Journal of Public Economics

Does Information Affect Homophily?

--Manuscript Draft--

Manuscript Number:	
Article Type:	Short Paper- General
Keywords:	homophily; mentorship; preference elicitation; gender
Corresponding Author:	Melanie Wasserman Los Angeles, California UNITED STATES
First Author:	Yana Gallen
Order of Authors:	Yana Gallen
	Melanie Wasserman
Abstract:	It is common for mentorship programs to use race, gender, and nationality to match mentors and mentees. Despite the popularity of these programs, there is little evidence on whether mentees value mentors with shared traits. Using novel administrative data from an online college mentoring platform connecting students and alumni, we document that female students indeed disproportionately reach out to female mentors. We investigate whether female students make costly trade-offs in order to access a female mentor. By eliciting students' preferences over mentor attributes, we find that female students are willing to trade off occupational match in order to access a female mentor. This willingness to pay for female mentors declines to zero when information on mentor quality is provided. The evidence suggests that female students use mentor gender to alleviate information problems, but do not derive direct utility from it. We discuss the implications of these results for the design of initiatives that match on shared traits.

Does Information Affect Homophily?*

Yana Gallen and Melanie Wasserman

September 2022

Abstract

It is common for mentorship programs to use race, gender, and nationality to match mentors and mentees. Despite the popularity of these programs, there is little evidence on whether mentees value mentors with shared traits. Using novel administrative data from an online college mentoring platform connecting students and alumni, we document that female students indeed disproportionately reach out to female mentors. We investigate whether female students make costly trade-offs in order to access a female mentor. By eliciting students' preferences over mentor attributes, we find that female students are willing to trade off occupational match in order to access a female mentor. This willingness to pay for female mentors declines to zero when information on mentor quality is provided. The evidence suggests that female students use mentor gender to alleviate information problems, but do not derive direct utility from it. We discuss the implications of these results for the design of initiatives that match on shared traits.

Keywords: homophily; mentorship; preference elicitation; gender

JEL Codes: J16, J24, J71

*Gallen: University of Chicago, Harris School of Public Policy. E-mail: yana@uchicago.edu. Wasserman: UCLA Anderson School of Management. 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, Neta Grossfeld, Mikaela Hassenzahl, Reigne Dadey, and Yash Srivastav for excellent research assistance, and Jacob Model for superb technical assistance. We have benefited from comments from Danny Yagan, Heather Royer, Corinne Low, Paola Giuliano, Trevor Gallen, Romain Wacziarg, and Steven Durlauf. This experiment was pre-registered on the AEA Registry under AEARCTR-0008674.

1 Introduction

Homophily—the tendency to associate with those who have traits similar to oneself—is a ubiquitous social phenomenon but its determinants are not well understood. In many settings, we may observe individuals making costly trade-offs in order to match on shared characteristics. For example, there is evidence that female patients prefer to stay on a long waitlist to see a female doctor even when male doctors are readily available (Reyes, 2008; McDevitt and Roberts, 2014). It may be natural to assume that the demand for shared characteristics reflects utility derived from interacting with someone similar to oneself. Indeed, homophily can arise because individuals obtain utility *directly* from interacting with someone like themselves (taste- or preference-based discrimination, as in Becker (1971)). However, it could also be the case that, in the absence of information on match quality, individuals rely on easily observed traits as signals of match quality (statistical discrimination based on various moments of the match distribution, as in Aigner and Cain (1977), or inaccurate statistical discrimination, as in Bohren et al. (2019)).

We study whether homophily by gender is driven by preferences for shared traits. A main prediction of Becker's model of taste-based discrimination is that people should be willing to pay to interact with members of their own group (Becker, 1971; Bertrand and Duflo, 2017; Charles and Guryan, 2018). We test this prediction in the context of mentorship. Mentorship is a setting where—unlike hiring or lending or renting—explicitly using race, gender, and nationality to determine matches is common, encouraged, and even considered best practice. Among the top 50 U.S. News colleges/universities, all but two host a mentorship program designed specifically for women in STEM fields, and 80% of the programs match students with a same-gender mentor.[1] Despite the popularity of these programs, as of yet, there is little evidence on whether mentees value same-gender mentors or whether demand for same-gender mentors arises due to a lack of information on mentor quality.

Using novel administrative data from an online college student/alumni mentoring platform serving eight colleges and universities, we document substantial homophily by gender in student-alumni interactions. Female students are 36% more likely to reach out to female mentors relative to male students, conditional on various observable characteristics including student major, alumni major, and alumni occupation. This propensity to reach out to female mentors may come at a cost: female mentors are 12% less likely than male mentors to respond to messages sent by female students.

Although these patterns are consistent with taste-based discrimination, that is, female students incurring a cost in order to access a female mentor, it is also possible that we as researchers are unable to control for all mentor attributes used in students' decisions; students could use information outside of the mentoring plat-

[1] Of the 24 programs that provide information on the nature of matching, 19 match female students with a female mentor.

form to decide whom to contact, leading to omitted variable bias. To causally identify students' preferences for mentor characteristics, we implement a hypothetical choice preference elicitation survey that incentivizes truthful responses. In the survey, students are shown pairs of hypothetical mentors' profiles and asked to select which mentor they prefer (Wiswall and Zafar, 2018). Students are informed that their answers to the survey will be used to provide personalized information on how to find mentors based on their preferences.

We find that female students strongly prefer female mentors, while male students exhibit a weak preference for male mentors. Furthermore, using the trade-offs students make between mentor gender and other mentor attributes, we estimate that female students are willing to give up access to a mentor with their preferred occupation in order to match with a mentor of the same gender.

Next we investigate whether female students' preference for female mentors reflects taste-based discrimination. Taste-based discrimination could arise from female students' affinity for interacting with women. Alternatively, it could arise from female students valuing an attribute that only female mentors possess, for example, first-hand knowledge of being a woman in STEM. We conduct a within-survey experiment to determine whether female students' willingness to pay for female mentors is only present in information-poor environments. The survey uses hypothetical choice preference elicitation with incentives for truthful reporting and randomizes students into (1) a basic profile condition, in which mentor profiles contain basic information about the mentor (name, job, graduation year, etc.) or (2) a ratings condition, in which profiles contain all basic information plus ratings from a past mentee. The ratings contain the past mentee's perception of the mentor's knowledge about job opportunities, friendliness/approachability, and the extent to which the mentor gave personalized advice. These attributes are often difficult to observe about mentors prior to contacting them. In addition to randomizing each of the ratings, the mentee's gender is randomized.

Female students are only willing to pay for female mentors when there is no information on mentor quality. In the basic profile condition, as discussed above, female students are willing to trade off a mentor with their preferred occupation in order to access a female mentor. In the ratings condition, we find that this willingness to pay declines to zero. Furthermore, the estimates imply that—when information on mentor quality is available—female students are unwilling to trade off *any* dimension of mentor quality in order to access a female mentor. We also find no evidence that female students' preferences for mentor quality differ from that of male students. All students—male and female—value the attributes described in the ratings, particularly a mentor's knowledge of job opportunities.

If female students' preference for female mentors is *not* due to taste-based discrimination, several alternative explanations are possible. Our survey reveals that female students believe that female mentors are more friendly/approachable than male mentors. In the absence of information on mentor approachability, female students' beliefs, whether they are accurate (Aigner and Cain, 1977) or inaccurate (Bohren et al., 2019), may lead them to gravitate to female mentors. Specifically, students may rely on the perception that women are more approachable than men, on average, which could stem from stereotypes that contain some truth but are often exaggerated (Bordalo et al., 2016). For example, Eyal and Epley (2017) find that though women are somewhat more socially sensitive than men, people believe that the average difference is larger than it actually is. Homophily could also arise from differences in other moments of the mentor quality distribution.[2] All of these explanations have in common that gender is valued for its information content and direct provision of that information would reduce students' valuations of mentor gender.

Our experimental design also allows us to investigate whether female students perceive gender-specific benefits (or costs) of same-gender pairings. Using the randomization of mentee gender to ratings, we find that female students similarly value ratings from male and female mentees and both types of ratings similarly attenuate female students' WTP for female mentors. These results suggest that female-specific experiences with mentors do not explain homophily by gender in our setting.

Our results have implications for initiatives that match on shared traits, such as mentorship programs that match on race/ethnicity, nationality, gender, and sexual orientation, or firms' efforts to increase diversity by asking underrepresented minority (URM) employees to conduct interviews with or otherwise help recruit URM applicants (Rivera, 2015). If shared traits are used as a signal of match quality, these initiatives—while well intentioned—could lead to efficiency losses relative to a scenario in which information on valued traits is used. As an example of this, ride-sharing platforms have opted to inform riders that their driver has been background checked rather than offer same-gender matching (Tang et al., 2021). In addition, since matching on shared traits often occurs in settings where individuals with the trait are scarce and the task has low promotability, shifting to matching based on quality metrics would alleviate the time burden of these initiatives on already underrepresented groups (Babcock et al., 2017).

Our paper contributes to a small literature that investigates the roots of homophily.[3] In contemporaneous work on patients' selection of physicians, Chan (2021) uses a survey-based preference elicitation and finds that homophily by gender is somewhat attenuated when information on physician quality is provided. Our paper examines homophily in a setting where matching on shared traits is considered best practice and considerable resources are devoted to initiatives that prioritize such matching. We also contribute to a broader literature that examines the determinants of discriminatory behavior—particularly focused on isolating the role of statistical discrimination—including papers that study coworker choice (Hedegaard and Tyran, 2018), manager choice (Alam, 2020), hiring (Agan and Starr, 2017; Kaas and Manger, 2012; Abel et al., 2020), and

[2] If students use a threshold crossing model of mentor quality to choose a mentor, then differences in the perceived variance of mentor quality (or match quality) by gender could lead to homophily (Heckman and Siegelman, 1993). For example, students could think that the variance of mentor quality differs by gender and female students could be more risk averse than male students, yielding different choices (Aigner and Cain, 1977).

[3] There is a large literature on social networks documenting homophily (Currarini et al., 2009; Bertrand and Duflo, 2017).

the take-up of advice (Ayalew et al., 2021).[4].

2 Observational Evidence: Homophily by Gender on an Online Mentoring Platform

Using administrative data from an online student-alumni mentoring platform, we provide descriptive evidence that college students tend to choose same-gender mentors.

2.1 Data

The online student-alumni mentoring website is designed to connect current undergraduates with alumni of their college/university in order to give students access to mentorship, career guidance, and professional connections as they search for jobs and internships. The site has more than 50,000 users across dozens of institutions ranging from small liberal arts colleges to large public universities. Students and alumni sign up for the site and create a profile with information about their academic background and their professional background. Users within the same university (students and alumni) can directly message one another on the platform. Our data include all messages sent between students and alumni, de-identified and linked to message sender and message recipient by a unique profile ID. Gender is assigned based on the first name of users. Our data also include information on the self-reported job title, degree, and graduation year of each alumna/alumnus user, as well as the intended degree of each student user. We manually classify college majors according to ACS 2016 general degree codes.[5] Occupations are derived from job title using O*NET-SOC AutoCoder.[6]

We observe 13,038 conversations on the site, where a conversation is defined as a series of messages between two people. In order to study the preferences of undergraduate students for contacting alumni for mentoring and advice, we restrict our analysis to the 6,325 conversations initiated by students and sent to alumni recipients, keeping only schools that had at least 100 student-initiated conversations. We also drop the 99th percentile most prolific student senders and restrict to conversations that pertain to the students' future careers. Dropped conversation topics include inquiries regarding interviews for a class project, invitations to speak to a class, thank you messages from prior interactions, and inquiries regarding housing/re-location. These restrictions yield a sample of 3,374 student-alumni interactions which we analyze

[4]Many papers additionally document differential effects by in-group status, e.g., in advising (Canaan and Mouganie, 2021; Porter and Serra, 2020), teaching (Carrell et al., 2010), social work (Behncke et al., 2010), and physician choice (Alsan et al., 2019; Cabral and Dillender, 2021; Zeltzer, 2020)

[5]There are 39 codes, available at: `https://usa.ipums.org/usa-action/variables/DEGFIELD#codes_section`

[6]See Online Appendix B for more details on data preparation.

in the next subsection.[7]

Appendix Table A1 provides summary statistics on the population, separately for students and for alumni. The student population is 50% female while the alumni users are 46% female. Users are primarily from research universities. Restricting to messages as described above, 12% of student users send at least one message on the site, and 11% of alumni respond to at least one such message on the site.

