Information Asymmetry in Job Search*

Michelle Jiang† Kai Zen‡

March 20, 2025

Click here for most recent version.

Abstract

Standard models of rational job search assume agents know the distribution of offered wages when deciding which jobs to accept. We test if incorrect beliefs about wages affect real-world job search behavior in a field experiment with 1100 senior-year undergraduate students in the graduating Class of 2023 at the University of California, Berkeley. Partnering with the Career Center, we present personalized information graphics on school-and-major-specific salary distributions to students in the treatment group. We first document novel evidence that even prior to labor market entry, errors exist in wage beliefs – some students overestimate the available distribution, while others underestimate the available distribution. Post-treatment, we find that students treated with correct information update their beliefs towards the truth, and this is reflected in changes in reservation wages. At the end of the school year, we find that in comparison to the control group, students who increased their reservation wage after treatment had higher total and base salaries conditional on employment, a result significant at the 5% level. However, these same students had a lower, but imprecisely estimated likelihood of being employed by June post-graduation. An opposite but symmetric effect occurred for students who decreased their reservation wage. Our results are consistent with job search models where workers with more optimistic expectations wait longer to accept a job, but accept higher wages. We compare our experimental estimates to simulated moments from the model and find that the mean experimental effect is close to the model in magnitude under reasonable parameters. Our paper suggests an economically important role for errors in beliefs about labor market conditions and shows the effectiveness of a light-touch information intervention on employment and earnings for first-time job seekers.

*We thank Sandra Black, Alessandra Casella, David Card, and Suresh Naidu for indispensable guidance and advice. We thank Hunt Allcott, Michael Best, Marianne Bitler, Laura Caron, Pierre-André Chiappori, Mark Dean, Léa Dousset, Tarikua Erda, Hannah Farkas, Émilien Gouin-Bonenfant, Jeffrey Guo, Abhi Gupta, Judd Kessler, Oliver Kim, Elizabeth Linos, Victoria Mooers, Kate Musen, Will Nober, Shin Oblander, Cristian Pop-Eleches, Roman Rivera, Evan Sadler, Susie Scanlan, Camilla Schneier, Miguel Urquiola, Homa Zarghamee, and seminar participants at Columbia University and UC Berkeley for detailed comments and feedback. We thank the UC Berkeley Career Center, in particular Elizabeth Davies, Santina Pitcher, and Melissa Han, for their generous time and help, and acknowledge with gratitude the financial support of the National Science Foundation (ID 2117566) and Columbia's Program for Economic Research. The project was completed under under IRB protocols CPHS 2021-10-14766 (UC Berkeley) and IRB-AAAT5934 (Columbia). All errors are our own.

†University of Massachusetts Boston, michelle.jiang@umb.edu

‡kaizen93@gmail.com

1 Introduction

How do individuals make choices about jobs? Search models posit that individuals with full information calculate a reservation wage, and from there, make optimal decisions when searching.[1] However, recent work suggests that individuals may not have correct information about the labor markets they face.[2] How does this incorrect information affect the outcomes of job search, and can the provision of information change search behavior? Ultimately, how does this affect our standard models of job search behavior?

In this paper, we examine job search behavior in the presence of incorrect information. To do so, we conduct a field experiment where we give individuals in the treatment group information about the wage distribution they face, and from there, follow their search behavior to see how it adapts to this new information. We ultimately assess workers' actual employment and earnings outcomes.

The context we use is graduating students from UC Berkeley, a top flagship university. We run a field experiment with 1100 senior-year undergraduate students in the graduating Class of 2023. Using exclusive UC Berkeley data under a partnership with the Career Center, we construct information graphics on school and major-specific salary distributions from the past three years, which are personalized and presented to students in the treatment group.[3] We are able to directly document students' beliefs about available wage distributions prior to the experiment, and then show how the treatment changes those beliefs and students' intended search wages and search behavior before the annual September Career Fair, where students historically launch their job search. At the end of the school year, we work with the Career Center to follow students' outcomes, and we observe their actual post-school employment and earnings outcomes. Our paper suggests an economically important role for subjective beliefs about labor market conditions and shows the effectiveness of a light-touch information intervention on employment and earnings for job seekers.

We first document a novel fact that prior to labor market entry, college students have imprecise beliefs about available wage distributions for their majors. We find that students tend to underestimate the 25th percentile salary and overestimate the 75th percentile salary, with 90% of students overestimating the interquartile range of the distribution. Notably, we

[1] McCall 1970, Pissarides 1985, Mortensen 1986, Mortensen and Pissarides 1994, Burdett and Mortensen 1998

[2] Cortes et al. 2023, Jager et al. 2023, Roussille 2023

[3] UC Berkeley undergraduate majors are split into several "colleges," such as the College of Engineering, the Rausser College of Natural Resources, and the College of Letters and Science. Students were asked to specify both their college and major. Information graphics were specific to both college and major.

are measuring students' beliefs about the whole within-major distribution, rather than their beliefs about their personal ability or personally attainable salaries - suggesting a difference driven by information about the whole distribution itself, rather than beliefs about personal ability. Across students, there is large variation both in the magnitude and direction of the error: students underestimate the 25th percentile and median by as much as -98% and overestimate by as much as +200%, while underestimating the 75th percentile by as much as -92% and overestimating by as much as +400%.

Second, we show that post-treatment, changing beliefs on the external wage distribution for the treatment group changes students' search behavior and outcomes. We characterize search behavior via workers' reservation wage: the minimum wage at which a worker would accept an offer over continuing search. Since each individual's reservation wage incorporates both information about the labor market and worker's individual preferences, the reservation wage is a sufficient statistic for workers' intended search behavior. We find that students given the information treatment adjust their reservation wages, with larger initial errors leading to larger adjustments. This result remains robust to inclusions of prior behavioral explanations such as confidence and risk aversion, which affect reservation wages by an order of magnitude lower than the information treatment. Next, we use a two-stage least squares strategy to show the impact of adjusting reservation wages via the information channel on graduating students' wages and employment probability. We find that each 1% increase in the reservation wage via the information channel leads to a 1.28% increase in total compensation conditional on accepting a job, a result statistically significant at the 5% level. We show that this result is due to both an increase in base salary and an increase in the likelihood of receiving a bonus salary. On the other hand, we find a negative but statistically non-significant effect on employment: each 1% increase in the reservation wage via the information channel leads to a 0.56% lower chance of being employed. Overall, our results suggest a stronger role for information mechanisms than previously believed.

Last, we study how our results impact standard models of job search behavior. We build a model of labor search with errors in beliefs and derive simulated results on how correcting errors in beliefs should affect the wage distribution. We find that for beliefs about the mean of the wage distribution, our experimental data fits within the interval of simulated results from the model with errors in beliefs. In addition, the model with errors in beliefs predicts higher variance in wages, which fits both our experimental data and accounts for a discrepancy between standard models and empirical data in prior literature.[4] We find that

[4]Hornstein, Krusell, Violante (2011) remark on how under standard job search models, wage dispersion is higher in data than predicted.

our information intervention compresses wages for the treatment group: post-treatment, the mean-min wage ratio, which measures wage dispersion, shrinks by 0.13 for students in treatment compared to the control group, a difference significant at the 5% level. We also find that across the sample, the distributional variance of the reservation wage is lower in treatment than control. However, this second difference is not statistically significant, potentially reflecting limited power. Ultimately, our paper suggests a theoretically important role for errors in beliefs about labor market conditions, and our results show the effectiveness of a light-touch information intervention on changing these errors.

Examining students entering the labor market provides information about job search behavior at a crucial time in an individual's work life. Research has shown that your first job can affect not only your lifetime earnings, but also your health outcomes, marriage outcomes, and middle-aged mortality.[5] However, while we know that the state of the economy matters, we know very little about the role of information in determining an individual's job.

