



HCEO WORKING PAPER SERIES

Working Paper



HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Informed Choices: Gender Gaps in Career Advice*

Yana Gallen and Melanie Wasserman

April 2021

Abstract

This paper estimates gender differences in access to informal information regarding the labor market. We conduct a large-scale field experiment in which real college students seek information from 10,000 working professionals about various career paths, and we randomize whether a professional receives a message from a male or a female student. We focus the experimental design and analysis on two career attributes that prior research has shown to differentially affect the labor market choices of women: the extent to which a career accommodates work/life balance and has a competitive culture. When students ask broadly for information about a career, we find that female students receive substantially more information on work/life balance relative to male students. This gender difference persists when students disclose that they are concerned about work/life balance. In contrast, professionals mention workplace culture to male and female students at similar rates. After the study, female students are more dissuaded from their preferred career path than male students, and this difference is in part explained by professionals' greater emphasis on work/life balance when responding to female students. Finally, we elicit students' preferences for professionals and find that gender differences in information provision would remain if students contacted their most preferred professionals.

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 Giacobasso 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, 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, MEA 2021, WEAI 2021, Upjohn Institute, University of Minnesota Carlson, Boston University Questrom, and Discrimination and Disparities Seminar. This study was registered in the AEA Social Sciences Registry under AEARCTR-0005464.

1 Introduction

Information is often obtained through informal and private means: a conversation with a professor in her office, an interaction with a prospective employer over email, or a text message with a friend (Montgomery, 1991; Duflo and Saez, 2003; Acemoglu et al., 2014). Because these communications take place person-to-person, an individuals' access to information may depend on their personal characteristics and who is willing to respond. Despite the fact that informal interactions are an everyday occurrence, the difficulty of observing these types of communications has limited research on their incidence, content, and impact on economic decisions. Using a large-scale field experiment, this paper provides a window into informal information exchanges and how they differ based on the characteristics of participants. To the extent there are disparities in access to informal information, these differences could help explain broader phenomena such as occupational segregation by gender, socioeconomic disparities in college enrollment, and race/ethnic differences in health behaviors (Bertrand, 2011; Blau and Kahn, 2017; Cortes and Pan, 2018; Porter and Serra, 2020; Hoxby and Turner, 2015; Alsan et al., 2019).

This paper focuses on informal exchanges between college students and professionals for the purpose of gathering information regarding their career decisions. We investigate whether student gender causally affects the information that students receive regarding various career paths. We recruit undergraduate students interested in learning about various careers to send messages on a popular online professional platform. We randomize whether each of the 10,000 professionals in our sample receives a message from a male or a female student. The messages sent by students are preformulated questions seeking information about the professional's career path. The use of the professional networking platform as well as the text of the messages are based on a university career center's guidance for informational interviews. We use four question templates, each intended to test a specific hypothesis regarding gender differences in information acquisition.

The first question asks broadly about the pros and cons of a given career path, which allows us to test whether professionals emphasize different career attributes to male and female students. The two career attributes we focus on are work/life balance and competitive culture, both of which differentially affect the labor market choices of women (Goldin, 2014; Wiswall and Zafar, 2018; Cubas et al., 2019; Niederle and Vesterlund, 2011; Flory et al., 2015). One reason professionals may differentially emphasize these career attributes to

female students is that the professionals statistically discriminate, that is, believe female students care more about these issues. Using the next two questions, we test whether gender differences in information provision persist when we control for professionals' perceptions of what students care about. These questions ask specifically for information about work/life balance or competitive culture. A remaining explanation for professionals' differential willingness to discuss these career attributes with female students is that professionals believe female students will experience more challenges with these issues during their careers. The fourth question is intended to control for the potentially different future experiences of male and female students by testing whether there are gender differences in responses to a fact-based question on the billable hours requirements in one of the career paths.

Our main finding is that the information professionals provide depends on student gender. While male and female students have similar response rates to the broad question that asks about the pros/cons of the professional's field, the text of the responses reveals substantial gender disparities. Professionals are more than two times as likely to provide information on work/life balance issues to female students relative to male students. The vast majority of these mentions are negative and increase students' concern about this issue. When students ask specifically about work/life balance, female students receive 28 percent more responses than do male students, meaning even when male students specifically request information on work/life balance, they obtain less of it. In addition, this means that the differential emphasis on work/life balance to female students in responses to the broad question is not entirely driven by perceptions that female students care more about this issue. Interestingly, there is no differential emphasis on workplace culture to female students, through either mentions in responses to the broad question or response rates to the specific competitive culture question. We also find that differences in the composition of professionals who respond to male and female students do not explain *any* of the greater emphasis on work/life balance to female students. To our knowledge, this is the first study to causally isolate how individuals' gender shapes the information they receive about career attributes.

To understand the ramifications of differential information provision in informal interactions, consider the analogous *formal* information provision scenario. Suppose students visit a university career planning office and express interest in a particular career. If gender-specific career guidance pamphlets are available to all students, some female students may

choose the male pamphlet, and some male students may choose the female pamphlet.¹ Some students may take both pamphlets. Differential information provision—akin to what we find in informal interactions—would imply that female students receive a pamphlet emphasizing certain career attributes, while male students receive a pamphlet emphasizing others, and no student is aware that there are two pamphlets available or that they were allocated based on gender. 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 conduct a follow-up survey to investigate whether gender gaps in information received translate into gender gaps in career outcomes. A growing literature has demonstrated that information impacts education and career choices (Jensen, 2010; Nguyen, 2008; Wiswall and Zafar, 2015a). In our survey, students are asked whether, relative to the start of the study, they are more or less likely to enter their preferred career path. Results of the survey indicate that female students become less likely to enter their preferred career path relative to male students. Moreover, nearly half of the gender gap in career plans can be explained by the fact that work/life balance is emphasized more to female students than to male students. The deterring effect of work/life balance mentions makes sense given the negative content of this information. We view this as suggestive evidence that gender differences in access to information may lead to occupational segregation by gender.²

Our experimental design resembles a traditional correspondence study in that the researchers maintain control over the format and content of the message sent, and student gender is orthogonal to professionals' characteristics. In contrast to a traditional correspondence study that uses fictitious individuals, our paper incorporates real students interested in career information, who seek advice from real professionals. While incorporating real students cedes precise control over student attributes, we take several

¹Gender-specific articles on careers are commonly available in newspapers, periodicals, and trade newsletters. In addition, books such as *Lean In*, by Sheryl Sandberg, provide career advice tailored to women.

²While this paper cannot provide direct evidence on the normative implications of gender differences in information received, we discuss a few possibilities. If there are gender differences in the importance of work/life balance information, then optimal information provision would also be differentiated by gender. A survey of students from the same university reveals that men demand *more* information on work/life balance than women, suggesting mismatch between information supplied and information demanded. However, 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 literature that estimates women's preference for temporal flexibility focuses on individuals at age 30 (Wiswall and Zafar, 2018). If professionals do not know or are biased in their beliefs about students' preferences, then gender differences information provision may be suboptimal and labor market inequities may be propagated from older to younger cohorts.

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.

The field experiment permits estimation of *average* gender differences in information received. Any average difference may be amplified or attenuated depending on students' preferences for whom to contact. While every professional contacted is a graduate of a top university working in one of the fields of interest to students, we find that students do not prefer all professionals equally. Before sending messages, students are asked to rank the professionals they are most interested in asking the types of questions in the study. We use these rankings to test whether preferred professionals give different answers than the average professional in the study, and whether female students select professionals who differentiate their answers less (or more) by student gender. After accounting for student selection into informal interactions, we find that responses to female students are still more likely to mention work/life balance issues.

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 *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). These papers establish that individuals, on average, are misinformed about fundamental economic parameters that guide decision making, such as the earnings associated with various college majors. Furthermore, individuals update their beliefs, stated choices, and actual choices when presented with accurate information. 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 omits the job-seeking motive of messages, 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 selection.

Finally, 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).³ 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 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

³See Lahey and Oxley (2018) for empirical estimates of time spent reviewing resumes.

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 to facilitate the experiment. The cost of incorporating real students is ceding control over some student characteristics. While we believe the online setting mitigates the concern that other student characteristics confound the effect of student gender, we test the sensitivity of our results to gender differences in the observable and unobservable characteristics of students.

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). A further innovation of the present paper is to incorporate the fact that, in reality, students are unlikely to contact professionals at random. We 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.⁴ 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 adjust 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 professionals on an online professional networking platform. Our design permits us to (1) estimate the causal effect of student gender on information provision, (2) assess how student preferences for professionals affect information provision, and (3) explore the consequences of differential information provision on student career choice.