2.2 Who contacts whom? Homophily by gender

Figure 1 Panel A characterizes homophily by gender by plotting the fraction of interactions that occur among same-gender members, against the availability of same-gender members on the platform (*inbreeding homophily*). Specifically, each dot represents the fraction of messages sent by female (male) students that are sent to female (male) alumni, on the y-axis, plotted against the fraction of alumni from that university who are female (male), on the x-axis, for each of the eight universities/colleges in the sample. The solid 45 degree line depicts the composition of student-alumni interactions that we would expect if students messaged alumni at random on the platform. The fraction of same-gender interactions on the site is higher than what would be expected by chance at almost all of the universities.

In Figure 1 Panel B, we further divide the students and alumni by their college major, and plot whether students tend to contact alumni of their same gender and major more than they would due to chance. The solid circles plot the fraction of male students in a given major who sent messages sent to alumni with the same gender and major against the fraction of alumni who are the same gender-major. The hollow diamonds plot the analogous data for female students. We again observe a strongly positive relationship and substantial deviation from the 45 degree line.

To probe whether the sorting patterns in Figure 1 are driven by other characteristics of alumni that are correlated with alumni gender, we estimate the following regression specification:

$$RecipientFemale_{ij} = \alpha + \beta StudentFemale_i + X_{ij}'\gamma + \epsilon_{ij} \tag{1}$$

where $RecipientFemale_{ij}$ is an indicator variable for whether the alumni recipient j is female and $StudentFemale_i$ is an indicator variable for whether the student sender i of the message is female. X_{ij} includes controls for sender and recipient characteristics. This specification tests whether students exhibit *relative homophily*: the difference in the rates at which female and male students to reach out to female mentors. The baseline results are reported in Table 1: without controls, the coefficient β is 0.193, indicating that female students

[7] See Gallen and Wasserman (2021a), Figure 1, for a complete description of initial message topics in this final subset of student-alumni interactions on the site.

are nearly 20 percentage points more likely to contact female mentors than male students. The differential pairing of female students and female alumni attenuates but remains significant when we add controls for school, student major, student graduation year, recipient major, and recipient occupation fixed effects, as well as a linear term for recipient graduation year.[8]

One reason that female students could be more likely to reach out to female mentors is that they expect in-group bias, that is, female mentors are more responsive or give better responses to female students than male mentors. However, we see little evidence for this explanation in our data. In fact, on the margins which we can directly measure in the data—the propensity of mentors to respond to messages from students, as well as the length on these responses—female students appear to be trading off responsiveness or response quality when messaging a female alumnus (Table 2).[9,10]

While female students' willingness to pay for same-gender mentors on this platform is consistent with taste-based discrimination, there is still a gap between what students observe about alumni when deciding whom to contact and what the researcher observes. For example, students can potentially glean additional information about alumni from an online search. Since alumni are bundles of characteristics, it is also difficult to ascertain which are valued by the students based on their choices. To address these issues, in the next section we implement a preference elicitation survey that isolates and quantifies students' willingness to pay to access a mentor of the same gender.

3 Estimating Willingness to Pay for Mentor Gender: Methodology

3.1 Preference elicitation survey

As discussed in the Introduction, a main prediction of Becker's taste-based discrimination model is that individuals should be willing to pay to access members of their own group. Are students willing to pay to access a mentor of the same gender by trading off other mentor characteristics (e.g. job market experience, availability, industry/occupation proximity)? Using a survey methodology to elicit willingness to pay for non-pecuniary job attributes developed by Wiswall and Zafar (2018) and used by Maestas et al. (2018), we estimate students' WTP for mentors of the same gender.

[8] Note that the attenuation of the coefficient on $StudentFemale_i$ when we add student and alumni controls suggests that observational measures of homophily may be driven by omitted variable bias. In Section 3 we formally elicit students' preferences for mentor characteristics in part to address this concern.

[9] In Appendix Table A2 we document that male students also receive slightly lower rates of response from female mentors and we cannot reject that the effect is different from zero or different from the effect for female students. We also note that female mentors' lower response rate is not explained by excess requests: they do not receive more messages from students than male mentors.

[10] In a field experiment that controls for all observable student characteristics and the wording of student messages, we show that female professionals are less responsive and give shorter replies to female students than male professionals (Gallen and Wasserman, 2021b).

Students taking the survey are shown 30 pairs of hypothetical mentors and asked to choose which professional they prefer within each pair.[11] Each mentor in the pair has a randomly assigned occupation, availability for mentoring (30 minutes or 60 minutes), first-generation college student status, graduation year (2015 or 2005), and name that unambiguously conveys gender. The characteristics of mentor profiles are sampled randomly and independently with equal probability across all possibilities both within and across profile pairs. By observing the choices of students in each mentor pair, we are able to estimate their preferences for each of the mentor attributes and use these estimates to compute their WTP for a mentor of the same gender.

Our recruitment and compensation procedures are designed to elicit students' true preferences over mentor characteristics (Becker et al., 1964). From November 2021 to January 2022, the study was advertised at UCLA using email lists from every undergraduate major, a handful of large undergraduate classes, and the career center newsletter. Study recruitment was targeted to students interested in career advice. Since the survey was advertised via email lists and accessed via a Qualtrics link, students were able to take the survey completely on their own without the supervision of a researcher. We think this setting guards against social desirability bias and social image concerns.

Once students began the study, they were informed: "We will use your responses in this section to give you personalized suggestions on how to find mentors. If you decide to receive these suggestions, you will receive these suggestions via email (which you will enter at the end of the survey). We will not contact any mentors on your behalf, we will only provide you with recommendations consistent with the choices you make in the next portion of this questionnaire." An example of the mentor targeting advice email is available in Appendix Figure A2. Students also received a $5 payment to their UCLA flexible spending card. A similar methodology is used by Kessler et al. (2019) to elicit employers' true preferences over employee characteristics. As an indication that students thoughtfully considered profiles, the median time to complete the survey was 11 minutes and 99.6 percent of students passed our attention check.

Because we recruited undergraduate students from all majors, the survey adapts to each student's preferences by only showing mentors with occupations of interest to the student. Before being shown the mentor pairs, each student is asked to select their preferred career path from a comprehensive set of 24 broad career paths.[12] To aid in the student's selection, we provided four examples of occupations associated with each career path. For example, if the student selected the broad career path "Marketing," then the student would see the following text: "Examples include: VP of Marketing, Business Analytics Lead, Brand Manager, and Sales Representative." In the preference elicitation, the mentor profiles are randomly assigned occupations

[11]Students were informed that "You should think of mentors as alumni of UCLA who have volunteered to help current students navigate their major choice, career choice, and to provide advice and answer questions related to these decisions."

[12]These coincide with the 24 broad career groups used by the UCLA online alumni-student mentoring platform, *UCLAOne*.

from the set of these same four occupations within the student's chosen career path. This customization ensures that students are only shown mentor profiles relevant to their interests.

3.2 Testing the effect of mentor quality information on willingness to pay

In order to test the effect of information provision on student WTP for same-gender mentors, before starting the survey, students are randomized to see one of two survey templates. Students randomized into the 'no ratings' template are shown only the information about mentors described above—gender, occupation, availability, first-generation status, and graduation year. Students randomized into the 'ratings' template received all of the information above, and additionally received ratings from a (hypothetical) past mentee. Appendix Figure A1 provides a screenshot of the mentor pairs shown to students during the survey in the 'ratings' template.[13] The 'no ratings' template is identical except the bottom box featuring ratings is omitted. We randomized the gender of the past mentee and the ratings. Ratings were either one star, three stars, or five stars (each with equal probability) in each of three evaluation categories: knowledgeable about job opportunities, easy to talk to/friendly, gave personalized advice. To select these attributes, in a pilot survey of the same population, we asked students why mentor gender is important. Two characteristics were by far the most cited: female mentors were more comfortable to interact with and better able to give advice "specifically for me." In the ratings, we also include a proxy for a mentor's general knowledge.

3.3 Econometric framework

In order to estimate students' preferences for mentor attributes, we assume student i of gender g has preferences over mentor j which can be approximated with a linear indirect utility function in mentor characteristics characteristics $\mathbf{x}$ in choice pair c:

$$V_{ijc} = \gamma^g + \mathbf{x}'_{ijc}\beta^g + \varepsilon_{ijc} \tag{2}$$

The probability that a student selects mentor a over mentor b in choice c is:

$$P^g\left(V_{iac} > V_{ibc}\right) = \alpha^g + (\mathbf{x}_{iac} - \mathbf{x}_{ibc})'\beta^g + \epsilon_{ic} \tag{3}$$

We estimate the following specification using a linear probability model (LPM):

$$C_{ic} = \alpha^g + (\mathbf{x}_{iac} - \mathbf{x}_{ibc})'\beta^g + \epsilon_{ic} \tag{4}$$

[13] Note that the location of the mentor was always Los Angeles.

where the dependent variable C_{ic} is an indicator for whether the student chose mentor a over mentor b in a given mentor pair. The independent variables are the differences in the characteristics of mentor a, $\mathbf{x}_{iac}$, and mentor b, $\mathbf{x}_{ibc}$ in choice pair c. The characteristics we control for are those observable to students: mentor gender, graduation year, availability, occupation, first-generation college student status, and when available, ratings and mentee gender. α^g captures the propensity to select the left profile (profile a) in a way that is unexplained by characteristics. In addition to the LPM, as robustness, we estimate a logit model.[14] This empirical specification is similar to those used by Maestas et al. (2018), Wiswall and Zafar (2018), and Mas and Pallais (2017). We do not adjust our results for inattention as in Mas and Pallais (2017) because in practice we find that 99.6 percent of students passed our attention check.

We use the estimates of students' preferences for mentor attributes to compute students' willingness to pay to access a mentor of the same gender. Willingness to pay metrics are traditionally denominated in monetary terms, for instance, the willingness to pay in hourly wages for a job with a higher fraction of coworkers who are female. Informal interactions for the purpose of information gathering seldom involve a monetary exchange.[15] For this reason, we use whether the student is willing to trade off a mentor with their preferred occupation in order to access a same-gender mentor, by computing the ratio of the two coefficients.

Note that, due to our survey design, mentor gender is randomly assigned to each profile and is, by construction, not correlated with other mentor characteristics. An additional benefit of the survey design is that we as researchers observe and control for all mentor attributes observed by students.

3.4 Summary statistics

Appendix Table A3 reports summary statistics for the 834 students who took the preference elicitation survey between November 2021 and January 2022. The survey respondents represent a diverse cross-section of UCLA undergraduates: 63% are female, 28% are first-generation college students, 54% are Asian American/Pacific Islander, and 14% are Hispanic/Latino. Students, on average, are sophomores, but freshmen through seniors are represented in the sample.[16] There are few differences between male and female students, aside from female students being slightly more likely to be first-generation college goers. We confirm that student demographics are balanced across the two survey templates in Appendix Table A4.

[14] The logit estimates the coefficients from $P^g\left(V_{iac} > V_{ibc}\right) = \frac{\exp\{(\mathbf{x}_{iac}-\mathbf{x}_{ibc})'\beta^g\}}{1+\exp\{(\mathbf{x}_{iac}-\mathbf{x}_{ibc})'\beta^g)\}}$.

[15] This stands in contrast to formal information gathering interactions, such as soliciting financial advice from a professional.

[16] Among currently enrolled UCLA undergraduates, 58% of students are female, 31% are first-generation students, 33% are Asian/Pacific Islander, and 21% are Hispanic/Latino.