Our paper contributes to three areas of research. First, our paper relates to a nascent but growing strand of the behavioral and labor literature on errors in beliefs. Jager et al. (2023) find that once workers have a job, workers anchor their beliefs to current wages, but will correct their beliefs in response to information about wages of similar workers. Roussille (2023) shows that on engineering job platform Hired.com, pre-filling job hunters' ask salary answer box with the median bid helped close the gender gap in asking wages. Alfonsi et al. (2023) show that providing mentorship to students in Ghana creates more realistic beliefs about available wages in the labor market. Our paper adds to this literature, showing that information about salaries plays a large role in students' first post-graduate job and salary outcome. By using an experiment to measure workers' pre- and post- treatment beliefs, we can causally isolate the impact of information conditional on prior beliefs. Under our results, workers overestimating versus underestimating would lead to opposite impacts. This may help contextualize results from prior literature, some of which finds workers decrease asking wages in the presence of more information and some of which finds workers increase asking wages in the presence of more information.[6] This additionally has implications for the pay transparency literature such as Baker et al. (2021) and Cullen and Pakzad-Hurson (2023), suggesting that differing results may be due to different prior beliefs.

[5]Kahn 2010, Oreopoulos et al. 2012, Schwandt and von Wachter 2019, and von Wachter 2020 provide discussion of the effect of first jobs on lifetime earnings, and Cutler et al. 2015 and Schwandt and von Wachter 2020 discuss the effect of first jobs on health, marriage, and mortality.

[6]As an example, Alfonsi et al. (2023), in which workers lower their asking wage once in a situation with more information, differs from Roussille (2023), in which workers increase their asking wage. Our paper suggests that these could be contrasting situations of prior beliefs where one group tends to overestimate, while the other tends to underestimate.

Second, we relate to a more general literature on behavioral inefficiencies in job search. Prior literature suggests a wide variety of alternative mechanisms for inefficiency in labor search behavior: differences in risk aversion, in confidence, in the valuation of non-wage amenities, in liquidity constraints, in the valuation of time, and more. Blau and Kahn (2017) provide an overall discussion. Alfonsi et al. (2023), Cortes et al. (2023), Reuben et al. (2017) , and Shurchkov and Eckel (2018) discuss the roles of risk aversion and confidence. Maestas et al. 2018 discuss the valuation on non-wage amenities. Gerards and Welters (2020) discuss liquidity constraints, and Smitizsky et al. (2021) discuss time valuation. The paper closest to ours is Cortes, Pan, Pilossoph, Reuben, and Zafar (2023), which surveys undergraduates from Boston University's Questrom School of Business. They find that women accept jobs earlier than men, accounting for 19-35% of the gender wage gap, due to women's higher risk aversion and men's higher confidence. A key strength of our research design is that the information intervention not only directly tests the information asymmetry hypothesis, but can account for other behavioral mechanisms which influence labor search behavior. Because students are surveyed both prior to the information intervention and after the information intervention, we can measure and control for standard measures of risk aversion in our results.

Finally, our paper relates to evidence from the education literature that information interventions can be powerful for changing school and major choice. Hoxby and Turner (2015) show that when high-achieving low-income students received information about colleges' curricula, prices, and employment outcomes, they were more likely to both apply and enroll. Conlon (2021) show that providing early-stage college students salary information can affect major and class choices, and Conlon and Patel (2023) show that reducing students' stereotypes about fields of study can also change students' classes and majors. We apply a similar intervention to the area of labor search.

The paper proceeds as follows. In Section 2, we describe the motivating theoretical foundation for how information asymmetry affects reservation wages. In Section 3, we establish the experimental design, data, and procedures. In Section 4, we discuss descriptive results from pre-treatment empirical findings and establish that workers have incorrect beliefs about externally available wage distributions. In Section 5, we describe our main econometric methodology and show how correcting incorrect beliefs changes students' personal wage expectations and subsequently their employment and earnings outcomes. In Section 6, we compare numerical predictions from theory to our empirical data results. In Section 7, we discuss heterogeneity across overestimating versus underestimating students and across student demographics. In Section 8, we conclude and provide next steps.

2 Illustrative Model

In Section 2, we establish a partial equilibrium job search model with subjective beliefs to establish how errors in beliefs can affect labor search behavior. We will show how the reservation wage is a "sufficient statistic" for an individual's job search behavior, which motivates our two stage least squares empirical strategy in Section 5. Our model builds upon both the original job search model from McCall (1970), as well as previous search models with subjective beliefs, specifically Cox and Oaxaca (2000), Cortes et al. (2023), and Alfonsi et al. (2023).

2.1 Setup

A worker seeks to maximize utility via rejecting or accepting a job offer. In each period t, a worker receives an employment offer with probability λ. If given an offer, the worker learns of the offer's attached wage w. The worker then chooses between (1) accepting the job and receiving wage w for the rest of time, or (2) rejecting the job, receiving a value of leisure b, and continuing the search next period.

The worker maximizes expected utility:

$$\max_{\{x_t\}} \sum_{t=0}^{\infty} \beta^t u(x_t), \beta \in (0,1)$$

where:

$$x_t = \begin{cases} w, & \text{accept job, receive wage w in all future time periods} \\ b, & \text{reject job, receive unemployment b and can search again} \end{cases}$$

The worker has subjective beliefs formed from a noisy signals of the truth. The worker does not observe the true distribution of wages, in which w is drawn from distribution $F(\log(w^*)) \sim N(\mu_w, \sigma_w)$. Instead, the worker observes wages as $\log(w) = \log(w^*) + \epsilon_i$, where ϵ_i is a perception error or bias drawn from a normal distribution with individual-specific mean $\mu_{\epsilon i}$ and variance $\sigma^2_{\epsilon i}$. Thus, the subjective offer distribution for worker i is:

$$log(w) \sim N(\mu_w + \mu_{\epsilon i}, \sigma^2_w + \sigma^2_{\epsilon i})$$

The value of employment for worker i is thus:

$$W(w|\mu_i, \sigma_i) = \frac{u(w)}{1-\beta}$$

And the value of unemployment is:

$$U(\mu_i, \sigma_i) = u(b) + \beta\lambda \int_w \max\{W_{t+1}(w|\mu_i, \sigma_i), U_{t+1}(\mu_i, \sigma_i)\} dF(w|\mu_i, \sigma_i) + \beta(1-\lambda)U_{t+1}(\mu_i, \sigma_i)$$

2.2 The Reservation Wage

We define the reservation wage w_R as the wage in which a worker is indifferent between accepting and rejecting the wage, where marginal benefit equals marginal cost:

$$W(w|\mu_i, \sigma_i) = U(\mu_i, \sigma_i)$$

Using our previous equations, we then know:

$$\frac{u(w_R)}{1-\beta} = u(b) + \beta\lambda \int_w \max\{W_{t+1}(w|\mu_i, \sigma_i), U_{t+1}(\mu_i, \sigma_i)\} dF(w|\mu_i, \sigma_i) + \beta(1-\lambda)U_{t+1}(\mu_i, \sigma_i)$$

For wages $w \leq w_R$, $U_{t+1}(\mu_i, \sigma_i) \geq W_{t+1}(w|\mu_i, \sigma_i)$ and the worker rejects the job. For $w \geq w_R$, $U_{t+1}(\mu_i, \sigma_i) \leq W_{t+1}(w|\mu_i, \sigma_i)$ and the worker accepts the job. Then,

$$\frac{u(w_R)}{1-\beta} = u(b) + \beta \int_0^{w_R} U_{t+1}(\mu_i, \sigma_i) dF(w|\mu_i, \sigma_i) + \beta\lambda \int_{w_R}^{\infty} W_{t+1}(w|\mu_i, \sigma_i) dF(w|\mu_i, \sigma_i)$$

Solving through, we find:

$$\Big(1 - \beta(1-\lambda)\Big)u(w_R) - u(b) = \beta\Big(\lambda u(E[w|\mu_i, \sigma_i]) - u(b)\Big) + \beta\lambda \int_0^{w_R} u'(w) F(w|\mu_i, \sigma_i) dw \quad (1)$$

A full proof is available in Appendix A (See: Section 10.1).