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

⁴The list of studies is available upon request.

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.

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.⁵ 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.⁶ 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. We provided students with the same photo of an iconic university building to use as a profile picture. We confirmed that students created a profile with the requisite restrictions through profile screenshots and independent verification on the platform.

The pool of professionals consists of approximately 10,000 individuals on the online professional platform 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 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. Professionals' profiles were checked to ensure they have work experience in one of the four fields.

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 online platform. We provided the text of the initial message that students sent to professionals. Each professional-

⁵Instructions for this process are available upon request.

⁶In 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.

student communication used one of four message types, which were designed to emulate a conventional request for career information during an informational interview.⁷ To test whether professionals emphasize different career attributes to male and female students, the broad message asked about the pros and cons of the professional’s field.⁸ To test whether men and women receive different advice conditional on raising a particular concern, we sent two message types that ask whether specific career attributes are concerns—work/life balance and competitive culture. We selected these concerns 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). Last, to test whether there are differences in the factual content of information given to men and women, we asked about the billable hours requirement for a lawyer at a large law firm. This message type was only sent to law professionals. 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 (law only): Asks about the billable hours requirements 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.

⁷These 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.

⁸Gallen 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.

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. The random assignment ensures that student characteristics are orthogonal to professional characteristics. The students sent the messages on weekdays during typical working hours.⁹ 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.

Students were asked to provide the initial responses they receive within 21 days of sending the messages.¹⁰ 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. The survey also asked whether students had follow-up interactions with professionals. Upon successful completion of this survey, students were paid \$75.

Methodological Advance and Identification: In several ways, our experimental design

⁹In 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.

¹⁰The vast majority of responses are received within two weeks of sending a message.

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. It is likely that student characteristics such as college major will be correlated with student gender. This means that the effect of student gender on the responses that students receive could be confounded by other student attributes.

The online setting serves to mitigate concerns that other student characteristics confound the effect of student gender. As discussed above, we ask students to limit the information provided on their profiles. But even on characteristics visible on the profile—such as their college majors, graduation year, and number of connections on the platform—male and female students may differ. Our regression specifications control for all student characteristics that are directly observed on the site. There still remains the possibility that there is publicly available online information on the student participants that would, for example, show up if a professional chose to conduct an online search of a student’s name and university affiliation. This additional student information could influence the professionals’ propensity to respond and the information that they impart. To address this concern, we test whether the effect of student gender is sensitive to the inclusion of student characteristics from the background survey and whether the student has an online presence aside from their profile. We also examine whether the results are robust to restricting the

sample to students without an online presence.

3 Data, Sample Restrictions, Summary Statistics

Data: We collect data on response rates and the text of initial responses. We analyze the text of initial responses 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. We code professionals' mentions of work/life balance 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."¹¹ 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 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, 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 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.

¹¹<https://dictionary.cambridge.org/us/dictionary/english/work-life-balance>

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.

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. 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 messages constitute sample attrition. We diligently followed up with all student participants and found that students who took a very long time to send 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.

We limit the sample to students whose first names unambiguously convey their true gender. We do so 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.¹²

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

¹²Our final sample of professionals was 10,003, so three students were assigned 101 professionals. Two of these students are in the final student sample.

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 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).¹³ 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 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 that the randomization was successful.

4 The Effect of Student Gender on Information Received

4.1 Econometric Framework

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

$$Y_{im} = \alpha + \beta StudentFemale_i + X'_{im} \gamma + \epsilon_{im} \quad (1)$$

where the dependent variable, Y_{im} , is an outcome such as an indicator for whether a message m sent by student i receives a response, or whether the response that student i receives to message m mentions a specific career attribute. The independent variables are an indicator for whether the student is female, $StudentFemale_i$, as well as a vector of message and student controls, X_{im} . 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 student characteristics 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

¹³This 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).

with the profile restrictions.¹⁴ Standard errors are clustered at the student level.

While the online setting limits the concern that student characteristics aside from those directly observable on the site confound the effect of student gender on response rate, it is possible that professionals find additional information on students through an online search. We test whether the coefficient on student gender is sensitive to the inclusion of additional student characteristics that may be available elsewhere online. These additional controls include 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 name, we consider this variable partially observed based on information in the student’s profile. In the regression results below, we 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.

4.2 Broad Question

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 and 2. 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 $StudentFemale_i$ 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).¹⁵ Based on these

¹⁴Some 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.

¹⁵Heckman 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

results, we conclude that professionals are just as willing to engage with male and female students who sent the broad message.

Work/life balance: 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 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. Overall, 11 percent of responses bring up work/life balance issues (Appendix Table A1). We observe, however, that responses to female students mention these issues more than twice as much as responses to male students. 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. When controls for student race/ethnicity are included in column 2, the coefficient on student female declines slightly to 0.072, indicating that student race/ethnicity could be conveyed through student names and is correlated with student gender. The further inclusion of student characteristics potentially observable elsewhere online marginally increases the coefficient to 0.076. 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 A3).

In Appendix Table A4 we investigate 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.¹⁶ We find that there are pronounced gender differences in mentions of the duration of a typical workweek. Responses to female students are 5.4 percentage points more likely to mention the duration of the typical workweek relative to male students. Given that the male mean for this category is 4.7 percent, female students hear about weekly hours at more than twice the rate of male students. It also appears that responses to female students are more likely to contain information about work schedule flexibility, but this variance of unobservable characteristics of male and female students is the same.

¹⁶These topics correspond to two O*NET work context categories and one supplemental category. For more information on these categories, see Online Appendix Table B.1.

gender difference is not statistically significant.

Mentions of work/life balance tend to be negative. 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. Consistent with our finding that female students receive more information on 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 A5).

Workplace culture: We also estimate gender differences in mentions of workplace culture in responses to the broad question. Only six of the responses to the broad question explicitly mention competitive aspects of workplace culture. Due to this low rate, we 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). Table 4 columns 4–6 report the results from estimating Equation (1), which tests whether responses to female students include more mentions of workplace culture. In column 4, the specification with the basic controls for message characteristics and student characteristics available on the profile, the point estimate for the coefficient on student female is close to zero. The coefficient declines a bit and remains statistically insignificant with the inclusion of additional student controls.

4.3 Specific Questions

One reason the professionals may differentially emphasize work/life balance issues to female students relative to male students is that the professionals statistically discriminate, that is, believe female students care more about these issues. In this section, we investigate whether gender differences in responses persist when students specifically ask whether work/life balance is a concern in the professional's field.

Response rates: In Table 3, columns 4–6, we find that, in contrast to the broad

question, student gender does affect professionals' propensity respond to the 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. Again this estimate is insensitive to the inclusion of additional student covariates. This means that even when male students specifically request information on work/life balance, they obtain less of it. 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.

Content: The text of the specific questions is designed to elicit a "yes" or "no" response. The message first describes that the student has heard work/life balance (cutthroat culture) is a challenge in the professional's field. Then the message asks directly whether the professional thinks this is a valid concern. Among the responses that answer the question in the text of the message, we manually classify the responses into those that say yes, it is a concern; those that say it depends on the workplace or specific career path in the field; and those that say no, it is not a concern.

Summary statistics from Appendix Table A1 show that overall the responses to the work/life balance question confirm that work/life is a concern in the professional's field. Only seven percent state that work/life balance is not a valid concern. Approximately half state that work/life balance depends on the firm, specific path, boss, etc., and 44 percent report that it is indeed a concern. In contrast, 30 percent of the responses to the competitive culture question state that it is not a concern in the professional's field. Only 16 percent confirm it is a concern, and the remaining 54 percent state that it depends.

Given that responses to the work/life balance question overwhelmingly validate the concern, it is possible that these responses make college students more concerned about work/life balance in the professional's field. Based on the evaluations of college students who were not study participants, more than half of the responses to the work/life balance question make students more concerned about work/life balance in the professional's field.

The results demonstrate that women receive more responses to the work/life balance question. Responses to both male and female students tend to make students more concerned about work/life balance in the professional's field. We additionally test whether there

are gender differences in the content and tone of the responses to the specific questions. The results are reported in Appendix Tables A5, A6, and A7. While the coefficients are imprecisely estimated, the responses to female students do not display meaningful content or tone differences relative to the responses to male students.

4.4 Factual Question

One reason that professionals may continue to differentially engage with female students in the specific work/life balance question is that they think the answer to the question is gender specific. In the factual question, we shut down this possibility. Recall that the factual question asks about the billable hours requirement for a lawyer at a large law firm. This question has an objective answer, which is invariant to the gender of the lawyer. If professionals nonetheless give different responses about a fact to male and female students, then we would conclude that this due to professionals' subjective determination of which topics are more important to discuss with male vs. female students. We find that professionals are 80 percent more likely to respond to female students than male students asking the factual question (Table 3, columns 10–12), though the coefficient is not consistently statistically significant. Appendix Table A8 tests for gender differences in hours quoted to students. While the point estimates suggest that women are quoted higher hours requirements, unfortunately, we are underpowered to detect large differences in hours quoted to male and female students.