4 Estimating Willingness to Pay for Mentor Gender: Results

In this section we use an incentive compatible preference elicitation survey to estimate students' preferences over mentor attributes. We find that female students have a strong preference for female mentors and are willing to trade off valuable mentor attributes in order to access a female mentor. In contrast, male students have a weak preference for male mentors. Female students' preference for female mentors is not driven by taste-based discrimination: when we provide students with information on mentor quality through ratings of mentors given by past mentees, female students are no longer willing to trade off valuable mentor characteristics in order to access a mentor of the same gender.

4.1 Female students are willing to pay for female mentors

We start off by estimating students' preferences for mentor characteristics in the 'no ratings' survey condition, separately for male and female students.[17] In Table 3 columns 1 and 2, we find that, all else equal, both male and female students value mentors whose occupation matches the student's preferred occupation (within the student's chosen broad career path).[18] In fact, students are 32-34 percentage points more likely to choose a mentor when the mentor's occupation switches from non-preferred to preferred.

We also find evidence of homophily: female students strongly and significantly prefer female mentors. Female students are 9.3 percentage points more likely to choose a mentor profile when the profile switches from male to female. In contrast, male students have a much weaker (and marginally significant) preference for male mentors. While mentor occupation and gender are both independently valued by female students, note that female students' preference for mentor occupation is substantially stronger than their preference for mentor gender.

Next we compute students' willingness to pay (WTP) to access a mentor of the same gender. While WTP metrics are traditionally denominated in monetary terms, informal interactions for the purpose of information gathering seldom charge a fee but often involve trade-offs. We calculate WTP for a female mentor as the ratio of the coefficients on female mentor and preferred occupation. The estimates indicate that female students are willing to give up a mentor with their preferred occupation 28 percent of the time in order to access a female mentor.[19] In contrast, the corresponding willingness to pay of male students for male mentors is just 5 percent. The results are nearly identical when using a logit specification (see

[17]We cannot separately analyze non-binary students due to their small sample size.

[18]After the preference elicitation, we ask students which of the four occupations in their chosen career path is their most preferred.

[19]This calculation depends on the linearity assumption in our econometric framework. If we limit our analysis to choice pairs in which female students are directly trading off their preferred occupation and whether the mentor is female, we find that female students make this trade off 21 percent of the time.

Appendix Table A5).[20]

4.2 Information on mentor quality eliminates willingness to pay

When additional information on mentor quality is available, are female students still willing to trade off valuable mentor attributes in order to access a female mentor? We investigate this question with use of the 'ratings' survey condition, in which we include information on mentor quality based on ratings from a past mentee. Specifically, in Table 3 columns 3 and 4, we estimate students' preferences for mentor characteristics in the 'ratings' survey condition, again by student gender. The inclusion of mentor ratings attenuates students' preferences for all original mentor attributes, but the attenuation is most pronounced for mentor gender. For both male and female students, the coefficients on mentor gender are now precisely estimated zeroes. Female students' willingness to pay for a female mentor—as measured by the trade-off of mentor gender relative to occupation match—declines by an order of magnitude and is now indistinguishable from zero. This means that when additional information on mentor quality is provided, students are no longer willing to trade off important mentor attributes such as occupation match in order to access a mentor of the same gender. Moreover, we can reject equality of female students' WTP estimates in the ratings and no ratings survey conditions.

We note that the attenuation of willingness to pay in the 'ratings' condition is not mechanically driven by the fact that profiles with ratings are longer, have more mentor attributes, or are in some other way distracting from the original attributes. When we analyze a pre-registered secondary outcome—the willingness to pay of first-generation college students for first-generation mentors—we find that including ratings *does not* attenuate their willingness to pay (Appendix Table A6).

When we examine students' valuation of mentor ratings, we find that students value all three categories, with knowledge about job opportunities valued a bit more than whether the mentor is easy to talk to/friendly and whether the mentor gives personalized advice. Furthermore, female students are not more sensitive to mentor quality than are male students: their respective coefficients on mentor quality are nearly identical.

4.3 Roots of homophily by gender

In the presence of information on mentor quality, female students are no longer willing to trade off valuable mentor characteristics in order to access a female mentor. This result implies that homophily is not driven by taste-based discrimination. Why does information provision affect female students' WTP for female mentors? Female students could be using mentor gender as a proxy for mentor quality. To shed light on female students'

[20]Note that the observational data shows much stronger homophily among male students than the preference elicitation survey, suggesting an important role for omitted variable bias in observational measures of homophily.

perceptions of how mentors gender shapes mentor quality, we asked students after the preference elicitation whether mentor gender was important to them and why. Fifty percent of female students and just 10% of male students reported that mentor gender is important.[21] Among the female students who stated that they valued a female mentor, 85% reported that it is because female mentors are friendlier/easier to talk to and and 53% reported that it is because female mentors are better at giving personalized advice. In contrast, only 9% reported that female mentors are more knowledgeable about job opportunities. Female students' perceptions that male and female mentors differ, on average, is consistent with statistical discrimination based on (accurate or inaccurate) beliefs. Student's beliefs could arise due to stereotypes that are partially accurate but exaggerated (Bordalo et al., 2016; Eyal and Epley, 2017).

We also explore whether the WTP for female mentors depends on the perception of gender-specific benefits (or costs). Using the randomization of the gender of the mentee who rates the mentor, we test whether (1) students value ratings from a same-gender mentee more and (2) whether the preference for female mentors is equally attenuated by male and female mentee ratings. In Table 4 we find that ratings from male and female mentees are equally valued by female students (as well as by male students). In addition, by limiting the analysis to pairs of profiles with only male or only female mentees, we find that both are equally effective in attenuating female students' WTP for female mentors. These results suggest that female students do not require information on the benefits that female mentees derived from female mentors, such as discussions of personal experiences being a woman in finance. Furthermore, female students do not require another woman to vouch for a mentor prior to mentor selection.

5 Implications for Program Design

Our results have implications for mentorship programs that match on race/ethnicity, nationality, gender, and sexual orientation. Optimal program design depends on the source of homophily. In some cases, matching based on shared traits may be optimal because students directly value that trait or get unique information from mentors with that trait. For example, as a pre-registered secondary outcome in our preference elicitation survey, we estimate that homophily by first-generation college student status is substantial and invariant to providing information on mentor quality (see Appendix Table A6).

If homophily is driven by lack of information on mentor quality, then resources could be better invested recruiting mentors based on quality rather than shared traits. For example, if recruiting female mentors requires sacrificing some dimension of mentor quality and female students are aware of the quality trade-off, then female students are unwilling to make that trade-off. Female students would rather have a mentor of a

[21] Students' stated preferences were strongly predictive of their revealed preferences from the preference elicitation.

different gender than sacrifice mentor quality.

How should mentorship programs incorporate participant preferences into their design? Given a matching rule, let $f(x)$ be the distribution of match quality for a given student when there is no screening of mentors. If the program restricts mentors to share traits with students (for example, by offering female students only female mentors), then it shifts the distribution of match quality to $f^g(x)$. For example, if match quality is on average higher in the population of female mentors, then $f^g(x) = f(x + a)$. An alternative policy is quality screening, which we can model as truncating the distribution $f(x)$ below some threshold, $f(x|x > q)$. Assuming that truncating based on quality is costly, and perhaps increasingly costly as the quality truncation threshold increases, programs may be better off restricting matches to shared traits.[22] If obtaining information on quality is straightforward, for example, through the use of existing surveys of mentee experiences, then the optimal policy would screen mentors on quality. See Appendix Figure A3 for a graphical example of match quality under the these policies.

More broadly, initiatives in employee recruitment, service-provider matching, and doctor-patient matching that commonly use shared traits as a coarse proxy for match quality—while well intentioned—could lead to efficiency losses relative to those that incorporate information on valued traits into the matching process. As an example of this, ride-sharing platforms have opted to inform riders that their driver has been background checked rather than offer same-gender matching. Finally, since matching on shared traits often occurs in settings where individuals with the trait are scarce, an additional benefit of shifting to matching based on quality metrics is it would alleviate the time burden of these initiatives on already underrepresented groups.

[22]We think of this cost abstractly. For example, in settings where there is selection of mentors into mentoring roles, the cost of truncating on quality may be that fewer mentors volunteer and the overall pool is worse or is scarce.

References

Abel, Martin, Rulof Burger, and Patrizio Piraino, "The Value of Reference Letters: Experimental Evidence from South Africa," *American Economic Journal: Applied Economics*, July 2020, *12* (3), 40–71.

Agan, Amanda and Sonja Starr, "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment," *The Quarterly Journal of Economics*, 08 2017, *133* (1), 191–235.

Aigner, Dennis J. and Glen G. Cain, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 1977, *30* (2), 175–187.

Alam, Md Moshi Ul, "Employee-Side Discrimination: Beliefs and Preferences," *Working paper*, 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.

Ayalew, Shibiru, Shanthi Manian, and Ketki Sheth, "Discrimination from below: Experimental evidence from Ethiopia," *Journal of Development Economics*, 2021, *151.*

Babcock, Linda, Maria P. Recalde, Lise Vesterlund, and Laurie Weingart, "Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability," *American Economic Review*, March 2017, *107* (3), 714–47.

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

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral Science*, 1964, *9* (3), 226–232.

Behncke, Stefanie, Markus Frölich, and Michael Lechner, "A Caseworker Like Me—Does The Similarity Between The Unemployed and Their Caseworkers Increase Job Placements?," *The Economic Journal*, 2010, *120* (549), 1430–1459.

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

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope, "Inaccurate Statistical Discrimination: An Identification Problem," Working Paper 25935, National Bureau of Economic Research June 2019.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Stereotypes*," *The Quarterly Journal of Economics*, 07 2016, *131* (4), 1753–1794.

Cabral, Marika and Marcus Dillender, "Gender Differences in Medical Evaluations: Evidence from Randomly Assigned Doctors," *Working paper*, 2021.

Canaan, Serena and Pierre Mouganie, "The Impact of Advisor Gender on Female Students' STEM Enrollment and Persistence," *Journal of Human Resources*, 2021.

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.

Chan, Alex, "Customer Discrimination and Quality Signals: A Field Experiment with Healthcare Shoppers," *Working Paper*, 2021.

Charles, Kerwin Kofi and Jonathan Guryan, "Taste-Based Discrimination," in "The New Palgrave Dictionary of Economics," London: Palgrave Macmillan UK, 2018, pp. 13425–13432.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin, "An Economic Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, 2009, *77* (4), 1003–1045.

Eyal, Tal and Nicholas Epley, "Exaggerating Accessible Differences: When Gender Stereotypes Overestimate Actual Group Differences," *Personality and Social Psychology Bulletin*, 2017, *43* (9), 1323–1336. PMID: 28903680.

Gallen, Yana and Melanie Wasserman, "Do Male and Female Students Use Networks Differently?," *AEA Papers and Proceedings*, May 2021, *111*, 154–58.

_ and _ , "Informed Choices: Gender Gaps in Career Advice," *IZA DP No. 14072*, 2021.

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.

Hedegaard, Morten Størling and Jean-Robert Tyran, "The Price of Prejudice," *American Economic Journal: Applied Economics*, January 2018, *10* (1), 40–63.

Kaas, Leo and Christian Manger, "Ethnic Discrimination in Germany's Labour Market: A Field Experiment," *German Economic Review*, 2012, *13* (1), 1–20.

Kessler, Judd B., Corinne Low, and Colin D. Sullivan, "Incentivized Resume Rating: Eliciting Employer Preferences without Deception," *American Economic Review*, November 2019, *109* (11), 3713–44.

Maestas, Nicole, Kathleen Mullen, David Powell, Till von Wachter, and Jeffrey Wenger, "The Value of Working Conditions in the United States and Implications for the Structure of Wages," *NBER Working Paper 25204*, 2018.

Mas, Alexandre and Amanda Pallais, "Valuing alternative work arrangements," *American Economic Review*, 2017, *107* (12), 3722–3759.