The reservation wage serves as the solution to the worker's partial equilibrium problem: a worker will accept any wage above the reservation wage, and will reject any wage below. It is the minimum wage that a worker would accept. The reservation wage thus captures any effects in the worker's decision calculus from preferences via the shape of the worker's utility function or via the value of leisure (b). We thus consider the reservation wage a "sufficient statistic" for the worker's individual behavior in a partial equilibrium model.

2.3 Model Implications

The search model leads to several implications. All proofs are contained in Appendix A (See: Section 10.3).

Proposition (1A). *Holding all else constant, reservation wages* $w_R(\mu_i, \sigma_i)$ *are increasing in* $\mu_{\epsilon i}$*, that is* $\frac{\partial w_R}{\partial \mu_{\epsilon i}} > 0$.

Proposition (1B). *Holding all else constant, reservation wages* $w_R(\mu, \sigma)$ *are increasing in* $\sigma^2_{\epsilon i}$*, that is* $\frac{\partial w_R}{\partial \sigma^2_{\epsilon i}} > 0$.

The propositions motivate the following testable implications. From Proposition 1A, we expect workers who overestimate the overall external wage distributions to have higher reservation wages, and workers who underestimate the overall external wage distribution to have lower reservation wages. Consequently, workers whose overestimation is corrected should lower their reservation wages, and workers whose underestimation is corrected should increase their reservation wages. From Proposition 1B, we expect workers who estimate a higher variance of available external wages to have higher reservation wages. Thus, in our experiment, we will focus on two statistics of external wage distributions: (1) "centrality" and (2) "spread." Centrality indicates the amount a worker is underestimating/overestimating the center of the distribution, while spread indicates how much a worker is overestimating/underestimating the range of a distribution.

Proposition (2). *The dispersion of reservation wages across agents is higher in a model with subjective beliefs than in a model where every worker has perfect information.*

Proposition 2 shows the effect of the model on the overall distribution of wages. From this proposition, we should expect higher wage dispersion in a labor search setting with subjective beliefs than in a labor search setting with perfect information. We discuss this implication further in Section 5.3.

3 Experimental Design, Data, and Procedures

Section 3 describes our research methodology: using a randomized controlled trial, we identify a causal effect on how individuals' beliefs about the available wage distribution affect their reservation wage, intended labor search behavior, and actual post-graduation job outcomes.

3.1 Setting, Recruitment, and Pre-registration

Setting. We conduct our experiment at the University of California, Berkeley. Berkeley is a public land-grant research university located in northern California and is considered the flagship of the University of California system. As of fall 2022, UC Berkeley enrolled over 32,000 undergraduate students, of which around 8000 are graduating seniors. The undergraduates are 55% women. By race, the undergraduates are 40% Asian American, 24% Underrepresented Minority (Black/Hispanic/Native), and 28% White/decline to state.[7] UC Berkeley's large population, major presence in America's most populous state, and common enrollment breakdown are advantages for the external validity of our study. Importantly for internal validity, we have a strong partnership with the UC Berkeley Career Center, who emailed students on our behalf, lent credibility to our experimental intervention, and provided data for the information intervention.

Recruitment. In partnership with the UC Berkeley Career Center, in August 2022 we recruited rising seniors who planned to graduate in Spring or Summer 2023. The recruitment email was sent by the Career Center on August 15th, a full month prior to UC Berkeley's annual September Career Fair, where students historically launch their job search.[8] In the email, students were invited to participate in a Qualtrics web-based survey on job search, which began with a consent form describing the purpose of the study. Students were informed that the survey contents would be subject to randomization, but to avoid experimenter demand effects, we did not specify that the randomization pertained to wage information.[9] To prevent contamination of the randomized controlled trial, we did not prime students that information about the labor market matters; in both the recruitment email and consent form, students were told: "The purpose of the study is research: in particular, to further our understanding of factors influencing labor search decisions." A full copy of both the recruitment email and consent form is in Online Appendix C.

To incentivize students to respond, we provided compensation for completing the baseline survey. Students were compensated $10 for each part of the survey they completed and entered into a drawing for three $150 Amazon gift cards. Of the over 8000 eligible students

[7]The remaining 12% are international students, who are uncategorized. Because UC Berkeley enrollment statistics are available for the entering class rather than the graduating class, we used statistics for the entering fall class of 2019, who are the majority of our graduating class of 2023. More statistics can be found in https://opa.berkeley.edu/campus-data/uc-berkeley-quick-facts and https://pages.github.berkeley.edu/OPA/our-berkeley/student-enrollments.html.

[8]Email reminders about the original email were sent weekly until September 15, 2022.

[9]We have explicit approval for this omission in both IRB protocols CPHS 2021-10-14766 (UC Berkeley) and IRB-AAAT5934 (Columbia).

in the study population, we received 1110 completed responses to our baseline survey (12.7% response rate).[10]

Pre-registration and Pre-analysis plan. Our experiment was pre-registered in the AEA RCT Registry under Trial 9782 and can be found at `https://www.socialscienceregistry.org/trials/9782`. Our pre-analysis plan can be found both on the linked website and in Online Appendix C.

3.2 Randomization

We use a stratified randomization design at the individual level to split students into treatment and control groups. Stratified randomization, in which subjects are first split into strata (groups) before being randomized into treatment and control, allows us to guarantee balance among the stratified dimensions. We use three dimensions: gender (male and female/non-binary), race (white and non-white), and major (STEM and non-STEM), resulting in eight strata. We discuss the reasoning and methodology for our strata in our pre-analysis plan, available in Online Appendix C. To account for our randomization methodology, all strata dimensions are included as controls in any regression specifications.

Randomization was done within Qualtrics. Within the first few pages of the survey, students were asked to self-report their gender and race demographics on a multiple choice grid. They were additionally asked to choose their major from a drop-down list of 69 UC Berkeley majors. The 69 majors were pre-sorted into "STEM" and "non-STEM" on the logistical backend of Qualtrics, but this delineation was not shown to students. After students submitted their answers to demographic questions, the Qualtrics survey would automatically randomize the student into "treatment" and "control," counting equal amounts of "treatment" and "control" within each of the 8 strata. To prevent contamination between treatment and control, students could not return to the demographic page after their answers to demographics and major were submitted.

To verify that randomization resulted in comparable treatment and control groups, we check whether or not baseline demographics and characteristics are similar across treatment and control. **Table 1** compares summary statistics between treatment and control across baseline demographics (gender, race, USA citizenship, family income, parental education), educational background (major choice, transfer status, financial aid status, GPA), behavioral characteristics (risk and confidence), beliefs about externally available wage distributions,

[10]For comparison, Cortes et al. (2023) survey recent college alumni at Boston University's Questrom School of Business about job search behavior and have a response rate of 20%.

and beliefs about personal wages. For each variable X (e.g. a dummy coefficient for female), we regress X on a 0-1 indicator for treatment. We report the coefficient on treatment in the second-to-last column of the table and the correspondent p-value in brackets in the last column. Across 22 pre-treatment variables, we find that parent's highest education is unbalanced at the 1% level, and external beliefs about the 75th percentile earnings are unbalanced at the 10% level. This is consistent with Glennerster and Takavarasha (2013)'s guidelines on randomization, which note that on average, one out of 10 variables are unbalanced at the 90% confidence level, while one out of 20 variables are unbalanced at the 95% confidence level. However, to increase precision, we will control for parent's highest education as a balance variable in all regressions.[11]

3.3 Survey Phases

To both measure the immediate impact of the information intervention on intended job search behavior **and** the long-term impact of the information intervention on post-school job outcomes, our experiment occurred in two stages: Phase I and Phase II. In Phase I, students were recruited, completed a baseline survey (within which they were sorted into treatment or control), and were given an information intervention if sorted into treatment. Immediately post-treatment, students were asked about their intended labor search behavior. However, to understand the actual impact of the information intervention on students' outcomes, we had to follow-up with students *after* job search was completed. To do so, we introduced a Phase II survey, which was incorporated into UC Berkeley Career Center's annual First Destination Survey (FDS), which collects data on student job outcomes. Phase II thus ran in the later half of the year, from January 1 - July 15, 2023, in order to coincide with annual FDS timing. Our experimental timeline occurred as follows:[12]

August 15 - September 18, 2022	Phase I: • Baseline Survey and Information Treatment
January 1 - July 15, 2023	Phase II: • First Destination Survey: job outcomes

[11] Because external beliefs on the 75th percentile are a variable of interest, they will be in the main specification and thus are not a control.

