Informed Choices: Gender Gaps in Career Advice*

Yana Gallen and Melanie Wasserman

August 2022

Abstract

This paper provides the first causal evidence that gender affects the information an individual receives about careers. We conduct a large-scale field experiment in which real college students seek career information from 10,000 working professionals. We randomize whether a professional receives a message from a male or a female student. When students ask broadly for information about a career, female students receive substantially more information on work/life balance relative to male students. This gender difference persists even when students specifically state an interest in learning about work/life balance. In contrast, professionals mention workplace culture to male and female students at similar rates. We develop a methodology to combine our experimental estimates with student preferences for professionals. Allowing students to choose which professionals they interact with does not reduce gender disparities in access to information, nor does it align the information students receive with the information they demand.

Keywords: career information; gender; discrimination; correspondence study

JEL Codes: C93, J16, J24, J71

*Gallen: University of Chicago, Harris School of Public Policy and IZA. E-mail: yana@uchicago.edu. Wasserman: UCLA Anderson School of Management and CEPR. E-mail: melanie.wasserman@anderson.ucla.edu. Funding is gratefully acknowledged from the Upjohn Institute Early Career Research Grant, UCLA Faculty Senate Grant, and the Initiative for the Study of Gender in the Economy at the Becker Friedman Institute at the University of Chicago. We thank Jack Landry, Reigne Dadey, Mikaela Hassenzahl, Amber Liu, Jessica Lyu, Jacob Kohlhepp, and Matias Giaccobasso for excellent research assistance. We have benefited from comments from Dan Black, Paola Giuliano, Romain Wacziarg, Brendan Price, Manasi Deshpande, Trevor Gallen, Jeff Grogger, Juanna Joensen, Matt Notowidigdo, Nico Voigtlaender, Joanna Lahey, Bruce Sacerdote, Heather Sarsons, Ulf Zölitz, and seminar participants at NBER Summer Institute 2020 Gender in the Economy workshop, University of Maryland Population Research Center, Chicago Federal Reserve Bank, University of Michigan, NBER Labor Studies Fall 2020 Meeting, UC Santa Barbara Applied Micro Lunch, University of Chicago Experimental Brownbag, Advances with Field Experiments, the University of Chicago Demography Workshop, ASSA 2021, the CEPR WE_ARE Seminar, UC Berkeley, University of Nebraska, LMU, MEA 2021, WEAI 2021, Upjohn Institute, University of Minnesota Carlson, Boston University Questrom, Discrimination and Disparities Seminar, University of Toronto Rotman, Bocconi â AXA Research Lab on Gender Inequality, CEBI - University of Copenhagen, University of Chicago BFI Conference on Discrimination in the 21st Century: Fostering Conversations Across Fields, and University of Zurich URPP Conference on Gender and Inequality. This study was registered in the AEA Social Sciences Registry under AEARCTR-0005464.

1 Introduction

Canonical models of occupational choice rely on the assumption that individuals have perfect information about the monetary and non-monetary attributes of jobs (Rosen, 1974, 1986; Roy, 1951). In such models, individuals choose occupations by weighing their own skills and preferences against the skill requirements and attributes of occupations. Without perfect information on job attributes, individuals may make sub-optimal training and career choices, which are difficult and costly to change (Neal, 1999; Papageorgiou, 2014). Despite the centrality of the perfect information assumption, there is limited empirical evidence on what determines an individual's access to information.

This paper provides the first causal evidence that gender affects the information an individual receives about careers. Although careers can be characterized by numerous attributes, a main driver of the gender pay gap is temporal demands in the form of long, irregular, and/or continuous hours worked (Polachek, 1981; McDowell, 1982; Goldin, 2014; Cortés and Pan, 2019).[1] A newer literature emphasizing the role of psychological traits has posited that women's aversion to competition may explain their lower wages and representation in jobs with pay contingent on competitive outcomes (Niederle and Vesterlund, 2011; Flory et al., 2015).[2] Our paper investigates whether an individual's gender causally affects the information they receive on these two attributes. Overall, we find that men receive less information than women on the temporal demands of jobs, even when they specifically ask for this information. Moreover, we provide evidence of a mismatch between the information received and the information demanded.

We conduct a large-scale field experiment in which real college students seek career information from 10,000 working professionals on an online professional networking platform. Among college students, conversations with professionals are ubiquitous and encouraged, making the setting well suited to studying career information.[3] In order to identify the causal effect of gender on information received, we randomize whether each of the 10,000 professionals in our sample is contacted by a male or female student and the preformulated question each professional is asked. The randomization ensures that student gender is unrelated to professionals' characteristics and the wording of

[1] Goldin (2014) argues that jobs disproportionately rewarding long hours amount to the last remaining hurdle in closing the gender pay gap.

[2] While the investigation of gender differences in preferences for competition has mostly been conducted through lab studies, survey evidence shows that these factors can account for around 16 percent of the gender pay gap (Blau and Kahn, 2017).

[3] University career centers often advise students to contact professionals (especially alumni) about career choice and job opportunities. 87 percent of students at the institution where we conduct our study have reached out to professionals for help with career choice.

questions. In addition, the online setting allows us to strictly limit which student characteristics are observed by professionals, ensuring that the students are perceived as otherwise similar, aside from their gender. Our experimental design additionally overcomes a key challenge that has prevented researchers from studying information transmission: information is often transmitted through informal, private, one-on-one conversations, which are inherently unobservable to the researcher. An innovation of our study is that the students provide us with the verbatim responses that they receive, resulting in a novel data set with the demographic characteristics of students, professionals, and the transcripts of their conversations.

Our main finding is that a student's gender affects the career information they receive. The experimental design and analysis focus on two career attributes that differentially affect the labor market choices of women: temporal demands, known colloquially as "work/life balance," and competitive culture (Goldin, 2014; Wiswall and Zafar, 2018; Cubas et al., 2019; Niederle and Vesterlund, 2011; Flory et al., 2015). When students ask for general information about a professional's career path, female students are more than twice as likely as male students to receive information on work/life balance. In contrast, professionals mention workplace culture to male and female students at similar rates. When professionals bring up work/life balance, the content is negative and increases students' concern about this issue. Notably, differences in the composition of professionals who respond to male and female students do not explain the greater emphasis on work/life balance to female students.

One natural explanation for professionals providing different information to male and female students is they are using student gender as a proxy for preferences (statistical discrimination). For example, professionals may believe that women, on average, care more about work/life balance relative to men, and provide information accordingly. In order to investigate whether professionals' perceptions of students' preferences drive the gender difference in information provision, students also sent messages that specifically stated their preference for discussing work/life balance issues in the professional's career path. We find that conditional on responding to this message, professionals answers to male and female students are similar. However, there is a large difference in response rates: female students are 28 percent more likely to receive a response. This response rate gap emerges only when students ask professionals questions about work/life balance, and not when students ask about workplace culture or ask the broad question. These results suggest that a straightforward intervention to reduce reliance on student observable characteristics does not equalize access to information for male and female students.

We test whether allowing students to select the professionals they are most interested in contacting attenuates the gender gap in information access. To do so, we develop a methodology to combine our experimental estimates with student preferences for the types of professionals they prefer to interact with. If students were to select professionals on their own, we find that there would be an even larger gender disparity in information received on work/life balance. In this scenario as well as in our main experiment, we note that the information received does not align with the information students demand. Evidence from a survey of the same student population reveals that male students demand significantly *more* information on work/life balance than female students.

Overall, our findings demonstrate that, relative to female students, male students receive less information on work/life balance, even when they specifically ask for it, and even though they want more of it. In contrast, female students receive substantially more information on work/life balance whether or not they ask for it. If students are aware of information gaps, then they can take steps to address them. A survey of the same student population reveals that 56 percent of students are unaware that women receive more information on work/life balance. The lack of awareness regarding information deficits and on what basis information is provided may give rise to long-term disparities in knowledge and outcomes.

We find that the information students receive matters for their career preferences. At the end of the study, we surveyed students regarding their career plans. Students who received information on work/life balance are more likely to be deterred from their preferred career path.

While this paper cannot provide direct evidence on the normative implications of gender differences in information received, our results suggest that there is, at a minimum, a mismatch between the information professionals supply and the information students demand.[4] Even in the presence of mismatch, it is possible that professionals' information provision is optimal, if students are unaware of changes to their preferences in the future and professionals respond with these future preferences in mind (Kuziemko et al., 2018). However, if professionals do not know or are biased in their beliefs about students' preferences, then gender differences information provision may be sub-optimal and labor market inequities may be propagated from older to younger cohorts.

This paper contributes to a number of literatures. First, this paper advances the robust literature investigating the role of informational frictions in shaping individuals' perceptions of key economic variables used in decision making. Narrowing the discussion to papers that pertain to educational and occupational decisions, this literature generally tests the effects of information

[4] A survey of students from the same university reveals that men demand *more* information on work/life balance than women.

provision on beliefs about the returns to education (Jensen, 2010; Nguyen, 2008; Dinkelman and Claudia Martínez, 2014; Hoxby and Turner, 2015), academic majors (Zafar, 2011; Wiswall and Zafar, 2015a,b), and occupations and jobs (Coffman et al., 2017; Wiswall and Zafar, 2018). Our paper provides a novel advance by investigating *access* to information. Specifically, we test whether there are gender differences in access to informal information about careers that can potentially contribute to disparities in expectations, preferences, and decisions regarding career paths.

Second, we contribute to the nascent literature that tests whether there is differential treatment of men and women in information-seeking settings. The most closely related papers are correspondence studies by Milkman et al. (2015) and Kalla et al. (2018). Milkman et al. (2015) uses fictitious prospective PhD students to send emails to faculty members asking about research opportunities, and finds that women and minorities are less likely to receive a reply than white men. The messages used in the study could be interpreted as seeking employment, information about employment, or both. Our study isolates the information seeking motive, by emphasizing that the student is not currently looking for job opportunities. Furthermore, our study analyzes the content of responses as well as response rates. Kalla et al. (2018) implements a large-scale experiment that uses fictitious students to send emails to local politicians seeking advice for a class project on how to become a politician. The study finds men and women are equally likely to receive a response. Our paper adds to this literature by focusing on gender differences in access to basic information about various career attributes. In addition, the messages are sent by real students, which emulates a realistic interaction that would occur on the professional platform, and allows us to explore the role of student preferences for professionals in generating gender disparities.

Last, this paper provides two advances to the literature that relies on correspondence studies to estimate discrimination. In a traditional resume study designed to estimate the effect of job applicant characteristics on callback rates, fictitious resumes with randomized applicant characteristics are sent to employers. One issue that has been raised regarding these studies is that—due to the fictional nature of the job applicants—employers are being deceived and their time is being wasted (Pager, 2007; Bertrand and Duflo, 2017; Kessler et al., 2019).[5] Our paper resolves this issue by incorporating real students interested in career information, seeking advice from real professionals. Recent work by Kessler et al. (2019) develops a new methodology called incentivized resume rating (IRR) for eliciting employer preferences for applicant characteristics, also without deceiving employers. In the IRR design, employers are asked by researchers to rate resumes and

[5]See Lahey and Oxley (2018) for empirical estimates of time spent reviewing resumes.

are incentivized to truthfully reveal their preferences. Both IRR and the methodology in this paper respect employers'/professionals' time. Our methodology additionally preserves the broad reach of a traditional correspondence study and does not require direct recruitment of employers/professionals.

Our methodology does require the recruitment of real students/job applicants. While incorporating real students cedes precise control over student attributes, we take several steps to ensure "all else is equal." First, we recruit students from similar majors and similar years in college, who are interested in career information in the four broad career paths. Second, we strictly limit other personal, academic, and professional information on the student's profile on the platform to ensure that the students are perceived as otherwise similar, aside from their gender. Third, in the regression specifications we control for all directly observable information on students' profiles. Fourth, we test the robustness of the results to the inclusion of student characteristics that could be inferred from the profile or observed elsewhere online. Finally, we find similar results when we limit the sample to students with no online presence aside from their profile. As we discuss in Section 4, our main results are robust to all of these specifications, as well as to running the main specification without any controls and to various classifications of student gender.

Because audit and correspondence studies rely on randomization to estimate discrimination, the estimates are only informative of discrimination on average (Heckman, 1998; Bertrand and Duflo, 2017). In reality, students are unlikely to contact professionals at random. A further innovation of the present paper is to incorporate information on student preferences for the professionals they want to interact with to quantify the role of selection into informal interactions in attenuating or amplifying average gender differences in information provision. To our knowledge, this paper is the first to incorporate agent selection into a correspondence-style study. A review of 80 audit and correspondence studies yields no other paper that accounts for agent selection.[6] The role of selection is potentially important in other settings, such as the labor market, retail market, and credit market. Future correspondence studies might consider incorporating the search behavior of real individuals to complement their estimates of average discrimination.

2 Experimental Design

To isolate the causal effect of student gender on the information students receive regarding careers, we implement a large-scale field experiment in which college students solicit information from

[6]The list of studies is available upon request.

professionals on an online professional networking platform.

Process: From February 2020 to June 2020, we recruited 100 college students at a large research university to send messages to 10,000 professionals. We advertised the study using email lists for the undergraduate economics, public policy, and math majors, extracurricular clubs related to economics, and undergraduate economics courses. The advertisement was targeted to students interested in career advice. Students interested in participating were asked to fill out a background survey, in which we asked for basic demographic information as well as whether the student was interested in receiving information on four career paths that undergraduate economics majors commonly choose post-graduation: finance, management consulting, data science, and law. We selected students who expressed interest in receiving information on the career attributes of these fields.[7]

In an in-person or virtual meeting, each student participant was guided through the process of creating a profile on a popular online professional networking site.[8] Almost 90 percent of students already had a profile on this platform and it is common for students to use this platform to reach out to professionals for career information/advice.[9] We asked that each student restrict their profile to minimal information, including their first name and last initial, student status, university affiliation, start year and anticipated year of graduation, college major, and the number of network connections they have on this platform. Students who already had a profile were asked to temporarily remove other information from their profile for the three-week duration of the study. Since it is not possible to control for the informational content embedded in students' own profile photos, which would have compromised the internal validity of the study. Our choice was therefore between between profiles without a photo—which could lead to professionals suspecting the profile is a bot/fake—and using a uniform photo across all student participants. In order to generate a realistic profile and increase response rates, we provided students with the same photo of an iconic university building to use as a profile picture (see Appendix Figure D1). We confirmed that students created a profile with the requisite restrictions through profile screenshots and independent verification on the site.

The pool of professionals consists of approximately 10,000 individuals on the site with work experience in the fields of finance, management consulting, law, or data science. The professionals were found through a search of the professional networking site for individuals who work in the

[7]See Appendix E.1 for the background survey.

[8]See Appendix E.2 for the instructions provided to students.

[9]In a survey of students from the same university interested in the four career paths, we find that 50 percent of students have used this platform to contact professionals for career information/advice.

students' metropolitan area, who have work experience in at least one of the four fields, and who have a degree from a U.S. News and World Report top-40 ranked university.[10] We used the list of 10,000 professionals to create sets of 100 randomly assigned professionals to provide to student participants.[11] Professionals were stratified by field. Within each field, professionals were randomly assigned a message type and the student who would contact them. Each student was given a list of 100 professionals to contact: 13 data scientists, 28 finance professionals, 33 lawyers, and 26 management consultants. These proportions reflect the composition of professionals that came up in a search of the site. The students chose neither the professionals whom they contacted nor the messages sent. Figure 1 provides a graphical depiction of the experimental design.

We provided the text of the initial message that students sent to professionals. Each professional-student communication used one of four message types, which were designed to emulate a conventional request for career information during an informational interview.[12] To test whether different career attributes are emphasized to male and female students, students sent a broad message that asked about the pros and cons of the professional's field.[13] To test whether there are gender differences in information received, conditional on bringing up a specific career attribute, student sent three message types: (a) a specific question asking whether work/life balance is a concern in the professional's field, (b) a specific question asking whether competitive culture is a concern in the professional's field, and (c) a factual question asking about billable hours at a large law firm (sent only to law professionals). We selected the above career attributes based on documented gender differences in preferences for competitive environments and temporal flexibility (Goldin, 2014; Wiswall and Zafar, 2018; Cubas et al., 2019; Niederle and Vesterlund, 2011; Flory et al., 2015). Note that the factual question differs from the specific questions in that the answer to this question should not depend on student gender, since billable hours are a contractual obligation invariant to employee characteristics. We designed the specific questions asking whether work/life balance or competitive culture is a concern to elicit a yes or no response, which we analyze below

[10]Professionals' profiles were checked to ensure they have work experience in one of the four fields.

[11]The random assignment took place prior to the recruitment of students, meaning that we could not add professionals to the experiment after the experiment started. Because we chose the 10,000 professionals who best matched our criteria, after the experiment started we could not add additional professional to the sample. If we had added to the professional sample after the experiment started, the characteristics of the new professionals would have been systematically different from the original 10,000.

[12]These messages were based on suggested wording from a university career center guide on informational interviews. See pages 10 and 11 of `https://career.ucla.edu/Portals/123/documents/career%20guide/UCLA_CareerGuide_2019-2021.pdf`.

[13]Gallen and Wasserman (2021) provides evidence from a student-alumni professional networking website that 64 percent of career-related messages ask broadly about the professional's career path. There is no gender difference in the propensity to ask this question.

in the response content.[14]

All message templates emphasize that the student is only seeking career information, as well as explicitly state that the student is not searching for a job. Message templates are in Appendix Figure A1. To summarize, the four message types are:

1. Broad: Asks broadly about the pros and cons of the professional's field.

2. Specific work/life balance: Asks if work/life balance is a concern in the field.

3. Specific competitive culture: Asks if competitive culture is a concern in the field.

4. Factual hours (law only): Asks what the billable hours requirements are at a large law firm.

Before sending any messages, students were asked to spend 20 minutes studying the profiles of professionals they would be messaging and provide three sets of rankings. Specifically, we ask them to rank the five professionals they would be most interested in asking about the pros and cons of the professional's field, work/life balance in the professional's field, and workplace culture in the professional's field. Students were informed that these rankings would not affect the next step of the study, in which students sent messages to all 100 professionals in their list.