McDevitt, Ryan C. and James W. Roberts, "Market structure and gender disparity in health care: preferences, competition, and quality of care," *The RAND Journal of Economics*, 2014, *45* (1), 116–139.

O*NET-SOC AutoCoder, "No Title," \url{https://onetsocautocoder.com/plus/onetmatch} 2020.

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.

Reyes, Jessica Wolpaw, "Gender Preference and Equilibrium in the Imperfectly Competitive Market for Physician Services," *Eastern Economic Journal*, 2008, *34*, 325–346.

Rivera, Lauren A., *Pedigree: How Elite Students Get Elite Jobs*, Princeton University Press, 2015.

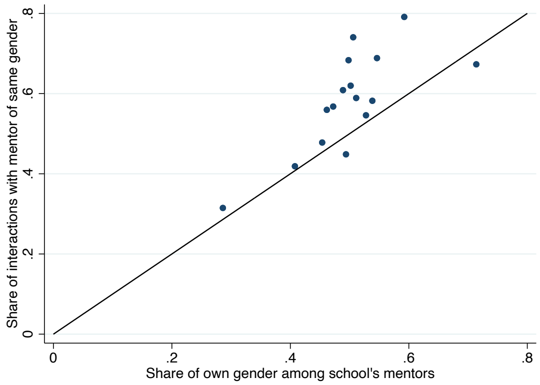
Tang, Yanli, Pengfei Guo, Christopher S. Tang, and Yulan Wang, "Gender-Related Operational Issues Arising from On-Demand Ride-Hailing Platforms: Safety Concerns and System Configuration," *Production and Operations Management*, 2021, *30* (10), 3481–3496.

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

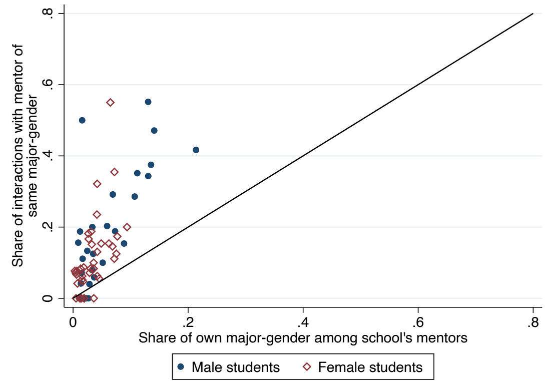
Zeltzer, Dan, "Gender homophily in referral networks: Consequences for the medicare physician earnings," *American Economic Journal: Applied Economics*, 2020, *12* (2), 169–197.

Figures and Tables

Figure 1: Homophily on an Online Mentoring Platform



(a) Homophily by Gender



(b) Homophily by Gender and College Major

Note: This figure uses data from eight universities/colleges to plot the share of messages initiated by students that were sent to an alumni with a shared trait. The left panel analyzes the fraction of conversations with a same-gender alumni and the right panel examines the fraction of conversations with a same-gender and same-major alumni.

Table 1: Relative Homophily by Gender

	(1)	(2)	(3)
Student Female	0.193***	0.136***	0.124***
	(0.020)	(0.018)	(0.018)
Mean among male students	0.333		
Mentor Controls	No	Yes	Yes
Student Controls	No	No	Yes
Observations	4144	4139	4139
R-squared	0.038	0.139	0.150

Note: This table displays coefficients β from a regression of the form $RecipientFemale_{imj} = \alpha + \beta StudentFemale_i + X'_{ij}\gamma + \epsilon_{imj}$ Controls include school, student major, student graduation year, recipient major, and recipient occupation fixed effects, as well as a linear term for recipient graduation year. Robust standard errors in parentheses, clustered at the student sender level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Responses to Female Students, by Mentor Gender

	(1) Response Received	(2) Length of Response	(3) Log Length of Response
Mentor is female	-0.083***	-38.820	-0.067
	(0.025)	(51.676)	(0.063)
Sample	Female Students	Female Students	Female Students
Mean among male mentors	0.667	539.566	5.767
Observations	1617	1039	1039
R-squared	0.119	0.133	0.174

Note: This table presents the results of a regression of the outcomes of messages sent by female students (labeled in each regression in columns 1-3) on an indicator for whether the message was sent to a female mentor. The mean outcome among messages sent to male mentors is listed in the bottom panel. All regressions include controls for school, student major, student graduation year, recipient major, and recipient occupation fixed effects, as well as a linear term for recipient graduation year. Robust standard errors clustered at the student level are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Student Preferences for Mentor Attributes: By Student Gender

	(1)	(2)	(3)	(4)
	No Ratings		Ratings	
	Female	Male	Female	Male
Mentor is female	0.093***	-0.016*	0.007	-0.002
	(0.008)	(0.009)	(0.006)	(0.009)
Mentor has preferred occ	0.335***	0.324***	0.130***	0.129***
	(0.012)	(0.016)	(0.011)	(0.016)
Mentor graduation year	0.007***	0.005***	0.001*	0.002**
	(0.001)	(0.002)	(0.001)	(0.001)
Availability (in 10 min increments)	0.031***	0.039***	0.003	0.010***
	(0.003)	(0.004)	(0.002)	(0.003)
Mentor first-gen	0.070***	0.037***	0.024***	0.018**
	(0.011)	(0.012)	(0.007)	(0.009)
Knowledgeable about job opportunities			0.091***	0.092***
			(0.003)	(0.004)
Easy to talk to/friendly			0.065***	0.067***
			(0.003)	(0.003)
Gave personalized advice			0.071***	0.067***
			(0.003)	(0.004)
Mentee is female			-0.008	0.010
			(0.007)	(0.009)
WTP for female mentor	0.278***	-0.051*	0.054	-0.012
	(0.027)	(0.029)	(0.049)	(0.068)
p-value $WTP_{noratings} = WTP_{ratings}$	0.000	0.601		
Observations	8100	4620	7710	3900
Number of students	270	154	257	130

Note: This table displays coefficients β from estimating the following linear probability model: $C_{ic} = \alpha^g + (\mathbf{x}_{iac} - \mathbf{x}_{ibc})'\beta^g + \epsilon_{ic}$. Willingness to pay is calculated as the ratio of the coefficients on female mentor and preferred occupation. Standard errors, clustered at the student level, are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

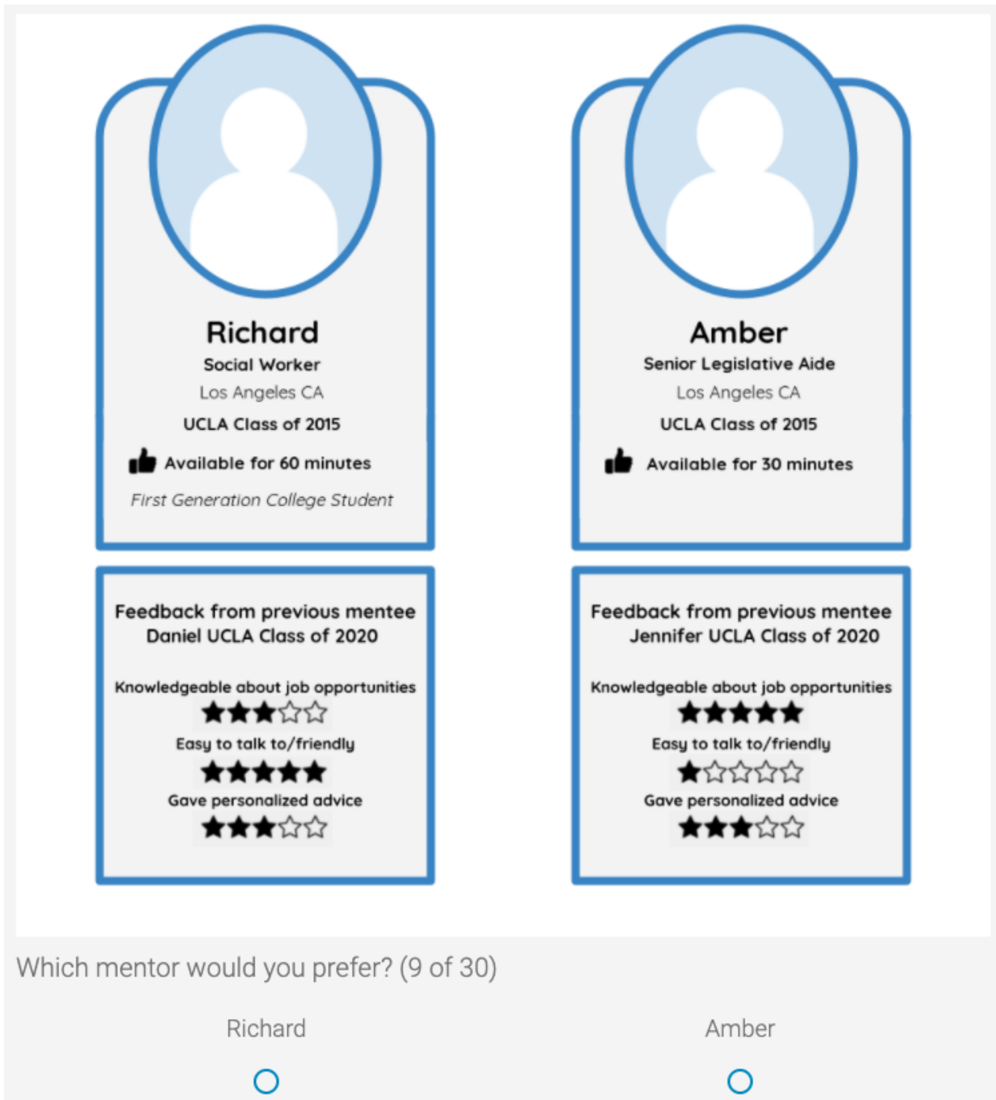
Table 4: Student Preferences for Mentor Attributes: Role of Mentee Gender

	(1) Female	(2) Male
Mentor is female	0.007	-0.002
	(0.006)	(0.009)
Mentor has preferred occ	0.130***	0.129***
	(0.011)	(0.016)
Mentor graduation year	0.001*	0.002**
	(0.001)	(0.001)
Availability (in 10 min increments)	0.003	0.010***
	(0.002)	(0.003)
Mentor first-gen	0.024***	0.018**
	(0.007)	(0.009)
Knowledgeable about job opportunities	0.091***	0.092***
	(0.003)	(0.005)
Easy to talk to/friendly	0.065***	0.068***
	(0.003)	(0.004)
Gave personalized advice	0.072***	0.065***
	(0.003)	(0.005)
Mentee is female × Knowledgeable about job opportunities	-0.000	0.000
	(0.004)	(0.006)
Mentee is female × Easy to talk to/friendly	0.000	-0.003
	(0.004)	(0.005)
Mentee is female × Gave personalized advice	-0.002	0.004
	(0.004)	(0.006)
WTP for female mentor	0.053	-0.013
	(0.049)	(0.068)
Observations	7710	3900
R-squared	0.399	0.394
Number of students	257	130

Note: Note: This table displays coefficients β from estimating the following linear probability model: $C_{ic} = \alpha^g + (\mathbf{x}_{iac} - \mathbf{x}_{ibc})'\beta^g + \epsilon_{ic}$. Willingness to pay is calculated as the ratio of the coefficients on female mentor and preferred occupation. Standard errors, clustered at the student level, are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A Appendix Figures and Tables

Figure A1: Example of Mentor Profiles



Note: This figure is a screenshot of a pair of profiles shown in the hypothetical choice preference elicitation survey administered among UCLA undergraduate students. The profiles are from the survey version with mentor ratings. The survey version without ratings omits the box below each profile. The profiles correspond to the career path, Community and Social Services. The full set of career paths is: Accounting; Administrative/Support; Arts and Design; Business Development; Community and Social Services; Consulting; Education; Engineering; Entrepreneurship; Finance; Healthcare Services; Human Resources; Information Technology; Legal; Marketing; Media and Communications; Military and Protective Services; Operations; Program and Product Management; Quality Assurance; Real Estate; Research; Sales; Purchasing.