[12] Pilot versions of the surveys - without the intervention - were run at Columbia University in July 2021 and again at UC Berkeley in 2022.

3.3.1 Phase I: Pre-Treatment Data Collection

Eligible[13] students received a Phase I survey invitation from the UC Berkeley Career Center on August 15, 2022 prior to any Career Center events - in particular, the early Fall career fair and on-campus interviews. The survey remained open from August 15th to September 18th. Phase I included both the baseline pre-treatment survey as well as the intervention. All students filled out a pre-treatment online survey, which collected data on their demographics (gender, race/ethnicity, parental education, family income quintile), risk attitudes, academic record, and major. As discussed in Section 3.2, randomization into treatment and control occurred after students submitted their demographic and background information.

In the next section, we acquire data on student beliefs about the external wage distribution. To do so, we asked students the average salary, 25th percentile salary, and 75th percentile salary for students graduating from their major at UC Berkeley in the past year.[14] We argue that for each individual student, the error in students' guesses from the actual statistic constitutes a measure of how informed students are: et ceteris paribus, a student who underestimates the 25th percentile by $20,000 is less informed than a student who underestimates the 25th percentile by only $1,000.

Third, we ask students about their expected personal wages. Expected personal wages differs from the above, which tests students' knowledge about the external wage distribution; questions about what students personally expect help us identify what wage each individual student intends to accept or reject jobs at. We have two important measures. First, students were asked for their reservation wage as: "What is the minimum annual salary you would require in order to accept a job?" This characterization of the reservation wage exactly matches that of our job search theory from Section 2, in which students reject job offers below the reservation wage and accept job offers above the reservation wage. Second, students were asked for their mean expected wage as: "What is the annual salary you expect to accept?" Together, the two questions establish the range of salaries that students are expecting and will allow us to measure the impact the information intervention has on pre-search reservation wages.

[13] All students registered within the UC Berkeley enrollment system to graduate in May 2023.

[14] Questions asked took the following format: "What do you believe was the 25th percentile (bottom 25%) annual salary for students graduating last year within your major at UC Berkeley? The 25th percentile is the salary of the bottom 25% of your major. Thus, if your major contains 100 people, and you rank them in order of salary from 1 (lowest) to 100 (highest), the 25th percentile salary would be that of the 25th person." The language we use to describe percentiles in these questions derives from comparable surveys on taxation from Hvidberg Kreiner and Stantcheva (2021), and we pre-tested this language on comparable undergraduates at Columbia University for comprehension.

3.3.2 Treatment: Information Intervention

After baseline questions, students sorted into treatment received a major-specific information infographic. The intervention provides the external wage distribution: the starting salary distributions of students from prior years of graduating UC Berkeley seniors in students' specific major.[15] Students were informed of the median, 25th percentile, and 75th percentile salaries for their major from the previous 3 years. Where available from the Career Center, data by sector, i.e. salaries for X major in Y sector, were given as well.[16] See **Figure 1** for an example for applied mathematics majors.

To follow-up, we check students' reactions to the information. We ask them to confirm the median salary of the major (based on the infographic they saw). We additionally re-ask students' reservation wages and expected wages, to measure how students incorporate the information into their personal expectations and intended job acceptance behavior.

The control group does not receive an infographic; instead, they only receive reminders about early Fall career fair and on-campus interviews (which are also included for the treatment group). However, we also re-ask students' reservation wages and expected wages. This serves as a check that students' personal wage expectations are stable. Additionally, it provides comparability to the treatment group: experimentalists may worry that survey fatigue or experimenter demand effects could lead to a second measurement of expected personal wages to be lower or higher than a first measure; a second measurement for the control group thus accounts for these possibilities and provides robustness.

3.3.3 Phase II: Post-Treatment Data Collection

In Phase II, we collect data on students' actual job outcomes. Because our questions were similar to the Career Center's annual "First Destination Survey" (FDS) collecting data on job outcomes of graduating seniors, we worked together to synthesize the surveys - our Phase II survey is thus completely incorporated into FDS. See Online Appendix C for a list of questions from the FDS. Phase II ran in the later half of the year, from January 1 - July 15, 2023, in order to coincide with annual FDS timing. We re-ask students about salary beliefs to measure if the information intervention led to persistently different beliefs for the treatment group compared to the control group. In addition, we gathered comprehensive

[15]Students were asked to self-report majors from a drop-down list; double-majors are thus classified based on which major the student themselves picked.

[16]There are five sectors: (1) Business (a for-profit organization that provides a service), e.g. financial consulting, (2) Industrial/Manufacturing (an organization that produces a product), e.g., computer manufacturer, (3) Nonprofit, (4) Education, and (5) Government. The Career Center collected sector data from prior years of students; students self-reported their job sector based on the descriptions above.

information about the final job offer students accept: the location, salary, amenities, and satisfaction with their offer.[17] To better understand earnings dynamics, we asked students to decompose their salary offer into base salary and extra compensation.

To incentivize students to respond, we provided compensation for completing the FDS if students had also completed the baseline survey. Students were compensated $10 and entered into a drawing for three $150 Amazon gift cards. Of the 1110 eligible students in the study population, we received 531 completed responses to our second survey (47.8% response rate compared to the first survey, 6.1% compared to all eligible students).[18]

Attrition between surveys is consistent across treatment and control. In pure numbers, 50.5% of our post-attrition sample is in control and 49.5% is in treatment. Across pre-treatment demographic variables, we find similar balance as in the baseline sample. **Table 2** compares summary statistics between treatment and control across the same baseline characteristics as **Table 1**. As before, for each variable X (e.g. a 0-1 indicator for female), we regress X on a 0-1 indicator for treatment and report the coefficient on treatment in the second-to-last column of the table, with the p-value in the last column. Across 22 variables, we find that only financial aid status is unbalanced at the 5% level. To increase precision, we thus control for financial aid status (in addition to parent's highest education from the original balance table) as a balance variable in all regressions.

3.4 Analysis Sample

We restrict our sample via several coherence questions. First, we required students answer all three questions on external beliefs and both questions on personal wage expectations with coherent numbers. Second, we required students to have coherent beliefs: beliefs about the 25th percentile needed to be lower than or equal to that of the 75th percentile, and the minimum wage one would accept needed to be lower than or equal to their expected wage. Third, we required that students needed to answer at least one of the wage expectation questions post-treatment. Finally, we winsorize all belief and wage variables at the 1% level.[19] Our resultant sample is thus 887 observations.

[17] While we also intended to ask students for their application, interviewing, and negotiation behavior - this was left off the FDS to due an administrative error, and so we only have this data for a very small subset of our sample.

[18] For comparison, Cortes et al. (2023) run their job search survey for college alumni in three phases. The second follow-up phase led to about 50% of the respondents in the previous phase continuing to respond, while the third led to 33% of the original group responding.

[19] For comparison, Jager et al. 2023 winsorizes at the 2% level, and Cortes et al. 2023 winsorize beliefs to between $25,000 and $175,000. Our winsorization results in a range of $10,000 and $250,000, suggesting a more generous use of the raw data than comparable papers.

3.5 Causality

A randomized controlled trial provides clear causality: any differences between treatment and control should be due only to the information intervention treatment. We can additionally provide more precision by controlling for baseline characteristics. A key strength of the research design is that the information intervention not only tests the information asymmetry hypothesis, but also measures other mechanisms – such as differences in risk aversion, confidence, valuation of non-wage amenities, liquidity constraints, and more – which allows us to examine heterogeneity on the impact of information across traditional behavioral differences. One concern may be that the treatment will percolate to the control group, if the treatment group widely shares the information. However, this would only lower any point estimates of the treatment effect, so our results can be interpreted as a lower bound on the effect of accurate information.

