.186)
Observations	28267	27387	1791	27387	1791

Note: Constant refers to the control group estimate. Each row above refers to the incremental estimate for each of the four treatment groups. Pr(Fav. Change) in Column (1) is the chance that a student changed his or her favorite occupation between second elicitation. ΔO_j^{Acutal} in Column (2) denotes the difference between the job opportunities of the student's favorite occupation at the second elicitation and the first elicitation. It is equal to 0 if the student did not change. ΔW_j^{Acutal} in Column (4) denotes the equivalent for the hourly wages. Columns (3) and (5) only contain observations where the student did switch favorite occupations between first and second elicitation. Standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual level.

Table 4: Treatment impact on profile choice

	(1) All students | News	(2) Theoretical students | News	(3) Practical students | News	(4) All students | Info	(5) Theoretical students | Info	(6) Practical students | Info
Job opp. news about profile	0.917***	0.903***	0.977			
	(0.0242)	(0.0326)	(0.0242)			
Wage news about profile	1.004	1.011	0.992			
	(0.00719)	(0.00897)	(0.0102)			
Job Opp. Info Treatment × Job opp. news about profile	1.093**	1.071	1.070*			
	(0.0412)	(0.0572)	(0.0387)			
Wage & Job Opp. Info Treatment × Job opp. news about profile	1.015	0.981	0.992			
	(0.0350)	(0.0496)	(0.0380)			
Job Opp. Info Treatment × Wage news about profile	0.999	0.992	1.012			
	(0.00998)	(0.0134)	(0.0136)			
Wage & Job Opp. Info Treatment × Wage news about profile	0.991	0.996	0.995			
	(0.0101)	(0.0130)	(0.0128)			
Average observed job opportunities of profile				0.740***	0.618***	0.923
				(0.0788)	(0.0939)	(0.120)
Average observed wages of profile				1.030	0.981	1.082
				(0.0275)	(0.0286)	(0.0545)
Job Opp. Info Treatment × Average observed job opportunities of profile				1.269*	1.161	1.312*
				(0.179)	(0.262)	(0.204)
Wage & Job Opp. Info Treatment × Average observed job opportunities of profile				1.272*	1.084	1.252
				(0.178)	(0.255)	(0.194)
Job Opp. Info Treatment × Average observed wages of profile				0.913*	1.041	0.801***
				(0.0444)	(0.0502)	(0.0631)
Wage & Job Opp. Info Treatment × Average observed wages of profile				0.924**	0.977	0.893*
				(0.0359)	(0.0493)	(0.0593)
Intended prior choice	7.461***	3.906***	10.25***	4.685***	2.556***	10.61***
	(0.909)	(0.842)	(1.398)	(0.565)	(0.345)	(1.478)
Occupation count	Yes	Yes	Yes	Yes	Yes	Yes
Observations	28871	12428	16443	7768	3238	4530
Chi-Sq Joint Sign. of Treatments × Job Opp.	6.582	3.137	4.579	3.541	0.438	3.154
P-value Joint Sign. of Treatments × Job Opp.	0.037	0.208	0.101	0.170	0.803	0.207
Chi-Sq Joint Sign. Wage & Job Opp. Treatment × Job Opp. and Wages	1.014	0.915	0.466	5.018	0.224	3.205
P-value Joint Sign. Wage & Job Opp. Treatment × Job Opp. and Wages	0.602	0.633	0.792	0.081	0.894	0.201

Note: Table shows odds ratios based on a conditional logit model with whether profile P was chosen as the outcome variable. *Job opp. news* and *Wage news* (Columns (1) to (3)) are defined as $\mathcal{O}_P = \sum_{j \in P} \left(O_j^{\text{Actual}} - O_{i,j}^{\text{Prior}} \right)$ and $\mathcal{W}_P = \sum_{j \in P} \left(W_j^{\text{Actual}} - W_{i,j}^{\text{Prior}} \right)$, respectively. Similarly, the *Average observed job opportunities* and *Average observed wages* (Columns (4) to (6)) are defined as $O_P = \frac{\sum_{j \in P} \left(O_j^{\text{Actual}} \right)}{N_P}$ and $W_P = \frac{\sum_{j \in P} \left(W_j^{\text{Actual}} \right)}{N_P}$. $j \in P$ denotes an occupation associated with the profile. N_P denotes the number of occupations the student picked that were associated with profile P. *Intended prior choice* is a dummy that indicates whether a student intended to choose the profile pre-treatment. Specifications include controls for the amount of information the student received about the profile (i.e., how many occupations associated with the profile were part of their choice set). The sample is restricted to second-year students only. For columns (1) to (3), it is restricted to those who took part in the experiment in the 2019/2020 school year. Columns (1) and (4) show results for all students, regardless of pathway. Columns (2) and (5) show results for students in the theoretical pathway and Columns (3) and (6) show results for students in the practical pathway. The χ^2-tests test the joint significance of the two treatments interacted with the job opportunities (i.e., did the job opportunities make treated students more or less likely to pick a profile) and the joint significance of the job opporunities and the wages for the treatment that received both. Standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 5: Treatment impact on early educational decisions

	(1) Ever retained	(2) Start gen. sec. ed.	(3) Start post-sec. voc. ed.
Job Opp. Info Treatment	-0.00560	-0.0142	0.00356
	(0.00897)	(0.0143)	(0.0154)
Wage & Job Opp. Info Treatment	0.00422	-0.00436	-0.00424
	(0.00941)	(0.0140)	(0.0140)
Took part in 2019/2020	0.00628	-0.0136*	-0.0435***
	(0.00544)	(0.00723)	(0.00808)
Male	0.0179***	0.00318	-0.0102
	(0.00487)	(0.00590)	(0.00618)
QOL Indicator	-0.0113***	-0.00399	0.00978***
	(0.00235)	(0.00323)	(0.00293)
Limited profile choice	0.0175*	0.0357**	-0.0401**
	(0.0102)	(0.0150)	(0.0173)
Unknown profile choice	0.0242***	0.0397***	-0.0448***
	(0.00764)	(0.0122)	(0.0147)
3rd year	-0.0291***	0.00940	0.0545***
	(0.00670)	(0.0116)	(0.0114)
4th year	-0.0623***	-0.0226*	0.0817***
	(0.00904)	(0.0124)	(0.0128)
Practical pathway	-0.0374***	-0.133***	0.145***
	(0.00605)	(0.00909)	(0.0101)
Constant	0.137***	0.154***	0.797***
	(0.0193)	(0.0282)	(0.0271)
Observations	11935	11935	11935
F-Stat Joint Sign. of Treatments	0.626	0.549	0.152
P-value Joint Sign. Treatments	0.536	0.578	0.859

Note: Tables shows results from ordinary least squares regression. Column (1) shows impact on grade repetition in secondary school. Column (2) shows the impact on the likelihood of a student moving on to general secondary education–a track preparing them for higher vocational education. Column (3) shows the impact on the likelihood of starting post-secondary vocational education. The F-tests and P-values reported at the bottom of the table are for the joint significance of the two treatments. Standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 6: Treatment impact on job opportunities and wages of chosen programs

	(1) Job opp. first choice	(2) Hourly wage first choice	(3) Job opp. final choice	(4) Hourly wage final choice
Job Opp. Info Treatment	0.0618**	0.190**	0.0574*	0.200**
	(0.0307)	(0.0851)	(0.0302)	(0.0780)
Wage & Job Opp. Info Treatment	0.0519*	0.131	0.0554**	0.120
	(0.0280)	(0.101)	(0.0273)	(0.102)
Took part in 2019/2020	-0.0106	-0.0773	0.00381	-0.0448
	(0.0223)	(0.0673)	(0.0222)	(0.0686)
Male	-0.0523	0.543***	-0.0431	0.541***
	(0.0356)	(0.0861)	(0.0329)	(0.0815)
QOL Indicator	0.00665	-0.0539**	0.00622	-0.0242
	(0.00727)	(0.0247)	(0.00668)	(0.0241)
Limited profile choice	0.0621*	0.224*	0.0459	0.216*
	(0.0364)	(0.120)	(0.0364)	(0.114)
Unknown profile choice	0.0385	0.0880	0.0339	0.0214
	(0.0353)	(0.105)	(0.0335)	(0.100)
3rd year	0.0213	0.221***	-0.00585	0.143*
	(0.0260)	(0.0783)	(0.0230)	(0.0738)
4th year	0.0616	0.191**	0.0406	0.0949
	(0.0413)	(0.0822)	(0.0396)	(0.0926)
Practical pathway	-0.130***	-0.805***	-0.129***	-0.756***
	(0.0234)	(0.0862)	(0.0239)	(0.0861)
Selected occupations fixed effects	Yes	Yes	Yes	Yes
Observations	10032	10032	10026	10026
F-Stat Joint Sign. of Treatments	2.521	2.529	2.711	3.304
P-value Joint Sign. Treatments	0.083	0.082	0.069	0.039