4.5 Separating Motivations for Gender Gaps in Information Provision

To summarize the main results, we find that professionals differentially emphasize work/life balance to female students. This emphasis is apparent both in responses to the question that asks broadly about the pros and cons of the professional's field, as well as in higher response rates to the specific work/life balance question, which generally confirm that work/life balance is a concern. In contrast, we find that professionals do not differentially emphasize workplace culture to female students.

Using our experimental design, we investigate why professionals emphasize work/life balance to female students. One possibility is that professionals want to talk to female students more than male students, about all topics (not just work/life balance). This explanation is unlikely, since response rates to female students are not consistently higher

across question types. A second explanation is that professionals believe female students care more about work/life balance and tailor their responses accordingly. The specific work/life balance question allows us to rule this out as the exclusive explanation, since professionals are still more willing to engage with female students when students bring up work/life balance as a concern. A third possibility is that the answer to the question is gender specific. For example, women may struggle more on average with work/life balance, and professionals simply report this gender-specific answer to students of the corresponding gender. While the results for the factual question are imprecisely estimated, we provide suggestive evidence that professionals continue to be more responsive to female students even when the answer to the question is objective and invariant to student gender. A remaining possibility is that professionals subjectively determine that work/life balance is more important to discuss with female students. We discuss the implications of this gender differentiation below.

5 Implications of Gender Differences in Information Received

In this section, we discuss the positive and normative implications of gender differences in information received. We empirically investigate two positive questions: (1) what are the ramifications of gender gaps in information provision for gender gaps in students' career choices? and (2) do professionals supply the information that students demand? We also discuss a few normative implications of gender gaps in information provision.

Career choice: A large literature establishes that information shapes individuals' college major, occupation, and job choices (Wiswall and Zafar, 2015a,b; Coffman et al., 2017; Subramanian, 2020; Porter and Serra, 2020; Delfino, 2021). Although our experiment is not designed to estimate the causal effect of information on students' career choices, we provide descriptive evidence from a survey administered at the conclusion of the study, which asks students about their career plans relative to the start of the study. Of the 76 students in the main sample, 73 completed the survey. We regress an indicator for whether a student was dissuaded from their preferred career path on student characteristics and

outcomes from the experiment:¹⁷

$$D_i = \gamma_0 + \gamma_1 StudentFemale_i + M_i' \gamma_2 + X_i' \gamma_3 + \varepsilon_i \quad (2)$$

where M_i include characteristics of the messages received by student i , such as whether the student received responses concerning work/life balance, and X_i is a vector of student characteristics, including the basic student controls, race, and student's preferred career path. We note that this analysis is descriptive. It is possible that female students update their career plans over the three-week study period differently than male students, for reasons unrelated to the study itself. That said, we view the following results as suggestive evidence of the role of information in generating gender gaps in career choices.

Relative to the start of the study, women are more dissuaded from their preferred career path than are men. Table 5, column 1, shows that female students are 9.5 percentage points more likely to be dissuaded from their preferred career path. When we add controls for other student characteristics as well as students' preferred career path, the gender difference increases to 11.4 percentage points. Including controls for whether a student received a response to the broad question that mentions work/life balance and whether the student received a response to the specific work/life balance question (Table 5, column 3) reduces the coefficient on student female by nearly half, indicating that part of the gender difference in career plans can be attributed to the different information female students received concerning work/life balance. In column 5, we include controls for work/life balance and workplace culture as well as for message length, and find that message characteristics beyond emphasis on work/life balance do not explain more of the gender gap in career plans. Our results suggest that the information transmitted through informal exchanges discourages women from their preferred career path, which may be relevant for the design of interventions that address potential information frictions that could lead to female underrepresentation in these fields (Bursztyn et al., 2020; Porter and Serra, 2020; Bayer and Rouse, 2016).

Mismatch: 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

¹⁷Students 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 measure dissuasion, D_i , as an indicator for responses of 4 or below to this question for student i 's preferred career path.

career-related topics. While all students are interested in discussing work/life balance with a professional, men allocate 14 percent of their time and women allocate 10 percent of their time to this topic, a statistically significant difference (Appendix Figure A3).¹⁸ This suggests that the information professionals supply does not match the information students demand. In particular, 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.

Normative implications: While this paper cannot provide direct evidence on the normative implications of gender differences in information received, we discuss a few possibilities. Gender differences in information provision could be optimal if there are gender differences in the importance of work/life balance information for career choice. Based on the survey results described above, male students demand *more* information on work/life balance than female students. However, 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 literature that estimates women’s preference for temporal flexibility focuses on individuals at age 30 (Wiswall and Zafar, 2018). Gender differences in information provision may be suboptimal if professionals do not know or are biased in their beliefs about students’ preferences. This could happen if there are secular changes in preferences for work/life balance or in the attributes of the labor market, and experienced professionals are unaware of these changes.

6 Other Mechanisms and Additional Results

Although we rely our experimental design to understand the sources that could generate the differential provision of information to male and female students, in this section, we test several other mechanisms, including the composition of professionals who respond and the differential treatment of students by certain subgroups of professionals.

Composition of professionals who respond: To investigate whether differences in the types of professionals who respond to male and female students contribute to gender differences in responses, we modify Equation (1) as follows:

¹⁸A 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>

$$Y_{imp} = \alpha + \beta StudentFemale_i + X'_{im}\gamma + X'_p\delta + \epsilon_{imp} \quad (3)$$

where the dependent variable is defined as above, but now we introduce the subscript p , which represents that a student contacts or receives a response from professional p . We test the sensitivity of β , the coefficient on $StudentFemale_i$, to the inclusion of controls for professionals' characteristics. Specifically, we include professional gender, undergraduate graduation year, undergraduate institution selectivity, whether the professional is an alumnus of the student's college, whether the professional has a graduate degree, and the extent of their network thickness. If different types of professionals respond to male and female students, and their willingness to bring up work/life balance issues is correlated with the professionals' characteristics, then we would expect the inclusion of professional controls to attenuate the coefficient on $StudentFemale_i$. This could happen if, for example, female professionals are more likely to respond to female students, and are also more inclined to bring up work/life balance. In Tables 6 and 7, across all outcomes, the coefficient on $StudentFemale_i$ is invariant to the inclusion of professional controls, supporting the notion that differences in the composition of professionals who respond to male and female students do not explain the greater emphasis on work/life balance to women.

Heterogeneity by professional characteristics: Next we investigate whether the differential provision of information to female students is concentrated among certain subgroups of professionals. The results are reported in Table 8, where each entry represents the coefficient on $StudentFemale_i$ from a separate regression, with the subgroup of professionals in the column title. Gender differences in response rates to the broad question exhibit little heterogeneity based on professional characteristics. In the specific work/life balance question, the higher response rates for female students are concentrated among younger professionals, those who are alumni of the students' college, and professionals in finance and management consulting. In the responses to the broad question, the greater emphasis on work/life balance issues to female students is concentrated among female professionals, alumni, those who are older, professionals with a degree from an Ivy League university, and those in finance and law (panel D).¹⁹

¹⁹While 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

It is worth noting that university career centers encourage college students to use this professional networking platform to reach out to university alumni for career advice. When we limit our analysis to professionals who are alumni of the students' university (Table 8 column 8), we find that these professionals are more responsive and exhibit more gender differentiation in their responses than professionals in the full sample.

Crowd out and other message content: Finally, we 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 length of replies, suggesting that the additional emphasis on work/life balance to female students may displace other information (Appendix Table A9). 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.

7 The Role of Selection

In order to estimate average gender differences in information received, we randomly assigned professionals to students. We find that, on average, professionals differentially emphasize work/life balance issues to female students. However, if in reality female students seek out different professionals than male students, average differences may not be informative of students' experiences. For example, 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).²⁰ It is reasonable to believe that individuals do not sample randomly from their full choice set. 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

for mentions of work/life balance.

²⁰In 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.

how equilibrium racial discrimination is affected by residential sorting. To our knowledge, this is the first paper to incorporate agent selection into a correspondence-style study.²¹

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:

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

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

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

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

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

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

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

Together, these moments can be used to compute β^s , which represents how responses to

²¹Another 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.

female students seeking information from female-preferred professionals differ from responses to male students seeking information from male-preferred professionals:

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

Our experimental design allows separate identification of α^g, β^g for $g \in \{m, f\}$. Note that $E(y_p(1)|p \in \mathbf{P}^f)$ and $E(y_p(0)|p \in \mathbf{P}^m)$ 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 $\mathbf{P}^m = \mathbf{P}^f$.