For data science, management consulting, and finance professionals, students sent half of the messages using the broad question and one-quarter of the messages using each specific question. For law professionals, each student sent 44 percent of the messages using the broad question, 22 percent using each specific question, and 12 percent using the factual question. Within each field, professionals were randomly assigned a message type. Each professional received only one message.

In order to estimate the causal effect of student gender on career information received, we randomized whether a professional was sent a message from a male or a female student as well as the specific message type (as depicted in Figure 1). The students sent the messages on weekdays during typical working hours.[15] When a message is sent to a professional, depending on the professional's site preferences, they receive an email notification, an app notification, and/or an alert on the website. After a few days, the site automatically generates a reminder email notification of the message if the professional has not yet responded to the request.

[14] Although the specific questions describe the career attributes in a negative light, we note that professionals were willing to refute the concern or say "it depends," especially in response to the competitive culture question (Appendix Table A1).

[15] In some cases, students were unable to send all 100 messages in one sitting. In these situations, we asked that the students send the messages as soon as they were able to do so. We recorded the actual date and time that each message was sent.

Students were asked to provide the initial responses they receive within 21 days of sending the messages.[16] In order to verify that we obtained all initial responses received, we asked students for screenshots of their message inbox as well as screenshots of each response. If a professional responded, the student could choose whether he or she would like to continue the interaction. We emphasized to students that we would not ask for detailed information on these follow-up interactions. As an indication that we selected students based on their genuine interest in career advice, 34 percent of students reported that they planned to stay in touch with at least one of the professionals who responded. Students were asked to not use the site for activities unrelated to the study for the three-week period. We independently verified that students did not change their profile or otherwise engage in site activity throughout the study period. Three weeks after sending the messages, we followed up with the students to ensure that we had received all of their initial responses. To assess the role of information received on students' future career choices, three weeks after sending messages, students filled out a survey with their career intentions. Upon successful completion of this survey, students were paid $75.

Methodological advance and identification: In several ways, our experimental design resembles a traditional correspondence study in which researchers send fictitious resumes to employers in order to estimate the causal effect of job applicant characteristics on callback rates. In these studies, the resume format, the information provided in the resume, and other aspects of the correspondence are controlled by the researcher. The advantages of creating fictitious applicants are numerous: the researcher has precise control over applicant attributes and avoids dealing with the complexities of the characteristics/behaviors of real people. By design, the applicant characteristic of interest is orthogonal to other applicant characteristics as well as to employer characteristics. In addition, the study is generally low cost and logistically straightforward to implement (Pager, 2007; Bertrand and Duflo, 2017).

In our study, we similarly maintain precise control over the text of the messages sent to professionals, and student characteristics are orthogonal to professional characteristics. In contrast to a traditional correspondence study, we incorporate real students who are interested in information on careers. The incorporation of real students poses some challenges with regard to identification of the causal effect of student gender, however, since we cede control over the attributes of students. In particular, we cannot ensure that other student characteristics are orthogonal to student gender.

In several ways, the online setting allows us to mitigate concerns that other student

[16]The vast majority of responses are received within two weeks of sending a message.

characteristics confound the effect of student gender: (1) as discussed above, we ask students to strictly limit the information provided on their profiles, (2) in our regressions, we control for all student characteristics that are directly observed on the site, (3) using information from the background survey and whether the student has an online presence aside from their profile, we test whether the effect of student gender is sensitive to the inclusion of student characteristics that could be inferred from the profile (e.g. race/ethnicity) or observed elsewhere online, and (4) we examine whether the results are robust to restricting the sample to students without an online presence.

3 Data and Econometric Framework

3.1 Data

We collect data on response rates and the text of initial responses. We analyze the text using manual classification, sentiment analysis, and natural language processing tools that characterize word distributions. For responses to the broad question, manual classification entails coding whether the response mentions work/life balance or competitive culture. To manually classify messages, we employed five research assistants and gave each batches of messages to code (11 batches total). Each batch contained only de-identified message text and *no other information* except a random code which would later allow us to merge the classifications with the characteristics of senders and recipients. Messages were de-identified prior to being put into batches by replacing any names with an X. The messages were coded by at least two RAs who were working as research assistants for course credit. The researchers then verified and reconciled the codes in case of disagreement. All coding and reconciliation was done without knowing the characteristics of message senders and recipients.

Professionals' mentions of work/life balance were coded using the following definition from the Cambridge dictionary: "The amount of time you spend doing your job compared with the amount of time you spend with your family and doing things you enjoy."[17] This includes explicit references to work/life balance, as well as discussion of the hours worked per week, extent of work-related travel, and conflict between/accommodation of work responsibilities and other life priorities. One paraphrased example of a work/life balance mention is: "Management Consulting can be considered a lifestyle since it requires travel, very long hours, always being on, and client-specific knowledge."

We code professionals' mentions of competitive culture when the response explicitly mentions

[17] https://dictionary.cambridge.org/us/dictionary/english/work-life-balance

competition within the workplace or among coworkers. Due to the low frequency of mentions of competitive culture, we also create a broader metric of workplace culture, which includes descriptions of interpersonal relations among colleagues, the work environment, or ethical issues in the workplace. One paraphrased example of a culture mention is: "Though this is changing, finance sometimes still depends on connections, bribes, or corruption."

For the responses to the specific questions, which were designed to elicit a yes or no, we manually classify whether the response confirms that work/life balance or culture is a concern, refutes that it is a concern, or says "it depends" on factors such as the company or more granular occupation. In addition, we hire undergraduates (who are not experiment participants) to provide their subjective evaluations of the tone of all responses, specifically whether the response would cause a typical undergraduate student to be more or less concerned about work/life balance or workplace culture in the professional's field. For the responses to the factual hours question, we manually extract the hours requirement, which is a numerical value of hours or numerical range of hours. For answers with a range of hours, we take the midpoint of the range.

To analyze the role of professional attributes in generating gender differences in information received, we collect publicly available information on professionals on this site, including their education, gender, and network thickness. We use profile pictures and textual information to assign the gender of each professional. In cases where a picture or text-based information on gender was not available on their profile, we assign gender based on the professional's first name using U.S. Census and Social Security Administration name files. This process successfully classified gender for 99.5 percent of professionals.

3.2 Sample restrictions

The study recruited 100 college students to send messages to approximately 10,000 professionals. One student (and 100 professionals) was used for a pilot and is excluded from the analysis. We discuss this pilot in detail in Appendix B. Five students withdrew due to unforeseen logistical issues with their profiles or with sending the messages. Of the 94 students who were able to successfully create a profile and send messages, 89 students provided data on the responses they received. The five students who dropped out after sending all of their messages constitute sample attrition. Once the experiment started, we could not add more professionals to the sample. Because we chose the 10,000 professionals who best matched our criteria, if we had added to the professional sample after the experiment started, then the characteristics of the new professionals would have

been systematically different from the original 10,000.[18] We diligently followed up with all student participants and found that students who took a very long time to provide responses (>4 months) had similar response rates to those students who completed the study promptly. This fact makes us less concerned that students who dropped out or who did not reply after sending messages did so because of the replies they received.

Since we intend to estimate the causal effect of student gender on information received, we limit the main analysis sample to students whose first names unambiguously convey their true gender. We note that all results are robust to including students with gender ambiguous names. We limit the main sample using the U.S. Census and Social Security Administration name files. If a student's name is at least 90 percent male or female, and coincides with the student's actual gender, then the student is included in the main analysis. This sample restriction drops 13 students. Our final sample for the analysis consists of 76 students who contacted 7,602 professionals across four career categories.[19]

3.3 Summary statistics

Summary statistics for the students in the final sample are reported in Table 1, overall and by student gender. The top panel presents student attributes that are visible on or can be easily inferred from the student's profile. Among all students, 58 percent are female. The students are primarily freshmen and sophomores, and 62 percent are economics majors. The substantial representation of economics majors is consistent with our recruiting strategy and the fact that the four career paths chosen are those that economics majors primarily enter post-graduation.

Many students already had profiles on this platform, which is reflected in their number of network connections. Other students—14 percent—created a profile for the first time through this study. We also record whether students had any information on their profile beyond what is listed in the top panel of this table or had another profile issue that precluded perfect compliance with the profile restrictions. In general, these profile issues were limited to minor deviations from protocol such as a few activities or skills being visible on the profile. We control for the deviations from profile restrictions in all regressions.

In a background survey that students filled out prior to sending messages, we collected

[18] Once a student started the message sending process, we also could not assign the professionals on their list to another student (we did not want to contact any professionals twice).

[19] Our final sample of professionals was 10,003, so three students were assigned 101 professionals. Two of these students are in the final student sample.

information on student attributes that are partially observed based on profile information, may be found elsewhere online, or correlated with information found online. For example, student race/ethnicity may be inferred from students' names and first generation college goer could be correlated with the extracurricular activities students are involved in (Jack, 2019). Students are evenly split between race/ethnicity categories and 22 percent are first generation college-goers. The majority of students have some online presence aside from their profile on this site. While male and female students are overall similar, we observe that female students are less likely to be economics or STEM majors, have fewer network connections, and are more likely to identify as Asian/Asian American.

Table 2 reports summary statistics for professionals, overall and by field. One-third of professionals are female, and this varies substantially across field, with representation the lowest in finance and the highest in law. The professionals are, on average, in their late 30s. Professionals were selected based on their attendance of a top-40 U.S. News and World Report university for some part of their education and this is reflected in the selectivity of undergraduate institutions and the substantial fraction who attended an Ivy League university. More than 20 percent of professionals are alumni of the student's college, with a lower fraction among lawyers. The majority have well-established networks on this site.

Appendix Table A1 presents summary statistics for the main outcomes, including response rates, response length, and mentions of work/life balance and workplace culture in responses to the broad question. The overall response rate across all question types is 12 percent, with a lower rate of response to the broad question (10 percent) and the highest rates of response to the specific work/life balance and competitive culture questions (14 and 15 percent, respectively).[20] In Appendix Figure A2, we observe that the distribution of response rates is centered around 12 percent. Responses are 435 characters, on average, with longer responses to the specific work/life balance question and shorter responses to the factual hours question. Among responses to the broad question that asks about the pros/cons of the professional's field, 11 percent bring up work/life balance issues and 12 percent mention workplace culture.

Appendix Table A2 reports results from tests of covariate balance. For each professional characteristic, we run a regression of this characteristic on whether the student who sent the professional a message is female. Professional characteristics are balanced across students, indicating

[20]This response rate is higher than a correspondence study that sends pitch emails to venture capitalists but lower than studies that send emails to politicians or academics (Gornall and Strebulaev, 2020; Kalla et al., 2018; Milkman et al., 2015).

that the randomization was successful.

3.4 Econometric framework

In order to estimate the causal effect of student gender on information received, we use the following regression specification:

$$Y_m = \alpha + \beta StudentFemale_m + X_m'\gamma + \epsilon_m \tag{1}$$

where the dependent variable, Y_m, is an outcome such as an indicator for whether message m received a response, or whether the response mentions a specific career attribute.[21] The independent variables are an indicator for whether the message was sent by a female student, $StudentFemale_m$, as well as a vector of message and student controls, X_m. In our baseline specification, we include controls for message characteristics: categorical variables for the day of the week and the time of day that the message was sent, a linear term for the date that the message was sent, and the field of the professional. We also include controls for the characteristics of the students who sent the messages that are directly observable on the site: college major (economics, STEM, other), expected college graduation year, number of network connections, and whether the student was completely compliant with the profile restrictions.[22] Standard errors are clustered at the student level.

Our selection on observables design is well suited to the online setting, where we control what the professionals see about students and can directly include these characteristics as controls in the regression. We also report specifications without any controls and find similar results. We do not use this as our main specification because there are observable differences between male and female students visible to professionals on students' profiles (Table 1).

We test whether the coefficient on student gender is sensitive to the inclusion of additional student characteristics that may be available elsewhere online: student race/ethnicity, college GPA, first generation student status, and an indicator for whether there is information publicly available on the student through an online search. College GPA and first generation student status are known to be correlated with the extracurricular activities students are involved in, which may be visible online (Jack, 2019). Since student race/ethnicity could also be conveyed through the student's

[21] We note that because each professional received exactly one message, indexing observations at the message level is interchangeable with indexing observations at the professional level.

[22] Some students who had profiles prior to the experiment were unable to completely remove all information from their profile. This extra information may include site activity, relevant labor market skills, and extracurricular activities.

name, we consider this variable partially observed and estimate a separate specification to test sensitivity of the main results to this specific control. As an additional check on whether students' online presence confounds the results, in the Appendix we limit the sample to students with no online presence aside from their profile on the professional platform and find similar results.

4 The Information Students Receive Depends on Their Gender

In this section, we document that student gender affects the information that students receive. When students ask professionals about the pros and cons of their career path, male and female students are equally likely to receive a response. Responses to female students, however, are substantially more likely to mention work/life balance issues.

4.1 Response rates

We start by testing whether student gender affects response rates to the broad question that asks about the pros/cons of the professional's field. We estimate Equation (1), and use as the dependent variable an indicator for whether a message received a response from the professional. The results are reported in Table 3, columns 1–3. Column 1 presents the results with the baseline message and student controls. We observe that response rates to male and female students are very similar; the coefficient on $StudentFemale_m$ is 0.011 and statistically insignificant. Consistent with the notion that the effect of student gender is not confounded by other student characteristics, when we include the supplemental student characteristics that may be observable elsewhere online, the coefficient on student female exhibits little change (columns 2 and 3).[23] The results are robust to the inclusion of students with ambiguously gendered names (Appendix Tables A3 and A4) and to not including controls (Appendix Table A5). Based on these results, we conclude that professionals are just as willing to engage with male and female students when they ask a broad question.

4.2 Response content

Given that response rates to male and female students are similar, we next analyze whether there are gender differences in the content of the responses to the broad question. As mentioned in

[23]Heckman and Siegelman (1993) raise the possibility that in correspondence studies, differences in the variance of unobservable productivity could explain differences in mean callback rates. We test for this in our setting using the methodology developed by Neumark (2012) and find that we cannot reject that the variance of unobservable characteristics of male and female students is the same.

the Introduction, we focus on two career attributes that are known to differentially affect the occupational and job choices of women relative to men: work/life balance and competitive culture. We restrict the sample to responses received, estimate Equation (1), and use as the dependent variable an indicator for whether the response mentions a work/life balance issue, including work hours, travel, lifestyle, or family/personal life considerations. Table 4 reports the results. Responses to female students are more than twice as likely as responses to male students to mention work/life balance issues. Among responses to male students, 6.7 percent mention a work/life balance issues. Using the estimates in column 1, the rate for female students is 8.7 percentage points higher. Controlling for student characteristics that are not directly observed on the site does not affect the results. When we account for the small gender differences in response rates to the broad question, the rates of mentioning work/life balance are similarly differentiated by student gender (see Appendix Table A6). The inclusion of students with ambiguously gendered names does not change the results (Appendix Tables A7 and A8), nor does the exclusion of controls (Appendix Table A9).

The greater emphasis on work/life balance to women is not driven by differences in the types of professionals who respond to male and female students. We modify Equation (1) to include controls for professionals' gender, undergraduate graduation year, undergraduate institution selectivity, network thickness, whether the professional is an alumnus of the student's college, and whether the professional has a graduate degree. In Table 5, across all outcomes, the coefficient on $StudentFemale_m$ is invariant to the inclusion of professional controls. Table 6 further explores whether the differential provision of information to female students is concentrated among certain subgroups of professionals. Each entry represents the coefficient on $StudentFemale_m$ from a separate regression, with the subgroup of professionals in the column title. Female professionals, alumni, those who are older, professionals with a degree from an Ivy League university, and those in finance and law are more likely to emphasize work/life balance to female students (panel D).[24] This heterogeneity will be an important ingredient in understanding how student preferences for certain types of professionals may impact the information they receive, as discussed in Section 5.2.

In Appendix Table A10 we estimate whether the additional information provided on work/life balance to female students is driven by three specific topics: (1) the duration of a typical workweek, (2) flexibility of work schedules, and (3) the ability (or inability) to work from home.[25] Responses

[24]While prior work has demonstrated that female students' educational and occupational outcomes are affected by the gender of their teacher, mentor, or role model (Carrell et al., 2010; Porter and Serra, 2020; Canaan and Mouganie, 2019), we do not find evidence that female professionals differentially respond to female students relative to male professionals. We are underpowered to detect same-gender match effects for mentions of work/life balance.

[25]These topics correspond to two O*NET work context categories and one researcher-defined category. For more

to female students are more than twice as likely (5.4 percentage points more likely) to mention the duration of the typical workweek relative to male students. It also appears that responses to female students are more likely to contain information about work schedule flexibility, but this gender difference is not statistically significant.

Mentions of work/life balance tend to be negative. Below are paraphrased examples of responses that mention work/life balance:

> [Law] A career in law opens many doors...and also offers long hours, hard work, firm deadlines, and many challenges.

> [Finance] Challenges can be the hours depending on the area of finance (corporate finance FPA, consulting, investment banking, or even accounting).

Using subjective evaluations from a team of college students who were not study participants, we characterize the anticipated effect of the responses. In particular, we ask the students to rate the extent to which a response would make a typical college student more or less concerned about work/life balance (workplace culture) in the professional's field. Based on the students' evaluations, responses containing mentions of work/life balance increased concern about this issue more than 75 percent of the time. Only three percent of such responses made students less concerned about work/life balance. When we consider all responses, we find that responses to female students are more likely to increase concern about work/life balance, but this contrast is not statistically significant (see Appendix Table A11).