Note: Tables shows results from ordinary least squares regression. Columns (1) and (3) show impact on the job opportunities of the chosen study program. Columns (2) and (4) show the impact on the hourly wage of the chosen study programs. Columns (1) and (2) focus on the first program we observe in the data, whereas Columns (3) and (4) focus on the last program we observe in the data. All specifications include fixed effects for each selected occupation (i.e., rank 1, 2, 3, 4, and 5). The analysis is necessarily restricted to students who chose to go into post-secondary vocational education and chose a program for which we can define labor market prospects. The F-tests and P-values reported at the bottom of the table are for the joint significance of the two treatments. Standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

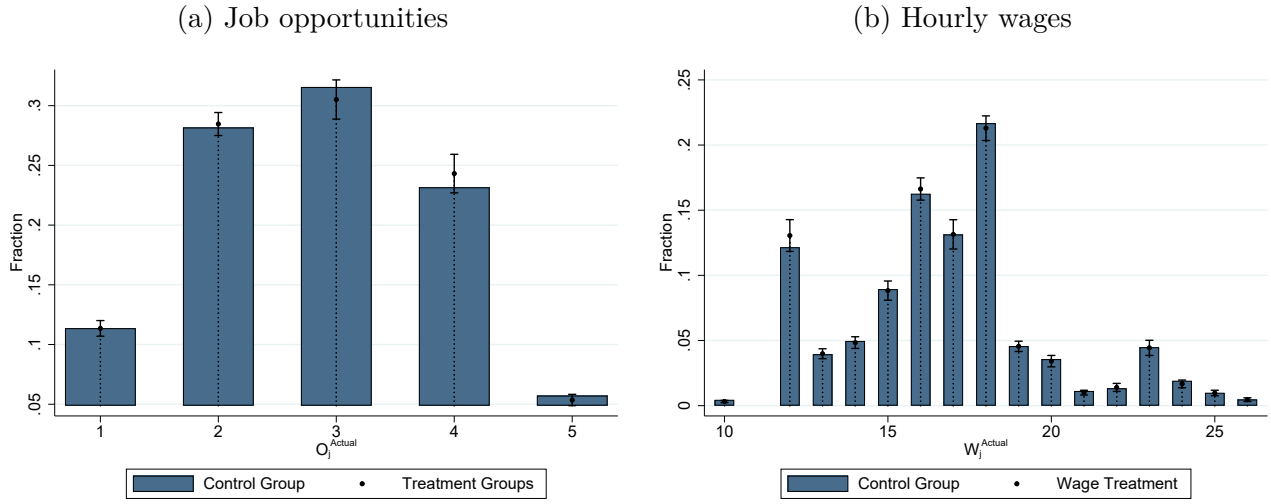
Table 7: Treatment impact on choice within initial set

	(1) First choice | News	(2) First choice | Info	(3) Final choice | News	(4) Final choice | Info
News about job opportunities	0.925		0.934	
	(0.0503)		(0.0492)	
News about wages	1.017		1.015	
	(0.0157)		(0.0170)	
Job Opp. Info Treatment × News about job opportunities	0.979		0.963	
	(0.0710)		(0.0702)	
Wage & Job Opp. Info Treatment × News about job opportunities	1.074		1.067	
	(0.0802)		(0.0797)	
Job Opp. Info Treatment × News about wages	1.029		1.027	
	(0.0209)		(0.0214)	
Wage & Job Opp. Info Treatment × News about wages	1.003		1.005	
	(0.0208)		(0.0214)	
Job opportunities of occupation		0.947		0.939
		(0.0429)		(0.0416)
Hourly wage of occupation		1.060***		1.064***
		(0.0158)		(0.0153)
Job Opp. Info Treatment × Job opportunities of occupation		1.076		1.070
		(0.0718)		(0.0664)
Wage & Job Opp. Info Treatment × Job opportunities of occupation		1.121		1.146**
		(0.0777)		(0.0788)
Job Opp. Info Treatment × Hourly wage of occupation		1.004		1.004
		(0.0211)		(0.0207)
Wage & Job Opp. Info Treatment × Hourly wage of occupation		1.032		1.009
		(0.0232)		(0.0221)
Prior rank 2	0.491***	0.476***	0.505***	0.492***
	(0.0294)	(0.0228)	(0.0301)	(0.0237)
Prior rank 3	0.320***	0.305***	0.341***	0.330***
	(0.0246)	(0.0178)	(0.0253)	(0.0185)
Prior rank 4	0.242***	0.222***	0.263***	0.250***
	(0.0235)	(0.0146)	(0.0242)	(0.0163)
Prior rank 5	0.159***	0.163***	0.160***	0.170***
	(0.0155)	(0.0117)	(0.0146)	(0.0114)
Observations	7917	15077	7728	14540
Chi-Sq Joint Sign. of Treatments × Job Opp.	1.847	2.877	2.012	3.955
P-value Joint Sign. of Treatments × Job Opp.	0.397	0.237	0.366	0.138
Chi-Sq Joint Sign. Wage & Job Opp. Treatment × Job Opp. and Wages	1.176	8.771	1.038	5.604
P-value Joint Sign. Wage & Job Opp. Treatment × Job Opp. and Wages	0.555	0.012	0.595	0.061

Note: Table shows odds ratios based on a conditional logit model with whether the student chosen a study program associated with occupation j as the outcome variable. Columns (1) and (2) consider the first choice we observe, and Columns (3) and (4) consider the final choice we observe. *News about job opportunities* and *wages* are defined as $\mathcal{O}_j = O_j^{Actual} - O_{i,j}^{Prior}$ and $\mathcal{W}_j = W_j^{Actual} - W_{i,j}^{Prior}$ (Columns (1) and (3)), respectively. That is, the higher this value, the better the 'news'. *Job opportunities* and *wage* are defined as usual: O_j^{Actual} and W_j^{Actual}. Dummies for the rank of the occupation in the student's first choice set are also included. The χ^2-tests test the joint significance of the two treatments interacted with the job opportunities (i.e., did the job opportunities make treated students more or less likely to pick a profile) and the joint significance of the job opportunities and the wages for the treatment that received both. Standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Figures

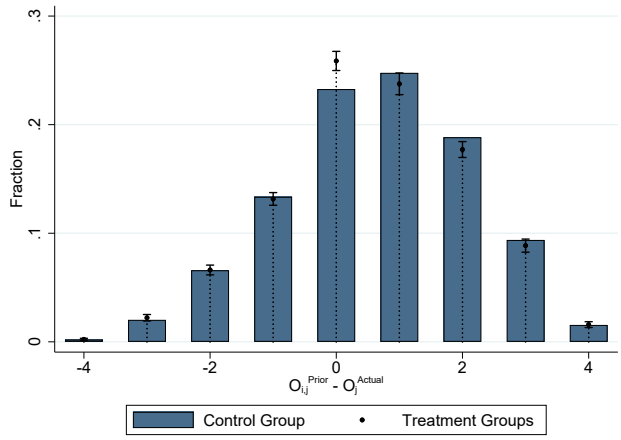
Figure 1: Job opportunities and hourly wages of selected occupations



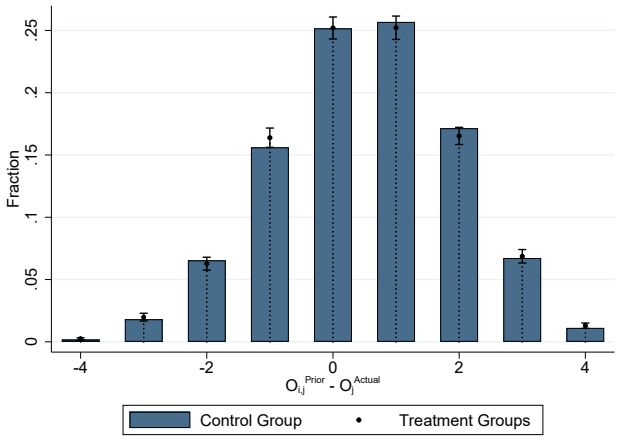
Note: graphic representation of multinomial logit estimation. Standard errors clustered at school level. Blue bars indicate level for control group. Black dots and error bars indicate level for treatment group and 95% confidence interval, respectively.