4 Errors in Beliefs Influence Reservation Wages: Descriptive Evidence

Prior to the intervention, we find via the baseline survey empirical facts about beliefs over available wage distributions. This builds upon a nascent but growing literature measuring workers' beliefs and perceptions about the labor market.

4.1 Empirical Finding 1: Prior to labor market entry, college students have errors in beliefs about available wage distributions.

We find that prior to labor market entry, students have errors in beliefs about available wage distributions. We asked students for the **major-specific** 25th percentile, average, and 75th percentile salaries for UC Berkeley students graduating last year from the student's chosen major. **Figure 3** displays histograms of the log deviation of beliefs from true salary statistics, while **Figure B1** in Appendix B provides the same graph in percent deviations.[20] Following our model, we use log deviations to isolate the error component, which controls both for outliers and for the fact that majors have widely variant salaries. Log deviations are calculated as follows:

$$\log(\text{Believed Percentile Salary}) - \log(\text{True Percentile Salary})$$

[20] One may ask why not just use percentiles. However, percentiles are still susceptible to outliers in the data - an individual who believes the 75th percentile salary for their major is \$200,000 when the true 75th percentile salary is \$62,000 is overestimating by over 1000% despite not having a particularly outlandish guess. This is an actual data point which occurs, and can skew the overall distribution.

Panel (a) shows the results for the 25th percentile. 79.3% of students underestimate the 25th percentile, with an average underestimation of 23.6%. Panel (b) shows results for the average. Since UC Berkeley does not have true salary data by average but only by median, we calculate: log(Believed Average) - log(True Median), thus making the assumption that the average and median are comparable. A bit over half (55.9%) of the students "underestimate," with the average student underestimating by 1.5%. Panel (c) shows results for the 75th percentile. 61.8% of students overestimate the 75th percentile, with the average student overestimating by 17.3%. Panel (d) shows the results for the interquartile range, calculated as the 75th percentile - the 25th percentile. We find that 90% of students overestimate the interquartile range, with the average student overestimating the interquartile range by 153%. Across students, there is large variation in error and direction of error. Students underestimate the 25th percentile and median by as much as -100% and overestimate by as much as +200%, while underestimating the 75th percentile by as much as -100% and overestimating by as much as +400%. Range estimations are incorrect by as much as -100% to +1500%.

One remaining question is whether the results are being driven (a) by students underestimating the 25th percentile and overestimating the 75th percentile, or (b) whether the same students who are underestimating the 25th percentile are also underestimating the 75th percentile. We find that (b) is more likely to be the case. To test this, we regress an indicator for overestimation of the 75th percentile on overestimation of the 25th percentile and separately for on overestimation of the median. At the 1% significance level, students who overestimate the 25th percentile are 27% more likely to overestimate the 75th percentile, and students who overestimate the median are 41% more likely to overestimate the 75th percentile. Results are qualitatively similar when we reverse the order and use overestimation of 75th percentile as a predictor for overestimation of the other percentiles. This suggests that from the perspective of first-order statistical moments, students who overestimate (underestimate) one percentile statistic are more likely to overestimate (underestimate) another percentile statistic. However, because 90% of students also had overestimations of the interquartile range, reversals of this trend can occur between the 25th and 75th percentile due to the second-order statistical moment (from higher variance). For example, a student who vastly underestimates the 25th percentile should also underestimate the 75th percentile. In contrast, a student who *barely* underestimates the 25th percentile could have their second-order range error be very large, leading to an overestimation of the 75th percentile even if, in general, underestimating the 25th percentile makes it more likely to underestimate the 75th percentile.

4.2 Empirical Finding 2: Student's personal wage expectations are positively correlated with their beliefs about the external wage distribution.

We find that students with higher beliefs about the 25th percentile, the average, and the 75th percentile for the external distribution also have higher reservation wages and higher expected wages. We measure personal wage expectations as (1) the reservation wage, or the minimum wage students would require to accept a job, and (2) the expected wage, or the wage students expect to accept.[21] In **Figure 3**, we graph bin scatters of log reservation wages (in light blue) over log external wage beliefs, and log expected (in dark blue) wages over log external wage beliefs. Slopes of personal expectations over the external belief range between 0.527 to 0.835, showing a positive correlation. On the other hand, external beliefs about the interquartile range are less correlated with personal wage expectation: for each log point increase in interquartile range, the log reservation wage only increases by 0.123 points, while the log expected wage increases by 0.164 points.[22]

In our next section, we turn to our experiment for causal results.

5 Correcting Errors in Beliefs Changes Labor Search Behavior: Experimental Evidence

In this section, we discuss our experimental results, and analyze how changing students' information about the available external wage distribution affects their personal wage expectations and later on, their actual job outcomes.

5.1 Information changes students' personal wage expectations.

First, we establish that the information treatment shifts students' personal wage expectations, conditional on their original error in overestimation or underestimation. Treatment-group students who underestimated initially should increase their reservation wages, while treatment-group students who overestimated should decrease their reservation wages. In the control group, the two measurements of the reservation wage should be the same.[23] In **Figure 4**, we graph the log change in reservation wage ($\log w_{R2} - \log w_{R1}$) against the original log estimation error (log (Believed Percentile Salary) - log (True Percentile Salary)). Since estimation error is represented via belief minus truth, overestimation is represented by pos-

[21] While the reservation wage is more economically useful because it is the equilibrium solution to the job search problem, the expected wage is potentially more intuitive to students.

[22] The same relationship occurs for levels rather than logs; see **Figure B2** in Appendix B.

[23] Recall we also took a second measurement of the reservation wage at the same time for the control group to account for survey fatigue effects.

itive numbers on the x-axis, while underestimation is represented by negative numbers on the x-axis. Thus, since treatment-group students who overestimated should decrease their reservation wages while treatment-group students who underestimated should increase their reservation wage in response to the information intervention, we expect a negative relationship on the graph. We find that this negative relationship is the case. For each of the 25th percentile, median, and 75th percentile graphs, the treatment group (in red) has a negative slope, while the control group (in blue) has a slope of close to 0. This suggests that as conjectured, treatment changed students' personal reservation wages.

In addition to first moment effects on overestimation and underestimation, we also expect a second moment effect: students in treatment who overestimated the interquartile range should shrink their reservation wage, while students in treatment who underestimated the interquartile range should increase their reservation wage (as noted in Proposition 2). Graphically, we see in **Figure 4** that a negative relationship holds as well. However, it is not as visually stark. To better clarify these results, we turn to our regression specification to quantify our results.

Our conjecture is that students with larger errors in the external wage distribution will change their reservation wages by more. To quantify this relationship, we use a learning rate regression, to measure the impact of information on our variable of interest, following similar literature in information provision experiments.[24] Our regression is as follows:

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i X_i + \gamma Controls_i + \epsilon_i \tag{2}$$

where Y_i is change in reservation wage, D_i is the treatment indicator, and X_i is a vector of student's estimation error: (log (Believed Percentile Salary) - log (True Percentile Salary)). Our X_i vector is composed of two parts:

1. **Centrality**: Centrality measures the first stastitical moment of students' overall underestimation or overestimation of the distribution. Our main specification will focus on estimation error of the median.[25] As before, overestimation is recorded as a positive value while underestimation is a negative value, so we expect $\beta_{3,Centrality}$ to be negative.

[24] Haaland et. al. 2023 provide a comprehensive summary of the literature in econometric measurement of information interventions.

[25] We use the other first moments (error in 25th percentile and error in 75th percentile) as robustness checks and find qualitatively similar results. The median is the main specification because it is the most intuitive as the center. Additionally, because the spread is the interquartile range, it encapsulates information about the 25th percentile and 75th percentile, so using the median incorporates all information we receive from each student.