Figure A2: Example of Advice Email

UCLA

Hello,

Thank you for taking the career advice survey. Please find below your personalized advice.

Did you know that you can search and review Alumni Mentor profiles on UCLAOne using the Directory tab: https://uclaone.com/ Alumni who are willing to provide mentorship have a "willing to help" banner.

Using the type of mentor you consistently chose in the mentor comparison portion of the survey, we evaluated whether you valued (1) the job title of a mentor (2) whether the mentor was a first-generation college goer and (3) the experience of a mentor (years since graduation).

Based on your choices, you seem to be interested in mentors who are first generation college students and are recent college graduates. Your choices did not suggest a strong preference for the other mentor characteristics.

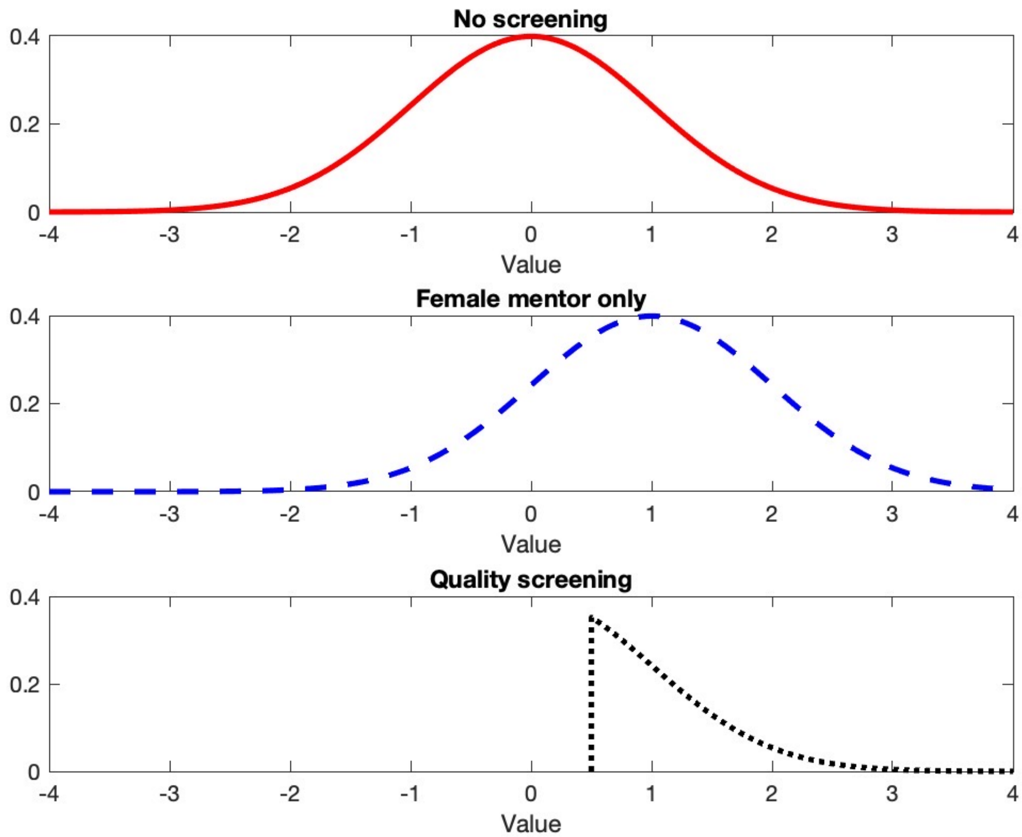
To find mentors with these characteristics, on the directory page of UCLAOne, filter your search using one or more of the following categories:

- UCLA: field of study, major, graduation year, communities (e.g. First-Generation)
- Keyword search (e.g. Senior Legislative Aide)
- Location: city, state, country

We hope this information was helpful to you!

Note: This figure is a screenshot of an advice email that students received, based on their survey responses.

Figure A3: Match quality under various policies



Note: This figure depicts the distribution of match quality for a given mentor-female student pair when there is no screening of mentors (top panel), when only female mentors are available to female students (middle panel), and when mentors are screened on quality (bottom panel). The distribution of match quality in these examples is normal. The distribution of match quality when only female mentors are available is assumed to have the same variance but a higher mean than the distribution of match quality when there is no screening.

Table A1: Mentoring Platform Summary Statistics: Student and Alumni Users

	Students		Alumni	
	Mean	SD	Mean	SD
Female	0.500	0.500	0.460	0.498
Graduation Year	2019	2.719	2005	14.130
Major unknown	0.532	0.499	0.106	0.308
Any Message Sent	0.116	0.321	0.093	0.29
Total Messages Sent	0.364	1.705	0.127	2.744
Liberal Arts College	0.337	0.473	0.463	0.499
Research University	0.663	0.473	0.537	0.499
Observations	9257		16113	

Note: This table displays summary statistics for student and alumni users of the mentoring platform among schools with substantial messaging between students and alumni in our data. The variable Any Message Sent is an indicator for whether a message was sent (or responded to, in the case of alumni) restricting to the set of conversations between students and alumni in which students initiated the conversation the topic of the conversation was job- or major- related.

Table A2: Responses to Male Students by Mentor Gender

	(1) Response Received	(2) Length of Response	(3) Log Length of Response
Mentor is female	-0.032	29.826	0.078
	(0.027)	(50.717)	(0.075)
Sample	Male Students	Male Students	Male Students
Mean among male mentors	0.570	438.032	5.579
Observations	1738	999	999
R-squared	0.109	0.140	0.177

Note: This table presents the results of a regression of the outcomes of messages sent by male students (labeled in each regression in columns 1-3) on an indicator for whether the message was sent to a female mentor. The mean outcome among messages sent to male mentors is listed in the bottom panel. All regressions include controls for school, student major, student graduation year, recipient major, and recipient occupation fixed effects, as well as a linear term for recipient graduation year. Robust standard errors clustered at the student level are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Survey Summary Statistics

	All Students	Female	Male	Non-binary	First-Gen	Non First-Gen
Female	0.63				0.68	0.61
	(0.48)				(0.47)	(0.49)
Non-binary	0.03				0.02	0.03
	(0.16)				(0.13)	(0.18)
First-generation college goer	0.28	0.30	0.25	0.17		
	(0.45)	(0.46)	(0.44)	(0.39)		
Asian/Pacific Islander	0.54	0.54	0.55	0.43	0.43	0.59
	(0.50)	(0.50)	(0.50)	(0.51)	(0.50)	(0.49)
Hispanic/Latino	0.14	0.13	0.14	0.17	0.38	0.04
	(0.34)	(0.34)	(0.35)	(0.39)	(0.49)	(0.20)
White/Caucasian	0.22	0.21	0.22	0.30	0.11	0.26
	(0.41)	(0.41)	(0.42)	(0.47)	(0.31)	(0.44)
Expected graduation year	2024	2024	2024	2024	2024	2024
	(1.11)	(1.12)	(1.06)	(1.31)	(1.11)	(1.11)
Observations	834	527	284	23	235	599

Note: This table reports summary statistics for the preference elicitation survey respondents. Students chose between three gender identities: male, female, and non-binary. Statistics are reported for all students, and separately by gender category and first-generation college student status. Standard deviations are in parentheses.

Table A4: Balance Table for Ratings vs. No Ratings

	No Ratings	Ratings	Difference	P-value
Fraction Female	0.619	0.646	0.027	0.416
Fraction First-Generation College Students	0.280	0.284	0.004	0.895
Fraction Asian/Pacific Islander	0.537	0.548	0.011	0.750
Fraction Hispanic/Latino	0.144	0.131	-0.014	0.563
Fraction White	0.218	0.221	0.003	0.911
Expected Graduation Year	2023.663	2023.721	0.058	0.448
Number of students	436	398		

Note: This table displays mean student characteristics for students who were randomized into the 'no ratings' preference elicitation template, the 'ratings' template, and provides the p-value for a t-test of the difference between the two groups.

Table A5: Student Preferences for Mentor Attributes Estimated with Logit: By Student Gender

	(1)	(2)	(3)	(4)
	No Ratings		Ratings	
	Female	Male	Female	Male
Mentor is female	0.480***	-0.081*	0.059	-0.030
	(0.043)	(0.046)	(0.043)	(0.059)
Mentor has preferred occ	1.738***	1.617***	0.909***	0.937***
	(0.087)	(0.118)	(0.078)	(0.105)
Mentor graduation year	0.038***	0.024***	0.008*	0.016**
	(0.005)	(0.008)	(0.005)	(0.007)
Availability (in 10 min increments)	0.159***	0.191***	0.031**	0.071***
	(0.015)	(0.021)	(0.015)	(0.021)
Mentor first-gen	0.359***	0.182***	0.167***	0.112**
	(0.056)	(0.058)	(0.047)	(0.056)
Knowledgeable about job opportunities			0.611***	0.629***
			(0.026)	(0.042)
Easy to talk to/friendly			0.446***	0.465***
			(0.024)	(0.031)
Gave personalized advice			0.489***	0.468***
			(0.025)	(0.035)
Mentee is female			-0.047	0.097
			(0.045)	(0.061)
WTP for female mentor	0.276***	-0.050*	0.065	-0.032
	(0.027)	(0.028)	(0.047)	(0.063)
p-value $\text{WTP}_{noratings} = \text{WTP}_{ratings}$	0.000	0.799		
Observations	8100	4620	7710	3900
Number of students	270	154	257	130

Note: This table displays coefficients β from estimating the following logit model: $C_{ic} = \alpha^g + (\mathbf{x}_{iac} - \mathbf{x}_{ibc})'\beta^g + \epsilon_{ic}$. Willingness to pay is calculated as the ratio of the coefficients on female mentor and preferred occupation. Standard errors, clustered at the student level, are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Student Preferences for Mentor Attributes: By First-Generation Status

	(1)	(2)	(3)	(4)
	No Ratings		Ratings	
	First-Gen	Non First-Gen	First-Gen	Non First-Gen
Mentor first-gen	0.159***	0.017**	0.070***	0.002
	(0.016)	(0.008)	(0.011)	(0.006)
Mentor is female	0.048***	0.053***	0.017*	0.001
	(0.012)	(0.008)	(0.010)	(0.006)
Mentor has preferred occ	0.306***	0.340***	0.090***	0.146***
	(0.018)	(0.011)	(0.014)	(0.011)
Mentor graduation year	0.006***	0.006***	0.001	0.002**
	(0.002)	(0.001)	(0.001)	(0.001)
Availability (in 10 min increments)	0.034***	0.033***	0.007**	0.005**
	(0.004)	(0.003)	(0.004)	(0.002)
Knowledgeable about job opportunities			0.094***	0.090***
			(0.004)	(0.003)
Easy to talk to/friendly			0.065***	0.066***
			(0.004)	(0.003)
Gave personalized advice			0.071***	0.069***
			(0.004)	(0.003)
Mentee is female			0.001	-0.003
			(0.011)	(0.006)
WTP for first-gen mentor	0.518***	0.051**	0.778***	0.017
	(0.065)	(0.025)	(0.170)	(0.042)
p-value $WTP_{noratings} = WTP_{ratings}$	0.153	0.483		
Observations	3660	9420	3390	8550
Number of students	122	314	113	285

Note: This table displays coefficients β from estimating the following linear probability model: $C_{ic} = \alpha^g + (\mathbf{x}_{iac} - \mathbf{x}_{ibc})'\beta^g + \epsilon_{ic}$. Willingness to pay is calculated as the ratio of the coefficients on first-generation mentor and preferred occupation. Standard errors, clustered at the student level, are in parentheses.