Estimating student preferences for professionals: To estimate students' preferences over 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.²² 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—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).²³ Using the estimates, we calculate the predicted probability that a professional is ranked first by the student, for each student-professional pair in the full sample. We use the predicted probabilities to form the preferred professional sets \mathbf{P} , \mathbf{P}^m , and \mathbf{P}^f from the full set of professionals in our sample. This method creates a data set that simulates an assignment process in which students chose

²²We did not ask students to rank the lawyers whom they would want to ask about billable hours requirements—there is no factual question ranking.

²³More specifically, for students we use the following characteristics: gender, network size, additional profile information, economics major indicator, STEM major indicator, graduation year, and an indicator for whether the student had any problem with their profile (for example, being slow to upload the profile picture). We interact these student characteristics with the following professional characteristics: 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. We also interact indicators for the student's preferred field with indicators of the professional's field.

the 100 professionals they would most prefer to contact from the full set of professionals in the sample (half asking a broad question, and a quarter devoted to each specific question).²⁴

Appendix Table A12 shows that the preferred sample of professionals differs in a number of ways from the full sample, and moreover, female- and male-preferred professionals differ.²⁵ Given these differences, the average gender gaps estimated in Section 4 may change when we incorporate student selection.

Econometric framework for incorporating selection: To estimate average gender differences in responses among preferred professionals, β^p , we use the following specification:

$$Y_p = \alpha + \beta^p FemaleStudentSent_p + \gamma X_p + \epsilon_p \quad (4)$$

where the sample is restricted to the set of professionals \mathbf{P} preferred by students.²⁶ Y_p is the outcome of a message sent to professional p . $FemaleStudentSent_p$ indicates whether in the main experiment the professional p actually received a message from a female or male student, X_p is a vector of professional-level controls, including characteristics of the student who sent the message to the professional.

We estimate β^f and β^m , the gender bias of professionals preferred by female and male students, respectively, as well as $\alpha^f - \alpha^m$, the difference in average response among professionals preferred by male and female students, with the following specification:

$$Y_{jp} = a + (\alpha^f - \alpha^m)FemaleStudentPreferred_j + \beta^m FemaleStudentSent_p + (\beta^f - \beta^m)FemaleStudentPreferred_j \times FemaleStudentSent_p + cX_p + dX_j + e_{jp} \quad (5)$$

where Y_{jp} is an outcome of a message sent to professional p who is preferred by student j , $FemaleStudentPreferred_j$ indicates whether the student j who prefers this professional is female, $FemaleStudentSent_p$ indicates whether in the main experiment the professional p

²⁴We limit the sample of professionals to the random subset who were contacted by students in our final sample and who were actually asked the question that parallels the question generating preferences. For example, when selecting the set of 48 professionals that a student would want to ask the broad question, we select the 48 professionals predicted to be that student's most preferred from the subset of professionals who were actually asked the broad question by a student in the main experiment.

²⁵Relative 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.

²⁶Note that a professional can be repeated in this sample if multiple students preferred the same professional.

actually received a message from a female student, and X_p is defined as above. We include a vector of additional student controls, X_j , that could be correlated with student gender. Since student j (almost) never contacted professional p in the main experiment, in the simulated data set, the gender of students who selected a particular professional is orthogonal to the gender of students who contacted the professional. We can therefore identify the gender bias of professionals preferred by students overall, and by female and male students (β^p , β^f , and β^m , respectively).

Gender differences in information received, incorporating selection: Table 9 reports the results for response rates. Recall that in the main experiment, response rates to female students were either similar to or higher than response rates to male students. Incorporating student selection of professionals reverses this pattern. Column 1 reports estimates of β^p from Equation (4), which measures gender bias among the preferred professionals \mathbf{P} . Although there are no gender differences in response rates among the preferred professionals, when we separately analyze the female- and male-preferred professionals, we find that female students tend to select professionals with lower response rates overall (in column 2 the coefficient on *FemaleStudentPreferred_j* is -0.041). The estimates of gender bias among male-preferred professionals and the differential bias among female-preferred professionals are negligible. Thus, β^s , the realized gender difference in response rates, incorporating student selection—the contrast between "Female Pref and Female Sent" and "Male Pref and Male Sent Mean" in the bottom half of the table—is also nearly four percentage points. The results are similar for the specific work/life balance and competitive culture questions.

These results also can be used to evaluate the efficacy of policies to reduce gender disparities in response rates. One policy we might consider is steering female students to the professionals preferred by male students, in an effort to increase female response rates. In the bottom half of Table 9, we calculate the counterfactual response rate if female students contacted male-preferred professionals (labeled "Male pref and female sent"). We can contrast this with the experienced response rate for female students (labeled "Female pref and female sent"). Across all question types, there is a substantial gap in the experienced and counterfactual response rates for female students, indicating that such a policy would be effective in increasing response rates to female students.

Table 10 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 2, we observe that male-preferred professionals are substantially more likely to mention work/life balance to female students (coefficient on *FemaleStudentSent_p* is 0.23).²⁷ This gender bias is similar for female-preferred professionals.²⁸ Using these estimates, we can consider the effects of a policy that steers male students to female-preferred professionals, in order to increase male students' access to information on work/life balance. If male students contacted female-preferred professionals, there would be no change in the gender gap in the frequency of work/life balance mentions. This is due to the fact that female-preferred professionals are slightly more gender-biased in their responses and also bring up work/life balance less frequently overall.

8 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

²⁷The lower response rates of female-preferred professionals reduces the gender disparity in hearing about work/life balance unconditional on response. Accounting for non-response, women are still more than three times as likely to hear about work/life balance issues relative to men among preferred professionals, which is a larger effect than in our main experimental results.

²⁸Although female-preferred professionals bring up work/life balance less overall to male students than male-preferred professionals (the coefficient on *FemaleStudentPreferred_j* is -0.030), female-preferred professionals also mention these issues slightly more to women than to men (the coefficient on the interaction term is 0.010).

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.
- Acemoglu, Daron, Kostas Bimpikis, and Asuman Ozdaglar**, "Dynamics of information exchange in endogenous social networks," *Theoretical Economics*, 2014, 9 (1), 41–97.
- Agan, Amanda Y and Sonja B Starr**, "Employer Neighborhoods and Racial Discrimination," Working Paper 28153, National Bureau of Economic Research November 2020.
- Alsan, Marcella, Owen Garrick, and Grant Graziani**, "Does Diversity Matter for Health? Experimental Evidence from Oakland," *American Economic Review*, December 2019, 109 (12), 4071–4111.
- Bayer, Amanda and Cecilia Elena Rouse**, "Diversity in the Economics Profession: A New Attack on an Old Problem," *Journal of Economic Perspectives*, November 2016, 30 (4), 221–42.
- 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**, “New Perspectives on Gender,” in Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Elsevier, 2011, chapter 17, pp. 1543–1590.
- **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.
- Burszty, Leonardo, Alessandra L. González, and David Yanagizawa-Drott**, “Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia,” *American Economic Review*, 2020, 110 (10), 2997–3029.
- 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 Becker’s 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.
- Cortes, Patricia and Jessica Pan**, “Occupation and Gender,” in Laura M. Argys Susan L. Averett and Saul D. Hoffman, eds., *The Oxford Handbook of Women and the Economy*, Oxford: Oxford University Press, 2018, chapter 18, pp. 425–461.

- 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.
- Delfino, Alexia**, “Breaking Gender Barriers: Experimental Evidence on Men in Pink-Collar Jobs,” Working Paper 14083, IZA 2021.
- 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.
- Duflo, Esther and Emmanuel Saez**, “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment,” *Quarterly Journal of Economics*, 2003, *118* (3), 815–842.
- 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.

- 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.
- 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.
- 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.
- 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.

- Montgomery, James D.**, “Social Networks and Labor-Market Outcomes: Toward an Economic Analysis,” *The American Economic Review*, 1991, 81 (5), 1408–1418.
- 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.
- 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.
- Subramanian, Nivedhitha**, “Workplace Attributes and Women’s Labor Supply Decisions: Evidence from a Randomized Experiment,” Working Paper 2020.
- 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.