Figure 2: Prior belief accuracy by relevant group

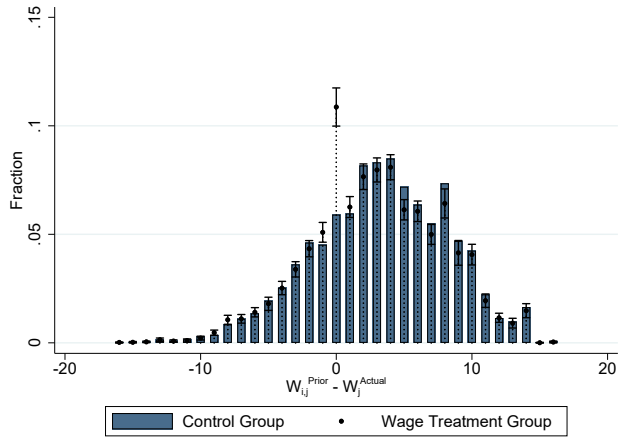
(a) Job opportunities 2018/2019



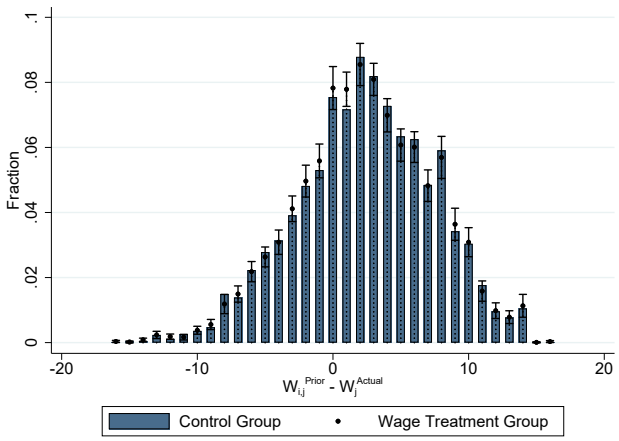
(b) Job opportunities 2019/2020



(c) Hourly wages 2018/2019



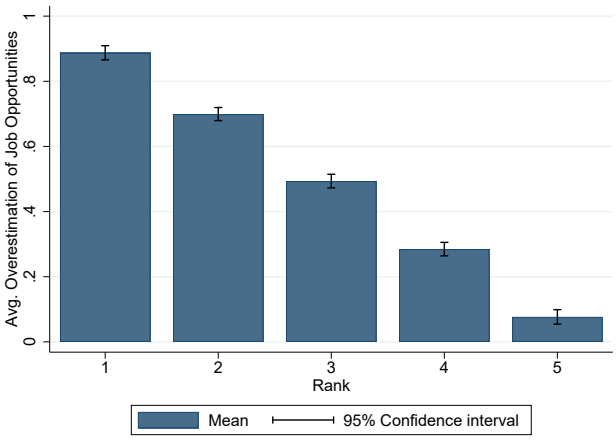
(d) Hourly wages 2019/2020



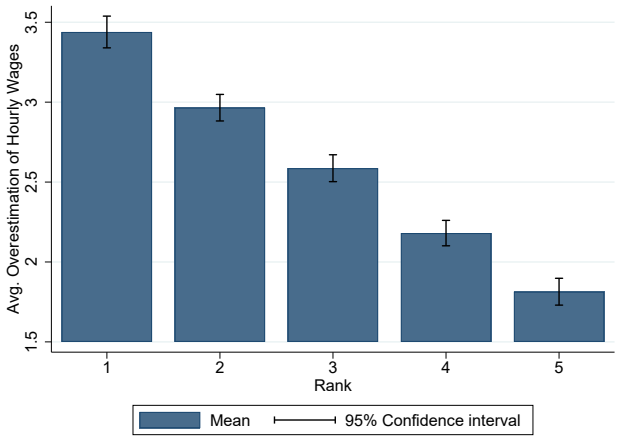
Note: graphic representation of multinomial logit estimation at occupation level. The x-axis displays the degree of overestimation. For the job opportunities, the numbers indicate the overestimation in categories (i.e., -2 denotes an underestimation of two categories, whereas +2 indicates an overestimation of two categories). For the hourly wages, the overestimation is displayed in Euros. Standard errors clustered at school level. Blue bars indicate level for control group. Black dots and error bars indicate level for treatment group and 95% confidence interval, respectively.

Figure 3: $O_{i,j}^{Prior} - O_j^{Actual}$ and $W_{i,j}^{Prior} - W_j^{Actual}$ by rank

(a) Job opportunities

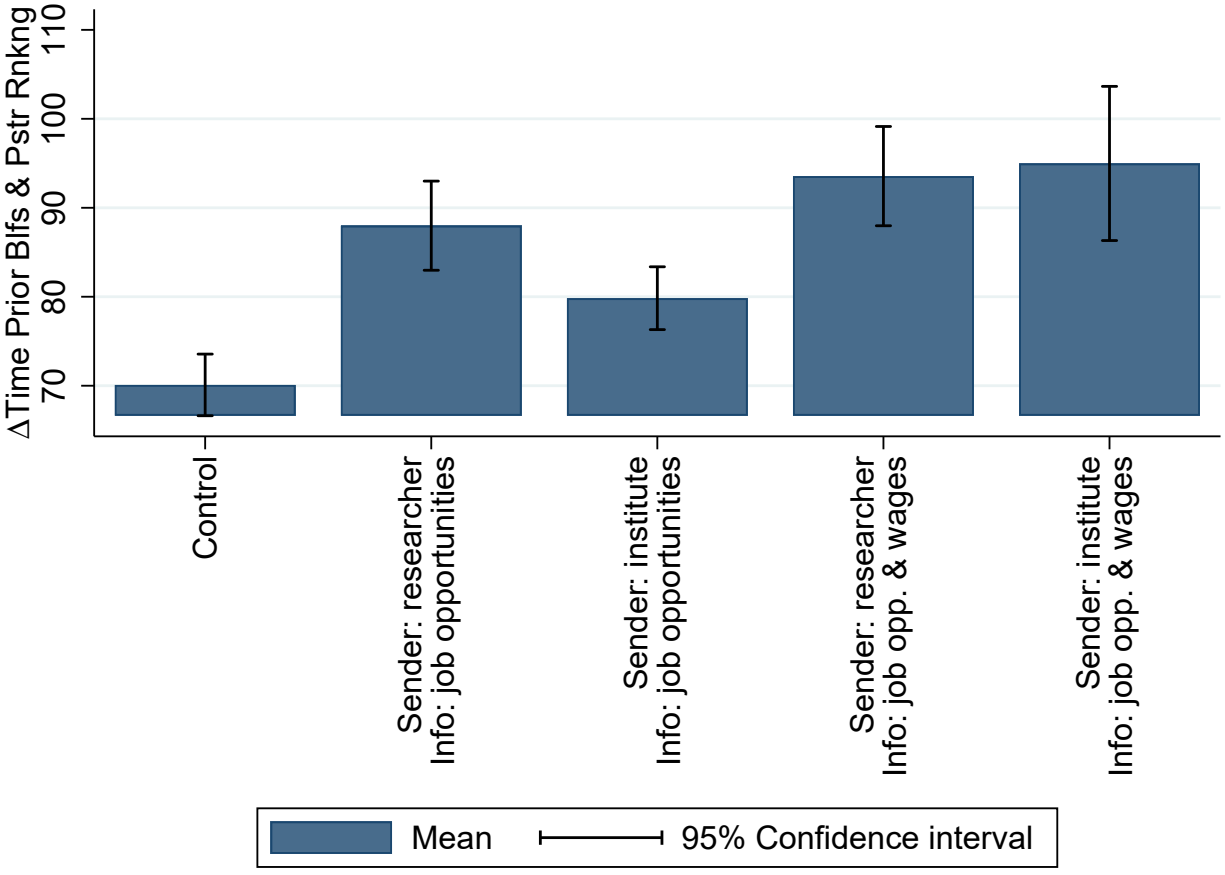


(b) Hourly wages



Note: The x-axis displays the initial rank of the occupation. The y-axis shows the degree of overestimation. For the job opportunities, the numbers indicate the overestimation in categories (i.e., -2 denotes an underestimation of two categories, whereas +2 indicates an overestimation of two categories). For the hourly wages, the overestimation is displayed in Euros. Standard errors clustered at school level. Estimation based on students from the 2019/2020 school year.

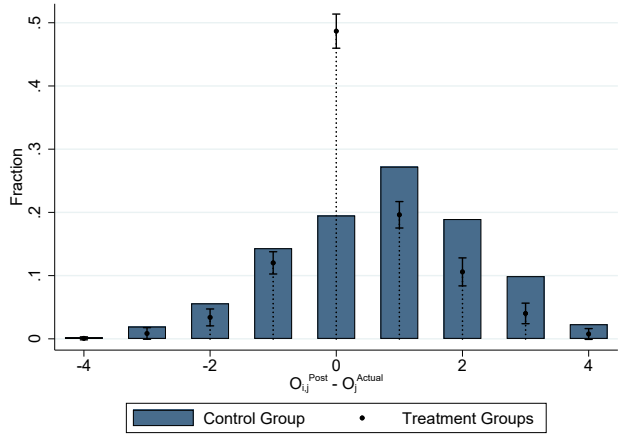
Figure 4: Time delta between prior beliefs and posterior ranking



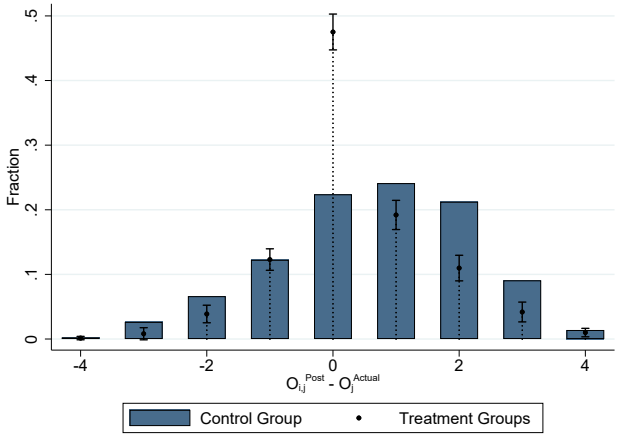
Note: Graphic representation of predictions based on ordinary least squares regression. Sample includes only students from the 2019/2020 school year. y-axis denotes the time spent on watching the video about work and processing the information in seconds. Standard errors are clustered at the school level in the regression.