We also investigate whether the additional information on work/life balance that women receive crowds out other, potentially useful, information on careers. We find no significant gender differences in the overall length of replies, suggesting that the additional emphasis on work/life balance to female students may displace other information (Appendix Table A12). Although our experimental design and analysis focus on two career attributes—work/life balance and workplace culture—in Online Appendix B, we also provide an exploratory analysis of gender differences in other message components, using manual classification, nonparametric natural language processing, and lexicon-based sentiment analysis.

Finally, we estimate gender differences in mentions of competitive culture in responses to the broad question. Competitive culture is mentioned in only six messages, but the rates of mentioning it are similar to male and female students (Appendix Table A13). Due to the extremely low frequency

information on these categories, see Online Appendix Table C.1.

of mentions of competitive culture, we also test for gender differences in mentions of workplace culture more generally, which includes descriptions of interpersonal relations among colleagues, the work environment, or ethical issues in the workplace. Twelve percent of all responses to the broad question mention workplace culture (Appendix Table A1). Across specifications, the point estimates for the coefficient on $StudentFemale_m$ are close to zero (Table 4 columns 4–6 and Table 5 columns 5 and 6).

To summarize, when students ask a broad question about the pros and cons of a career, we find that professionals differentially mention work/life balance issues to female students. In contrast, professionals bring up competitive culture (or culture more broadly defined) at similar rates to male and female students. Below, we investigate the sources of gender differences in access to work/life balance information. Before doing so, we ask: despite the well-documented gender differences in preferences for competitive environments, why don't professionals differentially bring up workplace culture issues to female students? There are several possibilities. One possibility is that even if professionals are aware that competitive culture disproportionately affects women, they may be reluctant to discuss sensitive or controversial issues (as well as issues that could make their firm vulnerable to legal action) during an initial written message exchange (Sockin and Sojourner, 2022). Consistent with these reasons, only three responses to the broad question mentioned workplace culture issues overtly related to women, such as sexism or sexual harassment. Another possibility is that the gender differences in competitiveness are less salient for professionals who selected into these competitive careers, and to students who seek to pursue them. Evidence for this possibility comes from Buser et al. (2014) and Reuben et al. (2015), which find that more competitive individuals are more likely to select into more prestigious tracks/high paying industries. In contrast, professionals may view work/life balance issues as tied to motherhood and expect that even women who select into these jobs may have work/life balance issues in the future.

5 Can Students Get the Information They Want?

5.1 Conditional on communicating the same preference, gender differences persist

When students seek basic information about the pros and cons of pursuing a particular career path, professionals are more than twice as likely to provide information on work/life balance to female students relative to male students. This differential information provision is consistent

with professionals lacking information about the preferences of individual students for career attributes, and thus statistically discriminating based on coarse observable student characteristics. Professionals may believe that women, on average, care more about work/life balance than men, and provide information accordingly. In this scenario, if professionals had information on each student's preferences for discussing career attributes, then the gender gap in information received—conditional on these preferences—would be eliminated. In contrast to this prediction, this section demonstrates that when students specifically state their interest in learning about work/life balance, male students continue to receive less information about work/life balance than female students.

Response rates

Recall that there are three message types that ask about specific career attributes: the message that asks whether work/life balance is a concern ("specific work/life balance"), the message that asks whether competitive culture is a concern ("specific competitive culture"), and the message that asks the minimum billable hours requirement for a first-year associate at a large law firm ("factual hours"). In Table 3, we investigate gender differences in response rates to these three questions. In columns 4–6, we find that, in contrast to the broad question, student gender does affect professionals' propensity respond to the specific work/life balance question. Considering the baseline specification from column 4, female students are 3.7 percentage points, or 28 percent, more likely to receive a response relative to male students. Furthermore, when students ask a fact-based question related to work/life balance ("factual hours"), female students receive 80 percent more responses than male students (Table 3, columns 10–12), though the coefficient is not consistently statistically significant. Taking these results together, we find that even when students specifically request information on work/life balance, male students receive less of it.[26]

Consistent with the result that workplace culture is not differentially emphasized to female students in the broad question, there is no gender difference in response rates to the specific question on competitive culture (columns 7–9). This result is not driven by professionals' unwillingness to engage with students on this topic; in fact, the specific competitive culture question had the highest response rate.[27]

[26]The composition of professionals also does not explain this gender difference (Appendix Table A14).

[27]These results are robust to the inclusion of students with ambiguously gendered names (Appendix Table A3).

Response content

Next, we characterize the content of responses to the questions on specific career attributes. Overall, the responses to the work/life balance question confirm that work/life balance is a concern in the professional's field and make students more concerned about this issue (Appendix Table A1). Only seven percent state that work/life balance is not a valid concern. The responses to female students do not display meaningful content or tone differences relative to the responses to male students (Appendix Tables A11, A15, and A16). Although the specific questions describe the career attributes in a negative light, we note that professionals were willing to refute the concern or say "it depends," especially when responding to the competitive culture question (Appendix Table A16).

Turning to the factual hours question, the point estimates suggest that women are quoted higher hours requirements for associates at large law firms. Unfortunately, we are underpowered to detect large differences (Appendix Table A17). We also note that because student gender affects response rates to the specific work/life balance question and the factual hours question, the effect of gender on the content of responses may be driven by the marginal respondents to female students or by differential treatment of male and female students.

To summarize, when students ask broadly about the pros and cons of a career path, we find that professionals differentially emphasize work/life balance to female students. If this gender disparity were rooted in statistical discrimination arising from professionals' limited information about students' preferences, then providing professionals with student preferences for career attributes would eliminate the gender gap. Instead, when students state their interest in learning about work/life balance, we find that male students continue to receive less information on this topic than female students. We conclude that a straightforward intervention to reduce reliance on student observable characteristics does not equalize the information that male and female students receive.

5.2 When students select professionals, gender differences persist

Conditional on asking professionals about work/life balance, male students receive less information. But what information do students want? Using a survey of students from the same university, we document gender differences in demand for career information. The survey asks students how they would allocate 15 minutes of time with a professional in their preferred career path among various career-related topics.[28] While all students are interested in discussing work/life balance with a

[28] See Appendix E.3 for the follow-up survey.

professional, men allocate 14 percent of their time and women allocate 10 percent of their time to this topic, a statistically significant difference (Figure 2).[29] This suggests that the information professionals supply does not match the information students demand.

If students could choose which professionals they contact, would male and female students gain access to the information they want? Our study—like every correspondence study—relies on random assignment to estimate discrimination, on average. However, it is reasonable to believe that individuals do not sample randomly from their full choice set.[30,31] To understand how student selection of professionals could alter conclusions about gender differences in access to information, we develop a methodology to combine our experimental estimates with student preferences for professionals. To our knowledge, this is the first paper to incorporate agent selection into a correspondence-style study.[32] More generally, correspondence studies can incorporate our methodology for combining experimental estimates with information on agent preferences to better understand whether average bias is reflected in agents' experiences.

Using our methodology, we find that allowing students to choose professionals does not align the information they receive with the information they demand.

Estimating student preferences for professionals

Before students sent messages, we asked them to rank professionals in terms of whom they would most prefer to ask the questions in the study.[33] Students were told that these rankings would not impact the messages they send in the study and were purely for the researchers to learn about their preferences over the professionals. Students were also told that they could choose the same five professionals for each of the three rankings, different professionals, or a combination thereof—

[29] A recent New York Times article also notes this gender reversal in preferences for work/life balance among 18-29 year olds. `https://www.nytimes.com/2019/09/17/style/generation-z-millennials-work-life-balance.html`

[30] For example, Pager and Pedulla (2015) documents that minority job seekers search more broadly for jobs than their nonminority counterparts. Abel (2017) further shows that immigrant job seekers are more likely to search for jobs farther away if they live in areas with higher levels of discrimination. Using a large-scale correspondence study, Agan and Starr (2020) find that Black-sounding job applicants experience more discrimination in less-Black neighborhoods, and simulate how equilibrium racial discrimination is affected by residential sorting.

[31] In the labor market, the extent to which minorities apply to discriminatory firms determines wage gaps, not the bias of the average employer (Becker, 1971; Heckman, 1998; Charles and Guryan, 2008). In a search framework, average discrimination may generate wage gaps (Black, 1995), but the magnitude of such differences depends on both supply-side and demand-side behavior.

[32] Another potential dimension of selection is selection into asking questions. In a survey of students from the same university, we find that male and female students have solicited information on work/life balance at equal rates in the past. In addition, Gallen and Wasserman (2021) document using data from an online student mentoring platform that there are few gender differences in questions asked.

[33] We did not ask students to rank the lawyers whom they would want to ask about billable hours requirements—there is no factual question ranking.

whatever reflected their preferences.

Using these student rankings, we estimate a rank-ordered logit choice model for student preferences over professional characteristics (Beggs et al., 1981).[34] We separately estimate male student preferences over professional characteristics and female student preferences over professional characteristics.

Appendix Table A20 shows that the professionals whom students rank differ in a number of ways from the full sample, and moreover, female and male students rank different types of professionals.[35] The table displays the average characteristics of all professionals ranked by students, those ranked by female students (column 2), by male students (column 4), as well as the average characteristics of professionals, weighted by the probability that a professional is preferred by a female student based on their characteristics according to a rank-order conditional logit model (column 3, titled "female predicted"), and the weighted average characteristics of professionals weighted by the probability that a professional is preferred by a male student based on their characteristics according to a rank-order conditional logit model (column 5, titled "male predicted"). The observation counts differ between the professionals actually ranked and those predicted to be preferred because we can do the prediction for all professionals in the sample, while each student only ranks five of the 100 professionals they were assigned.[36] Given these differences between the preferred sample and the full sample, the average gender gaps estimated in Section 4 may change when we incorporate student preferences. For example, preferred professionals are more likely to have attended an Ivy League university (16 percent of full sample and 20 percent of those ranked by students as preferred). Since those who received a degree from an Ivy League institution are more likely to differentiate their response to the broad question based on gender, we might expect the gender gap in mentions of work/life balance to expand when students choose professionals on their own.

[34]More specifically, we estimate student preferences over the following professional characteristics commonly observed on profiles: gender, binned undergraduate graduation year (1980s or earlier, 1990s, 2000s, 2010s+, or no information on graduation year available), connections (binned in low, medium, and high), an indicator for whether the professional is an alumnus of the student's college, undergraduate institution selectivity quartile, whether the professional has any graduate degree, and an indicator for whether the professional has any degree from an Ivy League institution.

[35]Relative to the full sample, preferred professionals are younger, more likely to be female, and more likely to have a degree from an Ivy League institution. The professionals preferred by female students relative to those preferred by male students are more likely to have a law or management consulting background, are more likely to be female, and are more likely to have a graduate degree.

[36]Here we display the results of the preference estimation based on rankings for the broad question. Results are similar if we use student rankings for the other questions.

Econometric framework for incorporating student preferences

We estimate the following specification:

$$Y_p = \alpha^g + \beta^g StudentFemale_p + \gamma^g X_p + e_{jp} \tag{2}$$

where Y_p is an outcome of a message sent to professional p, $StudentFemale_p$ indicates whether the professional p actually received a message from a female student in the main experiment, and X_p is a vector of characteristics of the student who sent the message to the professional as well as message-level controls such as time and date the message was sent, analogous to our main specification.[37]

We estimate two versions of this model. One in which the observations are weighted by the predicted probability that a professional is ranked by a male student, and another in which observations are weighted by the predicted probability that a professional is ranked by a female student. Weighting by the propensity to be preferred by a student of gender g gives us coefficients β^g where g indicates whether the regression is weighted by the preferences of male ($g = m$) or female ($g = f$) students. The predicted probabilities are estimated as described above using a rank-ordered conditional logit model over professional characteristics. This method is analogous to re-weighting non-representative surveys due to oversampling in order to obtain population level means and statistics, as described in Särndal et al. (1992). We can think of these estimates as being weighted so that the sample is non-representative of the overall sample of professionals, but instead representative of the sample of preferred professionals. In the same vein, the main estimates in the paper are weighted by the inverse propensity of being selected, yielding population-representative averages.

We emphasize that the weight given to a professional preferred by a student of gender g is orthogonal to the gender of the student who contacted the professional, since students were randomly assigned professionals to contact. We can therefore identify both the gender bias of the professionals preferred by male students (β^m) and the gender bias of the professionals preferred by female students (β^f). Note that with observational data, it is not possible to identify β^f and β^m unless all professionals are contacted by observationally equivalent male and female students. In addition, we can use the propensity weights to summarize the average response to male students among professionals preferred by male (α^m) and female (α^f) students. See Online Appendix D for a potential outcomes framework that formalizes identification.

[37] We note that because each professional received exactly one message, indexing observations at the professional level is interchangeable with indexing observations at the message level.

Gender differences in information received, incorporating student preferences

Allowing students to choose professionals does not eliminate gender differences in information received on work/life balance. Table 7 reports the results for mentions of work/life balance and workplace culture in responses to the broad question. Recall that in our main experiment, professionals were 7.2 percentage points more likely to mention work/life balance issues in their responses to female students relative to male students. Incorporating student selection amplifies this gender difference. In column 3, we observe that female-preferred professionals are substantially more likely to mention work/life balance to female students (coefficient on $StudentFemale_p$ is 0.097). This gender bias is similar for male-preferred professionals. Female-preferred professionals bring up work/life balance at approximately the same rate to male students as male-preferred professionals (0.057 vs. 0.051 in columns (3) and (4) of Table 7). In columns 1 and 2, we find that female students tend to select professionals with lower response rates overall, but the difference is negligible.[38]

6 Implications of Gender Differences in Information Received

In this section, we discuss potential implications of gender differences in information received. We empirically investigate the ramifications of information for students' career choices. We also discuss a few normative implications of gender gaps in information provision.

Effect of information on career choice: We provide suggestive evidence that the information students receive matters for their career preferences. Because professionals are randomly assigned to students, we are able to assess the effect of being randomly assigned a professional who brings up work/life balance (conditional on student observables) on students' career plans.

At the conclusion of the study, we surveyed students about their career plans. Of the 76 students in the main sample, 73 completed the survey. We measure whether a student indicates he/she is, relative to the start of the study, less likely to enter his/her preferred career path. Students were asked, "Relative to when you began sending messages for this study, are you, on a scale of 1–10, much less likely (1) ... much more likely (10) to go into [data science/finance/law/management consulting]?" We run the following regression:

$$L_i = \gamma_0 + WL_i'\gamma_1 + X_i'\gamma_2 + M_i'\gamma_3 + P_i'\gamma_4 + \varepsilon_i \tag{3}$$

[38] Appendix Table A21 presents analogous results for response rates to the specific work/life balance and competitive culture questions.

where L_i, is an indicator for responses of 4 or below to this question for student i's preferred career path or the reverse of the scale described above (where 10 indicates least likely). WL_i is vector including an indicator for whether a student received a response to the work/life balance question in their preferred career path, as well as an indicator for whether the student received a work/life balance mention in their preferred career path, X_i is a vector of student characteristics, including the basic student controls, race, and student's preferred career path, M_i includes characteristics of the messages received by student i, such as message length and whether the student received information about workplace culture, and P_i is a vector of the average characteristics of professionals the student contacted in their preferred career path.

Being randomly assigned a professional who brings up work/life balance deters students from their preferred career path. Table 8 shows that receiving receiving a response to the work/life balance question makes students less likely to go into their preferred career path. Similarly, point estimates suggest that receiving a response to the broad question which mentions work/life balance deters students from their preferred career path though this effect is insignificant. This result is similar whether the outcome variable is an indicator of being deterred or the continuous scale. These effects are robust to controlling for other message and professional characteristics (columns 2-3 and 5-6) and as well as other parameterizations of message content (Appendix Table A22).[39] Although the analysis is inconclusive due to the large standard errors stemming from the small number of students, the results suggests that the information professionals provide may matter for stated career choices. These results are consistent with a large literature demonstrating that individuals update their beliefs, stated choices, and actual choices when presented with information.[40]

Are students aware of the gender gap in information received? If students are aware that information is distributed based on gender, then it is possible for them to seek out information to address this disparity.[41] In our survey of the same student population, we find that 56 percent of students are unaware that women receive more information on work/life balance; male students are less likely to be aware than female students.[42] Because students are unaware of these information

[39]We note that most students are encouraged toward their preferred path. We focus on deterrence because the vast majority of work/life balance information that students receive is negative. When we look more holistically at the distribution of preferences, we can reject that the distributions of responses to the career plans question are the same for students who did vs. did not receive a mention of work/life balance (p-value of 0.07).

[40]See, for example: Nguyen (2008); Jensen (2010); Zafar (2011); Dinkelman and Claudia Martínez (2014); Wiswall and Zafar (2015a); Hoxby and Turner (2015); Wiswall and Zafar (2015b); Coffman et al. (2017); Wiswall and Zafar (2018).

[41]As Section 5.2 demonstrates, it may be difficult for students to access information by choosing which professionals they contact. But finding information may be possible using other methods, such as the internet.

[42]We thank Christine Exley for this helpful suggestion.

deficits, it is unlikely they are able to address them using other sources of information.

Normative implications: While this paper cannot provide direct evidence on the normative implications of gender differences in information received, our results suggest that there is, at a minimum, a mismatch between the information professionals supply and the information students demand. Based on the survey results described above, male students demand *more* information on work/life balance than female students. It is still possible that students are unaware of changes to their preferences in the future, and professionals respond with these future preferences in mind (Kuziemko et al., 2018). Indeed, the evidence on women's stronger preference for temporal flexibility focuses on individuals at age 30, while our sample spans ages 18-22 (Wiswall and Zafar, 2018). If professionals know students' preferences better than students know their own preferences, then gender differences in information provision could be optimal.

If professionals do not know or are biased in their beliefs about students' preferences, then gender differences information provision may be sub-optimal and labor market inequities may be propagated from older to younger cohorts. Another possibility is that professionals' responses to students are motivated by taste-based discrimination, which could lead to a sub-optimal allocation of talent in the labor market (Hsieh et al. (2019)).