Tables

Table 1: Student Summary Statistics

	All Students	Male	Female
<i>Profile Information</i>			
Female	0.58 (0.50)		
Expected Graduation Year	2022.24 (1.04)	2022.50 (0.95)	2022.05 (1.08)
Economics	0.62 (0.49)	0.69 (0.47)	0.57 (0.50)
STEM	0.22 (0.42)	0.25 (0.44)	0.20 (0.41)
0-49 Connections	0.46 (0.50)	0.44 (0.50)	0.48 (0.51)
50-249 Connections	0.28 (0.45)	0.25 (0.44)	0.30 (0.46)
250+ Connections	0.26 (0.44)	0.31 (0.47)	0.23 (0.42)
Profile Extra Info	0.47 (0.50)	0.56 (0.50)	0.41 (0.50)
Profile Issue	0.07 (0.25)	0.03 (0.18)	0.09 (0.29)
<i>Demographic Information</i>			
White/Caucasian	0.30 (0.46)	0.28 (0.46)	0.32 (0.47)
Asian/Pacific Islander	0.37 (0.49)	0.31 (0.47)	0.41 (0.50)
Other Race/Ethnicity	0.33 (0.47)	0.41 (0.50)	0.27 (0.45)
<i>Additional Student Information</i>			
GPA	3.64 (0.28)	3.62 (0.34)	3.65 (0.24)
First Generation College Student	0.22 (0.42)	0.25 (0.44)	0.20 (0.41)
Online Presence	0.71 (0.46)	0.66 (0.48)	0.75 (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 (0.47)	0.29 (0.45)	0.23 (0.42)	0.43 (0.49)	0.36 (0.48)
College Graduation Year	2003.62 (12.00)	2009.55 (7.59)	2003.83 (11.95)	1998.31 (11.84)	2007.45 (11.16)
College Selectivity - Admit Rate	0.25 (0.22)	0.39 (0.28)	0.25 (0.22)	0.20 (0.16)	0.28 (0.23)
Alumni of Student's College	0.21 (0.41)	0.25 (0.44)	0.27 (0.44)	0.15 (0.35)	0.21 (0.41)
Any Graduate Degree	0.70 (0.46)	0.72 (0.45)	0.50 (0.50)	1.00 (0.00)	0.51 (0.50)
Any Ivy Degree	0.16 (0.36)	0.07 (0.26)	0.16 (0.36)	0.19 (0.39)	0.15 (0.36)
0-249 Connections	0.11 (0.31)	0.13 (0.33)	0.10 (0.30)	0.15 (0.36)	0.05 (0.21)
250-499 Connections	0.21 (0.41)	0.24 (0.42)	0.21 (0.41)	0.27 (0.45)	0.13 (0.33)
500+ Connections	0.64 (0.48)	0.59 (0.49)	0.64 (0.48)	0.54 (0.50)	0.78 (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 (0.010)	0.014 (0.011)	0.013 (0.010)	0.037** (0.015)	0.041*** (0.015)	0.040** (0.016)	0.003 (0.015)	0.009 (0.017)	0.011 (0.017)	0.061* (0.032)	0.059 (0.036)	0.060 (0.039)
Finance	-0.055*** (0.018)	-0.055*** (0.018)	-0.055*** (0.018)	-0.118*** (0.037)	-0.118*** (0.037)	-0.118*** (0.037)	-0.094*** (0.032)	-0.094*** (0.032)	-0.093*** (0.032)			
Law	-0.084*** (0.017)	-0.084*** (0.017)	-0.084*** (0.017)	-0.134*** (0.032)	-0.134*** (0.033)	-0.134*** (0.033)	-0.107*** (0.033)	-0.107*** (0.033)	-0.108*** (0.033)			
Mgmt Consulting	-0.030* (0.018)	-0.030* (0.018)	-0.030* (0.017)	-0.071** (0.032)	-0.071** (0.032)	-0.071** (0.032)	-0.086*** (0.031)	-0.087*** (0.031)	-0.088*** (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	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*** (0.032)	0.072** (0.032)	0.076** (0.033)	-0.003 (0.034)	-0.024 (0.034)	-0.024 (0.034)
Finance	0.013 (0.028)	0.017 (0.029)	0.012 (0.029)	-0.127** (0.055)	-0.121** (0.055)	-0.123** (0.054)
Law	0.062 (0.043)	0.062 (0.043)	0.053 (0.042)	-0.172*** (0.053)	-0.170*** (0.054)	-0.173*** (0.054)
Mgmt Consulting	0.208*** (0.051)	0.209*** (0.052)	0.209*** (0.051)	-0.074 (0.060)	-0.071 (0.060)	-0.073 (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: Gender Differences in Career Plans: Relative to the Start of the Study, Is the Student Less Likely to Enter Preferred Career Path?

	(1)	(2)	(3)	(4)	(5)
Student Female	0.095** (0.046)	0.114* (0.064)	0.058 (0.051)	0.117* (0.068)	0.059 (0.058)
Response Mentioned Work/Life Balance			0.185 (0.136)		0.190 (0.159)
Received Response to Specific Work/Life Question			0.079* (0.043)		0.081* (0.044)
Response Mentioned Workplace Culture				0.045 (0.057)	0.002 (0.085)
Received Response to Specific Culture Question				0.010 (0.069)	-0.003 (0.069)
Male Mean	0.000				
Observations	73	73	73	73	73
Industry Controls		X	X	X	X
Student Controls		X	X	X	X
Message Controls					X

Note: Standard errors are clustered at the student level and are reported in parentheses.

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

Table 6: Role of Professional Composition in Gender Differences in Response Rates

	Broad		Work/Life Balance		Competitive Culture		Factual (Law Only)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Student Female	0.014 (0.011)	0.017 (0.011)	0.041*** (0.015)	0.041** (0.016)	0.009 (0.017)	0.008 (0.017)	0.059 (0.036)	0.055 (0.040)
Finance	-0.055*** (0.018)	-0.055*** (0.018)	-0.118*** (0.037)	-0.120*** (0.037)	-0.094*** (0.032)	-0.074** (0.034)		
Law	-0.084*** (0.017)	-0.060*** (0.020)	-0.134*** (0.033)	-0.099** (0.038)	-0.107*** (0.033)	-0.070* (0.037)		
Mgmt Consulting	-0.030* (0.018)	-0.032* (0.019)	-0.071** (0.032)	-0.079** (0.034)	-0.087*** (0.031)	-0.076** (0.032)		
Observations	3530	3530	1763	1763	1776	1776	298	298
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
Professional		X		X		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 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, and student race/ethnicity. Columns 1, 3, 5, and 7 report results from the preferred specification. Columns 2, 4, 6, and 8 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 7: Role of Professional Composition in Gender Differences in Responses to Specific Questions

	Work/Life Balance		Workplace Culture	
	(1)	(2)	(3)	(4)
Student Female	0.072** (0.032)	0.067** (0.033)	-0.024 (0.034)	-0.025 (0.038)
Finance	0.017 (0.029)	-0.007 (0.037)	-0.121** (0.055)	-0.112* (0.063)
Law	0.062 (0.043)	0.017 (0.062)	-0.170*** (0.054)	-0.212** (0.081)
Mgmt Consulting	0.209*** (0.052)	0.183*** (0.054)	-0.071 (0.060)	-0.065 (0.062)
Observations	363	363	363	363
Message Time/Date	X	X	X	X
Student Profile	X	X	X	X
Student Race/Ethnicity	X	X	X	X
Professional		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 response mentions work/life balance (columns 1 and 2) or workplace culture (columns 3 and 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, student profile characteristics, and student race/ethnicity. Columns 1 and 3 report results from the preferred specification. Columns 2 and 4 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 8: Heterogeneity by Professional Attributes