2. **Spread**: Spread measures the second statistical moment of whether students' overestimate the variance (or range) of the distribution. We thus use the students' estimation error of the interquartile range. As with centrality, overestimation is recorded as a positive value while underestimation is a negative value, so we expect $\beta_{3,Spread}$ to be negative.

We use both centrality and spread as one vector within the same regression to account for the student's error across the whole distribution. Since wages are log normal distributions, and log normal distributions are fully characterized by their first-moment mean (center) and second-moment variance (spread), having one measure of centrality and one measure of spread allows us to fully characterize student's estimation error. Our parameter of interest is thus the vector of $\beta_3 = \begin{pmatrix} \beta_{3,Centrality} & \beta_{3,Spread} \end{pmatrix}$, which indicates the log dollar change in reservation wage for each log dollar change in beliefs about the external wage distribution.

As controls, we include indicators for the three strata (male, white, STEM) and balance variables (parental education level and financial aid status). We also control for three behavioral covariates: willingness to take daily risk, willingness to take financial risk, and confidence, which have been shown to relate to inefficient search behavior (Cortes et al. 2023).[26]

In **Table 3**, we show the results of our regression, and find that $\beta_{3,Centrality}$ and $\beta_{3,Spread}$ are both negative and statistically significant. In column (1), we find that the information treatment's correction of a 1% overestimation of the center of the distribution decreases the reservation wage by -0.23%; conversely, the information treatment's correction of a 1% underestimation of the center of the distribution increases the reservation wage by +0.23% ($p < 0.05$). In column (1), we also find that we find that the information treatment's correction of a 1% overestimation of the spread of the distribution decreases the reservation wage by -0.044%; conversely, the information treatment's correction of a 1% underestimation of the spread of the distribution increases the reservation wage by +0.044% ($p < 0.01$). For robustness, we use the 25th percentile estimation error and the 75th percentile estimation

[26]Cortes et al. (2023) finds risk aversion and confidence as the two key behavioral drivers of inefficient search behavior, so we control for the same behavioral traits. As part of our survey, we collect data on risk aversion exactly as Cortes et al. measure it to derive a comparable stimate for our results. Risk aversion is measured as two items: (1) "How would you rate your willingness to take risks in daily activities?" and (2) "How would you rate your willingness to take risks in financial matters?". Each of these questions is answered via a 7-point scale from "Not willing at all" to "Very willing". Confidence is constructed using students' estimates of their GPA's rank compared to its true rank. Students are asked: "Using your best guess, how does your GPA compare to the rest of the students in your major in your graduating year? Select your rank in percentage terms relative to others. For instance, if your choice on the slider is 80, then you believe your GPA is higher than 80% of your major." Confidence is then equal to estimated rank - true rank, with true rank measured as the true GPA rank of students within the sample.

error as alternate measurements of the error in centrality and find similar results. Robustness results can be found in **Table B1**.

Are students' average expected wages affected differently than their reservation wages? We find that the same patterns hold for average expected wages as for reservation wages, but with slightly modified point estimates. In **Figure 6**, we graph the change in expected average wage ($\log E[w]_2 - \log E[w]_1$) against the original log estimation error (log (Believed Percentile Salary) - log (True Percentile Salary)). Similar to the reservation wage graphs in **Figure 4**, we find a negative relationship for the treatment group: treatment-group students who overestimated the center of the distribution decrease their average expected wages, while treatment-group students who underestimated the center of the distribution. In addition, we find a second moment effect: treatment-group students who overestimated the interquartile range shrink their reservation wage after treatment, while treatment-group students who underestimated the interquartile range increase their reservation wage. Column (2) of **Table 3** quantifies these results. We find that in comparison with the control group, the information treatment's correction of a 1% overestimation of the center of the distribution decreases the average expected wage by -0.37%; conversely, the information treatment's correction of a 1% underestimation of the center of the distribution increases the average expected wage by +0.37% ($p < 0.01$). In column (1), we also find that we find that the information treatment's correction of a 1% overestimation of the spread of the distribution decreases the reservation wage by -0.031% and vice versa for underestimation; however, this effect is not statistically significant. Thus, average expected wages respond similarly to reservation wages; however, as reservation wages are a more economically salient measurement of job search behavior (as noted in the theory), we will focus on reservation wages as our main parameter of interest.

Overall, how do reservation wages change in our experiment? Because slightly more students underestimated in our sample, we might expect that on net, the reservation wage for the treatment group should be higher than that of the control group. However, because 90% of students also overestimated the variance of the distribution, we might expect that the treatment group has a lower reservation wage overall. In **Figure 6**, we graph the cumulative distribution functions of the reservation wage by control and treatment. We find that the first effect dominates the second: a Kolmogorov-Smirnov test of the difference between distributions establishes that reservation wages are overall higher for the treatment group with 10% significance ($p = 0.053$).

Thus, we establish that the information treatment shifts students' personal wage expectations. This provides a first-stage to analyze students' post-search job outcomes.

5.2 Information changes earnings and employment.

A novel contribution of this paper is that we are able to observe actual earnings and employment outcomes in Phase II after treatment, rather than just intended search behavior. These results allow us to draw conclusions about how the treatment affects students' actual earnings and employment outcomes.

5.2.1 Estimation Strategy

We use a two-stage least squares (2SLS) identification strategy to identify a local average treatment effect of the information intervention on post-school earnings and employment outcomes. Our endogenous variable is the reservation wage, while our instrumental variable is the treatment and treatment interacted with initial estimation error. Because the reservation wage is a sufficient statistic for individual preferences and utility behavior, this econometric strategy allows us to effectively control for idiosyncratic differences across individuals. In addition, 2SLS gives us a more precise estimate of the true impact of information, since reduced form specifications in experiments can only recover an intention-to-treat (ITT) effect, which may be diluted if subjects don't internalize the information treatment.[27]

Our specification is:

$$\text{First Stage: } \tilde{w}_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i X_i + \gamma Controls_i + \epsilon_i \tag{3}$$

$$\text{Second Stage: } Outcome_i = \alpha_0 + \alpha_1 \tilde{w}_i + \alpha_2 X_i + \gamma Controls_i + \epsilon_i \tag{4}$$

where $\tilde{w}_i$ is the post-treatment log reservation wage[28], D_i is the treatment indicator, and X_i is a vector of student's estimation error: (log (Believed Percentile Salary) - log(True Percentile Salary)). As before, our X_i vector is composed of two parts:

1. **Centrality**: Centrality measures the first stastitical moment of students' overall underestimation or overestimation of the distribution. As before, our main specification will focus on estimation error of the median.

[27] We include the reduced form specification in the Appendix in **Table B2** for interested readers.

[28] For the control group, this is the second measurement of the reservation wage. We use the log of the reservation wage rather than the difference in logs because we expect that for our second stage, that the reservation wage itself should matter, rather than the change in reservation wage. For example, a student who shifts from a reservation wage of \$20,000 to \$22,000 experiences the same log difference as a student who shifts from a reservation wage of \$40,000 to \$44,000. However, we would expect the student with a \$44,000 reservation wage to have a higher accepted salary than the student with a \$22,000 reservation wage.

2. **Spread**: Spread measures the second statistical moment of whether students' overestimate the variance (or range) of the distribution. As before, we use the students' estimation error of the interquartile range.

The validity of the 2SLS as a local average treatment effect requires (1) exogeneity, (2) exclusion, (3) relevance, and (4) monotonicity. First, for exogeneity, our instrument must be uncorrelated with the error term. Since treatment is randomly assigned, this condition is fulfilled. Second, for the exclusion restriction, our instrument (the information treatment interacted with estimation error) can only affect the outcome via the endogenous variable (reservation wage). As shown via the job search model in Section 2, everything - from individual preferences for leisure to the information treatment itself - "passes through" the reservation wage to affect an agent's job outcomes, so this is fulfilled. Third, for relevance, we must show that the instrument affects the endogenous variable. In the previous section, we showed that treatment affects the reservation wage. To additionally formally test for weak instruments, we report at the bottom of each upcoming table the Anderson-Rubin χ^2 statistic and correspondent p-value. Fourth, for monotonicity, the treatment should affect identical individuals in the same way. Because the instrument uses Treatment combined with Estimation Error (which is positive for overestimation and negative for underestimation, both leading to a negative coefficient in the first stage), this is the case.