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

B Online Appendix - Preparation of Mentoring Platform Data

Assignment of mentor job titles to occupation codes

O*NET-SOC AutoCoder was created by R.M. Wilson Consulting, Inc. for the US Department of Labor for the purpose of assigning occupation codes to resumes, job descriptions, and job titles (O*NET-SOC AutoCoder, 2020). The AutoCoder, for example, returns SOC-2010 code 13-1111, "Management Analysts," for the job title "Analyst at Y Consulting." Because the job titles in our data are user-supplied, we cannot confidently assign occupation in some cases. O*NET-SOC AutoCoder provides a confidence estimate for every occupation-job title match. We accept any matches provided with confidence scores above 70 percent (on the advice of the developer) and then manually attempt to match occupation for all remaining job titles. In case after this process our data do not include information sufficient to assign occupation, we code that as missing.

Assignment of mentor and student gender

We first assign gender using the 1990 Census and 1940-1970 Social Security Administration (SSA) name files. For a given name, if 90 percent of individuals with this name are classified as either male or female, then the name is designated as such. The remaining names are left as unclassified. In cases where there is conflict between the Census and SSA assigned gender, a name is unclassified. Because our sample includes names uncommon in the US, we use the API genderize.io (available here `https://genderize.io`) to classify any names which are uncommon or unknown in the Census and SSA files, using the same 90 percent criteria for assigning names.

Melanie Wasserman
Assistant Professor of Economics
melanie.wasserman@anderson.ucla.edu
(310) 825-9169

UCLA Anderson School of Management
110 Westwood Plaza, C521
Los Angeles, CA 90095

September 29, 2022

Dear Editor,

Please consider the attached manuscript, "Does Information Affect Homophily?" for publication in the *Journal of Public Economics*.

We kindly request that our paper is handled by co-editor Ricardo Perez-Truglia.

The paper was previously submitted to (and rejected by) *The Review of Economics and Statistics*. We enclose the editor's decision letter as well as the three referee reports we received. We believe that we have addressed the referees' comments/suggestions and the paper is much improved. The main changes are summarized below:

- Discussion of the issue that the information treatment also makes profiles longer. We believe that this is the primary issue "difficult to overlook" in the previous submission. In retrospect, we see how the previous version of the paper did not directly explain why readers should not be concerned with a mechanical attenuation of preferences in the information treatment. In the new version of the paper, we are more clear and direct on this issue. First, it is important to note that the information treatment does not mechanically attenuate the willingness-to-pay metrics and we emphasize that this is why we focus our interpretation on the willingness-to-pay metrics rather than the preference coefficients. We also note that the treatment does not always attenuate willingness-to-pay, as in the case of homophily by first generation college goer status, mitigating the concern there is mechanical decline.
- Discussion of evidence that perceptions of gender differences in personality traits may be exaggerated
- Discussion of concerns regarding social desirability bias
- Clarification of willingness-to-pay metrics

We also include a version of the paper with all changes highlighted. There are a few responses to referees that include analysis or discussion that we did not incorporate into the paper (e.g. homophily by gender of first-generation male students, role of mentee gender for male mentor selection). We would be happy to include this material in the paper, if requested.

Thank you in advance for your consideration.

Best regards,

Melanie Wasserman

Yana Gallen

110 Westwood Plaza, C521
Los Angeles, California 90095

Dear Dr. Wasserman:

I have now received reports from three expert reviewers on your submission to the *Review of Economics and Statistics*. Unfortunately, I have bad news.

The reviewers agree that your paper addresses an important question and has the potential to make a nice contribution to the literature. However, the reviewers question the fit for a general interest journal and raise several important questions regarding the interpretation of the results. These issues are serious enough for two of the three reviewers to recommend rejection. Having read the paper and the reports, I agree with the more negative reviewers' view that these issues are difficult to overlook despite the importance of the research topic. I have, therefore, decided to reject the paper for publication in *REStat.* There is a lot to like in the paper, but I believe your manuscript is a better fit for a specialized field journal such as the *Journal of Public Economics.*

I realize that this is not welcome news. Please keep in mind that given the large volume of submissions to the journal, *REStat* rejects over 90 percent of the manuscripts it receives. As a result, many high-quality papers are not accepted. I hope that, in spite of the disappointing outcome on this occasion, you will find the excellent set of referee comments useful and will continue to consider *REStat* as an outlet for your future work.

Thank you for submitting your paper to the *Review of Economics and Statistics.* Sincerely,

Will Dobbie

for the Board of Editors

<u>Reviewer 1</u>

Referee Report for Review of Economics and Statistics MS 27574 "Does Information Affect Homophily?"

SUMMARY

This paper estimates female students' preference for gender concordance with an alum-mentor that they can connect with via an online platform. Besides the correlation data supporting a preference for gender and field-of-occupation concordances, the key causal exercise is preference elicitation using a survey that asks respondents to select a preferred mentor profile out of each pair of hypothetical mentor profiles. The authors attempt to incentivize the survey responses by generating recommendations based on the survey responses (although these are not directly consequential to the respondent's actual choice sets). The key result from the survey exercise is that while female students prefer a mentor who is also female, such preference for gender concordance disappears when a quality signal about the mentor was included in the mentor profiles. Beyond the key empirical results, the authors also offered a high-level discussion of the possible mechanisms while acknowledging limitations in interpreting the results.

OVERALL COMMENTS

This paper investigates an interesting and policy relevant question to education, in particular the design of matching mechanisms for mentorship programs that are increasing common in both academic and industry settings. The biggest contribution of this paper is the evidence it brings to the literature on mentorship, adding to the authors previous work in the area. The findings were interesting and surprising while pointing to clear actions for the design such mechanisms. For this valuable information alone, this paper deserves earnest consideration for publication in a good journal. The paper also thoughtfully exploits both observational and lab-in-the-field experimental data, despite being a short paper. This was backed up by its adoption of new methodology from the field of preference elicitation with respect to discrimination. Some shortcomings plague the application of these newer methodologies in the specific design adopted by the authors, which weakened the strength of this paper from the experimental economist's point of view. The authors can make this paper an even better paper by making a few expositional enhancements (e.g. clarify assumptions behind the discrete choice model) and more careful use of terminology (e.g. willingness to pay when no money-metric utility was estimated). My comments below are about the contribution, empirical design/results, interpretation, and exposition.

SUMMARY ON CONTRIBUTIONS

As noted above, the central contribution of the paper is the interesting facts regarding preferences for gender concordance in the market for mentors. Even with lingering concerns with identification and interpretation, these new data are of interest to many labor economists and policymakers: the main empirical findings deserve to be published. The key limitations of this work are (a) potential insufficient incentivization of the survey responses; and (b) threats to identification in both the observational and experimental data. While there is a solid contribution from the correlational evidence presented in the paper, the causal results are not as air-tight (especially due to (b1) below) making it hard to interpret the data and understand the true mechanism.

EMPIRICAL DESIGN AND RESULTS:

Overall, the main empirical results are clearly and effectively presented but do suffer from a few shortcomings. Here are some specific comments:

(a) Potential insufficient incentivization of the survey responses. The survey is incentivized by mapping choices over hypothetical mentors to an email with "personalized advice" (see Figure A2). However, this has no direct consequences to either the choice set of mentor or the cost of accessing any particular mentor. These emails effects summarizes what seems to matter for that student (e.g. "Based on your choices, you seem to be interested in mentors who are first generation college students and are recent college graduates. Your choices did not suggest a strong preference for the other mentor characteristics") but this type of "consequence" deviate from the other state-of-the-art designs where the choices would have been tied to a restricted choice set of real mentors. The authors have not convinced me that this "personalized advice" was viewed as something valuable to the students or that these advices even necessarily shifted the actual choices over real mentors or that the "personalized advice" (written the way it was) necessarily make truth telling a weakly dominant strategy for the students. Also, given that students can "glean additional information about alumni from an online search," it is not clear what choice-relevant information is provided by this email. Hypothetical bias should not be considered to be ruled out in this paper.

Thank you for this comment concerning our incentives. To guide our comments below, our exact phrasing to students was, noted on page 8:

> 'We will use your responses in this section to give you personalized suggestions on how to find mentors. If you decide to receive these suggestions, you will receive these suggestions via email (which you will enter at the end of the survey). We will not contact any mentors on your behalf, we will only provide you with recommendations consistent with the choices you make in the next portion of this questionnaire."

We see our design as closely related to Kessler et al. (2019). We similarly tell students that the choices they make in the survey will be used to provide personalized information on how to find mentors. In part, our incentive was to reward students who complete the survey with personalized information about how to navigate the process of finding mentors (thus lowering the cost of finding mentors that they prefer), and to interest students who are looking for mentorship to participate in the study.

Regarding the concern that the emails the student received summarized what mattered to the student, we note that *at the time the students took the survey*—before they received these emails—they did not know what the content of the email would be, but only saw the directions above. The description to students about the effect of their choices was general enough that we do not believe it would engender strategic reporting. Concerning alternatives, as suggested by the referee, which would make choices tied to a restricted set of real mentors, our setting did not allow us to provide students with actual mentors, unfortunately.

We also would like to push back a little on the notion that a more complex elicitation would more accurately measure student beliefs, compared to our simple formulation above. We note that the experimental economics literature has raised some flags about detailed incentivized mechanisms to elicit beliefs. Danz, Vesterlund, and Wilson (AER 2022) concludes that a low-information incentive similar to ours (in their case, simply informing subjects that guessing their true beliefs maximizes their earnings) generates *more* truthful reporting than state-of-the-art incentivized belief elicitations, *especially* compared to when the incentives for truth telling in these belief elicitations are made extremely explicit. In light of this, and given that there is little motivation not to tell the truth in our setting, we believe that our design likely maximized truth telling relative to a mechanism that would be more difficult and complex to explain.

(b1) Threats to identification for the experiment. The treatment and controls differ in multiple ways. One is the inclusion of a quality signal for the treatment group and not the control. But other things are also different (see Figure A1). First, the mentor profiles are substantially longer and larger for the treatment group and the "gender signal" (along with other signals like availability and class year) occupy one out of 10 or so rows of mentor information in the treatment group whereas the "gender signal" occupies one out of 6 or so rows of mentor information. There is clearly a lot more distraction away from gender and other attributes like availability and class year in the treatment group – this hypothesis that the key result is NOT rooted in statistical discrimination but rather due to the distraction (the stars grabbed all the attention) is actually consistent with the data in Table 3:

Looking at female students (columns 1 and 3 of Table 3), as the authors did for the main results, we can see that the estimated "WTP for female mentor" is not the only thing that is significantly lower with the ratings (i.e. # in column 3 significantly lower than # in column 1). You can see that the estimated coefficients for "Availability" (difference is

0.028 but s.e.'s for each estimate are 0.003 and 0.002), "Mentor first gen" (difference is 0.046 but s.e.'s for each estimate are 0.011 and 0.007), and "mentor graduation year" (difference is 0.006 but s.e.'s for each estimate are 0.001 and 0.001) are all significantly lower when there are ratings. If distraction was not the cause, that all the students did were to replace gender as a quality proxy with the previous mentee ratings, we should not see all the other coefficients also significantly drop.

Thank you for bringing up these important design choices. When making comparisons across treatments, we focus on the WTP metric–the ratio of the coefficient on mentor gender and mentor occupation–precisely due to concern that the inclusion of quality metrics in the 'ratings' condition affects all preferences for mentor characteristics (not just mentor gender). The coefficients measure the proportion of time a profile with a certain characteristic is chosen. If there is important information available (such as quality ratings) then the coefficients will indeed all fall. This is why it's important to consider more than just how often a profile is chosen, but how often it's chosen *relative to* other desirable mentor characteristics. In our setting, this is expressed as the ratio of coefficients on mentor gender and preferred occupation across both the 'ratings' and 'no ratings' conditions.

Fortunately, our design also builds a sort of "placebo" test for the possibility that longer profiles mechanically alter the tradeoffs students make. We are happy to have the opportunity to discuss this in more detail in the newest version of the paper. We have added the following description to the paper (p. 12) which acknowledges the important issue raised and which we hope is helpful in thinking about the magnitude of this problem:

> "We note that the attenuation of willingness to pay in the 'ratings' condition is not mechanically driven by the fact that profiles with ratings are longer, have more mentor attributes, or are in some other way distracting from the original attributes. When we analyze a pre-registered secondary outcome–the willingness to pay of first-generation college students for first-generation mentors–we find that including ratings *does not* attenuate their willingness to pay (Appendix Table A6)."