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Male	Female	Grad<2010	Grad>=2010	No Ivy	Ivy	Not Alum	Alum	Data Science	Finance	Law	Mgmt Consulting
0.015 (0.011) [0.114]	0.006 (0.019) [0.076]	0.009 (0.010) [0.082]	-0.002 (0.024) [0.145]	0.014 (0.013) [0.106]	0.012 (0.019) [0.075]	0.018 (0.011) [0.083]	0.008 (0.023) [0.160]	0.025 (0.032) [0.139]	0.014 (0.020) [0.096]	0.009 (0.015) [0.068]	0.009 (0.024) [0.124]
0.041* (0.021) [0.148]	0.045* (0.025) [0.094]	0.028 (0.018) [0.105]	0.085** (0.033) [0.165]	0.040** (0.018) [0.131]	0.036 (0.041) [0.129]	0.030* (0.016) [0.117]	0.072* (0.041) [0.182]	-0.027 (0.061) [0.245]	0.069* (0.037) [0.085]	0.034 (0.026) [0.104]	0.055* (0.028) [0.150]
0.019 (0.023) [0.154]	-0.016 (0.031) [0.108]	-0.012 (0.022) [0.125]	0.011 (0.039) [0.170]	0.018 (0.014) [0.135]	-0.034 (0.057) [0.155]	0.002 (0.018) [0.117]	0.022 (0.038) [0.223]	0.008 (0.061) [0.242]	0.046 (0.031) [0.109]	0.008 (0.032) [0.115]	-0.023 (0.040) [0.146]
0.064 (0.040) [0.063]	0.140** (0.066) [0.079]	0.098** (0.044) [0.050]	0.043 (0.052) [0.066]	0.054 (0.034) [0.076]	0.338** (0.132) [0.000]	0.030 (0.040) [0.096]	0.135*** (0.050) [0.018]	0.037 (0.026) [0.000]	0.084* (0.044) [0.000]	0.133* (0.073) [0.033]	0.064 (0.099) [0.180]
-0.021 (0.046) [0.126]	0.012 (0.089) [0.132]	-0.028 (0.052) [0.100]	0.090 (0.061) [0.092]	-0.037 (0.039) [0.137]	0.125 (0.189) [0.056]	-0.072 (0.048) [0.181]	0.065 (0.056) [0.036]	-0.062 (0.136) [0.185]	-0.093 (0.064) [0.119]	-0.066 (0.062) [0.100]	0.067 (0.087) [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 9: Effect of Student Gender on Response Rates, Inclusive of Student Selection

	Broad			Work/Life Balance			Competitive Culture		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female Student Sent	-0.001 (0.023) [0.029]	-0.008 (0.028) [0.041]	-0.008 (0.028) [0.041]	0.012 (0.039) [0.049]	0.014 (0.047) [0.061]	0.015 (0.046) [0.063]	-0.007 (0.051) [0.072]	-0.053 (0.062) [0.090]	-0.050 (0.062) [0.091]
Female Student Preferred		-0.041* (0.024) [0.044]	-0.024 (0.025) [0.045]		-0.041 (0.041) [0.059]	-0.048 (0.040) [0.062]		-0.112** (0.053) [0.067]	-0.113** (0.054) [0.066]
Female Student Pref x Female Student Sent		0.011 (0.031) [0.051]	0.012 (0.031) [0.050]		-0.004 (0.052) [0.064]	-0.004 (0.052) [0.065]		0.082 (0.068) [0.089]	0.081 (0.069) [0.090]
Male Pref and Male Sent Mean		0.137			0.150			0.232	
Male Pref and Female Sent ($\beta^m - \alpha^m$)		0.129	0.129		0.164	0.165		0.179	0.181
Female Pref and Female Sent ($\alpha^f + \beta^f$)		0.100	0.117		0.119	0.112		0.150	0.150
P-value for Experienced F-M diff		0.163	0.452		0.511	0.450		0.225	0.241
Observations	3648	3648	3648	1824	1824	1824	1824	1824	1824
Message Time/Date	X	X	X	X	X	X	X	X	X
Sender Profile, Race	X	X	X	X	X	X	X	X	X
Student Profile, Race			X			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (4) (columns 1, 4, and 7) and Equation (5) (columns 2, 3, 5, 6, 8, and 9), in which the dependent variable is an indicator for whether the message received a response, by question type. Columns 1, 4, and 7 report results from a specification that includes an indicator for whether the professional received a message from a female student as well as controls for characteristics available on the students' profile and race/ethnicity of the student. Columns 2, 5, and 8 report results from the regression specification outlined in Equation (5), where the independent variables are an indicator for whether the student who sent the message is female, an indicator for whether the professional was in the preferred set of a female student, and the interaction of these terms, as well as the professional's field, message time/date characteristics, and student profile characteristics. Columns 3, 6, and 9 report results from a specification that additionally includes controls for the characteristics of students selecting the professionals. The bottom panel of the table provides the mean outcome for male students among professionals preferred by male students, as well as estimates of the mean outcome for female students if they had the same preferences as male students, the mean outcome for female students among the professionals preferred by female students, and p-values for a test of whether the mean outcome for female students among the professionals preferred by female students equals the mean outcome for male students among the professionals preferred by male students. Two sets of standard errors are reported: standard errors clustered at the professional level and bootstrapped standard errors are reported in brackets. The bootstrapped standard errors are based on 50 iterations and cluster at the student-ranker in the step when the model of student preferences is estimated to account for estimation error, and also cluster at the professional-response level when the most-preferred professionals are selected.

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

Table 10: Effect of Student Gender on Mentions of Work/Life Balance and Workplace Culture, Inclusive of Student Selection

	Work/Life Balance			Workplace Culture		
	(1)	(2)	(3)	(4)	(5)	(6)
Female Student Sent	0.237*** (0.075) [0.092]	0.233*** (0.081) [0.118]	0.229*** (0.082) [0.121]	-0.073 (0.063) [0.094]	-0.076 (0.072) [0.130]	-0.073 (0.070) [0.140]
Female Student Preferred		-0.029 (0.048) [0.094]	-0.044 (0.051) [0.109]		-0.003 (0.073) [0.118]	-0.021 (0.081) [0.137]
Female Student Pref x Female Student Sent		0.010 (0.075) [0.154]	0.026 (0.075) [0.153]		0.005 (0.078) [0.165]	0.007 (0.080) [0.162]
Male Pref and Male Sent Mean		0.043			0.170	
Male Pref and Female Sent ($\beta^m - \alpha^m$)		0.276	0.271		0.094	0.097
Female Pref and Female Sent ($\alpha^f + \beta^f$)		0.257	0.253		0.096	0.083
P-value for Experienced F-M diff		0.012	0.011		0.281	0.227
Observations	392	392	392	392	392	392
Message Time/Date	X	X	X	X	X	X
Sender Profile, Race	X	X	X	X	X	X
Student Profile, Race			X			X

Note: This table reports the results of the estimation of the regression specification outlined in Equation (4) (columns 1 and 4) and Equation (5) (columns 2, 3, 5, and 6), 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). Columns 1 and 4 report results from a specification that includes an indicator for whether the professional received a message from a female student as well as controls for characteristics available on the students' profile and race/ethnicity of the student. Columns 2 and 5 report results from the regression specification outlined in Equation (5), where the independent variables are an indicator for whether the student who sent the message is female, an indicator for whether the professional was in the preferred set of a female student, and the interaction of these terms, as well as the professional's field, message time/date characteristics, and student profile characteristics. Columns 3 and 6 report results from a specification that additionally includes controls for the characteristics of students selecting the professionals. The bottom panel of the table provides the mean outcome for male students among professionals preferred by male students, as well as estimates of the mean outcome for female students if they had the same preferences as male students, the mean outcome for female students among the professionals preferred by female students, and p-values for a test of whether the mean outcome for female students among the professionals preferred by female students equals the mean outcome for male students among the professionals preferred by male students. Two sets of standard errors are reported: standard errors clustered at the professional level and are reported in parentheses, and bootstrapped standard errors are reported in brackets. The bootstrapped standard errors are based on 50 iterations and cluster at the student-ranker in the step when the model of student preferences is estimated to account for estimation error, and also cluster at the professional-response level when the most-preferred professionals are selected.

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

A Online Appendix Figures and Tables

Figure A1: Message Templates

Broad Question

Dear Mr. x,
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 [data science/ finance/ law/ management consulting] so that I have a realistic grasp of the field. Could you share your quick thoughts on the advantages and challenges in [data science/ finance/ law/management consulting]?

Specific WL Balance

Dear Mr. x,
As of right now I'm not actively searching for a job, but I'm really drawn to a career in [data science/ finance/ law/management consulting]. I've heard that work-life balance in [data science/ finance/ law/ management consulting] is challenging. Could you share your quick thoughts on whether this is a valid concern?