7 Conclusion

Information transmission through informal interactions is an everyday, routine occurrence. This paper provides a window into informal exchanges and additionally sheds light on a subtle form of disparate treatment of individuals based on their demographic characteristics. In a large-scale field experiment with college students interested in career advice, we estimate the causal effect of student gender on information provided by professionals regarding career paths. The experimental design also advances the correspondence study methodology by incorporating real individuals who are genuinely interested in the interactions being studied. We find that professionals differentially emphasize work/life balance to female students, even when students do not specifically ask about this issue. When students ask specifically about work/life balance, professionals are more willing to engage with female students on this topic. We combine the results of the field experiment with student preferences for professionals, and find that average gender differences persist when we incorporate student selection. Looking beyond our setting, correspondence and audit studies are powerful tools to study average discrimination in the labor market. Incorporating the preferences

of real individuals complements estimates of average discrimination and enriches our understanding of the realized experiences of individuals.

If information access depends on an individual's gender, absent knowing the nature of the missing, inaccurate, or emphasized information, it may be difficult for individuals to correct these disparities. Indeed, there are substantial gender gaps in knowledge of fundamental economic parameters, which are used to inform consumption, financial, and labor market choices (D'Acunto et al., 2019; Dwyer et al., 2002). Our work points to disparate information access as one plausible determinant of these knowledge and behavior gaps.

References

Abel, Martin, "Labor market discrimination and sorting: evidence from South Africa," *World Bank Policy Research Working Paper*, 2017.

Agan, Amanda Y and Sonja B Starr, "Employer Neighborhoods and Racial Discrimination," Working Paper 28153, National Bureau of Economic Research November 2020.

Becker, Gary, *The Economics of Discrimination*, 2 ed., University of Chicago Press, 1971.

Beggs, S., S. Cardell, and J. Hausman, "Assessing the potential demand for electric cars," *Journal of Econometrics*, September 1981, *17* (1), 1–19.

Bertrand, Marianne and Esther Duflo, "Field Experiments on Discrimination," *Handbook of Economic Field Experiments*, 2017, *1*, 309–393.

Black, Dan A., "Discrimination in an Equilibrium Search Model," *Journal of Labor Economics*, 1995, *13* (2), 309–334.

Blau, Francine D and Lawrence M Kahn, "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, sep 2017, *55* (3), 789–865.

Bohren, Aislinn, Alex Imas, and Michael Rosenberg, "The Language of Discrimination: Using Experimental versus Observational Data," *AEA Papers and Proceedings*, 2018, *108*, 169–74.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek, " Gender, Competitiveness, and Career Choices *," *The Quarterly Journal of Economics*, 05 2014, *129* (3), 1409–1447.

Canaan, Serena and Pierre Mouganie, "Female Science Advisors and the STEM Gender Gap," *SSRN Electronic Journal*, 2019.

Carrell, Scott E, Marianne E Page, and James E West, "Sex and Science: How Professor Gender Perpetuates the Gender Gap," *The Quarterly Journal of Economics*, 2010, *125* (3), 1101–1144.

Charles, Kerwin Kofi and Jonathan Guryan, "Prejudice and Wages: An Empirical Assessment of Beckerâs The Economics of Discrimination," *Journal of Political Economy*, 2008, *116* (5), 773–809.

Coffman, Lucas C, Clayton R Featherstone, and Judd B Kessler, "Can Social Information Affect What Job You Choose and Keep?," *American Economic Journal: Applied Economics*, January 2017, *9* (1), 96–117.

Cortés, Patricia and Jessica Pan, "When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap," *Journal of Labor Economics*, 2019, *37* (2), 351–398.

Cubas, German, Chinhui Juhn, and Pedro Silos, "Coordinated Work Schedules and the Gender Wage Gap," Working Paper 26548, National Bureau of Economic Research December 2019.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber, "Exposure to Daily Price Changes and Inflation Expectations," Working Paper 26237, National Bureau of Economic Research September 2019.

Dinkelman, Taryn and A Claudia Martínez, "Investing in Schooling in Chile: The Role of Information about Financial Aid for Higher Education," *The Review of Economics and Statistics*, May 2014, *96* (2), 244–257.

Dwyer, Peggy D., James H. Gilkeson, and John A. List, "Gender differences in revealed risk taking: Evidence from mutual fund investors," *Economics Letters*, 2002, *76* (2), 151–158.

Flory, Jeffrey A, Andreas Leibbrandt, and John A List, "Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job-Entry Decisions," *The Review of Economic Studies*, 2015, *82* (1), 122–155.

Gallen, Yana and Melanie Wasserman, "Do Male and Female Students Use Networks Differently?," *American Economic Association Papers and Proceedings*, 2021, *forthcoming.*

Goldin, Claudia, "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, April 2014, *104* (4), 1091–1119.

Gornall, Will and Ilya A. Strebulaev, "Gender, Race, and Entrepreneurship: A Randomized Field Experiment on Venture Capitalists and Angels," *SSRN Electronic Journal*, 2020.

_ , Ilya A Strebulaev, Stella Chen, Michelle Gan, Marco Soohyeong Lee, Justin Norden, Alexander Qu, Ashutosh Synghal, Eric Tolocka, Brian Triana, and Andy Wardle,

"Gender, Race, and Entrepreneurship: A Randomized Field Experiment on Venture Capitalists and Angels * and participants at the Stanford Financing of Innovation Summit for comments and suggestions on the project," 2018.

Heckman, James and Peter Siegelman, "The Urban Institute Audit Studies: Their Methods and Findings," in Fix and Struyk, eds., *Clear and Convincing Evidence: Measurement of Discrimination in America*, The Urban Institute Press, 1993, pp. 187–258.

Heckman, James J., "Detecting Discrimination," *Journal of Economic Perspectives*, 1998, *12* (2), 101–116.

Hoxby, Caroline M and Sarah Turner, "What High-Achieving Low-Income Students Know about College," *American Economic Review*, May 2015, *105* (5), 514–517.

Hsieh, Chang-Tai, Erik Hurst, Charles Jones, and Peter Klenow, "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Jack, Anthony Abraham, *The Privileged Poor: How Elite Colleges Are Failing Disadvantaged Students*, Harvard University Press, 2019.

Jensen, Robert, "The (Perceived) Returns to Education and the Demand for Schooling," *The Quarterly Journal of Economics*, 2010, *125* (May), 515–548.

Kalla, Joshua, Frances Rosenbluth, and Dawn Langan Teele, "Are You My Mentor? A Field Experiment on Gender, Ethnicity, and Political Self-Starters," *The Journal of Politics*, 2018, *80* (1), 337–341.

Kalla, Joshua L. and David E. Broockman, "Campaign Contributions Facilitate Access to Congressional Officials: A Randomized Field Experiment," *American Journal of Political Science*, 2016, *60* (3), 545–558.

Kessler, Judd B., Corinne Low, and Colin D. Sullivan, "Incentivized resume rating: Eliciting employer preferences without deception," *American Economic Review*, 2019, *109* (11), 3713–3744.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington, "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?," Working Paper 24740, National Bureau of Economic Research June 2018.

Lahey, Joanna N and Douglas R Oxley, "Discrimination at the Intersection of Age, Race, and Gender: Evidence from a Lab-in-the-field Experiment," Working Paper 25357, National Bureau of Economic Research December 2018.

McDowell, John, "Obsolescence of Knowledge and Career Publication Profiles: Some Evidence of Differences among Fields in Costs of Interrupted Careers," *American Economic Review*, 1982, *72* (4), 752–68.

Milkman, Katherine L., Modupe Akinola, and Dolly Chugh, "What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations.," *Journal of Applied Psychology*, 2015, *100* (6), 1678–1712.

Neal, Derek, "The Complexity of Job Mobility among Young Men," *Journal of Labor Economics*, April 1999, *17* (2), 237–261.

Neumark, David, "Detecting Discrimination in Audit and Correspondence Studies," *The Journal of Human Resources*, 2012, *47* (4), 1128–1157.

Nguyen, Trang, "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar," 2008.

Niederle, Muriel and Lise Vesterlund, "Gender and Competition," *Annual Review of Economics*, 2011, *3*, 601–630.

Pager, Devah, "The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future," *Annals of the American Academy of Political and Social Science*, 2007, *609* (1), 104–133.

_ **and David Pedulla**, "Race, Self-Selection, and the Job Search Process," *American Journal of Sociology*, 2015, *120* (4), 1005–54.

Papageorgiou, Theodore, "Learning Your Comparative Advantages," *The Review of Economic Studies*, 02 2014, *81* (3), 1263–1295.

Polachek, Solomon, "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure," *The Review of Economics and Statistics*, 1981, *63* (1), 60–69.

Porter, Catherine and Danila Serra, "Gender Differences in the Choice of Major: The Importance of Female Role Models," *American Economic Journal: Applied Economics*, July 2020, *12* (3), 226–54.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales, "Taste for Competition and the Gender Gap Among Young Business Professionals," Working Paper 21695, National Bureau of Economic Research November 2015.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 1974, *82* (1), 34–55.

_ , "Chapter 12 The theory of equalizing differences," in "in," Vol. 1 of *Handbook of Labor Economics*, Elsevier, 1986, pp. 641–692.

Roy, A. D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Särndal, Carl-Erik, Bengt Swensson, and Jan Wretman, "Model assisted survey sampling.," 1992.

Sockin, Jason and Aaron Sojourner, "What's the Inside Scoop? Challenges in the Supply and Demand for Information on Employers," *Journal of Labor Economics*, 2022, *0* (ja), null.

Wiswall, Matthew and Basit Zafar, "Determinants of College Major Choice: Identification using an Information Experiment," *Review of Economic Studies*, 2015, *82* (2), 791–824.

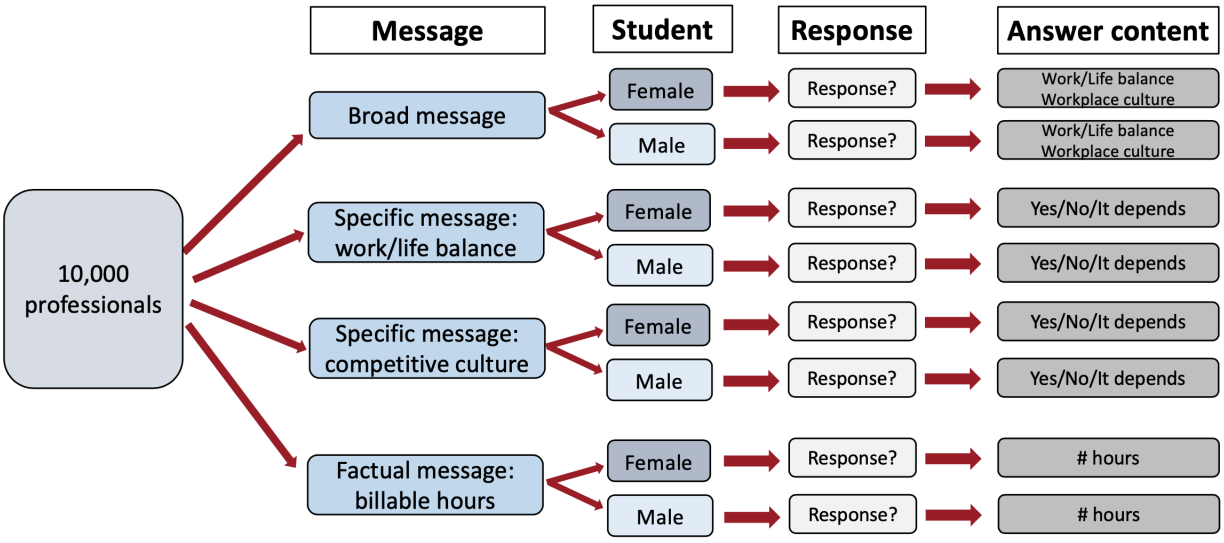
_ **and** _ , "How do College Students Respond to Public Information about Earnings?," *Journal of Human Capital*, 2015, *9* (2), 117–169.

_ **and** _ , "Preference for the Workplace, Investment in Human capital, and Gender," *Quarterly Journal of Economics*, 2018, *133* (1), 457–507.

Zafar, Basit, "How Do College Students Form Expectations?," *Journal of Labor Economics*, 2011, *29* (2), 301–348.

Figures and Tables

Figure 1: Experimental Design and Main Outcomes



Note: This figure depicts the experimental design and main outcomes.

Figure 2: Student Demand for Career Information, by Student Gender

Note: This figure uses the follow-up survey of 151 undergraduates at the same university to depict student demand for career information, by topic and student gender. Each student was asked how they would allocate 15 minutes of time spent with a professional in their preferred career path among 8 career-related topics. The figure plots the average percentage of the 15 minutes allocated to each topic, separately for female and male students. The following gender contrasts are statistically significant at the 5 percent level: different types of jobs within the field and work/life balance.

Table 1: Student Summary Statistics

	All Students	Male	Female
Profile Information			
Female	0.58		
	(0.50)		
Expected Graduation Year	2022.24	2022.50	2022.05
	(1.04)	(0.95)	(1.08)
Economics	0.62	0.69	0.57
	(0.49)	(0.47)	(0.50)
STEM	0.22	0.25	0.20
	(0.42)	(0.44)	(0.41)
0-49 Connections	0.46	0.44	0.48
	(0.50)	(0.50)	(0.51)
50-249 Connections	0.28	0.25	0.30
	(0.45)	(0.44)	(0.46)
250+ Connections	0.26	0.31	0.23
	(0.44)	(0.47)	(0.42)
Profile Extra Info	0.47	0.56	0.41
	(0.50)	(0.50)	(0.50)
Profile Issue	0.07	0.03	0.09
	(0.25)	(0.18)	(0.29)
Demographic Information			
White/Caucasian	0.30	0.28	0.32
	(0.46)	(0.46)	(0.47)
Asian/Pacific Islander	0.37	0.31	0.41
	(0.49)	(0.47)	(0.50)
Other Race/Ethnicity	0.33	0.41	0.27
	(0.47)	(0.50)	(0.45)
Additional Student Information			
GPA	3.64	3.62	3.65
	(0.28)	(0.34)	(0.24)
First Generation College Student	0.22	0.25	0.20
	(0.42)	(0.44)	(0.41)
Online Presence	0.71	0.66	0.75
	(0.46)	(0.48)	(0.44)
Observations	76	32	44

Note: This table reports means for each student characteristic, with standard deviations in parentheses.

Table 2: Professionals Summary Statistics

	All Professionals	Data Science	Finance	Law	Mgmt Consulting
Data Science	0.13				
	(0.33)				
Finance	0.28				
	(0.45)				
Law	0.33				
	(0.47)				
Mgmt Consulting	0.26				
	(0.44)				
Female	0.34	0.29	0.23	0.43	0.36
	(0.47)	(0.45)	(0.42)	(0.49)	(0.48)
College Graduation Year	2003.62	2009.55	2003.83	1998.31	2007.45
	(12.00)	(7.59)	(11.95)	(11.84)	(11.16)
College Selectivity - Admit Rate	0.25	0.39	0.25	0.20	0.28
	(0.22)	(0.28)	(0.22)	(0.16)	(0.23)
Alumni of Student's College	0.21	0.25	0.27	0.15	0.21
	(0.41)	(0.44)	(0.44)	(0.35)	(0.41)
Any Graduate Degree	0.70	0.72	0.50	1.00	0.51
	(0.46)	(0.45)	(0.50)	(0.00)	(0.50)
Any Ivy Degree	0.16	0.07	0.16	0.19	0.15
	(0.36)	(0.26)	(0.36)	(0.39)	(0.36)
0-249 Connections	0.11	0.13	0.10	0.15	0.05
	(0.31)	(0.33)	(0.30)	(0.36)	(0.21)
250-499 Connections	0.21	0.24	0.21	0.27	0.13
	(0.41)	(0.42)	(0.41)	(0.45)	(0.33)
500+ Connections	0.64	0.59	0.64	0.54	0.78
	(0.48)	(0.49)	(0.48)	(0.50)	(0.41)
Observations	7602	970	2156	2522	1954

Note: This table reports summary statistics for the sample of professionals, overall and by professional field. Means for each professional characteristic are reported, with standard deviations in parentheses.