(b2) Threats to identification for the observational data results (Section 2). As the authors pointed out on page 7 3rd para., results in Table 1 should not be casually be interpreted as causal. The authors cannot observe the choice sets or the attributes that the students evaluated when they make their mentor choices. One thing that was not mentioned is that endogeneity might also be present (correlations between gender and other relevant qualities). The authors are clear and careful to not overclaim here. My suggestion is to investigate further to see if more insights can come out of this observational data, given the weakness outlined above regarding the experiment. If the authors can identify natural experiments or exogenous variations in variables of interest, this can push the paper to a higher tiered journal.

Thank you for these suggestions. We acknowledge that the observational data has limitations, which is why we use it as descriptive evidence to motivate our survey experiment. Unfortunately, we have not been able to identify natural experiments or other sources of exogenous variation in the observational data.

Misuse of Willingness to pay (WTP). The authors describe the results throughout the paper as an estimation of "willingness to pay" and the shifts in WTP. When one talks about WTP in the discrete choice model (DCM) estimation literature, one is referring to money-metric utility – something that is calculated by dividing the DCM estimated coefficient for each attribute by the price coefficient. In the authors' choice experiment, price or money is nowhere presented within each option (in other words, there cannot be any variation in the generated data to estimate price coefficient(s) or utility for money.). The authors acknowledge this lack of data/variation on p. 10 2nd para. to a certain extent. It is therefore reckless to claim that WTP has been estimated or to state (e.g. in the abstract) that "[t]his willingness to pay for female mentors declines to zero...". In truth, while one can infer outcomes expressed in terms of WTP but this would require additional assumptions (this is most clear under the logit specification – it is not appropriate to compare coefficient from logit estimation of DCMs with different model specifications [one of them including ratings, and the other not]). The authors should replace result statements mentioning WTP with choice probability shifts (which is indeed estimated). While the effort to make the results more interpretable is laudable, the shorthand of WTP might mislead the readers regarding what was actually estimated/estimable.

Thank you for this discussion of willingness to pay metrics. We endeavor to be as transparent as possible with our usage of WTP. To that end, we include the following sentences in section 3.3, the econometric framework:

> We use the estimates of students' preferences for mentor attributes to compute students' willingness to pay to access a mentor of the same gender. Willingness to pay metrics are traditionally denominated in monetary terms, for instance, the willingness to pay in hourly wages for a job with a higher fraction of coworkers who are female. Informal interactions for the purpose of information gathering seldom involve a monetary exchange. For this reason,we use whether the student is willing to trade off a mentor with their preferred occupation in order to access a same-gender mentor, by computing the ratio of the two coefficients.

We reiterate our usage of WTP in section 4.1, where we discuss the results:

> While WTP metrics are traditionally denominated in monetary terms, informal interactions for the purpose of information gathering seldom charge a fee but often involve trade-offs. We calculate WTP for a female mentor as the ratio of the coefficients on female mentor and preferred occupation.

Finally, we clarify that we are not interpreting coefficients from "logit estimation of DCMs with different model specifications." Our primary specification is OLS, and the coefficients summarize the difference in the proportion of time a profile with given characteristics is chosen, conditional on other characteristics. Of course, this proportion depends on the other information available, which is why we take ratios of coefficients to obtain the WTP metric discussed above. We do provide estimates from a logit specification in the appendix, and interpret the coefficients at the mean in that case, but still relative to each other (the WTP metric).

Results for male students. Why is the preference for gender concordance of male students so much weaker than that of female students? Can you dig a bit more? Does it apply for all male students or only non-minorities or only high SES?

Thank you for this suggestion to dig a bit more into the null gender concordance results for male students. When we look at first-generation male students, we find that there is a slightly greater preference for male mentors, but this effect is not statistically distinguishable from zero. Due to space constraints, we have not included these results in the main text, but would be happy to revisit this if you like.

<u>Male student preferences for mentor attributes: by first-gen status</u>

Mentor is female	-0.012
	(0.011)
X Student first-gen	-0.019
	(0.021)
Mentor has preferred occ	0.332***
	(0.018)
X Student first-gen	-0.037
	(0.042)
Mentor graduation year	0.005***
	(0.002)
X Student first-gen	-0.001
	(0.004)
Availability (in 10 min increments)	0.039***
	(0.005)
X Student first-gen	-0.002
	(0.009)
Mentor first-gen	0.019
	(0.012)
X Student first-gen	0.075**
	(0.030)
Student first-gen	-0.009
	(0.018)
p-value WTP first gen = WTP non-first-gen:	0.338
Observations	4620
Number of students	154

EXPOSITION

Mathematical notation can be improved. For example, on page 9, Section 3.3 first para. Third line "x" should be a vector. Also, on the same section, it would be helpful to define variables in the equations (e.g. what is gamma?). Finally, some equations are labelled and some not (e.g. eq (2) is labelled but not the one right above it).Make clear modelling assumptions. For example, in the first equation in the bottom of page 9: what distribution does epsilon follow?

Thank you for these suggestions. In section 3.3, we revised 'x' to be a vector. All equations are now numbered.

Reviewer 2

Thank you for the opportunity to read your interesting work! You explore an important research question, using a combination of observational and experimental methods. I found your paper to be clear, concise, and well-focused; the analysis is well-chosen and well-explained. Below, I offer my comments and suggestions.

1. I think the paper could be better connected to other work on discrimination, in-group preferences, role models, and homophily. While some papers in these literatures are cited, I wanted to understand better how the questions and findings in the current paper are advancing this literature (beyond applying these questions to the mentorship context). Making these connections and synthesizing work on the channels through which discrimination in these types of contexts seems to operate would increase the contribution of the paper in my view.

Thank you for these suggestions. Since this is a short paper, we have limited space for extensive discussion of the prior literature on discrimination. Below we detail our response to this comment, but have not yet added this to the paper. We would be happy to do so if it would be helpful.

We see our paper as contributing to the literature that uses information to distinguish between taste-based and statistical discrimination. In the context of hiring, correspondence studies vary the information available to employers in order to disentangle these two types of discrimination. For example, Oreopoulos (2011) uses a correspondence study to estimate discrimination against immigrants in the Canadian labor market and finds that providing information on the language skills, educational background, and firm experience does not attenuate differences in callback rates, suggesting a strong role for taste-based discrmination. Similarly, Hedegaard and Tyran (2019) find that Danish students discriminate against immigrants even when information on productivity is provided. In contrast, Agan and Starr (2017) find that removing information about criminal convictions from men's resume's increases racial discrimination in the U.S. These information issues are the root of the Heckman and Siegelman (1993) critique. Neumark and Rich (2019) find a role for these biases in estimates of average discrimination in audit/correspondence studies, but the change in the implied degree of discrimination is sometimes positive and sometimes negative.

Our paper similarly estimates students' propensity to choose a same-gender mentor with and without additional information on mentor quality. In addition, we disentangle taste-based from statistical discrimination using a key prediction from Becker's model: in the presence of taste-based discrimination, individuals should be willing to pay to access a mentor of the same gender. Specifically, we compute how much students are willing to trade off in order to access a mentor of the same gender, with and without additional information on mentor quality.

Unlike hiring, we investigate a setting in which discrimination–in the form of same-gender preference–is actually encouraged, but as of yet, there is little evidence on its roots and the tradeoffs involved. It is often presumed that there is an additional benefit associated with matching on shared traits. In the healthcare context with doctor-patient matching and in the employee recruitment context through interviewer-interviewee matching, it may also be important to know what drives observed preferences for same-gender matches.

There are two potential downsides to encouraging matching on shared traits without scrutinizing the source of the preference. Using the example of the healthcare setting, if female doctors are scarce, then implementing same-gender matching may result in longer wait times for female patients. In addition, using the example of mentorship, if female mentors are scarce, then initiatives focused on female mentees can result in an undue burden on female mentors. Ultimately, understanding the source of observed homophily can help inform the design of these important initiatives. These issues are quite distinct from the issues that arise with discrimination in the hiring context.

2. I think the biggest limitation of the paper is your inability to go beyond outreach/preferences to look at other outcomes (conversation quality or quantity, impact on major choice or job search behavior, satisfaction of mentee/mentor, etc.). Linking your results to these more downstream outcomes would take the paper from good to great, and could help us to really understand the extent to which these preferences reflect “mistakes” or misperceptions and the costs of these mistakes in terms of educational and labor market outcomes.

Thank you for these comments. We agree that one limitation of the experiment is that we cannot go beyond students’ preferences and WTP. In Section 2, we provide descriptive evidence from the online mentoring platform that female students’ disproportionately outreach to female mentors may come at a cost: female students receive lower response rates from female mentors relative to male mentors. Conditional on response, female students receive responses of similar length.

3. I wanted more discussion and interpretation of the results. In particular, what seems to explain the patterns that are found, among both female and male students? What is the particular misperception that these students seem to hold, and is the misperception different among female students than male students, or do male students just have offsetting biases that lead them to not prefer female mentors absent quality ratings? Is it the case that only female students have a gender preference, or is it just that this is more socially acceptable for female students to declare compared to male students? Section 4.3 touches on these questions but I wanted to see them explored in a more thorough, rigorous way. I expected this type of analysis and exploration to be a much bigger piece of the paper. Detailed, incentivized, quantitative survey evidence on beliefs (beyond the questions asked as follow-ups in the experiment) might really help to disentangle different stories for what seems to be going on.

Thank you for these suggestions. We provide discussion and interpretation of the results in section 4.3 After the preference elicitation, we ask students (1) whether they have a gender preference and (2) if so, why? Half of female students state that they have a preference for female mentors, while just ten percent of male students state they prefer male mentors. The declaration of mentor gender preference is strongly predictive of their revealed preferences from the preference elicitation.

In the paper we write:

> Among the female students who stated that they valued a female mentor, 85% reported that it is because female mentors are friendlier/easier to talk to and 53% reported that it is because female mentors are better at giving personalized advice. In contrast, only 9% reported that female mentors are more knowledgeable about job opportunities. Female students’ perceptions that male and female mentors differ, on average, is consistent with statistical discrimination based on (accurate or inaccurate) beliefs. Student’s beliefs could arise due to stereotypes that are partially accurate but exaggerated (Bordalo et al., 2016; Eyal and Epley, 2017).

We also have added to the paper on pg. 4 , “Eyal and Epley (2017) find that though women are somewhat more socially sensitive than men, people believe that the average difference is larger than it actually is.”

We hope that this evidence is compelling but would be happy to discuss further other steps we could take to provide additional evidence.

4. The paper does not currently discuss a role for image concerns or social desirability bias in contributing to the results. I wonder about this particularly in the experimental context, where students are making explicit choices between pairs of profiles where gender information is somewhat salient. Students, particularly male students, may be worried about revealing a preference for male mentors. One small piece of evidence that might be suggestive of this type of behavior is your finding that male students' preference for male mentors is stronger in the observational data than in the experimental data. While this may be due to the unobserved variables in the observational data as you suggest, it could also be a result of image concerns playing more of a role in the experimental context. These image concerns could also play a role in only 10% of male students claiming the gender matters in their choice of mentor in the survey part.

While it is highly unlikely that image concerns are the only factor in producing the experimental results, it still seems worthwhile to discuss how these types of concerns could contaminate or bias your results. You might also explore whether there is any empirical evidence that image concerns are not a major concern.

Thank you for bringing up social desirability bias. We added the following sentences to address this possibility on pg. 8:

> Since the survey was advertised via email lists and accessed via a Qualtrics link, students were able to take the survey completely on their own without the supervision of a researcher. We think this setting guards against social desirability bias and social image concerns.