Figure 5: Posterior belief accuracy by relevant group

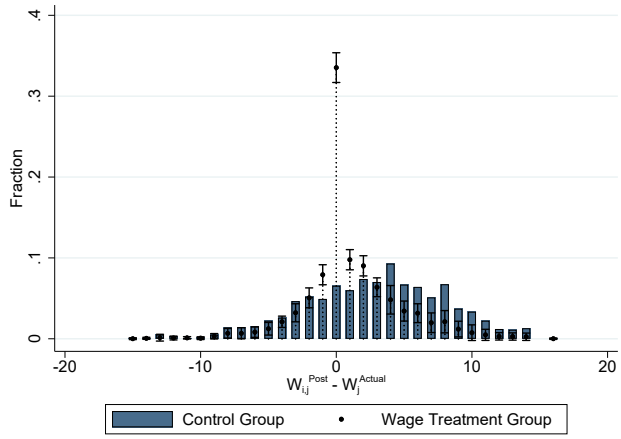
(a) Job opportunities 2018/2019



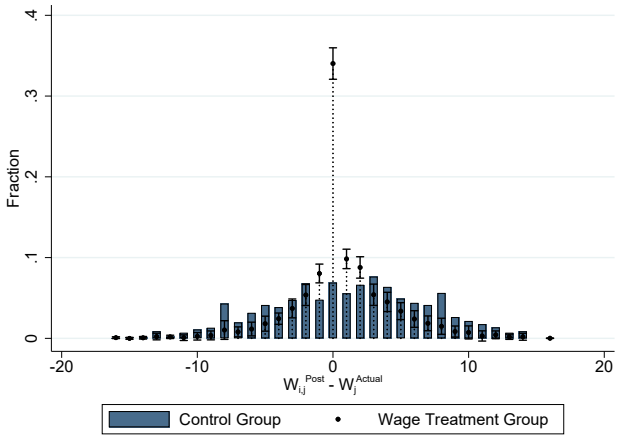
(b) Job opportunities 2019/2020



(c) Hourly wages 2018/2019



(d) Hourly wages 2019/2020



Note: graphic representation of multinomial logit estimation at occupation level. The x-axis displays the degree of overestimation. For the job opportunities, the numbers indicate the overestimation in categories (i.e., -2 denotes an underestimation of two categories, whereas +2 indicates an overestimation of two categories). For the hourly wages, the overestimation is displayed in Euros. Standard errors clustered at school level. Blue bars indicate level for control group. Black dots and error bars indicate level for treatment group and 95% confidence interval, respectively.

Appendix A: Recruitment Text

Dutch

ROA (Researchcentrum voor Onderwijs en Arbeidsmarkt aangesloten bij Universiteit Maastricht) en Qompas zijn samen door het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW) gevraagd om onderzoek uit te voeren naar de invloed van arbeidsmarktinformatie op de keuze van vmbo-leerlingen voor een studie.

Door middel van een A/B-test in de lesmethode Qompas VMBO/Mavo gaan we onderzoeken of vmbo'ers bij het maken van hun studiekeuze letten op informatie over baankans en of die informatie ertoe bijdraagt dat zij een betere keuze maken. Met deze informatie kan Qompas haar lesmethode doorontwikkelen om scholieren in de toekomst nog beter te kunnen helpen met hun studiekeuze.

Wij hopen dat uw school meewerkt aan dit onderzoek. Alle gegevens worden anoniem verwerkt. Voor meer informatie kunt u contact opnemen met [REDACTED].

English

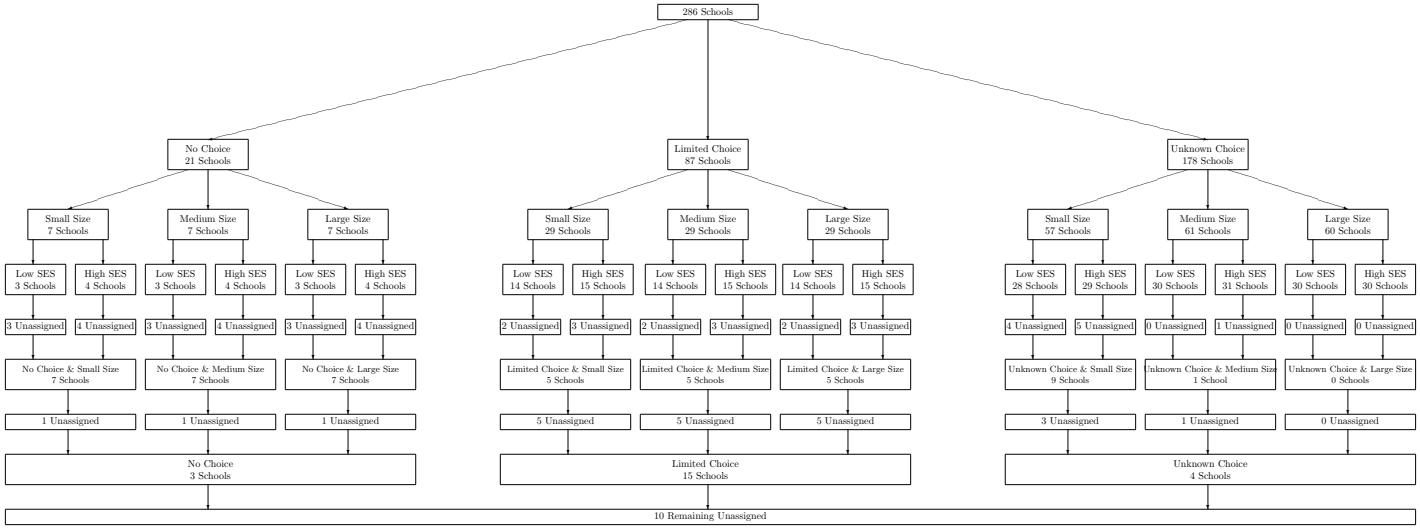
ROA (The Research Center for Education and the Labor Market, part of Maastricht University) and Qompas have been asked by the Ministry of Education, Culture and Science (OCW) to do a study on the influence of labor market information on the education choices of intermediate vocational education students.

Through an A/B-test in the Qompas system we will research whether intermediate vocational education students take information about job opportunities into account when making education choices and whether this information helps them make a better choice. With this information, Qompas can further develop its platform, making it even more helpful for students with making their education choice.

We hope your school will participate in this study. All details will be processed anonymously. For more information, you can contact [REDACTED].

Appendix B: Additional figures

Figure B1: Graphical Representation of Randomization



Appendix C: Additional tables

Table C1: Balance check survey respondents

	(1) Answered Survey	(2) Age	(3) Grade	(4) Male	(5) QOL score
Sender: researcher					
Info: job opportunities	0.0148				
	(0.0178)				
Sender: institute					
Info: job opportunities	0.00580				
	(0.0184)				
Sender: researcher					
Info: job opp. & wages	-0.00294				
	(0.0157)				
Sender: institute					
Info: job opp. & wages	-0.0116				
	(0.0125)				
Answered Survey		-0.0840*	-0.0161	-0.192***	-0.0251
		(0.0500)	(0.0343)	(0.0273)	(0.0680)
Constant	0.0960***	14.88***	3.242***	0.563***	6.710***
	(0.00848)	(0.0482)	(0.0400)	(0.0147)	(0.0759)
Observations	4389	4012	4389	4388	4292
F-Stat Treatments	0.637				
P-value F-Stat Treatments	0.637				

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual level.

Table C2: Balance table Administrative Data

	Control		Treatment 1		Treatment 2		Treatment 3		Treatment 4		
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	P-value joint sign.
Nmbr of Students	171.03	134.05	194.51	146.92	182.36	119.97	182.06	113.27	150.10	110.95	0.55
Nmbr of Profiles Available	3.35	1.01	2.92	1.32	3.28	1.15	3.19	1.06	3.14	1.20	0.48
Age	13.99	0.91	14.01	0.95	14.09	0.92	13.93	0.89	14.03	0.94	0.80
Male	0.53	0.50	0.52	0.50	0.57	0.49	0.52	0.50	0.51	0.50	0.60
Grade	2.47	0.63	2.50	0.66	2.52	0.68	2.38	0.60	2.51	0.67	0.75
QOL Score	6.63	1.35	6.62	1.27	6.88	1.17	6.77	1.46	6.51	1.47	0.42

Note: No. of students and no. of profiles available are school level variables. Age, male, grade and QOL score are individual level variables. QOL score refers to the quality of life indicator score of the neighborhood the student lives in (see Section 3 in the paper). The last column of the Table indicates whether a joint significance tests shows a significant difference between the treatment groups and the control group for each of the variables considered.