Our parameter of interest is α_1, which indicates the log dollar impact on outcome (e.g. final accepted wage) for each log dollar of the reservation wage. We expect α_1 to be positive: those who increased their reservation wage should increase their actual earnings.

5.2.2 Results

In **Table 4**, we report α_1 for a variety of job outcomes. We first discuss the impact on earnings, before discussing the impact on employment. In theory, we should expect students whose reservation wages increased to accept jobs with higher wages but wait longer to accept jobs (less likely to be employed by graduation). Conversely, students whose reservation wages decreased should accept jobs with lower wages but be more likely to have accepted a job (more likely to be employed by graduation).

5.2.2.1 Earnings

We find that consistent with theory, increasing the reservation wage via the information channel increases Log Total Salary, conditional on being employed. In **Table 4** Column (1), we examine the effect on Log Total Salary and find a coefficient of 1.283, significant

at the 5% level. Under this coefficient, each 1% increase in reservation wage due to the information intervention leads to 1.283% increase in Log Total Salary. For robustness, we compute Anderson-Rubin 95% confidence intervals, and find that the coefficient is between [0.30,3.73]. This rules out both a negative effect of reservation wage changes on total salary, as well as a null effect.

Is this increase in earnings due to an increase in base salary numbers or due to an increase in bonus compensation (e.g. signing bonus)?[29] We find that both are the case. In **Table 4** Column (2), we examine the effect of a 1% increase in reservation wage on log base salary. We find a coefficient of 1.342, significant at the 10% level. Under this coefficient, each 1% increase in reservation wage due to the information intervention leads to 1.34% increase in Log Base Salary. Anderson-Rubin 95% confidence intervals, however, cannot rule out negative values (but are primarily in the positive domain). In **Table 4** Column (3), we examine the effect of a 1% increase in reservation wage on the likelihood of receiving bonus compensation, and find a coefficient of 1.262, significant at the 10% level. Under this coefficient, each 1% increase in reservation wage due to the information intervention leads to 1.26% increase in the likelihood of receiving bonus compensation. The associated Anderson-Rubin 95% confidence interval is [0.61,4.05], suggesting that we can rule out negative and null effects. In **Table 4** Column (4), we examine the effect of a 1% increase in reservation wage on the *amount* of bonus salary. We find a non-significant coefficient of 2.000, so a 1% increase in reservation wage leads to a 2% increase in the amount of signing bonus, conditional on receiving a bonus. However, the Anderson-Rubin 95% confidence intervals are very wide ([-13.6, 17.7]) and relatively centered around 0, suggesting that this result is tenuous. Thus, we conclude that the increase in total earnings is due both to an increase in base salary numbers and to an increase in the likelihood of receiving a bonus, but that the effects on the amount of the bonus is ambiguous.

5.2.2.2 Employment

We find that consistent with theory, increasing the reservation wage via the information channel decreases the likelihood of employment by graduation (although this result is not statistically significant). In **Table 4** Column (5), we examine the effect of a 1 point increase in log reservation wage on the likelihood of employment (measured via a 0-1 indicator). We find that a 1% increase in reservation wage via the information treatment leads to a -0.564% lower likelihood of being employed. Although the result is not statistically significant,

[29] In data collection, the relevant question is "How are you compensated?" There are three fill-in blanks: (1) Base Salary, (2) Signing Bonus (not required), (3) Other guaranteed compensation (not required). We report both (2) and (3) as one combined bonus income variable.

our computed Anderson-Rubin 95% confidence interval is [-2.28,0.11], which is primarily negative. We can thus rule out strong effects in the opposite direction of theory, which predicts that higher reservation wages should lead to longer periods of unemployment.

We also examine the effects of increasing the reservation wage via the information channel on the likelihood of attending graduate school.[30] In **Table 5** Column (6), we examine the impact of a 1 point increase in log reservation wage on the likelihood of employment (measured via a 0-1 indicator). We find that a 1% increase in reservation wage via the information treatment leads to a -0.256% lower likelihood of attending graduate school. This result is also not statistically significant, but the Anderson-Rubin confidence interval of [-0.93,-0.56], which is a purely negative range. The negative effects on graduate school are consistent with a job search framework, as graduate school tends to pay lower salary (or no salary) compared to traditional post-college jobs in the private sector.

5.2.2.3 Summary and Robustness

Our results are consistent with a reservation wage job search model, in which students with higher reservation wages reject jobs more - leading to a higher likelihood of unemployment but also a higher expected final wage. For additional robustness, we use alternative specifications with the 25th percentile and 75th percentile as the measures of centrality, respectively. **Table B3** in the Appendix shows those results are qualitatively similar for both earnings and employment.

5.3 Subjective Beliefs Lead to Higher Wage Dispersion

What is the overall impact of errors in beliefs on the wage distribution? In **Figure 7**, we see a more compressed final wage distribution for the treatment group: while some students, whose reservation wages were increased, rejected more jobs and accepted higher final wages, other students, whose reservation wages were decreased, accepted lower final wages (and thus increased their likelihood of accepting earlier). The distribution we observe is consistent with empirical papers showing that pay transparency leads to a more compressed wage distribution (Cullen 2023).

We can econometrically test this hypothesis. First, we use Bartlett's equal variances test to compare the distributional variance of the reservation wage between treatment and control.[31] We find a χ^2 value of 1.26 that the variances are different, with a p-value of 0.26. We also find

[30] Please note that we received our data in July, so there is a chance students reneged prior to attending their graduate programs.

[31] Here, we use the second measurement of the reservation wage, post the timing of treatment.

that for log total salary, a χ^2 value of 0.97 that the variances are different, with a p-value of 0.32. This suggests that statistically, there is no difference. However, our power (and sample size) may be too low to draw conclusions.

How are students' personal wage distributions affected? Previous job search model papers (Hornstein et al. 2011) define the mean-min wage ratio, as

$$Mm \equiv \frac{E[w]}{w_R}$$

where $E[w]$ is a worker's expected wage, and w_R is a worker's reservation wage. Hornstein et al. 2011 prove that the mean-min wage ratio is a sufficient statistic for wage dispersion of a worker's personal wage distribution. Because we measure both the worker's expected wage $E[w]$ and the reservation wage w_R post-treatment, we can test the impact of perfect information on wage dispersion. To do so, we regress the mean-min wage ratio post-treatment on an indicator for treatment, controlling as before for strata, balance variables, and behavioral covariates. We find in **Figure 9** that treatment leads to a 0.13 unit decrease in the mean-min wage ratio, which is indicative of a decrease in general wage dispersion. Thus, workers' personal wage distributions have lower variance under perfect information than under subjective beliefs.

6 Model Estimation: Numerical Comparative Statics

How do our empirical results compare to what our theoretical model predicts? We use our Section 2 model to simulate a grid of potential values for the impact of centrality and spread on reservation wages, and compare to the empirical values we find in our data.

To solve for reservation wages under our model, we set some starting parameters. First, we use the CRRA utility function:

$$u(w) = \frac{w^{1-\eta}}{1-\eta}$$

to reflect that agents can be risk averse, with a risk aversion parameter of $\eta = 1.5$.[32] We additionally set a discount factor $\beta = 0.95$ and an offer arrival rate $\lambda = 0.3$.[33] Finally, we set unemployment earnings to \$10,000, reflecting a lower value since college students are not typically immediately eligible for unemployment insurance but may have other means of

[32] Cortes et al. (2023) estimate model parameters of risk aversion $\eta = 1.487$ for men and $\eta = 1.738$ for women.

[33] Arcidiacono et al. (2022), which estimates a continuous time job search model, find an offer arrival rate of between 0.294 and 0.495.

income that serve as an outside option.