Table 3: Effect of Student Gender on Response Rates, By Question Type

	Broad			Work/Life Balance			Competitive Culture			Factual (Law Only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Student Female	0.011	0.014	0.013	0.037**	0.041***	0.040**	0.003	0.009	0.011	0.061*	0.059	0.060
	(0.010)	(0.011)	(0.010)	(0.015)	(0.015)	(0.016)	(0.015)	(0.017)	(0.017)	(0.032)	(0.036)	(0.039)
Finance	-0.055***	-0.055***	-0.055***	-0.118***	-0.118***	-0.118***	-0.094***	-0.094***	-0.093***			
	(0.018)	(0.018)	(0.018)	(0.037)	(0.037)	(0.037)	(0.032)	(0.032)	(0.032)			
Law	-0.084***	-0.084***	-0.084***	-0.134***	-0.134***	-0.134***	-0.107***	-0.107***	-0.108***			
	(0.017)	(0.017)	(0.017)	(0.032)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)			
Mgmt Consulting	-0.030*	-0.030*	-0.030*	-0.071**	-0.071**	-0.071**	-0.086***	-0.087***	-0.088***			
	(0.018)	(0.018)	(0.017)	(0.032)	(0.032)	(0.032)	(0.031)	(0.031)	(0.031)			
Male Mean	0.101			0.130			0.139			0.073		
Observations	3530	3530	3530	1763	1763	1763	1776	1776	1776	298	298	298
Message Time/Date	X	X	X	X	X	X	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X	X	X	X	X	X	X
Student Race/Ethnicity		X	X		X	X		X	X		X	X
Additional Student			X			X			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a message received a response, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. Separate regressions are estimated for each question type: broad, specific - work/life balance, specific - competitive culture, and factual. Columns 1, 4, 7, and 10 report results from the baseline specification. Columns 2, 5, 8, and 11 report results from a specification that additionally includes controls for student race/ethnicity. Columns 3, 6, 9, and 12 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Effect of Student Gender on Mentions of Work/Life Balance and Workplace Culture

	Work/Life Balance			Workplace Culture		
	(1)	(2)	(3)	(4)	(5)	(6)
Student Female	0.087***	0.072**	0.076**	-0.003	-0.024	-0.024
	(0.032)	(0.032)	(0.033)	(0.034)	(0.034)	(0.034)
Finance	0.013	0.017	0.012	-0.127**	-0.121**	-0.123**
	(0.028)	(0.029)	(0.029)	(0.055)	(0.055)	(0.054)
Law	0.062	0.062	0.053	-0.172***	-0.170***	-0.173***
	(0.043)	(0.043)	(0.042)	(0.053)	(0.054)	(0.054)
Mgmt Consulting	0.208***	0.209***	0.209***	-0.074	-0.071	-0.073
	(0.051)	(0.052)	(0.051)	(0.060)	(0.060)	(0.061)
Male Mean	0.067			0.128		
Observations	363	363	363	363	363	363
Message Time/Date	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X
Student Race/Ethnicity		X	X		X	X
Additional Student			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response mentions work/life balance (columns 1-3) or workplace culture (columns 4-6), and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. Columns 1 and 4 report results from the baseline specification. Columns 2 and 5 report results from a specification that additionally includes controls for student race/ethnicity. Columns 3 and 6 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Role of Professional Composition in Gender Differences in Responses to the Broad Question

	Broad Response Rate		Work/Life Balance Mention		Workplace Culture Mention	
	(1)	(2)	(3)	(4)	(5)	(6)
Student Female	0.014	0.018	0.072**	0.057*	-0.024	-0.022
	(0.011)	(0.011)	(0.032)	(0.033)	(0.034)	(0.039)
Finance	-0.055***	-0.056***	0.017	-0.014	-0.121**	-0.116*
	(0.018)	(0.018)	(0.029)	(0.035)	(0.055)	(0.062)
Law	-0.084***	-0.061***	0.062	0.020	-0.170***	-0.222**
	(0.017)	(0.020)	(0.043)	(0.060)	(0.054)	(0.082)
Mgmt Consulting	-0.030*	-0.033*	0.209***	0.186***	-0.071	-0.067
	(0.018)	(0.019)	(0.052)	(0.053)	(0.060)	(0.063)
Observations	3530	3530	363	363	363	363
Message Time/Date	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X
Student Race/Ethnicity	X	X	X	X	X	X
Professional		X		X		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a message received a response (columns 1 and 2) or mentions of work/life balance (column 3 and 4) or workplace culture (columns 5 and 6), and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, student profile characteristics, and student race/ethnicity. Separate regressions are estimated for each question type: broad, specific - work/life balance, specific - competitive culture, and factual. Columns 1, 3, and 5 report results from the preferred specification. Columns 2, 4, and 6 report results from a specification that additionally includes controls for professional characteristics. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Heterogeneity by Professional Attributes

(1) Male	(2) Female	(3) Grad<2010	(4) Grad>=2010	(5) No Ivy	(6) Ivy	(7) Not Alum	(8) Alum	(9) Data Science	(10) Finance	(11) Law	(12) Mgmt Consulting
Panel A. Response Rate, Broad Question											
0.015	0.006	0.009	-0.002	0.010	0.012	0.015	0.008	0.025	0.014	0.009	0.009
(0.011)	(0.019)	(0.010)	(0.024)	(0.014)	(0.019)	(0.012)	(0.023)	(0.032)	(0.020)	(0.015)	(0.024)
[0.114]	[0.076]	[0.082]	[0.145]	[0.114]	[0.075]	[0.087]	[0.160]	[0.139]	[0.096]	[0.068]	[0.124]
Panel B. Response Rate, Specific Question - Work/Life Balance											
0.041*	0.045*	0.028	0.085**	0.051***	0.036	0.038**	0.072*	-0.027	0.069*	0.034	0.055*
(0.021)	(0.025)	(0.018)	(0.033)	(0.019)	(0.041)	(0.018)	(0.041)	(0.061)	(0.037)	(0.026)	(0.028)
[0.148]	[0.094]	[0.105]	[0.165]	[0.132]	[0.129]	[0.115]	[0.182]	[0.245]	[0.085]	[0.104]	[0.150]
Panel C. Response Rate, Specific Question - Competitive Culture											
0.019	-0.016	-0.012	0.011	0.003	-0.034	-0.016	0.022	0.008	0.046	0.008	-0.023
(0.023)	(0.031)	(0.022)	(0.039)	(0.019)	(0.057)	(0.024)	(0.038)	(0.061)	(0.031)	(0.032)	(0.040)
[0.154]	[0.108]	[0.125]	[0.170]	[0.145]	[0.155]	[0.125]	[0.223]	[0.242]	[0.109]	[0.115]	[0.146]
Panel D. Broad Question, Mention of Work/Life Balance											
0.064	0.140**	0.098**	0.043	0.053	0.338**	0.027	0.135***	0.037	0.084*	0.133*	0.064
(0.040)	(0.066)	(0.044)	(0.052)	(0.036)	(0.132)	(0.045)	(0.050)	(0.026)	(0.044)	(0.073)	(0.099)
[0.063]	[0.079]	[0.050]	[0.066]	[0.083]	[0.000]	[0.107]	[0.018]	[0.000]	[0.000]	[0.033]	[0.180]
Panel E. Broad Question, Mention of Workplace Culture											
-0.021	0.012	-0.028	0.090	-0.007	0.125	-0.044	0.065	-0.062	-0.093	-0.066	0.067
(0.046)	(0.089)	(0.052)	(0.061)	(0.040)	(0.189)	(0.053)	(0.056)	(0.136)	(0.064)	(0.062)	(0.087)
[0.126]	[0.132]	[0.100]	[0.092]	[0.116]	[0.056]	[0.155]	[0.036]	[0.185]	[0.119]	[0.100]	[0.120]

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is listed in each panel title, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, student profile characteristics, and student race/ethnicity. The column titles list the subsample used for estimation. Each entry in the table reports the estimated coefficient on student female from a separate specification. Standard errors are clustered at the student level and are reported in parentheses. Dependent variable means for male students are in brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Effect of Student Gender on Mentions of Work/Life Balance and Workplace Culture, Inclusive of Student Selection

	(1)	(2)	(3)	(4)	(5)	(6)
	Broad Response Rate		Work/Life Balance		Workplace Culture	
	Female Pref.	Male Pref.	Female Pref.	Male Pref.	Female Pref.	Male Pref.
Student female	0.010	0.020	0.097**	0.110**	-0.002	-0.017
	(0.018)	(0.020)	(0.046)	(0.054)	(0.052)	(0.067)
Male mean	0.095	0.121	0.057	0.051	0.111	0.141
	(0.012)	(0.013)	(0.023)	(0.043)	(0.033)	(0.055)
Observations	3530	3530	363	363	363	363

Note: This table reports the results of the estimation of the regression specification outlined in Equation (2), in which the dependent variable is an indicator for whether a professional responded to the broad message (columns 1–2), for whether a response mentions work/life balance (columns 3–4) or workplace culture (columns 5–6). Columns 1, 3, and 5 report results from the regression specification outlined in Equation (2) in which each observation is weighted by the propensity of the professional to be preferred by female students according to a rank-ordered conditional logit model of female student preferences over professional characteristics. Columns 2, 4, and 6 report results from the regression specification outlined in Equation (2) in which each observation is weighted by the propensity of the professional to be preferred by male students according to a rank-ordered conditional logit model of male student preferences over professional characteristics. In all regressions, the independent variables are an indicator for whether the student who sent the message is female, message time/date characteristics, and student profile characteristics. The bootstrapped standard errors are based on 1000 iterations and cluster at the student-ranker in the step when the model of student preferences is estimated to account for estimation error.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Effect of Information Received on Career Plans:
Is the Student Less Likely to Enter Preferred Career Path?

	Less Likely to Enter (binary)			Less Likely to Enter (continuous)		
	(1)	(2)	(3)	(4)	(5)	(6)
Response Mentioned W/L Balance	0.185	0.189	0.226	0.340	0.600	0.810
	(0.136)	(0.160)	(0.150)	(0.933)	(0.967)	(0.996)
Received Response to W/L Question	0.079*	0.079*	0.116	1.018**	1.008**	1.103*
	(0.043)	(0.044)	(0.078)	(0.444)	(0.443)	(0.603)
Response Mentioned Workplace Culture		-0.004	0.008		-0.264	-0.033
		(0.095)	(0.112)		(0.841)	(1.129)
Received Response to Culture Question		-0.007	-0.044		-0.922*	-1.015
		(0.071)	(0.075)		(0.496)	(0.615)
Male Mean	0.000			3.710		
Observations	73	73	73	73	73	73
Industry Controls	X	X	X	X	X	X
Student Controls	X	X	X	X	X	X
Message Controls		X	X		X	X
Professional Controls			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (3), in which the dependent variable is an indicator for whether a student is dissuaded from her preferred career path, relative to the start of the study, and the independent variables are whether the student received any information on work/life balance in her preferred career path, characteristics listed on the student's profile, and the student's preferred career path. Column 2 includes whether the student received any information on workplace culture in her preferred career path and response length. Column 3 includes all previously listed controls as well as the characteristics of the professionals. Robust standard errors are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A Online Appendix Figures and Tables

Figure A1: Message Templates

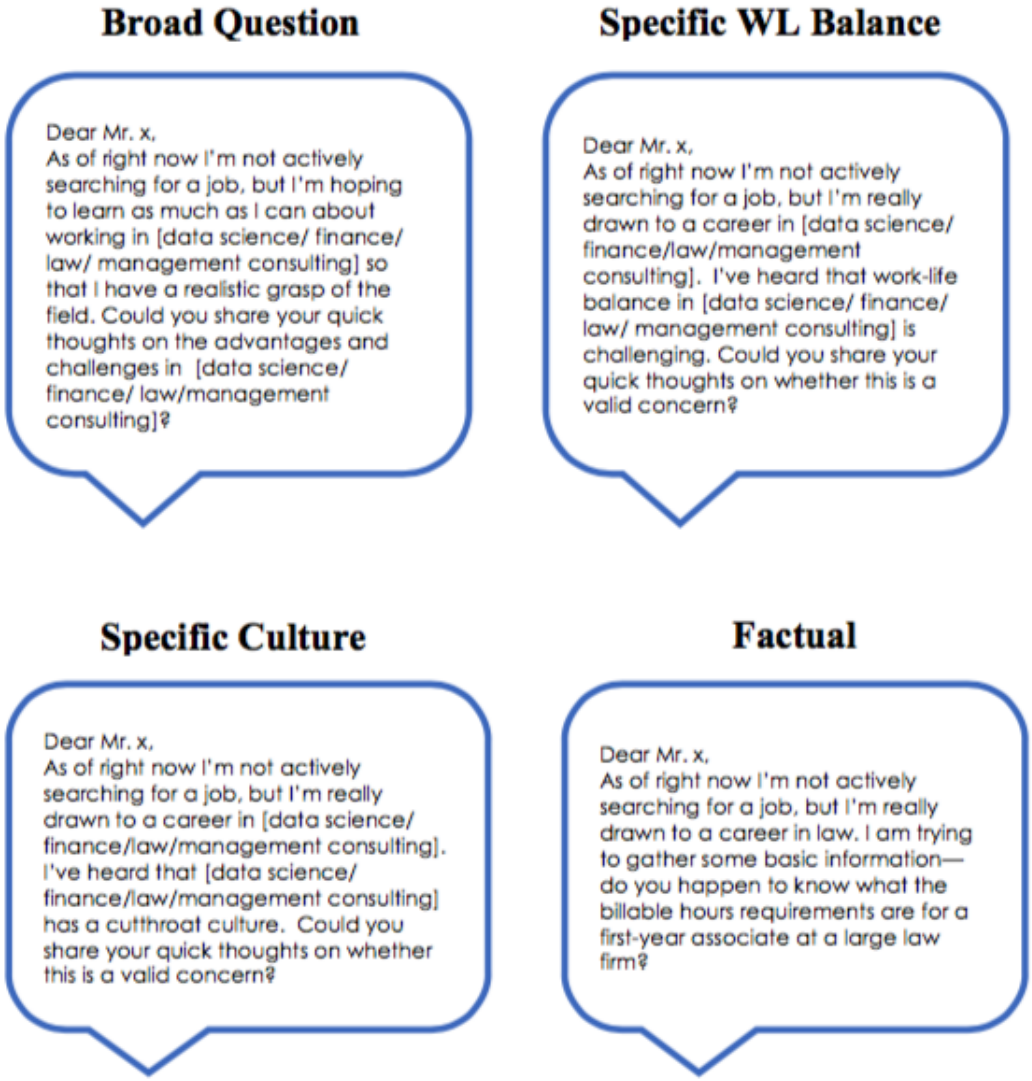
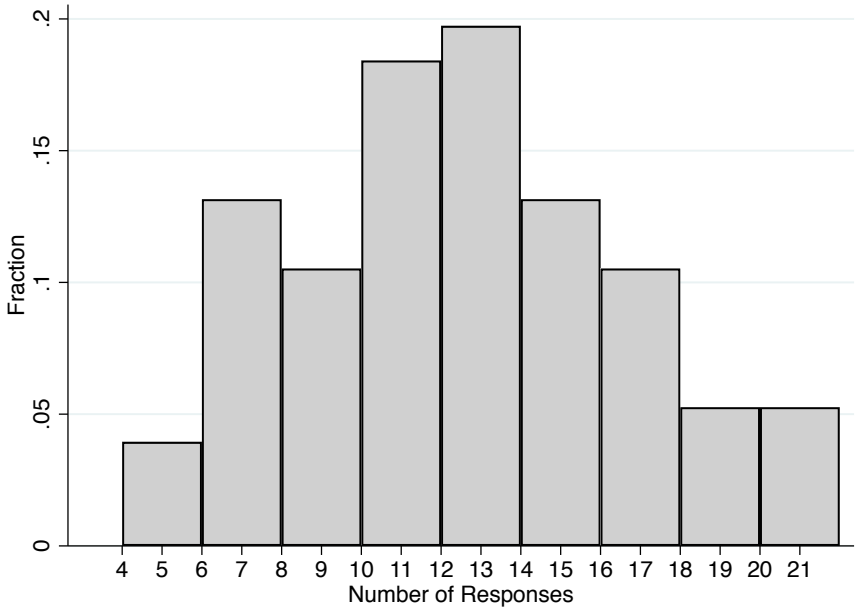


Figure A2: Distribution of Number of Responses



Note: This figure plots the distribution of the number of responses received across the 76 students in our analysis sample.

Table A1: Outcome Summary Statistics

	All Messages	Broad	Specific - Work/Life	Specific - Culture	Factual
Response Rate	0.12	0.10	0.14	0.15	0.11
	(0.33)	(0.30)	(0.35)	(0.36)	(0.32)
Response Character Count	434.73	414.39	486.64	429.95	304.18
	(558.77)	(687.34)	(492.37)	(396.61)	(553.03)
Work/Life Balance Mentioned		0.11			
		(0.32)			
Workplace Culture Mentioned		0.12			
		(0.33)			
Valid concern?					
Yes			0.44	0.16	
			(0.50)	(0.37)	
It depends			0.49	0.54	
			(0.50)	(0.50)	
No			0.07	0.30	
			(0.26)	(0.46)	
Billable Hours Quoted					1989.00
					(77.42)
Observations	7367	3530	1763	1776	298

Note: This table reports summary statistics for the main outcomes, overall and by question type. Means for each outcome are reported, with standard deviations in parentheses.