Table C3: Job opportunities of selected occupations by treatment group and year

		(1) Mean value	(2) Rank 1	(3) Rank 2	(4) Rank 3	(5) Rank 4	(6) Rank 5
Sender: researcher							
Info: job opportunities		-0.00181	-0.00589	-0.0207	-0.0105	0.0253	-0.000699
		(0.0223)	(0.0317)	(0.0288)	(0.0286)	(0.0294)	(0.0317)
Sender: institute							
Info: job opportunities		0.0174	-0.0187	0.0376	0.0205	0.0482	0.00394
		(0.0272)	(0.0330)	(0.0335)	(0.0354)	(0.0391)	(0.0437)
Sender: researcher							
Info: job opp. & wages		0.00634	-0.0190	0.0302	0.0441*	0.00259	-0.0236
		(0.0194)	(0.0273)	(0.0344)	(0.0239)	(0.0284)	(0.0296)
Sender: institute							
Info: job opp. & wages		-0.0285	-0.0288	0.000159	-0.0253	-0.0411	-0.0507*
		(0.0211)	(0.0365)	(0.0288)	(0.0264)	(0.0305)	(0.0281)
2019/2020		0.0160	0.0450*	0.0575***	-0.00715	0.0200	-0.0298*
		(0.0145)	(0.0250)	(0.0212)	(0.0230)	(0.0223)	(0.0173)
Sender: researcher							
Info: job opportunities	× 2019/2020	0.00402	-0.0156	-0.0202	0.0586	-0.0542	0.0415
		(0.0237)	(0.0389)	(0.0355)	(0.0378)	(0.0348)	(0.0364)
Sender: institute							
Info: job opportunities	× 2019/2020	-0.0168	-0.0180	-0.0629*	-0.0153	-0.0312	0.0353
		(0.0251)	(0.0365)	(0.0370)	(0.0343)	(0.0406)	(0.0357)
Sender: researcher							
Info: job opp. & wages	× 2019/2020	0.0395	0.0508	-0.0384	0.0233	0.0590	0.0904**
		(0.0272)	(0.0426)	(0.0447)	(0.0384)	(0.0429)	(0.0412)
Sender: institute							
Info: job opp. & wages	× 2019/2020	0.00884	-0.0310	-0.0408	0.00896	0.0160	0.0853**
		(0.0308)	(0.0491)	(0.0406)	(0.0452)	(0.0395)	(0.0345)
Constant		2.826***	2.875***	2.798***	2.813***	2.814***	2.843***
		(0.0116)	(0.0215)	(0.0174)	(0.0160)	(0.0151)	(0.0151)
Observations		28267	27805	27811	27801	27715	27598
F-Stat Non-interacted Treatments		0.799	0.230	0.848	1.927	1.289	0.955
P-value F-Stat Non-interacted Treatments		0.527	0.922	0.496	0.107	0.275	0.433
F-Stat Treatments + Treatments x 19/20		1.825	1.421	0.602	1.845	2.291	1.675
P-value F-Stat Treatments + Treatments x 19/20		0.125	0.228	0.661	0.121	0.060	0.156

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual level. Discrepancies in observations are caused by the fact that a few occupations could not be assigned to a level of job opportunities. We do calculate an average score for the other occupations a student selected in this case.

Table C4: Hourly wages of selected occupations by treatment group and year

		(1) Mean value	(2) Rank 1	(3) Rank 2	(4) Rank 3	(5) Rank 4	(6) Rank 5
Sender: researcher							
Info: job opportunities		-0.194**	-0.266**	-0.152	-0.236**	-0.197**	-0.0812
		(0.0906)	(0.124)	(0.107)	(0.106)	(0.0923)	(0.111)
Sender: institute							
Info: job opportunities		-0.0628	-0.0616	-0.0123	0.0110	-0.141	-0.106
		(0.0764)	(0.109)	(0.0866)	(0.0951)	(0.0916)	(0.0824)
Sender: researcher							
Info: job opp. & wages		-0.0284	-0.0886	0.0328	-0.00867	-0.0544	-0.0343
		(0.101)	(0.150)	(0.0991)	(0.109)	(0.104)	(0.120)
Sender: institute							
Info: job opp. & wages		-0.0441	-0.0531	0.126	-0.0951	-0.110	-0.0585
		(0.119)	(0.161)	(0.129)	(0.123)	(0.118)	(0.127)
2019/2020		-0.0111	-0.0542	0.0884	0.00624	-0.0298	-0.0636
		(0.0547)	(0.0774)	(0.0662)	(0.0806)	(0.0615)	(0.0806)
Sender: researcher							
Info: job opportunities	× 2019/2020	0.172	0.225*	0.0570	0.240	0.109	0.177
		(0.108)	(0.130)	(0.145)	(0.146)	(0.139)	(0.133)
Sender: institute							
Info: job opportunities	× 2019/2020	0.0426	0.0362	-0.0325	0.0129	0.0359	0.190*
		(0.0737)	(0.116)	(0.0993)	(0.113)	(0.1000)	(0.115)
Sender: researcher							
Info: job opp. & wages	× 2019/2020	-0.0672	0.0798	-0.176	-0.139	-0.0173	-0.0347
		(0.107)	(0.138)	(0.128)	(0.136)	(0.125)	(0.153)
Sender: institute							
Info: job opp. & wages	× 2019/2020	-0.0786	-0.0604	-0.396***	-0.0962	-0.0110	0.126
		(0.0939)	(0.139)	(0.122)	(0.127)	(0.112)	(0.124)
Constant		16.79***	16.86***	16.61***	16.72***	16.81***	16.85***
		(0.0573)	(0.0778)	(0.0597)	(0.0737)	(0.0633)	(0.0640)
Observations		28267	27805	27811	27801	27715	27598
F-Stat Non-interacted Treatments		1.224	1.215	1.065	1.974	1.320	0.442
P-value F-Stat Non-interacted Treatments		0.301	0.305	0.374	0.099	0.263	0.778
F-Stat Treatments + Treatments x 19/20		0.512	0.152	1.514	1.275	0.685	0.688
P-value F-Stat Treatments + Treatments x 19/20		0.727	0.962	0.199	0.280	0.603	0.601

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual level. Discrepancies in observations are caused by the fact that a few occupations could not be assigned to a level of hourly wages. We do calculate an average score for the other occupations a student selected in this case.

Table C5: Heterogeneity job opportunities of selected occupations

	(1)	(2)	(3)	(4)	(5)	(6)
	Mean value	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
Age	-0.00225	-0.00221	0.00638	-0.00332	-0.00607	-0.00811
	(0.00612)	(0.00957)	(0.00893)	(0.00931)	(0.00946)	(0.00932)
3rd year	0.00306	0.0181	-0.00476	-0.00380	-0.00341	0.0139
	(0.0126)	(0.0198)	(0.0194)	(0.0190)	(0.0181)	(0.0188)
4th year	0.0134	0.0476	-0.0187	0.0176	-0.0111	0.0405
	(0.0209)	(0.0315)	(0.0349)	(0.0318)	(0.0331)	(0.0273)
Male	0.231***	0.0707***	0.187***	0.277***	0.320***	0.323***
	(0.0114)	(0.0163)	(0.0160)	(0.0167)	(0.0155)	(0.0153)
QOL score	0.00114	-0.000470	0.00565	0.000602	0.00830	-0.00786
	(0.00402)	(0.00610)	(0.00596)	(0.00600)	(0.00587)	(0.00586)
No. of Profiles Available=3	0.00808	0.0669***	0.0250	-0.0157	-0.0152	-0.0161
	(0.0193)	(0.0251)	(0.0262)	(0.0268)	(0.0245)	(0.0289)
No. of Profiles Available=4	0.0189	0.0772***	0.0380	0.0222	-0.0252	-0.00991
	(0.0180)	(0.0226)	(0.0249)	(0.0247)	(0.0231)	(0.0268)
Constant	2.725***	2.815***	2.573***	2.709***	2.715***	2.839***
	(0.0908)	(0.140)	(0.129)	(0.128)	(0.141)	(0.142)
Observations	25979	25555	25555	25548	25458	25362

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual level. Only includes control group students. 2nd year, female students in schools where only 1 profile is available are baseline.

Table C6: Heterogeneity hourly wages of selected occupations by treatment group and year

	(1)	(2)	(3)	(4)	(5)	(6)
	Mean value	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
Age	-0.00225	-0.00221	0.00638	-0.00332	-0.00607	-0.00811
	(0.00612)	(0.00957)	(0.00893)	(0.00931)	(0.00946)	(0.00932)
3rd year	0.00306	0.0181	-0.00476	-0.00380	-0.00341	0.0139
	(0.0126)	(0.0198)	(0.0194)	(0.0190)	(0.0181)	(0.0188)
4th year	0.0134	0.0476	-0.0187	0.0176	-0.0111	0.0405
	(0.0209)	(0.0315)	(0.0349)	(0.0318)	(0.0331)	(0.0273)
Male	0.231***	0.0707***	0.187***	0.277***	0.320***	0.323***
	(0.0114)	(0.0163)	(0.0160)	(0.0167)	(0.0155)	(0.0153)
QOL score	0.00114	-0.000470	0.00565	0.000602	0.00830	-0.00786
	(0.00402)	(0.00610)	(0.00596)	(0.00600)	(0.00587)	(0.00586)
No. of Profiles Available=3	0.00808	0.0669***	0.0250	-0.0157	-0.0152	-0.0161
	(0.0193)	(0.0251)	(0.0262)	(0.0268)	(0.0245)	(0.0289)
No. of Profiles Available=4	0.0189	0.0772***	0.0380	0.0222	-0.0252	-0.00991
	(0.0180)	(0.0226)	(0.0249)	(0.0247)	(0.0231)	(0.0268)
Constant	2.725***	2.815***	2.573***	2.709***	2.715***	2.839***
	(0.0908)	(0.140)	(0.129)	(0.128)	(0.141)	(0.142)
Observations	25979	25555	25555	25548	25458	25362

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual level. Only includes control group students. 2nd year, female students in schools where only 1 profile is available are baseline.