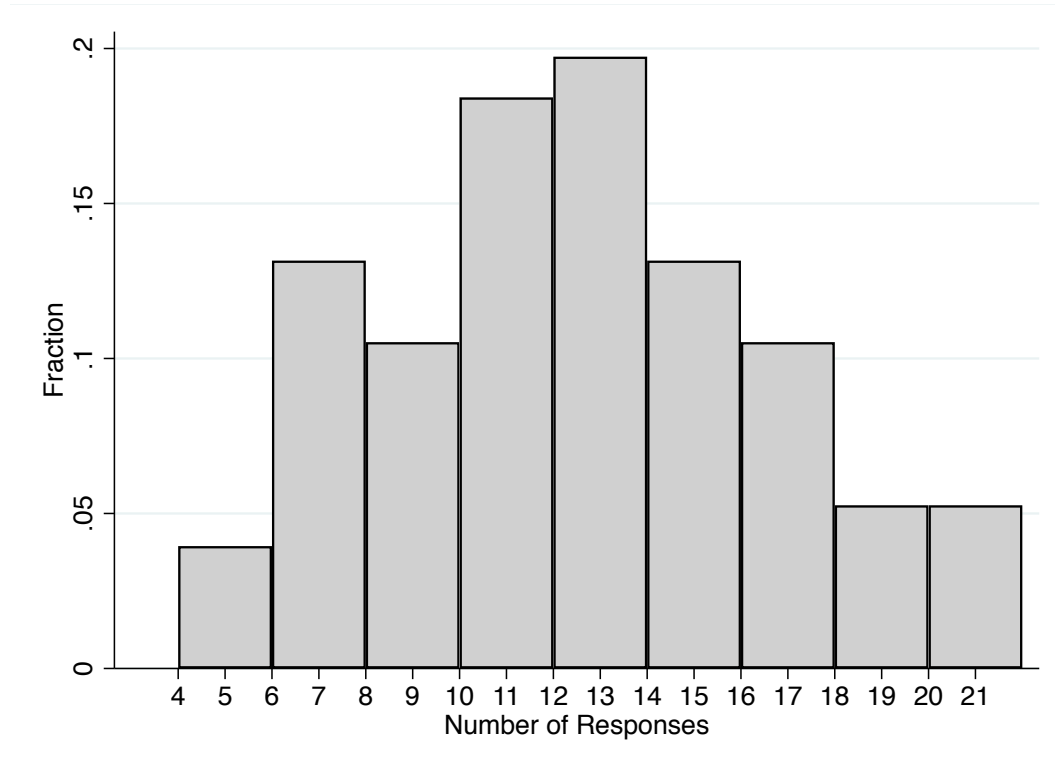
Specific Culture

Dear Mr. x,
As of right now I'm not actively searching for a job, but I'm really drawn to a career in [data science/ finance/ law/management consulting]. I've heard that [data science/ finance/ law/management consulting] has a cutthroat culture. Could you share your quick thoughts on whether this is a valid concern?

Factual

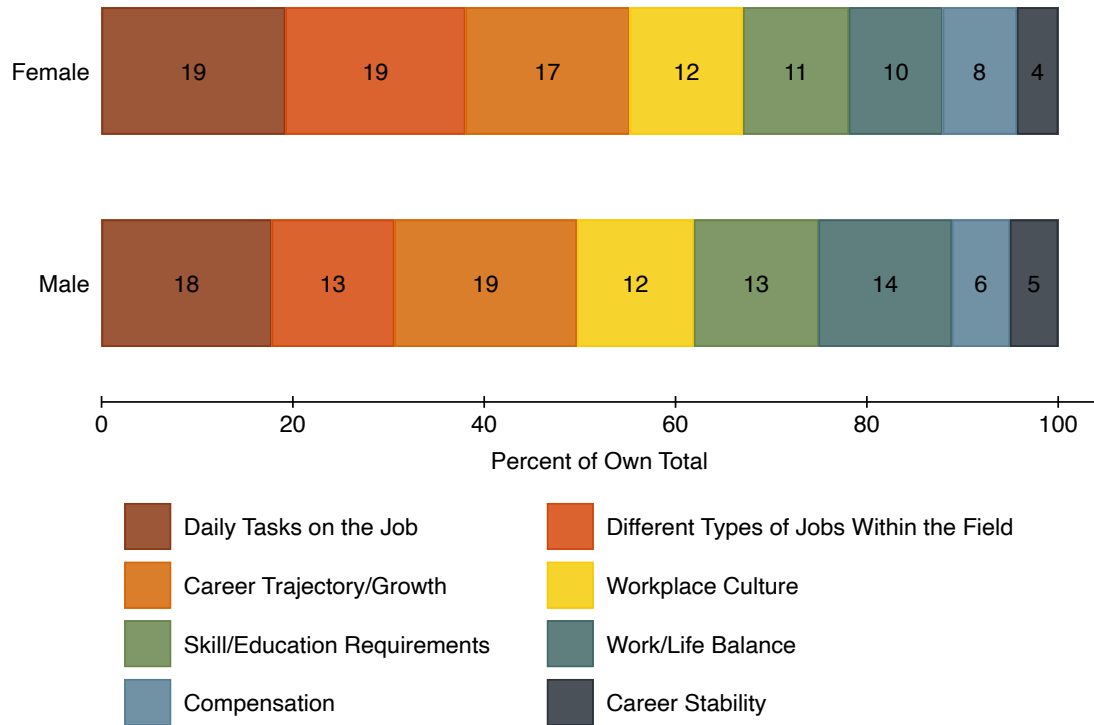
Dear Mr. x,
As of right now I'm not actively searching for a job, but I'm really drawn to a career in law. I am trying to gather some basic information—do you happen to know what the billable hours requirements are for a first-year associate at a large law firm?

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.

Figure A3: 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 A1: Outcome Summary Statistics

	All Messages	Broad	Specific - Work/Life	Specific - Culture	Factual
Response Rate	0.12 (0.33)	0.10 (0.30)	0.14 (0.35)	0.15 (0.36)	0.11 (0.32)
Response Character Count	434.73 (558.77)	414.39 (687.34)	486.64 (492.37)	429.95 (396.61)	304.18 (553.03)
Work/Life Balance Mentioned		0.11 (0.32)			
Workplace Culture Mentioned		0.12 (0.33)			
<i>Valid concern?</i>					
Yes			0.44 (0.50)	0.16 (0.37)	
It depends			0.49 (0.50)	0.54 (0.50)	
No			0.07 (0.26)	0.30 (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)	(2)
	All Messages	Sent Messages Only
Data Science	0.000 (0.001)	-0.000 (0.002)
Finance	0.000 (0.001)	-0.003 (0.004)
Law	-0.000 (0.001)	0.008 (0.008)
Mgmt Consulting	-0.001 (0.001)	-0.005 (0.004)
Professional Female	0.004 (0.012)	0.004 (0.012)
0-249 Connections	0.004 (0.007)	0.004 (0.008)
250-499 Connections	0.017* (0.010)	0.021** (0.010)
500+ Connections	-0.015 (0.012)	-0.018 (0.012)
College graduation year	0.311 (0.321)	0.208 (0.320)
Alumni of Student's College	-0.003 (0.010)	-0.005 (0.011)
Undergraduation Selectivity Quartile 1	-0.009 (0.008)	-0.012 (0.008)
Undergraduation Selectivity Quartile 2	0.008 (0.009)	0.008 (0.009)
Undergraduation Selectivity Quartile 3	0.005 (0.011)	0.008 (0.012)
Undergraduation Selectivity Quartile 4	0.007 (0.008)	0.006 (0.008)
Any Graduate Degree	0.008 (0.010)	0.011 (0.011)
Any Ivy Degree	0.003 (0.008)	0.002 (0.008)
<i>N</i>	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.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ 50

Table A3: 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*** (0.003)	0.008** (0.004)	0.001 (0.003)	-0.002 (0.003)
Finance	-0.001 (0.003)	-0.001 (0.003)	-0.023*** (0.008)	-0.023*** (0.008)
Law	0.002 (0.004)	0.002 (0.004)	-0.028*** (0.008)	-0.028*** (0.008)
Mgmt Consulting	0.025*** (0.006)	0.025*** (0.006)	-0.014 (0.009)	-0.014 (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 A4: Effect of Student Gender on Mentions of Specific Work/Life Balance Issues

	(1)	(2)	(3)	(4)	(5)	(6)
	Duration	Typical Workweek	Work Schedule Flexibility	Extent of Travel/Work from Home		
Student Female	0.054* (0.030)	0.057* (0.030)	0.023 (0.022)	0.026 (0.023)	0.015 (0.022)	0.016 (0.023)
Finance	0.026 (0.025)	0.022 (0.024)	0.021 (0.013)	0.018 (0.015)	0.001 (0.018)	0.001 (0.018)
Law	0.062 (0.040)	0.057 (0.038)	0.043* (0.022)	0.036* (0.021)	-0.004 (0.017)	-0.004 (0.017)
Mgmt Consulting	0.142*** (0.042)	0.143*** (0.042)	0.089*** (0.029)	0.089*** (0.030)	0.170*** (0.045)	0.171*** (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 A5: Gender Differences in Response Tone (Student Ratings)

	Student more concerned about work/life balance		Student more concerned about workplace culture	
	(1) Broad	(2) Work/Life Balance	(3) Broad	(4) Competitive Culture
Student Female	0.033 (0.027)	-0.032 (0.048)	-0.018 (0.015)	-0.004 (0.034)
Finance	-0.023 (0.023)	0.312*** (0.053)	-0.090*** (0.024)	0.101* (0.057)
Law	0.051 (0.034)	0.365*** (0.061)	-0.076*** (0.025)	0.110** (0.053)
Mgmt Consulting	0.161*** (0.042)	0.492*** (0.039)	-0.056* (0.029)	-0.038 (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 A6: Gender Differences in Responses to "Is work/life balance a concern?"