Table A2: Tests of Randomization

	(1) All Messages	(2) Sent Messages Only
Data Science	0.000	-0.000
	(0.001)	(0.002)
Finance	0.000	-0.003
	(0.001)	(0.004)
Law	-0.000	0.008
	(0.001)	(0.008)
Mgmt Consulting	-0.001	-0.005
	(0.001)	(0.004)
Professional Female	0.004	0.004
	(0.012)	(0.012)
0-249 Connections	0.004	0.004
	(0.007)	(0.008)
250-499 Connections	0.017*	0.021**
	(0.010)	(0.010)
500+ Connections	-0.015	-0.018
	(0.012)	(0.012)
College graduation year	0.311	0.208
	(0.321)	(0.320)
Alumni of Student's College	-0.003	-0.005
	(0.010)	(0.011)
Undergraduation Selectivity Quartile 1	-0.009	-0.012
	(0.008)	(0.008)
Undergraduation Selectivity Quartile 2	0.008	0.008
	(0.009)	(0.009)
Undergraduation Selectivity Quartile 3	0.005	0.008
	(0.011)	(0.012)
Undergraduation Selectivity Quartile 4	0.007	0.006
	(0.008)	(0.008)
Any Graduate Degree	0.008	0.011
	(0.010)	(0.011)
Any Ivy Degree	0.003	0.002
	(0.008)	(0.008)
N	7602	7367

Note: This table reports the results of the estimation of a regression specification, in which the dependent variable is a professional characteristic, listed in the rows, and the independent variable is indicator for whether the student who sent the message to the professional is female. Each entry represents the estimated coefficient from a separate specification. Column 1 reports the results for all messages that were assigned to students. Column 2 reports the results for the subset of messages that students actually sent. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Effect of Student Gender on Response Rates, By Question Type
Including Students with Ambiguously Gendered Names

	Broad			Work/Life Balance			Competitive Culture			Factual (Law Only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Student Female	0.011	0.015	0.016	0.040**	0.046***	0.046***	0.001	0.008	0.008	0.072**	0.092***	0.088**
	(0.010)	(0.010)	(0.010)	(0.015)	(0.016)	(0.017)	(0.014)	(0.015)	(0.015)	(0.030)	(0.035)	(0.035)
Finance	-0.051***	-0.051***	-0.051***	-0.120***	-0.120***	-0.120***	-0.092***	-0.092***	-0.092***			
	(0.016)	(0.016)	(0.016)	(0.034)	(0.034)	(0.034)	(0.030)	(0.030)	(0.030)			
Law	-0.081***	-0.082***	-0.081***	-0.134***	-0.134***	-0.134***	-0.098***	-0.099***	-0.099***			
	(0.015)	(0.015)	(0.015)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)			
Mgmt Consulting	-0.031*	-0.032**	-0.031**	-0.072**	-0.072**	-0.073**	-0.077***	-0.078***	-0.079***			
	(0.016)	(0.016)	(0.016)	(0.031)	(0.031)	(0.031)	(0.028)	(0.028)	(0.028)			
Male Mean	0.099			0.130			0.139			0.071		
Observations	4147	4147	4147	2063	2063	2063	2081	2081	2081	350	350	350
Message Time/Date	X	X	X	X	X	X	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X	X	X	X	X	X	X
Student Race/Ethnicity		X	X		X	X		X	X		X	X
Additional Student			X			X			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a message received a response, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. The sample is expanded to include the 13 students with ambiguously gendered names. Separate regressions are estimated for each question type: broad, specific - work/life balance, specific - competitive culture, and factual. Columns 1, 4, 7, and 10 report results from the baseline specification. Columns 2, 5, 8, and 11 report results from a specification that additionally includes controls for student race/ethnicity. Columns 3, 6, 9, and 12 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Effect of Student Gender on Response Rates, By Question Type
Including Students with Ambiguously Gendered Names: Additional Specifications

	Broad			Work/Life Balance			Competitive Culture			Factual (Law Only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Student Female	0.003	0.009	0.011	0.025	0.030*	0.040**	0.011	0.012	0.001	0.073**	0.075**	0.072**
	(0.011)	(0.009)	(0.010)	(0.017)	(0.016)	(0.015)	(0.015)	(0.014)	(0.014)	(0.032)	(0.031)	(0.030)
Finance		-0.051***	-0.051***		-0.120***	-0.120***		-0.092***	-0.092***			
		(0.016)	(0.016)		(0.034)	(0.034)		(0.030)	(0.030)			
Law		-0.081***	-0.081***		-0.135***	-0.134***		-0.098***	-0.098***			
		(0.015)	(0.015)		(0.030)	(0.030)		(0.031)	(0.030)			
Mgmt Consulting		-0.031*	-0.031*		-0.073**	-0.072**		-0.076***	-0.077***			
		(0.016)	(0.016)		(0.031)	(0.031)		(0.028)	(0.028)			
Male Mean	0.099		0.099	0.130		0.130	0.139		0.139	0.071		0.071
Observations	4147	4147	4147	2063	2063	2063	2081	2081	2081	350	350	350
Message Time/Date		X	X		X	X		X	X		X	X
Student Profile			X			X			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a message received a response, and the independent variable is an indicator for whether the student who sent the message is female. Column 2 adds controls for the professional's field, message time/date characteristics. Column 3 adds controls for student profile characteristics and is the baseline specification in Appendix Table A3. The sample is expanded to include the 13 students with ambiguously gendered names. Separate regressions are estimated for each question type: broad, specific - work/life balance, specific - competitive culture, and factual. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Effect of Student Gender on Response Rates, By Question Type: Additional Specifications

	Broad			Work/Life Balance			Competitive Culture			Factual (Law Only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Student Female	0.003	0.009	0.011	0.022	0.028*	0.037**	0.017	0.018	0.003	0.070**	0.072**	0.061*
	(0.011)	(0.010)	(0.010)	(0.017)	(0.016)	(0.015)	(0.016)	(0.015)	(0.015)	(0.033)	(0.033)	(0.032)
Finance		-0.055***	-0.055***		-0.118***	-0.118***		-0.094***	-0.094***			
		(0.017)	(0.018)		(0.037)	(0.037)		(0.032)	(0.032)			
Law		-0.084***	-0.084***		-0.135***	-0.134***		-0.106***	-0.107***			
		(0.017)	(0.017)		(0.033)	(0.032)		(0.033)	(0.033)			
Mgmt Consulting		-0.030*	-0.030*		-0.072**	-0.071**		-0.084***	-0.086***			
		(0.018)	(0.018)		(0.033)	(0.032)		(0.031)	(0.031)			
Male Mean	0.101		0.101	0.130		0.130	0.139		0.139	0.073		0.073
Observations	3530	3530	3530	1763	1763	1763	1776	1776	1776	298	298	298
Message Time/Date		X	X		X	X		X	X		X	X
Student Profile			X			X			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a message received a response, and the independent variable is an indicator for whether the student who sent the message is female. Column 2 adds controls for the professional's field, message time/date characteristics. Column 3 adds controls for student profile characteristics and is the baseline specification in Table 3. Separate regressions are estimated for each question type: broad, specific - work/life balance, specific - competitive culture, and factual. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Effect of Student Gender on Mentions of Work/Life Balance and Workplace Culture Accounting for Non-response

	Work/Life Balance		Workplace Culture	
	(1)	(2)	(3)	(4)
Student Female	0.009***	0.008**	0.001	-0.002
	(0.003)	(0.004)	(0.003)	(0.003)
Finance	-0.001	-0.001	-0.023***	-0.023***
	(0.003)	(0.003)	(0.008)	(0.008)
Law	0.002	0.002	-0.028***	-0.028***
	(0.004)	(0.004)	(0.008)	(0.008)
Mgmt Consulting	0.025***	0.025***	-0.014	-0.014
	(0.006)	(0.006)	(0.009)	(0.009)
Male Mean	0.007		0.013	
Observations	3530	3530	3530	3530
Message Time/Date	X	X	X	X
Student Profile	X	X	X	X
Student Race/Ethnicity	X	X	X	X
Additional Student		X		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response mentions work/life balance (columns 1–2) or workplace culture (columns 3–4), and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. Messages that do not receive a response are coded as not mentioning these career attributes. Columns 1 and 3 report results from the baseline specification, which includes controls for student race/ethnicity. Columns 2 and 4 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Effect of Student Gender on Mentions of Work/Life Balance and Workplace Culture
Including Students with Ambiguously Gendered Names

	Work/Life Balance		Workplace Culture	
	(1)	(2)	(3)	(4)
Student Female	0.074***	0.068**	-0.005	-0.021
	(0.028)	(0.028)	(0.032)	(0.033)
Finance	0.015	0.010	-0.096*	-0.098**
	(0.026)	(0.026)	(0.049)	(0.049)
Law	0.076*	0.070*	-0.128**	-0.131**
	(0.039)	(0.039)	(0.051)	(0.051)
Mgmt Consulting	0.201***	0.202***	-0.038	-0.041
	(0.045)	(0.046)	(0.056)	(0.056)
Male Mean	0.066		0.126	
Observations	420	420	420	420
Message Time/Date	X	X	X	X
Student Profile	X	X	X	X
Student Race/Ethnicity	X	X	X	X
Additional Student		X		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response mentions work/life balance (columns 1–2) or workplace culture (columns 3–4), and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. The sample is expanded to include the 13 students with ambiguously gendered names. Columns 1 and 3 report results from the baseline specification, which includes controls for student race/ethnicity. Columns 2 and 4 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Effect of Student Gender on Mentions of Work/Life Balance and Workplace Culture
Including Students with Ambiguously Gendered Names: Additional Specifications

	Work/Life Balance			Workplace Culture		
	(1)	(2)	(3)	(4)	(5)	(6)
Student Female	0.068**	0.061**	0.074***	-0.011	-0.018	-0.005
	(0.028)	(0.025)	(0.028)	(0.034)	(0.031)	(0.032)
Finance		0.017	0.015		-0.092*	-0.096*
		(0.024)	(0.026)		(0.049)	(0.049)
Law		0.083**	0.076*		-0.122**	-0.128**
		(0.041)	(0.039)		(0.052)	(0.051)
Mgmt Consulting		0.204***	0.201***		-0.033	-0.038
		(0.045)	(0.045)		(0.055)	(0.056)
Male Mean	0.066	0.066	0.066	0.126	0.126	0.126
Observations	420	420	420	420	420	420
Message Time/Date		X	X		X	X
Student Profile			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response mentions work/life balance (columns 1-3) or workplace culture (columns 4-6), and the independent variable is an indicator for whether the student who sent the message is female. Column 2 adds controls for the professional's field, message time/date characteristics. Column 3 adds controls for student profile characteristics and is the baseline specification in Appendix Table A7. The sample is expanded to include the 13 students with ambiguously gendered names. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A9: Effect of Student Gender on Mentions of Work/Life Balance and Workplace Culture: Additional Specifications

	Work/Life Balance			Workplace Culture		
	(1)	(2)	(3)	(4)	(5)	(6)
Student Female	0.073**	0.068**	0.087***	-0.006	-0.017	-0.003
	(0.031)	(0.028)	(0.032)	(0.036)	(0.033)	(0.034)
Finance		0.013	0.013		-0.125**	-0.127**
		(0.027)	(0.028)		(0.055)	(0.055)
Law		0.068	0.062		-0.165***	-0.172***
		(0.044)	(0.043)		(0.054)	(0.053)
Mgmt Consulting		0.204***	0.208***		-0.072	-0.074
		(0.050)	(0.051)		(0.060)	(0.060)
Male Mean	0.067		0.067	0.128		0.128
Observations	363	363	363	363	363	363
Message Time/Date		X	X		X	X
Student Profile			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response mentions work/life balance (columns 1-3) or workplace culture (columns 4-6), and the independent variable is an indicator for whether the student who sent the message is female. Column 2 adds controls for the professional's field, message time/date characteristics. Column 3 adds controls for student profile characteristics and is the baseline specification in Table 4. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A10: Effect of Student Gender on Mentions of Specific Work/Life Balance Issues

	Duration Typical Workweek		Work Schedule Flexibility		Extent of Travel/Work from Home	
	(1)	(2)	(3)	(4)	(5)	(6)
Student Female	0.054*	0.057*	0.023	0.026	0.015	0.016
	(0.030)	(0.030)	(0.022)	(0.023)	(0.022)	(0.023)
Finance	0.026	0.022	0.021	0.018	0.001	0.001
	(0.025)	(0.024)	(0.013)	(0.015)	(0.018)	(0.018)
Law	0.062	0.057	0.043*	0.036*	-0.004	-0.004
	(0.040)	(0.038)	(0.022)	(0.021)	(0.017)	(0.017)
Mgmt Consulting	0.142***	0.143***	0.089***	0.089***	0.170***	0.171***
	(0.042)	(0.042)	(0.029)	(0.030)	(0.045)	(0.044)
Male Mean	0.047		0.027		0.034	
Observations	363	363	363	363	363	363
Message Time/Date	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X
Student Race/Ethnicity	X	X	X	X	X	X
Additional Student		X		X		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response mentions a specific work/life balance issue, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. Columns 1, 3, and 5 report results from the preferred specification, which also controls for student/race ethnicity. Columns 2, 4, and 6 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A11: Gender Differences in Response Tone (Student Ratings)

	Student more concerned about work/life balance		Student more concerned about workplace culture	
	(1)	(2)	(3)	(4)
	Broad	Work/Life Balance	Broad	Competitive Culture
Student Female	0.033	-0.032	-0.018	-0.004
	(0.027)	(0.048)	(0.015)	(0.034)
Finance	-0.023	0.312***	-0.090***	0.101*
	(0.023)	(0.053)	(0.024)	(0.057)
Law	0.051	0.365***	-0.076***	0.110**
	(0.034)	(0.061)	(0.025)	(0.053)
Mgmt Consulting	0.161***	0.492***	-0.056*	-0.038
	(0.042)	(0.039)	(0.029)	(0.051)
Male Mean	0.114	0.523	0.087	0.293
Observations	3717	2626	3717	2682
Message Time/Date	X	X	X	X
Student Profile	X	X	X	X
Student Race/Ethnicity	X	X	X	X
Additional Student				

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response would make a typical college student more concerned about work/life balance or workplace culture, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, student profile characteristics, and student race/ethnicity. The question type is listed in each column title. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A12: Gender Differences in Response Length

	Broad		Work/Life Balance		Competitive Culture		Factual (Law Only)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Character Count	Log(Count)	Character Count	Log(Count)	Character Count	Log(Count)	Character Count	Log(Count)
Student Female	-8.747	-0.125	10.755	0.094	41.875	0.007	-115.259	-0.323
	(39.126)	(0.107)	(52.191)	(0.156)	(48.090)	(0.113)	(102.349)	(0.597)
Finance	-261.644***	-0.595***	-20.247	0.020	-55.515	-0.171		
	(74.237)	(0.180)	(54.440)	(0.153)	(66.544)	(0.152)		
Law	-210.993***	-0.381**	18.193	0.039	26.743	0.027		
	(78.255)	(0.176)	(57.283)	(0.154)	(70.103)	(0.164)		
Mgmt Consulting	-163.236**	-0.320*	134.036**	0.389***	14.302	0.092		
	(76.321)	(0.177)	(59.732)	(0.143)	(66.023)	(0.141)		
Male Mean	359.547	5.402	414.608	5.691	367.980	5.661	259.889	5.170
Observations	359	359	249	249	262	262	33	33
Message Time/Date	X	X	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X	X	X
Student Race/Ethnicity	X	X	X	X	X	X	X	X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is the length of the response, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, student profile characteristics, and student race/ethnicity. Columns 1, 3 5, and 7 analyze the response's character count, while columns 2, 4, 6, and 8 analyze the natural logarithm of the character count. Responses to each question are analyzed separately. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A13: Effect of Student Gender on Mentions of Competitive Culture

	Competitive Culture		
	(1)	(2)	(3)
Student Female	0.020	0.016	0.016
	(0.016)	(0.016)	(0.016)
Finance	0.038**	0.039**	0.040**
	(0.019)	(0.019)	(0.019)
Law	-0.005	-0.007	-0.009
	(0.006)	(0.008)	(0.008)
Mgmt Consulting	0.023	0.023*	0.019
	(0.014)	(0.013)	(0.012)
Male Mean	0.007	0.007	
Observations	363	363	363
Message Time/Date	X	X	X
Student Profile	X	X	X
Student Race/Ethnicity		X	X
Additional Student			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response mentions competitive culture, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. Messages that do not receive a response are coded as not mentioning these career attributes. Column 1 reports results from the baseline specification. Column 2 includes controls for student race/ethnicity and 3 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A14: Role of Professional Composition in Gender Differences in Response Rates

	Work/Life Balance		Competitive Culture		Factual (Law Only)	
	(1)	(2)	(3)	(4)	(5)	(6)
Student Female	0.041***	0.042***	0.009	0.009	0.059	0.056
	(0.015)	(0.016)	(0.017)	(0.017)	(0.036)	(0.039)
Finance	-0.118***	-0.117***	-0.094***	-0.072**		
	(0.037)	(0.037)	(0.032)	(0.033)		
Law	-0.134***	-0.100**	-0.107***	-0.070*		
	(0.033)	(0.039)	(0.033)	(0.038)		
Mgmt Consulting	-0.071**	-0.078**	-0.087***	-0.075**		
	(0.032)	(0.034)	(0.031)	(0.032)		
Observations	1763	1763	1776	1776	298	298
Message Time/Date	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X
Student Race/Ethnicity	X	X	X	X	X	X
Professional		X		X		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a message received a response, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, student profile characteristics, and student race/ethnicity. Separate regressions are estimated for each question type: broad, specific - work/life balance, specific - competitive culture, and factual. Columns 1, 3, and 5 report results from the preferred specification. Columns 2, 4, and 6 report results from a specification that additionally includes controls for professional characteristics. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A15: Gender Differences in Responses to "Is work/life balance a concern?"

	Yes		It depends		No	
	(1)	(2)	(3)	(4)	(5)	(6)
Student Female	0.002	-0.008	-0.023	-0.015	0.020	0.023
	(0.070)	(0.074)	(0.076)	(0.078)	(0.034)	(0.035)
Finance	0.217**	0.215**	-0.030	-0.034	-0.186***	-0.181***
	(0.088)	(0.087)	(0.108)	(0.107)	(0.069)	(0.068)
Law	0.315***	0.314***	-0.076	-0.063	-0.240***	-0.251***
	(0.091)	(0.093)	(0.111)	(0.115)	(0.065)	(0.067)
Mgmt Consulting	0.674***	0.672***	-0.427***	-0.422***	-0.247***	-0.250***
	(0.062)	(0.064)	(0.089)	(0.091)	(0.060)	(0.061)
Male Mean	0.427		0.512		0.061	
Observations	211	211	211	211	211	211
Message Time/Date	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X
Student Race/Ethnicity	X	X	X	X	X	X
Additional Student		X		X		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response to the specific work/life balance question is one of the categories in the column titles, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. Columns 1, 3, and 5 report results from the preferred specification, which also controls for student/race ethnicity. Columns 2, 4, and 6 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A16: Gender Differences in Responses to "Is cutthroat culture a concern?"