Table C7: Heterogeneity in prior beliefs

	(1) $O_{i,j}^{Prior} - O_j^{Actual}$	(2) $\lvert O_{i,j}^{Prior} - O_j^{Actual} \rvert$	(3) $O_{i,j}^{Prior} - O_j^{Actual} = 0$	(4) $W_{i,j}^{Prior} - W_j^{Actual}$	(5) $\lvert W_{i,j}^{Prior} - W_j^{Actual} \rvert$	(6) $W_{i,j}^{Prior} - W_j^{Actual} = 0$
Age	0.00421	0.000280	0.00645**	0.0371	0.0268	0.00108
	(0.00885)	(0.00744)	(0.00317)	(0.0328)	(0.0255)	(0.00224)
3rd year	0.00935	0.00441	-0.0109**	-0.130	-0.252***	0.00432
	(0.0213)	(0.0134)	(0.00547)	(0.0875)	(0.0595)	(0.00473)
4th year	0.0785***	-0.0192	-0.00269	-0.436***	-0.527***	0.0111
	(0.0287)	(0.0213)	(0.0103)	(0.150)	(0.107)	(0.00720)
Male	0.0677***	0.0471***	-0.0151***	0.644***	0.422***	-0.0167***
	(0.0147)	(0.0110)	(0.00529)	(0.0798)	(0.0487)	(0.00384)
QOL score	0.00345	-0.00286	0.00120	-0.0400	-0.0424*	0.00266**
	(0.00562)	(0.00408)	(0.00159)	(0.0315)	(0.0220)	(0.00125)
No. of Profiles Available=3	0.0419	-0.0412**	0.0209**	0.107	-0.119	0.00500
	(0.0331)	(0.0166)	(0.00861)	(0.234)	(0.0927)	(0.00537)
No. of Profiles Available=4	0.0419	-0.0294**	0.0156**	-0.0550	-0.165*	0.0122**
	(0.0271)	(0.0146)	(0.00767)	(0.223)	(0.0832)	(0.00545)
Rank=2	-0.235***	-0.105***	0.0271***	-0.584***	-0.344***	0.00623
	(0.0114)	(0.0113)	(0.00517)	(0.0356)	(0.0323)	(0.00398)
Rank=3	-0.425***	-0.174***	0.0382***	-0.902***	-0.499***	0.0113**
	(0.0131)	(0.0130)	(0.00647)	(0.0367)	(0.0314)	(0.00508)
Rank=4	-0.611***	-0.211***	0.0511***	-1.192***	-0.546***	0.0110**
	(0.0155)	(0.0128)	(0.00686)	(0.0500)	(0.0348)	(0.00466)
Rank=5	-0.818***	-0.200***	0.0438***	-1.485***	-0.583***	0.0187***
	(0.0200)	(0.0153)	(0.00759)	(0.0529)	(0.0427)	(0.00434)
Constant	-0.241	1.138***	0.111	-3.519***	4.856***	-0.0272
	(0.190)	(0.129)	(0.0678)	(0.712)	(0.540)	(0.0354)
Observations	41842	41842	41842	41825	41825	41825

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual-occupation level. Regressions include occupation dummies. Only includes control group students.

Table C8: Treatment effect on posterior beliefs job opportunities by prior belief accuracy

	(1) $\lvert O_{i,j}^{Post} - O_j^{Actual} \rvert$	(2) $O_{i,j}^{Post} - O_j^{Actual} = 0$
Treated	0.0319***	0.0154*
	(0.00824)	(0.00860)
$\lvert O_{i,j}^{Prior} - O_j^{Actual} \rvert$	0.911***	-0.241***
	(0.00456)	(0.00361)
Treated × $\lvert O_{i,j}^{Prior} - O_j^{Actual} \rvert$	-0.112***	0.00920*
	(0.00925)	(0.00478)
$\left(O_{i,j}^{Prior} - O_j^{Actual} > 0\right)$	-0.0242***	-0.339***
	(0.00600)	(0.00559)
Treated × $\left(O_{i,j}^{Prior} - O_j^{Actual} > 0\right)$	-0.0574***	0.0823***
	(0.0115)	(0.00782)
Wage information	0.0205**	-0.00861
	(0.00961)	(0.00949)
Wage information × $\lvert O_{i,j}^{Prior} - O_j^{Actual} \rvert$	-0.0197	0.00949*
	(0.0137)	(0.00548)
Wage information × $\left(O_{i,j}^{Prior} - O_j^{Actual} > 0\right)$	-0.0293*	0.0168
	(0.0174)	(0.0102)
Constant	0.120**	0.557***
	(0.0522)	(0.0368)
Observations	64579	64579

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual-occupation level. Regressions contain occupation dummies. Treated = All treatment groups. Wage info = Treatments 3 & 4.

Table C9: Treatment effect on posterior beliefs hourly wages by prior belief accuracy

	(1) $\lvert W_{i,j}^{Post} - W_j^{Actual} \rvert$	(2) $W_{i,j}^{Post} - W_j^{Actual} = 0$
Treated	0.188***	-0.0250***
	(0.0306)	(0.00909)
$\lvert W_{i,j}^{Prior} - W_j^{Actual} \rvert$	0.914***	-0.0222***
	(0.00406)	(0.000609)
Treated × $\lvert W_{i,j}^{Prior} - W_j^{Actual} \rvert$	-0.0481***	0.00294***
	(0.00660)	(0.000859)
$\left(W_{i,j}^{Prior} - W_j^{Actual} > 0\right)$	0.00122	-0.173***
	(0.0179)	(0.00466)
Treated × $\left(W_{i,j}^{Prior} - W_j^{Actual} > 0\right)$	-0.0337	0.0151**
	(0.0307)	(0.00668)
Wage information	0.0232	0.0928***
	(0.0468)	(0.0114)
Wage information × $\lvert W_{i,j}^{Prior} - W_j^{Actual} \rvert$	-0.156***	-0.00325***
	(0.0127)	(0.00121)
Wage information × $\left(W_{i,j}^{Prior} - W_j^{Actual} > 0\right)$	-0.114**	0.0293***
	(0.0536)	(0.00998)
Constant	0.458***	0.196***
	(0.159)	(0.0227)
Observations	64565	64565

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual-occupation level. Regressions contain occupation dummies. Treated = All treatment groups. Wage info = Treatments 3 & 4.

Table C10: Detailed treatment effect on posterior beliefs

	(1) $O_{i,j}^{Post} - O_j^{Actual}$	(2) $\lvert O_{i,j}^{Post} - O_j^{Actual} \rvert$	(3) $O_{i,j}^{Post} - O_j^{Actual} = 0$	(4) $W_{i,j}^{Post} - W_j^{Actual}$	(5) $\lvert W_{i,j}^{Post} - W_j^{Actual} \rvert$	(6) $W_{i,j}^{Post} - W_j^{Actual} = 0$
Sender: researcher						
Info: job opportunities	-0.103***	-0.135***	0.0721***	-0.173**	-0.127**	0.00334
	(0.0135)	(0.0159)	(0.00821)	(0.0722)	(0.0585)	(0.00265)
Sender: institute						
Info: job opportunities	-0.103***	-0.139***	0.0762***	-0.0627	-0.0367	0.00271
	(0.0171)	(0.0120)	(0.00627)	(0.0726)	(0.0591)	(0.00309)
Sender: researcher						
Info: job opp. & wages	-0.135***	-0.138***	0.0714***	-0.870***	-0.914***	0.113***
	(0.0137)	(0.0124)	(0.00563)	(0.0843)	(0.0685)	(0.00590)
Sender: institute						
Info: job opp. & wages	-0.127***	-0.163***	0.0854***	-0.799***	-0.920***	0.110***
	(0.0155)	(0.0140)	(0.00704)	(0.0758)	(0.0690)	(0.00488)
Constant	-0.559***	1.005***	0.249***	-3.247***	4.486***	0.0204
	(0.0706)	(0.0539)	(0.0309)	(0.287)	(0.229)	(0.0142)
Observations	136721	136721	136721	136707	136707	136707
F-Stat Researcher; Job Opp. = Institute; Job Opp.	0.000	0.039	0.185			
P-value Researcher; Job Opp. = Institute; Job Opp.	0.999	0.844	0.667			
F-Stat Researcher; Job Opp. & Wage = Institute; Job Opp. & Wage	0.249	2.038	2.824	0.568	0.005	0.219
P-value F-Stat f2 F-Stat Researcher; Job Opp. & Wage = Institute; Job Opp. & Wage	0.618	0.155	0.094	0.452	0.944	0.640

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual-occupation level. Regressions contain occupation dummies.