While we cannot completely rule out social image concerns, if these concerns were operative, it is likely they would affect more than one group of students. Instead, we find that first-generation college students are quite comfortable expressing a preference for first-generation mentors, and likewise with female students and female mentors.

5. While you mention the idea of inaccurate statistical discrimination, it might be worth digging deeper into stereotypes as a particularly relevant form of inaccurate statistical discrimination. It could be that female students misperceive or exaggerate the differences between female and male mentors, in line with the model of stereotyping presented in Bordalo et al (2016) and as applied in papers like Arnold, Dobbie, and Yang (2018) on bail decisions or the closely-related Chan (2021) paper on discrimination among healthcare shoppers. In line with trying to dig deeper into mechanisms, it may be worthwhile to explore these types of channels.

Thank you for these suggestions. While we do not have direct evidence on whether statistical discrimination is accurate or inaccurate in our setting, we discuss the possibility of inaccurate discrimination on pg. 12.

> Among the female students who stated that they valued a female mentor, 85% reported that it is because female mentors are friendlier/easier to talk to and and 53% reported that is it because female mentors are better at giving personalized advice. In contrast, only 9% reported that female mentors are more knowledgeable about job opportunities. Female students' perceptions that male and female mentors differ, on average, is consistent with statistical discrimination based on (accurate or inaccurate) beliefs. Student's beliefs could arise due to stereotypes that are partially accurate but exaggerated (Bordalo et al., 2016; Eyal and Epley, 2017).

6. In Table 4 you explore whether the gender of the person who provides the ratings matters for mitigating the preference for female mentors. I might be missing something, but I wonder whether it would make sense to try to look at this separately for male mentors. In particular, the idea that female students would value ratings from female mentees to rule out concerns about sexual harassment would seem to be primarily an issue for male mentors. Asking whether they value female mentees' ratings differently in general (for both male and female mentors) may mask a stronger effect for just male mentors?

We provide the requested analysis below, in which female students are deciding between male mentors. If anything, when female mentees provide the rating, female students are less likely to choose a mentor, but this coefficient is insignificant. There is no indication that female students place more weight on ratings from female mentees. This evidence further confirms that the gender of the mentee does not matter much for mentor choice.

Female student preferences for mentor attributes when both mentors are male

Mentor has preferred occ	0.125***
	(0.017)
Mentor graduation year	0.002
	(0.001)
Availability (in 10 min increments)	0.002
	(0.004)
Mentor first-gen	0.031**
	(0.014)
Knowledgeable about job opportunities	0.092***
	(0.005)
Easy to talk to/friendly	0.064***
	(0.006)
Gave personalized advice	0.068***
	(0.006)
Mentee is female×Knowledgeable about job opportunities	0.006
	(0.008)
Mentee is female×Easy to talk to/friendly	0.005
	(0.008)
Mentee is female×Gave personalized advice	-0.003
	(0.008)
Mentee is female	-0.033
	(0.041)
Observations	1852
Number of students	257

Reviewer 3

This paper investigates homophily in mentoring networks using two different data sets: (1) an administrative data set from an online college mentoring platform that connects students with alumni, and (2) data from a hypothetical choice experiment. Using the first data set, the authors show that there is homophily in networking behavior by gender. Female students initiate conversations with female mentors at higher rates than male students. Interestingly, this behavior is costly for female students because female mentors are less likely to reply to female students than male mentors. The authors then use the hypothetical choice experiment to investigate why homophily exists in the first data set. They recruit 834 students from UCLA and ask them to make 30 binary choices between hypothetical mentors. While choices are not directly incentivized, students know that the study investigates how and from whom individuals seek advice and are told that the authors will use their responses to give them personalized suggestions on how to find mentors.

The hypothetical choice experiment consist of two treatments and uses a between-subject design. Participants make decisions either in a control (no ratings) condition, or in a ratings condition. In both treatments, students see the first name, occupation, location, class year, and availability of the hypothetical mentor as well as whether they are a first generation college student. In the ratings treatment, students also see three ratings the hypothetical mentor has received from a hypothetical mentee. These ratings include the hypothetical mentee's first name, class year, and ratings along three dimensions (on a 5-star scale): how knowledgeable the mentor is about job opportunities, how easy to talk to or friendly the mentor is, and the extent to which the mentor gave them personalized advice. Hypothetical mentor and mentee characteristics as well as ratings are randomized by the researchers.

Results from the hypothetical choice experiment show that female students are more likely to choose female over male mentors in the control treatment. Male students also have a slight preference for mentors of their own gender, but the coefficient is much smaller and only marginally significant. Importantly, the homophily results documented in the control treatment disappear in the ratings treatment. The authors interpret this result as suggesting that homophily is driven by information problems that are alleviated with the provision of quality ratings.

I found the paper interesting and enjoyed reading it. I believe it contributes to the literature. My main comment is that the methodology used by the authors has some limitations which the authors could acknowledge in the paper. I make some comments that I hope help the authors improve the paper below.

Main comments:

1) While the authors investigate a very interesting question with the hypothetical choice experiment, whether mentors value mentors with shared traits, the extent to which results are generalizable and reflect preferences for real mentors is an open question. The authors could be upfront about this methodological limitation and at the very least discuss this as a direction for future research. While similar instruments and methodological approaches have been used by important papers in the literature, many of these other papers

typically go to great lengths to show that results from the hypothetical choice experiment are externally valid and map onto the outcomes we care about in the field (e.g. Mas and Pallais 2017, Wiswall and Zafar 2018). Furthermore, some of the previous papers in the literature provide more direct incentives for respondents to reveal their true preferences in hypothetical choice experiments than what the authors provide in this study. For example, the recruiters rating hypothetical resumes in Kessler et al (2019) get access to 10 resumes of job seekers who match the preferences they report. In this study, the only incentive students get from the hypothetical choices they make is suggestions on how to find mentors, not the time or contact information of mentors based on the choices they make. This doesn't invalidate the study but is a limitation that the authors could discuss in the paper.

> Thanks for bringing up these important methodological issues. We see our design as closely related to Kessler et al. (2019). We similarly tell students that the choices they make in the survey will be used to provide personalized information on how to find mentors.
>
> We note also that Danz, Vesterlund, and Wilson (AER 2022) concludes that a low-information incentive similar to ours (in their case, simply informing subjects that guessing their true beliefs maximizes their earnings) generates *more* truthful reporting than state-of-the-art incentivized belief elicitations, *especially* compared to when the incentives for truth telling in these belief elicitations are made extremely explicit. In light of this, and given that there is little motivation not to tell the truth in our setting, we believe that our design likely maximized truth telling relative to a mechanism that would be more difficult and complex to explain.

2) I could not help but wonder whether experimenter demand effects / social desirability bias interacts with treatment in the hypothetical choice experiment and if any bounds could be imposed on this demand effect if present (see De Quidt et al 2018). For example, maybe it becomes socially unacceptable to display a preference for a given gender when quality ratings are provided and this eliminates the homophily result, not because the information problem is alleviated but rather because displaying such a preference is harder to justify in a setting without direct incentives to report true preferences. This concern is probably unwarranted, nevertheless, the authors should probably discuss it. For example, by using results from the analysis of homophily by first generation college student status. Table A6 shows that the ratings treatment reduces the preference of first-generation mentors by first generation college students by approximately 9 percentage points but does not eliminate it. Since the difference between the preference for mentors with a shared gender is similar among female students in columns 1 and 3 of Table 3, this suggests that the magnitude of the information effect is around 9 percentage points on average, rather than experimenter demand effects being this large since social desirability bias is presumably smaller or even absent for first generation college student status.

> Thank you for raising this point that perhaps a reason that student preferences decline in the ratings condition is that homophily is seen as less socially acceptable in that setting. We have expanded our discussion of why social desirability is an unlikely mechanism in the paper. The first-generation college student preferences are actually supportive that this is not likely a mechanism. It is important not to interpret levels of coefficients in these regressions, but the ratio of coefficients, which reflects the WTP for one characteristic in terms of another. When we study the WTP for first-generation college mentors among first-generation college students, we see that this if anything *rises* in the information condition (the level of the coefficient falls, but this reflects that fact that valuable information has been added in the ratings condition. Relative to other attributes, first-generation status remains important to first-generation students). We have added the following description to the paper (p. 12) which acknowledges the important issue raised and which we hope is helpful in thinking about the magnitude of this problem:

> "We note that the attenuation of willingness to pay in the 'ratings' condition is not mechanically driven by the fact that profiles with ratings are longer, have more mentor attributes, or are in some other way distracting from the original attributes. When we analyze a pre-registered secondary outcome–the willingness to pay of first-generation college students for first-generation mentors–we find that including ratings *does not* attenuate their willingness to pay (Appendix Table A6)."

3) Based on comment 1 made above, I found the conclusions made from the analysis presented in Table 4 somewhat far-fetched. The authors are assuming that student choices reflect the value they put on ratings provided by hypothetical male and female mentees/raters. While the students may have the incentive to reveal their true preferences over hypothetical mentors in the experiment, if they think doing so will help them get better personalized information on how to find mentors, they may not be taking into account the rating a fictitious rater provides about a hypothetical mentor because it is not perceived to affect the information they get.

Thank you for this discussion. As a reminder, our exact phrasing to students was, as on page 8:

> "We will use your responses in this section to give you personalized suggestions on how to find mentors. If you decide to receive these suggestions, you will receive these suggestions via email (which you will enter at the end of the survey). We will not contact any mentors on your behalf, we will only provide you with recommendations consistent with the choices you make in the next portion of this questionnaire."

We understand the source of this comment from a reader who saw exactly what type of guidance was (ex-post) generated for students. However, at the time students took the survey, it was fairly ambiguous exactly what kind of guidance they would receive concerning how to find mentors. If quality was important to students, they would not want to forgo the opportunity to be guided to quality mentors.

Other comments:

4) The authors seem to argue several times throughout the paper that matching mentors to mentees based on quality ratings rather than gender could be welfare enhancing. While this could work well within organizations, where mentoring may be considered part of someone's job description and thus ratings are expected/acceptable, they may not work well in settings where there is selection of mentors into mentoring roles. Couldn't ratings reduce the likelihood that mentors volunteer their time to mentor others? If so, what is the overall welfare effect we can expect from adopting quality ratings?

This is an excellent point. We have discussed considerations for when providing additional information is optimal rather than relying on coarse characteristics-based matching primarily in terms of the cost of truncating available mentors based on quality in section 5. We write:

> "Assuming that truncating based on quality is costly, and perhaps increasingly costly as the quality truncation threshold increases, programs may be better off restricting matches to shared traits. If obtaining information on quality is straightforward, for example, through the use of existing surveys of mentee experiences, then the optimal policy would screen mentors on quality."

To address this concern, we have added footnote 22 on page 14:

> "We think of this cost abstractly. For example, in settings where there is selection of mentors into mentoring roles, the cost of truncating on quality may be that fewer mentors volunteer and the overall pool is worse or is scarce."

5) Tables 3 and 4 could include a table note that describes how the WTP measure is calculated (as the ratio of the female mentor to preferred occupation coefficient). I had to go back and look for this in the text. The same comment probably applies to any appendix table that has a WTP measure reported.

Thank you for these suggestions. We now include the sentence "Willingness to pay is calculated as the ratio of the coefficients on female mentor and preferred occupation." in all of our tables that have this calculation.

Declaration Statement

"Does Information Affect Homophily?"

Yana Gallen and Melanie Wasserman

September 2022

Declarations of interest: none

Manuscript with Changes Highlighted

Click here to access/download
Companion Files
GallenWasserman-Manuscript-ChangesHighlighted.pdf

Original Editor Decision Letter

Click here to access/download
Companion Files
RESTAT MS27574-1 Decision letter.pdf

Original R1 Report

Click here to access/download
Companion Files
RESTAT MS27574-1 R1s report.pdf

Original R2 Report

Click here to access/download
Companion Files
RESTAT MS27574-1 R2s report.txt

Original R3 Report

Click here to access/download

Companion Files

RESTAT MS27574-1 R3s report.txt