	Yes		It depends		No	
	(1)	(2)	(3)	(4)	(5)	(6)
Student Female	0.002	0.024	0.052	0.010	-0.053	-0.033
	(0.051)	(0.050)	(0.071)	(0.066)	(0.069)	(0.068)
Finance	0.072	0.082	0.198*	0.205*	-0.269***	-0.287***
	(0.079)	(0.081)	(0.111)	(0.115)	(0.096)	(0.100)
Law	0.135*	0.133*	0.093	0.098	-0.228**	-0.231**
	(0.080)	(0.078)	(0.113)	(0.109)	(0.087)	(0.088)
Mgmt Consulting	-0.035	-0.035	-0.050	-0.037	0.085	0.071
	(0.056)	(0.055)	(0.108)	(0.106)	(0.109)	(0.110)
Male Mean	0.134		0.512		0.354	
Observations	215	215	215	215	215	215
Message Time/Date	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X
Student Race/Ethnicity	X	X	X	X	X	X
Additional Student		X		X		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response to the specific competitive culture question is one of the categories in the column titles, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. Columns 1, 3, and 5 report results from the preferred specification, which also controls for student/race ethnicity. Columns 2, 4, and 6 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A17: Gender Differences in Hours Quoted in Response to Factual Question

	(1)	(2)
Student Female	81.611	154.402***
	(73.971)	(2.621)
Male Mean	1937.500	
Observations	25	25
Message Time/Date	X	X
Student Profile	X	X
Student Race/Ethnicity	X	X
Additional Student		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is the hours quoted in responses to the factual question, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, student profile characteristics, and student race/ethnicity. We only analyze only responses that include a numeric value or range. Column 1 reports results from the preferred specification, which also controls for student/race ethnicity. Column 2 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A18: Effect of Student Gender on Response Rates, By Question Type
Restricting to Students with No Online Presence

	Broad			Work/Life Balance			Competitive Culture			Factual (Law Only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Student Female	0.031	0.043**	0.040*	0.048*	0.066**	0.087**	0.051	0.045	0.027	0.105***	0.077***	0.062***
	(0.018)	(0.020)	(0.023)	(0.027)	(0.030)	(0.040)	(0.040)	(0.044)	(0.050)	(0.035)	(0.016)	(0.021)
Finance	-0.113***	-0.114***	-0.114***	-0.056	-0.055	-0.053	-0.070	-0.070	-0.070			
	(0.036)	(0.036)	(0.036)	(0.062)	(0.062)	(0.062)	(0.072)	(0.072)	(0.072)			
Law	-0.119***	-0.119***	-0.119***	-0.115*	-0.115*	-0.117*	-0.085	-0.085	-0.082			
	(0.039)	(0.039)	(0.039)	(0.057)	(0.058)	(0.058)	(0.065)	(0.065)	(0.064)			
Mgmt Consulting	-0.074*	-0.076*	-0.076*	-0.058	-0.063	-0.068	-0.053	-0.052	-0.047			
	(0.041)	(0.041)	(0.042)	(0.061)	(0.059)	(0.060)	(0.053)	(0.052)	(0.052)			
Male Mean	0.095			0.115			0.125			0.023		
Observations	1030	1030	1030	511	511	511	520	520	520	87	87	87
Message Time/Date	X	X	X	X	X	X	X	X	X	X	X	X
Student Profile	X	X	X	X	X	X	X	X	X	X	X	X
Student Race/Ethnicity		X	X		X	X		X	X		X	X
Additional Student			X			X			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a message received a response, and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. The sample is restricted to messages sent by students who do not have an online presence. Separate regressions are estimated for each question type: broad, specific - work/life balance, specific - competitive culture, and factual. Columns 1, 4, 7, and 10 report results from the baseline specification. Columns 2, 5, 8, and 11 report results from a specification that additionally includes controls for student race/ethnicity. Columns 3, 6, 9, and 12 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A19: Effect of Student Gender on Mentions of Work/Life Balance and Workplace Culture
Restricting to Students with No Online Presence

	Work/Life Balance		Workplace Culture	
	(1)	(2)	(3)	(4)
Student Female	0.095**	0.138***	0.052	0.061
	(0.037)	(0.030)	(0.042)	(0.047)
Finance	-0.021	-0.027	-0.136	-0.134
	(0.055)	(0.056)	(0.106)	(0.109)
Law	-0.025	-0.040	-0.254*	-0.263*
	(0.070)	(0.077)	(0.129)	(0.140)
Mgmt Consulting	0.068	0.054	-0.110	-0.115
	(0.086)	(0.087)	(0.116)	(0.121)
Male Mean	0.061		0.102	
Observations	110	110	110	110
Message Time/Date	X	X	X	X
Student Profile	X	X	X	X
Student Race/Ethnicity	X	X	X	X
Additional Student		X		X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (1), in which the dependent variable is an indicator for whether a response mentions work/life balance (columns 1–2) or workplace culture (columns 3–4), and the independent variables are an indicator for whether the student who sent the message is female, the professional's field, message time/date characteristics, and student profile characteristics. The sample is restricted to messages sent by students who do not have an online presence. Columns 1 and 3 report results from the baseline specification, which includes controls for student race/ethnicity. Columns 2 and 4 additionally include controls for student characteristics that may be observable elsewhere online. The omitted field is data science. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A20: Attributes of Preferred Professionals

	All Ranked	Female Ranked	Female Predicted	Male Ranked	Male Predicted
Data Science	0.18	0.14	0.14	0.24	0.25
	(0.38)	(0.34)	(0.34)	(0.43)	(0.44)
Finance	0.33	0.28	0.24	0.39	0.41
	(0.47)	(0.45)	(0.43)	(0.49)	(0.49)
Law	0.28	0.38	0.38	0.14	0.14
	(0.45)	(0.49)	(0.48)	(0.35)	(0.34)
Mgmt Consulting	0.22	0.21	0.25	0.23	0.20
	(0.41)	(0.41)	(0.43)	(0.42)	(0.40)
Female	0.43	0.58	0.60	0.20	0.27
	(0.50)	(0.49)	(0.49)	(0.40)	(0.44)
College Graduation Year	2005.20	2005.04	2005.01	2005.45	2006.04
	(10.90)	(10.44)	(10.91)	(11.67)	(10.87)
College Selectivity - Admit Rate	0.23	0.22	0.24	0.23	0.25
	(0.19)	(0.18)	(0.22)	(0.20)	(0.23)
Alumni of Student's College	0.23	0.19	0.22	0.29	0.30
	(0.42)	(0.40)	(0.41)	(0.45)	(0.46)
Any Graduate Degree	0.70	0.73	0.75	0.65	0.65
	(0.46)	(0.45)	(0.44)	(0.48)	(0.48)
Any Ivy Degree	0.20	0.21	0.20	0.19	0.15
	(0.40)	(0.41)	(0.40)	(0.40)	(0.35)
0-249 Connections	0.12	0.11	0.11	0.13	0.10
	(0.33)	(0.32)	(0.31)	(0.34)	(0.30)
250-499 Connections	0.19	0.18	0.16	0.19	0.18
	(0.39)	(0.39)	(0.37)	(0.40)	(0.38)
500+ Connections	0.64	0.69	0.71	0.57	0.65
	(0.48)	(0.47)	(0.46)	(0.50)	(0.48)
Observations	209	125	3647	84	3647

Note: This table reports summary statistics for the preferred sample of professionals, overall and by student gender. Means for each professional characteristic are reported, with standard deviations in parentheses.

Table A21: Effect of Student Gender on Response Rates, Inclusive of Student Selection

	(1)	(2)	(3)	(4)	(5)	(6)
	Broad Response Rate		Work/Life Response Rate		Culture Response Rate	
	Female Pref	Male Pref	Female Pref	Male Pref	Female Pref	Male Pref
Student female	0.010	0.020	0.047*	0.047*	-0.014	0.023
	(0.018)	(0.020)	(0.028)	(0.027)	(0.027)	(0.028)
Male mean	0.095	0.121	0.116	0.142	0.128	0.151
	(0.012)	(0.013)	(0.018)	(0.017)	(0.017)	(0.017)
Observations	3530	3530	1763	1763	1776	1776

Note: This table reports the results of the estimation of the regression specification outlined in Equation (2), in which the dependent variable is an indicator for whether a a message received a response, by question type as indicated in the columns. Columns 1, 3, and 5 report results from the regression specification outlined in Equation (2) in which each observation is weighted by the propensity of the professional to be preferred by female students according to a rank-ordered conditional logit model of female student preferences over professional characteristics. Columns 2, 4, and 6 report results from the regression specification outlined in Equation (2) in which each observation is weighted by the propensity of the professional to be preferred by female students according to a rank-ordered conditional logit model of female student preferences over professional characteristics. In all regressions, the independent variables are an indicator for whether the student who sent the message is female, message time/date characteristics, and student profile characteristics. The bootstrapped standard errors are based on 1000 iterations and cluster at the student-ranker in the step when the model of student preferences is estimated to account for estimation error.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A22: Effect of Information Received on Career Plans:
Is the Student Less Likely to Enter Preferred Career Path?
Received Any Information on Work/Life Balance

	Less Likely to Enter (binary)			Less Likely to Enter (continuous)		
	(1)	(2)	(3)	(4)	(5)	(6)
Received Info on Work/Life Balance	0.077*	0.075*	0.157**	0.665	0.723	1.006
	(0.042)	(0.043)	(0.066)	(0.496)	(0.495)	(0.622)
Received Info on Workplace Culture		0.014	-0.044		-0.797	-0.966
		(0.063)	(0.074)		(0.499)	(0.614)
Male Mean	0.000			3.710		
Observations	73	73	73	73	73	73
Industry Controls	X	X	X	X	X	X
Student Controls	X	X	X	X	X	X
Message Controls		X	X		X	X
Professional Controls			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (3), in which the dependent variable is an indicator for whether a student is dissuaded from her preferred career path, relative to the start of the study, and the independent variables are whether the student received any information on work/life balance in her preferred career path, characteristics listed on the student's profile, and the student's preferred career path. Column 2 includes whether the student received any information on workplace culture in her preferred career path and response length. Column 3 includes all previously listed controls as well as the characteristics of the professionals. Robust standard errors are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Online Appendix: Additional Details on Experimental Design

In this Appendix, we provide additional details on the experimental design, including the choice of occupational attributes, the design and results of two pilots, and power calculations.

B.1 Vignette Pilot with MBA Students

In order to pilot the specific wording for the messages as well as to get a sense for the types of responses we could expect in the actual experiment, in Summer 2019, we conducted a vignette-study pilot among first-year UCLA Anderson MBA students. The survey asked students the career path that they had the most experience in among management consulting, finance, marketing, and tech startups, which are four popular career paths among this student population. The survey included the following broad question:

> Suppose you receive the following message from a student at your undergraduate alma mater with a question on their career path choice.
>
> Hello,
>
> My name is [Scott/Stephanie] and I'm a junior trying to figure out my plans after graduation. As of right now I'm not actively searching for a job, but I'm hoping to learn as much as I can about working in [finance/management consulting/marketing/tech startups] so that I have a realistic grasp of the field. Could you share your quick thoughts on the challenges and opportunities in [finance/management consulting/marketing/tech startups]? Thank you in advance for your thoughts!
>
> [Scott/Stephanie]
>
> Please pretend that you actually received this message and answer naturally. Please provide your response below.

Each survey respondent was randomized to receive one of three specific questions about career attributes: work/life balance, workplace culture, and job instability:

> Hello, My name is [Scott/Stephanie] and I'm a junior trying to figure out my plans after graduation. I'm considering going into [finance/management consulting/marketing/tech startups] and would appreciate hearing your thoughts on whether this is the best option for me. I'm really drawn to virtually all aspects of the job. I've heard that work/life balance is challenging, which makes me worried about choosing [finance/management consulting/marketing/tech startups] as a long-term career. Do you think this is a valid concern? Thank you in advance for your thoughts! [Scott/Stephanie]

> Hello, My name is [Scott/Stephanie] and I'm a junior trying to figure out my plans after graduation. I'm considering going into [finance/management consulting/marketing/tech startups] and would appreciate hearing your thoughts on whether this is the best option for me. I'm really drawn to virtually all aspects of the job. I've heard that the culture is very cutthroat, which makes me worried about choosing [finance/management consulting/marketing/tech startups] as a long-term career. Do you think this is a valid concern? Thank you in advance for your thoughts! [Scott/Stephanie]
>
> Hello, My name is [Scott/Stephanie] and I'm a junior trying to figure out my plans after graduation. I'm considering going into [finance/management consulting/marketing/tech startups] and would appreciate hearing your thoughts on whether this is the best option for me. I'm really drawn to virtually all aspects of the job. However, I'm worried about job instability. Do you think this is a valid concern? Thank you in advance for your thoughts! [Scott/Stephanie]

We randomized whether the survey respondent received messages from Scott or Stephanie.

We coded whether responses to the broad question mentioned work/life balance, workplace culture, and job instability. We also used the responses to create a detailed rubric to comprehensively categorize each component of the response, starting with the O*NET classification system.

95 students responded to the survey. In responses to the broad messages, the rate of mentioning work/life balance was 13 percent, workplace culture was 6 percent, and job instability was 5 percent. For the career paths that overlap with our main experiment (management consulting and finance), the rates were 16 percent for work/life balance, 7 percent for workplace culture, and 2 percent for job instability. These survey respondents were 10.7 percentage points more likely to bring up work/life balance issues to Stephanie relative to Scott. There was no gender difference in rates of bringing up workplace culture issues. Survey respondents were slightly less likely to bring up job instability to Stephanie.

Due to the lower rate of bringing up job instability in response to the broad message and statistical power concerns, we chose two career attributes for the main experiment: work/life balance and workplace culture.

B.2 Response Rate Pilot on Networking Platform

In our power calculations, we assumed an overall 15% response rate, based which is in between response rates in studies of cold emails to politicians (Kalla and Broockman (2016)) and cold emails to venture capitalists (Gornall et al. (2018)). This response rate, combined with the rate of mentioning work/life balance from the pilot above, led us to a sample size of 10,000.

In order to assign message-types to professionals, we used 100 of our 10,000 professionals to gauge

relative response rates on our professional networking platform, and to finalize the relative number of broad, specific, and factual questions we would have, and primarily to work through logistical details of the experimental procedure. We did this pilot two months before we finalized our full professional sample, so the sample of professionals does differ systematically from the remaining 9,900 in the sample (it includes more alumni, for example). We randomized the remaining professional after this pilot was completed. This student had a 14% response rate overall, similar to what we saw overall, and a higher response rate in the factual question than what we saw ultimately in the study.

B.3 Profile photo

Figure D1: Photo of Iconic University Building

C Online Appendix: Details of Text Analysis

In this Appendix, we provide details of the manual classification, Kullback-Leiber Divergence metric, and the lexicon-based sentiment analysis.

C.1 Manual Classification

We provide an exploratory analysis of the remaining message components. We manually code each component of the response. Specifically, we construct a rubric based on the O*NET classification of occupations' work contexts and activities. Our rubric supplements the O*NET classification with additional fields that are mentioned in the messages (such as job search advice or compensation), to ensure we categorize the vast majority of the message text. Since many O*NET categories are used infrequently, we group related categories together. For example, we group interpersonal career attributes such as "communicate with persons outside the organization," "communicate with supervisors, peers, or subordinates," and "deal with external customers." The Online Appendix provides information on each grouping used, as well as the remaining ungrouped but frequently used categories.[43] This rubric allows us to test whether the responses to male and female students exhibit other content differences.

A description of each frequently used category is in Online Appendix Table C.1 below.

Category	Description	Source
Analytical Aspects		
Estimate Quantifiable Characteristics of Products, Events, or Information	Estimating sizes, distances, and quantities; or determining time, costs, resources, or materials needed to perform a work activity	O*NET Work Activity
Get Information	Observing, receiving, and otherwise obtaining information from all relevant sources	O*NET Work Activity
Analyze Data or Information	Identify the underlying principles, reasons, or facts of information by breaking down information or data into separate parts	O*NET Work Activity
Evaluate Information to Determine Compliance with Standards	Using relevant information and individual judgment to determine whether events or processes comply with laws, regulations, or standards.	O*NET Work Activity
Process Information	Compile, code, categorize, calculate, tabulate, audit, or verify information or data	O*NET Work Activity
Interact with Computers	Use computers and computer systems (including hardware and software) to program, write software, set up functions, enter data, or process information	O*NET Work Activity

[43] We consider frequent usage to be attributes that appear in more than 5 percent of messages.