	Yes			It depends			No		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Student Female	0.002 (0.070)	-0.008 (0.074)	-0.023 (0.076)	-0.015 (0.078)	0.020 (0.034)	0.023 (0.035)			
Finance	0.217** (0.088)	0.215** (0.087)	-0.030 (0.108)	-0.034 (0.107)	-0.186*** (0.069)	-0.181*** (0.068)			
Law	0.315*** (0.091)	0.314*** (0.093)	-0.076 (0.111)	-0.063 (0.115)	-0.240*** (0.065)	-0.251*** (0.067)			
Mgmt Consulting	0.674*** (0.062)	0.672*** (0.064)	-0.427*** (0.089)	-0.422*** (0.091)	-0.247*** (0.060)	-0.250*** (0.061)			
Male Mean	0.427				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 A7: Gender Differences in Responses to "Is cutthroat culture a concern?"

	Yes			It depends			No		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Student Female	0.002 (0.051)	0.024 (0.050)	0.052 (0.071)	0.010 (0.066)	-0.053 (0.069)	-0.033 (0.068)			
Finance	0.072 (0.079)	0.082 (0.081)	0.198* (0.111)	0.205* (0.115)	-0.269*** (0.096)	-0.287*** (0.100)			
Law	0.135* (0.080)	0.133* (0.078)	0.093 (0.113)	0.098 (0.109)	-0.228** (0.087)	-0.231** (0.088)			
Mgmt Consulting	-0.035 (0.056)	-0.035 (0.055)	-0.050 (0.108)	-0.037 (0.106)	0.085 (0.109)	0.071 (0.110)			
Male Mean	0.134		0.512						
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 A8: Gender Differences in Hours Quoted
in Response to Factual Question

	(1)	(2)
Student Female	81.611 (73.971)	154.402*** (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 A9: Gender Differences in Response Length

	Broad		Work/Life Balance		Competitive Culture		Factual (Law Only)	
	(1) Character Count	(2) Log(Count)	(3) Character Count	(4) Log(Count)	(5) Character Count	(6) Log(Count)	(7) Character Count	(8) Log(Count)
Student Female	-8.747 (39.126)	-0.125 (0.107)	10.755 (52.191)	0.094 (0.156)	41.875 (48.090)	0.007 (0.113)	-115.259 (102.349)	-0.323 (0.597)
Finance	-261.644*** (74.237)	-0.595*** (0.180)	-20.247 (54.440)	0.020 (0.153)	-55.515 (66.544)	-0.171 (0.152)		
Law	-210.993*** (78.255)	-0.381** (0.176)	18.193 (57.283)	0.039 (0.154)	26.743 (70.103)	0.027 (0.164)		
Mgmt Consulting	-163.236** (76.321)	-0.320* (0.177)	134.036** (59.732)	0.389*** (0.143)	14.302 (66.023)	0.092 (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 A10: 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 (0.018)	0.043** (0.020)	0.040* (0.023)	0.048* (0.027)	0.066** (0.030)	0.087** (0.040)	0.051 (0.040)	0.045 (0.044)	0.027 (0.050)	0.105*** (0.035)	0.077*** (0.016)	0.062*** (0.021)
Finance	-0.113*** (0.036)	-0.114*** (0.036)	-0.114*** (0.036)	-0.056 (0.062)	-0.055 (0.062)	-0.053 (0.062)	-0.070 (0.072)	-0.070 (0.072)	-0.070 (0.072)	-0.070 (0.072)	-0.070 (0.072)	-0.070 (0.072)
Law	-0.119*** (0.039)	-0.119*** (0.039)	-0.119*** (0.039)	-0.115* (0.057)	-0.115* (0.058)	-0.117* (0.058)	-0.085 (0.065)	-0.085 (0.065)	-0.082 (0.064)	-0.085 (0.065)	-0.085 (0.065)	-0.082 (0.064)
Mgmt Consulting	-0.074* (0.041)	-0.076* (0.041)	-0.076* (0.042)	-0.058 (0.061)	-0.063 (0.059)	-0.068 (0.060)	-0.053 (0.053)	-0.052 (0.052)	-0.047 (0.052)	-0.052 (0.052)	-0.052 (0.052)	-0.047 (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	X	X	X	X
Additional Student			X	X	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 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 A11: 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** (0.037)	0.138*** (0.030)	0.052 (0.042)	0.061 (0.047)
Finance	-0.021 (0.055)	-0.027 (0.056)	-0.136 (0.106)	-0.134 (0.109)
Law	-0.025 (0.070)	-0.040 (0.077)	-0.254* (0.129)	-0.263* (0.140)
Mgmt Consulting	0.068 (0.086)	0.054 (0.087)	-0.110 (0.116)	-0.115 (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 A12: Attributes of Preferred Professionals

	All Preferred Professionals	Female Preferred	Male Preferred
Data Science	0.16 (0.36)	0.09 (0.29)	0.25 (0.43)
Finance	0.33 (0.47)	0.25 (0.43)	0.44 (0.50)
Law	0.30 (0.46)	0.41 (0.49)	0.16 (0.36)
Mgmt Consulting	0.21 (0.41)	0.25 (0.43)	0.16 (0.36)
Female	0.60 (0.49)	0.88 (0.33)	0.21 (0.41)
College Graduation Year	2005.16 (10.80)	2004.36 (10.47)	2006.34 (11.17)
College Selectivity - Admit Rate	0.21 (0.20)	0.22 (0.21)	0.21 (0.19)
Alumni of Student's College	0.34 (0.47)	0.31 (0.46)	0.39 (0.49)
Any Graduate Degree	0.75 (0.43)	0.80 (0.40)	0.68 (0.47)
Any Ivy Degree	0.23 (0.42)	0.27 (0.44)	0.18 (0.38)
0-249 Connections	0.12 (0.32)	0.09 (0.28)	0.16 (0.37)
250-499 Connections	0.13 (0.34)	0.11 (0.31)	0.17 (0.38)
500+ Connections	0.66 (0.47)	0.76 (0.43)	0.53 (0.50)
Observations	3648	2112	1536

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.

B 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.

B.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.²⁹ 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 B.1 below.

Category	Description	Source
<i>Analytical Aspects</i>		
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

²⁹We 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
<i>Decision-making Aspects</i>		
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
<i>Excitement and Impact Aspects</i>		
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
<i>Interpersonal Aspects</i>		
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
<i>Work/Life Balance Aspects</i>		
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
<i>Individual Categories that Appear in >5% of Responses</i>		
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.

B.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* (0.043)	-0.078* (0.043)
Explains Career Paths	-0.037 (0.045)	-0.037 (0.043)
Mentions Analytical Aspects of Career	0.050 (0.042)	0.051 (0.041)
Mentions Decision-Making/Responsibility Aspects of Career	0.012 (0.028)	0.013 (0.028)
Mentions Excitement/Impact Aspects of Career	0.024 (0.040)	0.029 (0.038)
Mentions Interpersonal Aspects of Career	0.022 (0.045)	0.025 (0.044)
Compensation	0.052 (0.039)	0.051 (0.038)
Job Stability	0.035 (0.028)	0.036 (0.028)
Short v. Long Term Considerations	-0.022 (0.032)	-0.023 (0.032)
Qualities of Individuals who Like/Succeed	0.043 (0.026)	0.040 (0.024)
Broadness of Question	-0.026 (0.032)	-0.027 (0.033)
Info on Job Search	-0.035 (0.029)	-0.037 (0.029)
Offer to Discuss Further	0.064 (0.051)	0.064 (0.051)
Decision is Person Specific	-0.025 (0.025)	-0.025 (0.024)
States Qualifications for Answering	-0.082* (0.045)	-0.084* (0.043)
Education Requirements/Environment	0.030 (0.032)	0.029 (0.031)
Other attribute	0.010 (0.034)	0.008 (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

B.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.³⁰ The NRC lexicon contains 6,468 unique words, and each word can have multiple sentiments/emotions (categories are not unique).³¹ The Bing lexicon contains 6,785 unique words, and all but three words are uniquely classified as either positive or negative.

³⁰The website with supporting information is here: www.cs.uic.edu/~liub/FBS/sentiment-analysis.html

³¹The 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 -wyoming-, with 1,000 bootstrap replications. Column -Clust- reports WY adjusted standard errors clustered at the student level.