Table C11: Sender effect on posterior beliefs

	(1) $O_{i,j}^{Post} - O_j^{Actual}$	(2) $\lvert O_{i,j}^{Post} - O_j^{Actual} \rvert$	(3) $O_{i,j}^{Post} - O_j^{Actual} = 0$	(4) $W_{i,j}^{Post} - W_j^{Actual}$	(5) $\lvert W_{i,j}^{Post} - W_j^{Actual} \rvert$	(6) $W_{i,j}^{Post} - W_j^{Actual} = 0$
Sender: Female - Low status	-0.00271	-0.0100	-0.00256	0.388**	0.0337	-0.00948
	(0.0250)	(0.0198)	(0.00918)	(0.188)	(0.129)	(0.0166)
Sender: Male - High status	-0.00166	0.0103	-0.00616	0.133	-0.0303	0.00212
	(0.0263)	(0.0238)	(0.0108)	(0.122)	(0.119)	(0.0154)
Sender: Male - Low status	0.00609	0.0236	-0.00504	0.0675	-0.104	0.0188
	(0.0277)	(0.0250)	(0.0106)	(0.225)	(0.165)	(0.0151)
Male	0.100***	0.105***	-0.0412***	0.583***	0.391***	-0.0330**
	(0.0278)	(0.0240)	(0.0104)	(0.196)	(0.120)	(0.0160)
Sender: Female - Low status × Male	-0.0102	0.0163	0.00345	-0.479*	-0.214	0.0294
	(0.0405)	(0.0315)	(0.0154)	(0.244)	(0.153)	(0.0232)
Sender: Male - High status × Male	0.0312	-0.0128	0.0104	-0.223	-0.127	0.0227
	(0.0438)	(0.0400)	(0.0167)	(0.214)	(0.185)	(0.0206)
Sender: Male - Low status × Male	0.0356	-0.0196	0.0195	0.140	0.0185	-0.000800
	(0.0402)	(0.0376)	(0.0162)	(0.236)	(0.215)	(0.0217)
Constant	-0.690***	0.696***	0.373***	-3.798***	3.643***	0.101*
	(0.118)	(0.0921)	(0.0606)	(0.810)	(0.635)	(0.0587)
Observations	44964	44964	44964	20466	20466	20466

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual-occupation level. Regressions contain occupation dummies. Female students with female high status sender are baseline. Regressions (1), (2) and (3) contain treatments 1 & 3. Regressions (4), (5) and (6) conly contain treatment 3.

Table C12: Heterogeneous treatment effects on posterior beliefs job opportunities

	(1) $O_{i,j}^{Post} - O_j^{Actual}$	(2) $\lvert O_{i,j}^{Post} - O_j^{Actual} \rvert$	(3) $O_{i,j}^{Post} - O_j^{Actual} = 0$
Treated	-0.0783**	-0.154***	0.0759***
	(0.0374)	(0.0228)	(0.0117)
Age (demeaned)	0.00340	0.000383	0.00659**
	(0.00877)	(0.00725)	(0.00313)
Treated × Age (demeaned)	0.00432	0.00924	-0.0106**
	(0.0116)	(0.00948)	(0.00461)
3rd year	0.00995	0.00359	-0.0115**
	(0.0210)	(0.0129)	(0.00570)
4th year	0.0902***	-0.0236	-0.00475
	(0.0272)	(0.0202)	(0.00975)
Treated × 3rd year	-0.00166	-0.0449**	0.0368***
	(0.0257)	(0.0194)	(0.00941)
Treated × 4th year	-0.0585	-0.0897***	0.0587***
	(0.0385)	(0.0330)	(0.0167)
Male	0.0632***	0.0452***	-0.00691
	(0.0145)	(0.0107)	(0.00506)
Treated × Male	0.0511***	0.0566***	-0.0287***
	(0.0183)	(0.0129)	(0.00617)
QOL score (demeaned)	0.00438	-0.00309	0.000764
	(0.00609)	(0.00436)	(0.00175)
Treated × QOL score (demeaned)	-0.00702	-0.00333	0.00244
	(0.00736)	(0.00549)	(0.00239)
No. of Profiles Available=3	0.0462	-0.0433**	0.0201**
	(0.0370)	(0.0169)	(0.00983)
No. of Profiles Available=4	0.0504	-0.0330**	0.0129
	(0.0315)	(0.0159)	(0.00906)
Treated × No. of Profiles Available=3	-0.0511	0.0191	-0.0113
	(0.0421)	(0.0243)	(0.0131)
Treated × No. of Profiles Available=4	-0.0787**	-0.0178	0.00822
	(0.0362)	(0.0226)	(0.0120)
Constant	-0.676***	0.972***	0.257***
	(0.0729)	(0.0569)	(0.0319)
Observations	125647	125647	125647

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual-occupation level. Only includes control group students. Treated = All Treatments. 2nd year, female students in schools where only 1 profile is available are baseline.

Table C13: Heterogeneous treatment effects on posterior beliefs hourly wages

	(1) $W_{i,j}^{Post} - W_j^{Actual}$	(2) $\lvert W_{i,j}^{Post} - W_j^{Actual} \rvert$	(3) $W_{i,j}^{Post} - W_j^{Actual} = 0$
Wage information	-0.712***	-1.013***	0.117***
	(0.252)	(0.109)	(0.00849)
Age (demeaned)	0.0285	0.0182	0.00130
	(0.0337)	(0.0257)	(0.00217)
Wage information × Age (demeaned)	-0.0235	-0.00156	-0.00640
	(0.0508)	(0.0404)	(0.00433)
3rd year	-0.119	-0.249***	0.00474
	(0.0871)	(0.0553)	(0.00471)
4th year	-0.419***	-0.484***	0.00865
	(0.158)	(0.106)	(0.00709)
Wage information × 3rd year	0.117	-0.00494	0.0167
	(0.141)	(0.110)	(0.0114)
Wage information × 4th year	0.289	0.00619	0.0305*
	(0.220)	(0.182)	(0.0162)
Male	0.618***	0.370***	-0.0173***
	(0.0771)	(0.0493)	(0.00460)
Wage information × Male	-0.0446	0.0921	-0.00908
	(0.0995)	(0.0710)	(0.00765)
QOL score (demeaned)	-0.0390	-0.0389*	0.00185
	(0.0338)	(0.0227)	(0.00123)
Wage information × QOL score (demeaned)	0.0360	0.0216	-0.00283
	(0.0407)	(0.0322)	(0.00234)
No. of Profiles Available=3	0.0950	-0.164*	0.0119**
	(0.258)	(0.0984)	(0.00495)
No. of Profiles Available=4	-0.0871	-0.204**	0.0163***
	(0.247)	(0.0839)	(0.00506)
Wage information × No. of Profiles Available=3	-0.170	0.0753	-0.00454
	(0.280)	(0.137)	(0.00995)
Wage information × No. of Profiles Available=4	-0.261	-0.0408	-0.00568
	(0.262)	(0.115)	(0.00918)
Constant	-3.743***	4.765***	-0.00183
	(0.397)	(0.349)	(0.0213)
Observations	80042	80042	80042

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual-occupation level. Only includes control group students. Wage info = Treatments 3 & 4. Treatments 1 & 2 are excluded from this analysis. 2nd year, female students in schools where only 1 profile is available are baseline.

Table C14: Medium-term treatment effect on posterior beliefs

		(1) $O_{i,j}^{Survey} - O_j^{Actual}$	(2) $\lvert O_{i,j}^{Survey} - O_j^{Actual} \rvert$	(3) $O_{i,j}^{Survey} - O_j^{Actual} = 0$	(4) $W_{i,j}^{Survey} - W_j^{Actual}$	(5) $\lvert W_{i,j}^{Survey} - W_j^{Actual} \rvert$	(6) $W_{i,j}^{Survey} - W_j^{Actual} = 0$
Sender: researcher							
Info: job opportunities		-0.0173	0.00867	-0.0590**	-0.490	0.198	-0.0138
		(0.125)	(0.0612)	(0.0282)	(0.601)	(0.276)	(0.0191)
Sender: institute							
Info: job opportunities		0.284**	0.166**	-0.0415*	0.167	0.336	-0.0238
		(0.140)	(0.0713)	(0.0244)	(0.519)	(0.260)	(0.0200)
Sender: researcher							
Info: job opp. & wages		0.269*	0.0508	-0.0530*	1.189*	0.353	-0.0344
		(0.157)	(0.0760)	(0.0317)	(0.604)	(0.254)	(0.0235)
Sender: institute							
Info: job opp. & wages		-0.207	-0.0380	0.00785	0.121	0.0692	0.0103
		(0.161)	(0.0668)	(0.0322)	(0.543)	(0.258)	(0.0318)
2019/2020		0.238	0.184*	-0.120***	-0.464	0.643	-0.0231
		(0.208)	(0.0962)	(0.0454)	(1.224)	(0.424)	(0.0254)
Sender: researcher							
Info: job opportunities	× 2019/2020	-0.00801	-0.397**	0.205***	-2.186	0.0274	0.115*
		(0.265)	(0.153)	(0.0709)	(1.656)	(0.885)	(0.0674)
Sender: institute							
Info: job opportunities	× 2019/2020	-0.181	-0.389***	0.0716	0.336	-1.107**	0.0777*
		(0.234)	(0.118)	(0.0659)	(1.287)	(0.485)	(0.0396)
Sender: researcher							
Info: job opp. & wages	× 2019/2020	-0.136	-0.302*	0.217**	-0.846	-0.765	0.0378
		(0.265)	(0.165)	(0.0846)	(1.368)	(0.693)	(0.0380)
Sender: institute							
Info: job opp. & wages	× 2019/2020	0.471	-0.286**	0.152**	0.0694	-0.675	-0.00231
		(0.293)	(0.122)	(0.0616)	(1.497)	(0.760)	(0.0428)
Constant		0.216**	1.154***	0.284***	-0.741**	3.857***	0.0887***
		(0.0859)	(0.0422)	(0.0171)	(0.368)	(0.151)	(0.0129)
Observations		2079	2057	1928	1928	1909	1928
F-Stat Non-interacted (Wage) Treatments		2.907	2.044	1.950	2.142	0.992	1.286
P-value F-Stat Non-interacted (Wage) Treatments		0.025	0.093	0.107	0.122	0.374	0.281
F-Stat (Wage) Treatments + (Wage) Treatments x 19/20		0.570	2.622	3.496	0.066	0.805	0.222
P-value F-Stat (Wage) Treatments + (Wage) Treatments x 19/20		0.685	0.039	0.010	0.936	0.450	0.801