Interpret Meaning of Information for Others	Translating or explaining what information means and how it can be used	O*NET Work Activity
Decision-making Aspects		
Develop Objectives and Strategies	Establishing long-range objectives and specifying the strategies and actions to achieve them	O*NET Work Activity
Make Decisions or Solve Problems	Analyze information and evaluate results to choose the best solution and solve problems	O*NET Work Activity
Organize, Plan, and Prioritize Work	Developing specific goals and plans to prioritize, organize, and accomplish your work.	O*NET Work Activity
Excitement and Impact Aspects		
Think Creatively	Developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions	O*NET Work Activity
Update and Use Relevant Knowledge	Keep up-to-date technically and apply new knowledge to your job	O*NET Work Activity
Responsibility for Outcomes	How responsible is the worker for work outcomes and results of other workers	O*NET Work Context
Consequence of Error	How serious would the result usually be if the worker made a mistake that was not readily correctable?	O*NET Work Context
Freedom to Make Decisions	How much decision making freedom, without supervision, does the job offer?	O*NET Work Context
Impact of Decisions on Coworkers/Company Results	What results do your decisions usually have on other people or the image or reputation or financial resources of your employer?	O*NET Work Context
Importance of Being Exact or Accurate	How important is being very exact or highly accurate in performing this job?	O*NET Work Context
Structured v. Unstructured Work	To what extent is this job structured for the worker, rather than allowing the worker to determine tasks, priorities, and goals?	O*NET Work Context
Importance of Repeating Same Tasks	How important is repeating the same physical activities (e.g., key entry) or mental activities (e.g., checking entries in a ledger) over and over, without stopping, to performing this job?	O*NET Work Context
Projects Monotonous/Constantly Changing		Supplemental Category
Interpersonal Aspects		
Communicate with Persons Outside Org.	Communicate with people outside the organization, represent the organization to customers, the public, government, and other external sources. This information can be exchanged in person, in writing, or by telephone or e-mail	O*NET Work Activity

Communicate with Supervisors, Peers, Subordinates	Providing information to supervisors, co-workers, and subordinates by telephone, in written form, e-mail, or in person	O*NET Work Activity
Coordinate Work Activities of Others	Getting members of a group to work together to accomplish tasks	O*NET Work Activity
Developing and Building Teams	Encouraging and building mutual trust, respect, and cooperation among team members	O*NET Work Activity
Establish and Maintain Personal Relationships	Developing constructive and cooperative working relationships with others, and maintaining them over time	O*NET Work Activity
Guide, Direct, and Motivate Subordinates	Providing guidance and direction to subordinates, including setting performance standards and monitoring performance	O*NET Work Activity
Provide Consultation and Advice to Others	Providing guidance and expert advice to management or other groups on technical, systems-, or process-related topics	O*NET Work Activity
Resolve Conflicts and Negotiate with Others	Handling complaints, settling disputes, and resolving grievances and conflicts, or otherwise negotiating with others	O*NET Work Activity
Sell or Influence Others	Convincing others to buy merchandise/goods or to otherwise change their minds or actions	O*NET Work Activity
Coordinate or Lead Others	How important is it to coordinate or lead others in accomplishing work activities in this job?	O*NET Work Context
Deal with External Customers	Job entails work with external customers or the public	O*NET Work Context
Deal with Unpleasant or Angry People	How frequently does the worker have to deal with unpleasant, angry, or discourteous individuals as part of the job requirements?	O*NET Work Context
Face-to-face Discussions	How often do you have to have face-to-face discussions with individuals or teams in this job?	O*NET Work Context
Frequency of Conflict	How often are there conflict situations the employee has to face in this job?	O*NET Work Context
Work with Work Group or Team	How important is it to work with others in a group or team in this job?	O*NET Work Context
Work/Life Balance Aspects		
Duration of Typical Workweek	Number of hours typically worked in one week	O*NET Work Context
Work Schedule Flexibility	Timing of work is flexible/inflexible	Supplemental Category
Extent of Travel/Work from Home	Location of work is flexible/inflexible, including work-related travel	Supplemental Category
Individual Categories that Appear in >5% of Responses		
Explains Paths within Field	Explains various paths within the field	Supplemental Category

Compensation	Mentions pay including salary or bonus	Supplemental Category
Job Stability	Jobs within career path stable/unstable	Supplemental Category
Short v. Long term Considerations	Any time dimension to career path, including whether it positions one well for future jobs or has changing attributes as one gains experience	Supplemental Category
Qualities of Individuals who Like/Succeed	Attributes of people who do well in this career path	Supplemental Category
Broadness of Question	Statement that the question is broad	Supplemental Category
Info on Job Search	Information on how to find a job within the field	Supplemental Category
Implicit/Explicit Offer to Discuss Further	Statement to discuss further (over message, email, phone, etc.) or asks a follow-up question	Supplemental Category
Decision is Person-Specific	Statement that the career decision depends on the person and their attributes/preferences	Supplemental Category
States Qualifications for Answering	Statement of experience in career path with intention of demonstrating that one is/isn't equipped to answer	Supplemental Category
Education Requirements and Environment	Statement of degree requirements and/or description of the attributes of those requirements (e.g. law school is grueling)	Supplemental Category

Online Appendix Table B2 reports the results. Responses to female students are less likely to offer any type of advice and less likely to state the professional's qualifications for answering the question. Responses to female students are also less likely to explain career paths and provide information on how to find a job, but these differences are not statistically significant. Responses to female students are more likely to discuss the analytical aspects of a career, compensation, and qualities of individuals who like/succeed in the field, as well as provide an offer to discuss further, but again these contrasts are not statistically significant. A joint test of significance indicates that we can reject that the gender differences are jointly zero. Combining the results on response length with the gender differences in other response content, we find evidence consistent with work/life balance crowding out other career information.

C.2 Kullback-Leiber Divergence Metric

Throughout this section, we use the term "female corpus" to refer to the set of words (with frequencies) used in all responses to female students. We use the term "male corpus" to refer to the set of words (with frequencies) used in all responses to male students.

When we refer to the distribution of words in a corpus, we refer to the distribution over unique words, where the probability of word j is given by:

Table B2: Gender Differences in Other Response Components

	(1) Main Specification	(2) Additional Student Controls
Offers Advice of Any Type	-0.075*	-0.078*
	(0.043)	(0.043)
Explains Career Paths	-0.037	-0.037
	(0.045)	(0.043)
Mentions Analytical Aspects of Career	0.050	0.051
	(0.042)	(0.041)
Mentions Decision-Making/Responsibility Aspects of Career	0.012	0.013
	(0.028)	(0.028)
Mentions Excitement/Impact Aspects of Career	0.024	0.029
	(0.040)	(0.038)
Mentions Interpersonal Aspects of Career	0.022	0.025
	(0.045)	(0.044)
Compensation	0.052	0.051
	(0.039)	(0.038)
Job Stability	0.035	0.036
	(0.028)	(0.028)
Short v. Long Term Considerations	-0.022	-0.023
	(0.032)	(0.032)
Qualities of Individuals who Like/Succeed	0.043	0.040
	(0.026)	(0.024)
Broadness of Question	-0.026	-0.027
	(0.032)	(0.033)
Info on Job Search	-0.035	-0.037
	(0.029)	(0.029)
Offer to Discuss Further	0.064	0.064
	(0.051)	(0.051)
Decision is Person Specific	-0.025	-0.025
	(0.025)	(0.024)
States Qualifications for Answering	-0.082*	-0.084*
	(0.045)	(0.043)
Education Requirements/Environment	0.030	0.029
	(0.032)	(0.031)
Other attribute	0.010	0.008
	(0.034)	(0.033)
p-value from joint test M=F	0.002	0.001
N	363	363

Note: This table reports the results of the estimation of Equation (1), in which the dependent variable is an indicator for whether a response mentions the categories listed in the rows. Standard errors are clustered at the student level and are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

$$p_j = \frac{\#\text{ of occurrences of j}}{\text{total word-occurrences in corpus}}$$

Note that in this sense, the point estimates do not distinguish between words that occur once in many messages and words that occur many times in a single message: only the total number of occurrences across all messages matter.

Measure of Divergence

In order to compare the differences in language used to respond to male students and female students, we define a measure of divergence, which compares the distribution of words in the female corpus to the distribution of words in the male corpus.

Before defining the measure, we must deal with one critical issue: how to treat words which occur in one corpus but not the other. In our application, the set of words that are not shared across corpi is actually quite large. This can be seen in Table B3. Of the total 3,855 unique words in responses to female students, nearly half are not found in the male responses.

Table B3: Vocabulary Overlap of Responses to Female and Male Students

Analysis	Total Words	Shared Words	Female Only Words	Male Only Words
All	4,817	1,928	1,927	962
Broad	3,045	1,093	1,195	757
Factual	557	135	365	57
Specific Cutthroat	2,444	835	1,123	486
Specific Work-Life	2,402	926	1,020	456

To accommodate this feature of our data, we follow Bohren et al. (2018) and use what we define as the smoothed Kullback-Leiber (K-L) divergence of two corpi. This is the K-L divergence between the two distributions with Lidstone smoothing applied. We use a smoothing parameter of 0.5. The formal definition of our smoothed K-L divergence is given below.

Definition 1 *Given corpus F and corpus M, let V_i denote the vocabulary in corpus i and $C_i()$ denote a function giving the count of a word in corpus i. Then the* ***smoothed K-L divergence*** *of the distributions of F from M is given by:*

$$D_{KLS}(F, M) := \sum_{w \in V_F \cup V_M} p(w) log\left(\frac{p(w)}{q(w)}\right)$$

where:

$$p(w) := \frac{C_F(w) + 0.5}{\sum_{s \in V_F} C_F(s) + 0.5|V_F \cup V_M|}$$

$$q(w) := \frac{C_M(w) + 0.5}{\sum_{s \in V_M} C_M(s) + 0.5|V_F \cup V_M|}$$

We can interpret this measure as the expectation of the logarithmic difference of the distributions, where the expectation uses the female word distribution. In this sense, we are measuring how likely it is that the male observations were taken from the female distribution.

Estimation Procedure

To estimate the K-L divergence metric, we use the definition and replace all probability distributions with their sample analogues. To perform inference we use the bootstrapping procedure outlined in Bohren et al. (2018). This procedure consists of the following: (1) Count the number of responses to male students (N_M) and the number to female students (N_F). (2) For each bootstrap iteration, randomly sample without replacement N_F responses from the full set of responses. Call these responses the placebo female group. (3) Call the remaining N_M responses the placebo male group. (4) Calculate the relevant divergence metric using the placebo groups instead of the true gender. (5) The p-value is the percentage of bootstrap estimates which are less than the point estimate.

To derive what we call p-values clustered at the student level, we perform the following block bootstrap procedure: (1) Count the number of unique male students (N_M) and the number of unique female students (N_F). (2) For each bootstrap iteration, randomly sample without replacement N_F students from the full set of students. Call these students the placebo female group. (3) Call the remaining N_M students the placebo male group. (4) Calculate the relevant divergence metric using the placebo groups instead of the true gender. (5) The p-value is the percentage of bootstrap estimates which are larger than the point estimate.

Consistent with the prior literature, 1,000 bootstrap replications were performed to calculate p-values for each K-L divergence estimate (1,000 replications per p-value).

Data Preparation and Analysis Tools

The sample restrictions are the same as in the main analysis: the 76 students whose names unambiguously convey their gender and who completed the study. The response sample is limited those received within 21 days.

The K-L divergence analyses were conducted using R 3.5.3. The text responses are processed using the packages "stringr" and "quanteda." The command "textstat_frequency" is the main command used to compute word frequencies. Words are defined to be sets of letters separated by spaces. The only processing performed on message text is the removal of punctuation and the removal of the word "x." "x" was used to manually redact messages of identifying information

like company and person names. Other than these two processing steps, no other processing was performed. Words are not stemmed and stop words are not removed.

Results

The K-L divergence metric is reported in Table B4. In addition to the overall analysis (denoted "All"), the analysis is performed by question type: broad, specific work/life balance, specific competitive culture, and factual.

Table B4 reports point-estimates of the K-L divergence of the male response corpus from the female response corpus. It utilizes the smoothed K-L divergence metric given in Definition 1; p-values are computed using bootstrapping responses. Clustered p-values are computed from bootstrapping students.

Considering all of the responses received, the responses to female students are not drawn from a different word distribution than the responses to male students (p-value=0.745). When we look by question type, we also do not see significant gender differences in the word distributions used for messages to male vs. female students.

Table B4: Smoothed Kullback-Leiber Divergence: Male vs. Female Students

	Responses	K-L Divergence	p-value	Clust. p-value
All	913	0.102	0.688	0.745
Broad	363	0.159	0.817	0.848
Factual	34	0.299	0.399	0.394
Narrow Cutthroat	264	0.168	0.476	0.554
Narrow Work-Life	252	0.158	0.387	0.458

C.3 Sentiment Analysis

To measure the tone and emotional content of the messages, we utilize lexicon-based sentiment analysis. Lexicon analyses rely on human-coded databases of words mapped to emotions. The two we utilize are the National Research Council Canada (NRC) lexicon and the Bing lexicon. The NRC lexicon provides eight emotional categories and two sentiment categories (positive or negative). The Bing lexicon provides two sentiment categories only.[44]. The NRC lexicon contains 6,468 unique words, and each word can have multiple sentiments/emotions (categories are not unique).[45] The Bing lexicon contains 6,785 unique words, and all but three words are uniquely classified as either positive or negative.

[44]The website with supporting information is here: `www.cs.uic.edu/~liub/FBS/sentiment-analysis.html`

[45]The website with supporting information is here: `https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm`

For each lexicon and for each sentiment/emotional category in each lexicon, we compute a sentiment score that we call the "sentiment fraction." The sentiment fraction of sentiment j and response i is given by:

$$SF_{i,j} = \frac{\#\text{ words of sentiment } j \text{ in message } i}{\text{total words in message } i}$$

This normalizes sentiment score with respect to message length, and provides a measure of the emotion/sentiment per word in the message. All word counts are counting the number of occurrences of words, not the number of unique words. In this analysis only, we exclude a list of words that are industry related that happen to have sentiment connotation. These include words like "lawyer", which in normal conversation would have a negative connotation, but because our experiment involved discussing a career in law, it has a neutral connotation. As a result, these words are excluded from both from the sentiment count (numerator) and the count of words (denominator).

In terms of vocabulary coverage of the lexicons, out of a total of 4,804 words, 691 words are classified under the Bing lexicon and 936 are classified under the NRC lexicon.

The sentiment analysis, which includes t-tests (adjusted for multiple hypothesis testing) of the difference of means, is presented in Table B5. The table compares the mean fraction of words of each sentiment within responses to male and female students. Overall, there are no significant gender differences in the sentiment of responses, nor are there differences in the sentiment of responses to any particular question.

Table B5: Gender Differences in Sentiments of Responses

	All			Broad			Factual			Narrow (Cutthroat)			Narrow (Balance)		
	Diff	p-val	Clust.	Diff	p-val	Clust.	Diff	p-val	Clust.	Diff	p-val	Clust.	Diff	p-val	Clust.
Bing Lexicon															
Negative	0.002	0.505	0.529	0.002	0.631	0.571	0.012	0.554	0.418	-0.003	0.559	0.566	0.004	0.163	0.188
Positive	0.000	0.870	0.889	0.002	0.668	0.702	0.017	0.554	0.418	0.002	0.624	0.579	-0.004	0.455	0.462
NRC Lexicon															
Anger	-0.001	0.950	0.969	0.000	0.990	0.974	0.002	0.688	0.515	-0.003	0.567	0.691	0.000	0.991	0.995
Anticipation	0.001	0.993	0.993	0.005	0.822	0.883	0.023	0.432	0.401	-0.002	0.987	0.989	-0.004	0.810	0.884
Disgust	0.000	0.959	0.970	0.000	0.990	0.974	-0.004	0.236	0.515	0.000	0.987	0.989	-0.001	0.890	0.905
Fear	0.000	0.993	0.993	0.003	0.581	0.433	0.002	0.598	0.449	-0.003	0.728	0.835	-0.002	0.794	0.874
Joy	0.002	0.950	0.949	0.004	0.933	0.903	0.006	0.942	0.949	0.005	0.568	0.633	-0.004	0.808	0.874
Negative	-0.001	0.990	0.993	0.002	0.983	0.961	0.004	0.633	0.515	-0.004	0.490	0.633	0.000	0.991	0.995
Positive	-0.004	0.826	0.752	-0.004	0.965	0.947	0.016	0.834	0.791	0.002	0.987	0.989	-0.012	0.530	0.390
Sadness	0.000	0.993	0.993	0.001	0.990	0.974	0.000	0.995	0.997	0.001	0.987	0.989	-0.002	0.794	0.861
Surprise	0.002	0.746	0.752	0.001	0.990	0.974	0.015	0.688	0.515	0.001	0.987	0.989	0.002	0.890	0.905
Trust	-0.001	0.993	0.993	0.002	0.990	0.974	-0.006	0.942	0.949	0.002	0.978	0.975	-0.006	0.681	0.571

Note: This table displays t-tests comparing the fraction of words from a particular sentiment across female and male students. Negative differences indicate female students received a higher fraction of words from the sentiment than males. p-values are corrected for multiple comparisons (within lexicon). Westfall-Young corrections are performed using the STATA package -wyoung-, with 1,000 bootstrap replications. Column -Clust- reports WY adjusted standard errors clustered at the student level.

D Online Appendix: Potential outcomes framework for incorporating selection

Let $y_p(1)$ be the response of professional p to a female student and $y_p(0)$ be the response of the same professional to an equivalent male student, asking the same question. Our main experiment described in Section 2 allows us to estimate β, the average gender difference in the responses of professionals in our sample:

$$\beta = E\left(y_p(1) - y_p(0) | p \in 1, ..., P\right)$$

However, students may prefer some professionals over others. Let the set of professionals preferred by students be $\mathbf{P} \subseteq \{1, ..., P\}$. Given information about student preferences over professionals, we can define β^p to be the average gender difference in responses among the preferred professionals:

$$\beta^p = E\left(y_p(1) - y_p(0) | p \in \mathbf{P}\right)$$

We also define gender bias among professionals preferred by male and female students:

$$\beta^g = E\left(y_p(1) - y_p(0) | p \in \mathbf{P^g}\right)$$

where $g \in \{m, f\}$ indicates gender and $\mathbf{P^m}$ is the set of professionals preferred by male students and $\mathbf{P^f}$ is the set of professionals preferred by female students. β^m is the average gender bias of professionals preferred by male students, and β^f is the average gender bias of professionals preferred by female students.

Even if $\beta^f = \beta^m$, professionals preferred by female students may have different levels of y than the professionals preferred by male students. We define the average responses to male students in the set of professionals preferred by males and females, respectively, as

$$\alpha^g = E\left(y_p(0) | p \in \mathbf{P^g}\right), \quad g \in \{m, f\}$$

Together, these moments can be used to compute β^s, which represents how responses to female students seeking information from female-preferred professionals differ from responses to male students seeking information from male-preferred professionals:

$$\beta^s = \underbrace{\beta^f + \alpha^f}_{E\left(y_p(1) | p \in \mathbf{P^f}\right)} - \underbrace{\alpha^m}_{E(y_p(0) | p \in \mathbf{P^m})}$$

Our experimental design allows separate identification of α^g, β^g for $g \in \{m, f\}$. Note that $E\left(y_p(1) | p \in \mathbf{P^f}\right)$ and $E\left(y_p(0) | p \in \mathbf{P^m}\right)$ are available in observational data, meaning that one can

estimate $\beta^s = \beta^f + \alpha^f - \alpha^m$. However, using only observational data on student selection of professionals and the outcomes of these interactions, the average bias of professionals contacted by male and female students (β^m and β^f, respectively) are not identified unless $\mathbf{P^m} = \mathbf{P^f}$.