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. F-Stat Non-interacted Treatments is compared to 2018/2019 control group. F-Stat Treatments + Treatments x 19/20 is compared to 2019/2020 control group.

Table C15: Heterogeneous treatment effect on favorite occupation's job opportunities

	(1) Pr(Fav. Change)	(2) ΔO_j^{Actual}	(3) ΔO_j^{Actual} (Changed)
Treated	0.0249***	-0.00210	-0.226
	(0.00939)	(0.0107)	(0.183)
Age (demeaned)	-0.00159	0.00862*	0.221**
	(0.00257)	(0.00441)	(0.110)
Treated × Age (demeaned)	0.00300	-0.0111*	-0.270**
	(0.00371)	(0.00567)	(0.120)
3rd year	-0.00226	-0.00539	-0.199
	(0.00584)	(0.00781)	(0.161)
4th year	-0.0264***	-0.0103	-0.254
	(0.00896)	(0.0117)	(0.345)
Treated × 3rd year	-0.00773	0.00867	0.303
	(0.00776)	(0.0103)	(0.186)
Treated × 4th year	0.00653	0.0155	0.471
	(0.0129)	(0.0162)	(0.383)
Male	0.0160***	-0.00129	-0.0415
	(0.00528)	(0.00739)	(0.129)
Treated × Male	-0.00538	0.00515	0.0406
	(0.00675)	(0.00933)	(0.151)
QOL score (demeaned)	-0.00357	-0.00153	-0.0126
	(0.00216)	(0.00255)	(0.0401)
Treated × QOL score (demeaned)	0.0000950	0.00280	0.0445
	(0.00267)	(0.00338)	(0.0495)
No. of Profiles Available=3	0.0185*	-0.00702	-0.255
	(0.00945)	(0.00967)	(0.161)
No. of Profiles Available=4	0.0171**	-0.0185**	-0.447***
	(0.00824)	(0.00827)	(0.136)
Treated × No. of Profiles Available=3	-0.00217	0.00296	0.128
	(0.0114)	(0.0132)	(0.209)
Treated × No. of Profiles Available=4	-0.00255	0.0289***	0.528***
	(0.00971)	(0.0110)	(0.176)
Constant	0.0328***	0.0192**	0.530***
	(0.00790)	(0.00791)	(0.146)
Observations	25174	25174	1634

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Treated = All Treatments. Regressions at individual level. 2nd year, female students in schools where only 1 profile is available are baseline.

Table C16: Heterogeneous treatment effect on favorite occupation's hourly wages

	(1) Pr(Fav. Change)	(2) ΔW_j^{Actual}	(3) ΔW_j^{Actual} (Changed)
Wage information	0.0207*	0.0101	0.205
	(0.0118)	(0.0405)	(0.764)
Age (demeaned)	-0.00159	0.00141	0.0468
	(0.00257)	(0.0120)	(0.305)
Wage information × Age (demeaned)	0.00254	-0.00235	-0.133
	(0.00487)	(0.0234)	(0.410)
3rd year	-0.00226	-0.0238	-0.468
	(0.00585)	(0.0222)	(0.441)
4th year	-0.0264***	0.0352	1.472
	(0.00897)	(0.0327)	(1.118)
Wage information × 3rd year	0.00239	0.0657	1.112*
	(0.0101)	(0.0441)	(0.661)
Wage information × 4th year	0.00676	0.0222	-0.0980
	(0.0150)	(0.0611)	(1.397)
Male	0.0160***	-0.0161	-0.375
	(0.00529)	(0.0204)	(0.366)
Wage information × Male	0.00181	0.0552	0.549
	(0.00827)	(0.0363)	(0.517)
QOL score (demeaned)	-0.00357	0.00636	0.103
	(0.00216)	(0.00694)	(0.121)
Wage information × QOL score (demeaned)	0.000869	-0.0222*	-0.251
	(0.00327)	(0.0117)	(0.157)
No. of Profiles Available=3	0.0185*	0.0370	0.514
	(0.00946)	(0.0297)	(0.651)
No. of Profiles Available=4	0.0171**	0.00310	-0.00864
	(0.00825)	(0.0246)	(0.554)
Wage information × No. of Profiles Available=3	0.00197	-0.00110	-0.336
	(0.0141)	(0.0431)	(0.786)
Wage information × No. of Profiles Available=4	-0.00126	0.0472	0.461
	(0.0126)	(0.0418)	(0.739)
Constant	0.0328***	0.00990	0.346
	(0.00791)	(0.0228)	(0.581)
Observations	16036	16036	1034

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Wage info = Treatments 3 & 4. Treatments 1 & 2 are excluded from this analysis. are excluded Regressions at individual level. 2nd year, female students in schools where only 1 profile is available are baseline.

Table C17: Effect of sender on likelihood favorite occupation changing and its prospects

	(1) Pr(Fav. Change)	(2) ΔO_j^{Actual}	(3) ΔO_j^{Actual} (Changed)	(4) ΔW_j^{Actual}	(5) ΔW_j^{Actual} (Changed)
Sender: Female - Low status	-0.00203	-0.000304	0.00430	0.00653	0.126
	(0.0106)	(0.0163)	(0.252)	(0.0675)	(0.954)
Sender: Male - High status	0.00575	-0.00312	-0.0667	0.0896	1.105
	(0.0118)	(0.0169)	(0.242)	(0.0722)	(0.840)
Sender: Male - Low status	-0.00370	0.0151	0.262	0.111	2.233
	(0.0128)	(0.0146)	(0.227)	(0.0822)	(1.326)
Male	0.000274	0.00166	0.0241	0.0736	0.934
	(0.0102)	(0.0150)	(0.232)	(0.0811)	(1.047)
Sender: Female - Low status × Male	0.00811	-0.00507	-0.105	-0.0000434	-0.0565
	(0.0147)	(0.0194)	(0.293)	(0.115)	(1.493)
Sender: Male - High status × Male	-0.00110	0.0146	0.209	-0.0299	-0.482
	(0.0157)	(0.0238)	(0.334)	(0.118)	(1.383)
Sender: Male - Low status × Male	0.0239	0.00338	-0.119	-0.0904	-2.264
	(0.0172)	(0.0214)	(0.288)	(0.114)	(1.568)
Constant	0.0655***	0.0186	0.284	0.0170	0.229
	(0.00807)	(0.0121)	(0.182)	(0.0480)	(0.642)
Observations	9016	9016	629	4099	315

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Regressions at individual level. Regressions contain occupation dummies. Female students with female high status sender are baseline. Regressions (1), (2) and (3) contain treatments 1 & 3. Regressions (4), (5) and (6) conly contain treatment 3.

Table C18: Long-term treatment effect on prospects favorite occupation

	(1) ΔO_j^{Actual} (Experiment)	(2) ΔO_j^{Actual} (Survey)	(3) ΔW_j^{Actual} (Experiment)	(4) ΔW_j^{Actual} (Survey)
Sender: researcher				
Info: job opportunities	0.0416	-0.0825	0.0525	0.0150
	(0.0676)	(0.128)	(0.0865)	(0.294)
Sender: institute				
Info: job opportunities	-0.0334	0.0211	-0.0432	-0.0688
	(0.0310)	(0.126)	(0.0478)	(0.367)
Sender: researcher				
Info: job opp. & wages	0.00715	0.136	0.00715	-0.455
	(0.0353)	(0.160)	(0.0726)	(0.478)
Sender: institute				
Info: job opp. & wages	-0.0236	-0.0768	-0.0236	0.157
	(0.0213)	(0.132)	(0.0432)	(0.334)
Constant	0.0236	0.126	0.0236	0.0394
	(0.0213)	(0.0867)	(0.0432)	(0.175)
Observations	447	447	447	447

Note: standard errors clustered at the school level in brackets; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. F-Stat Non-interacted Treatments is compared to 2018/2019 control group. F-Stat Treatments + Treatments x 19/20 is compared to 2019/2020 control group.