As in our model, we assume that wage distributions are log normally distributed $N(\mu, \sigma^2)$. Using a grid[34] of $\mu \in [10.5, 11]$ and $\sigma \in [0.2, 0.3]$, we simulate reservation wages using Equation (1) from Section 2.2, and solve for the slope of the reservation wage with respect to μ, $\frac{\partial \log w_R}{\partial \mu}$, and the slope of the reservation wage with respect to σ, $\frac{\partial \log w_R}{\partial \sigma}$. As an example, we find that for $\mu = 11$ and $\sigma = 0.25$, the calculated annual reservation wage is \$79,082.53, and that $\frac{\partial \log w_R}{\partial \mu} = 0.877$ and $\frac{\partial \log w_R}{\partial \sigma} = 0.982$. **Figure 9** displays our results for the simulation grid. Overall, we find that $\frac{\partial \log w_R}{\partial \mu}$ ranges between [0.87,0.91], while $\frac{\partial \log w_R}{\partial \sigma}$ ranges between [0.98,1.21].

How does this compare to our empirical estimates? Using students' beliefs about the 25th percentile and 75th percentile earnings, denoted w_{P25} and w_{P75} respectively, we are able to estimate beliefs over μ and σ. Under log normality,

$$\mu = \frac{\log(w_{P75}) + \log(w_{P25})}{2}$$

$$\sigma = \frac{\log(w_{P75}) - \mu}{\sqrt{2}\text{erf}^{-1}(0.5)}$$

We can estimate μ and σ for both believed wage distributions and true wage distributions. For our treatment students, then, $\delta\mu = \mu_{true} - \mu_{believed}$ and $\delta\sigma = \sigma_{true} - \sigma_{believed}$. Then, we can calculate:

$$\frac{\partial \log w_R}{\partial \mu} = \frac{\log w_{R,2} - \log w_{R,1}}{\mu_{true} - \mu_{believed}}$$

$$\frac{\partial \log w_R}{\partial \sigma} = \frac{\log w_{R,2} - \log w_{R,1}}{\mu_{true} - \sigma_{believed}}$$

where $w_{R,1}$ is the reservation wage measured pre-treatment and $w_{R,2}$ is the reservation wage measured post-treatment. We can then calculate an individual slope (i.e. responsiveness of reservation wages to centrality μ and spread σ) for each student. We take the assumption that to have a relevant slope, students must have responded the information in some way, so we drop in this exercise any student who had a reservation wage change in 0. In the opposite direction, we trim the slopes at the 1% level to avoid outliers of students who greatly changed their reservation wages. Under these assumptions, we find the following average values:

$$\frac{\partial \log w_R}{\partial \mu} = 0.868$$

[34]These are approximations of the average μ and σ that we see in the true salary data.

$$\frac{\partial \log w_R}{\partial \sigma} = 0.265$$

While $\frac{\partial \log w_R}{\partial \mu} = 0.868$ is very close to the predicted $\frac{\partial \log w_R}{\partial \mu} \in [0.87, 0.91]$, our empirical $\frac{\partial \log w_R}{\partial \sigma}$ estimate is lower than our predicted estimate. This may reflect our assumption of using CRRA utility in the theory, or alternatively may reflect the fact that students were not directly given the variance in our information treatment. Additionally, the data showed much more heterogeneity than the model would predict. **Figure 10** shows our empirical estimates of $\frac{\partial \log w_R}{\partial \mu}$ and $\frac{\partial \log w_R}{\partial \sigma}$.[35] We find that slope estimates for students are largely variant, suggesting that students have potentially large heterogeneity on preferences parameters such as utility functions, risk aversion, values of unemployment, and more.[36]

7 Heterogeneity

In Section 7, we discuss heterogeneity in results. We first explore differences in results between subjects who overestimated or underestimated, before examining differences in results between subjects across different backgrounds.

7.1 Overestimation vs. Underestimation

Do overestimators respond to information equally as underestimators? The behavioral reference point literature suggest possibly not; agents have been shown to be loss averse, in particular not updating beliefs to unfavorable information (Kahneman et al. 1991, Tversky and Kahneman 1991, Villas-Boas 2023). Thus, due to loss aversion, overestimators may be less responsive to an information intervention than underestimators. Because we have both overestimators and underestimators in our sample, we can test this hypothesis.

First, we split our sample into overestimators and underestimators, categorized by whether students underestimated or overestimated the center of the distribution. **Table 5** shows results for a regression in log change reservation wage on change in reservation wage, treatment, and treatment interacted with estimation error.[37] Column (1) shows results for overestimators; Column (2) shows results for underestimators. At the 95% confidence level, our results suggest there is no significant difference in response to the information treatment between students who underestimated and students who overestimated.

[35] For visual purposes, we trim the axes at the 10% level, but the same graph is available with the full data in Appendix Figure B5.

[36] Some of the values are negative. We note that because these are partials (holding all else constant), but we measure the impact of the intervention after a change in **both** μ and σ, students still respond in the correct direction overall, but not in the context of partial derivatives regarding one component.

[37] Refer back to Equation 1 in Section 5.1 for the exact specification.

Second, to test the impact of information on actual job outcomes, we split our 2SLS first stage such that in the instruments, our treatment indicator is always only interacted with overestimators or underestimators. The specification is as follows:

$$\begin{aligned}\text{First Stage: } \tilde{w}_i &= \beta_0 + \beta_1 D_i OC_i + \beta_2 D_i UC_i + \beta_3 D_i OS_i + \beta_4 D_i OC_i|M_i| + \beta_5 D_i UC_i|M_i| \\ &+\beta_6 D_i OS_i|R_i| + \beta_7 D_i US_i|R_i| + \beta_8 OC_i + \beta_9 OS_i + \beta_{10} M_i + \beta_{11} R_i \\ &+\gamma Controls_i + \epsilon_i\end{aligned}$$

$$\text{Second Stage: } Outcome_i = \alpha_0 + \alpha_1 \tilde{w}_i + \alpha_2 OC_i + \alpha_3 OS_i + \alpha_4 M_i + \alpha_5 R_i + \gamma Controls_i + \epsilon_i$$

where OC_i is an indicator for overestimating the center of the distribution and UC_i is an indicator for underestimating the center of the distribution, OS_i is an indicator for overestimating the spread of the distribution and US_i is an indicator for underestimating the spread of the distribution, D_i is the treatment indicator, M_i is the estimation error of the center[38], and R_i is the estimation error in the range. Our parameter of interest is α_1. In essence, we use the same 2SLS regression as in Section 5.2, but allow flexibility in variation between overestimators and underestimators in the first stage.

The added flexibility increases the power of our first stage, and consequently the statistical significance of our results. We report results in **Table 6**.

For earnings, we find similar coefficients to our main results, but with more statistical significance. In Column (1), we find that a 1% increase in reservation wage due to the information intervention leads to 1.065% increase in Log Total Salary ($p < 0.05$), which is very comparable to our coefficient of 1.283 in the main results. In Columns (2)-(4), we examine again whether this increase in Log Total Salary is due to increases in base salary or extra compensation. We find in Column (2) that a 1% increase in reservation wage due to the information intervention leads to a 1.174% increase in Log Base Salary ($p < 0.01$), again comparable to our coefficient of 1.342 in the main results. In Column (3), we find that a 1% increase in reservation wage via the information channel leads to a 0.824% increase in the likelihood of receiving bonus compensation ($p < 0.05$). As with our main results, results on the amount of bonus compensation are inconclusive and statistically indistinguishable from zero.

For employment, we find similar coefficients (and significance levels) to our main results. In Column (5), we observe that a 1% increase in reservation wage via the information treatment leads to a -0.313% lower likelihood of being employed, comparable with our original coefficient of -0.564. Additionally, our Anderson-Rubin 95% confidence interval. Although the result

[38] As before, (log (Believed Average) - log (True Median)).

is not statistically significant, our computed Anderson-Rubin 95% confidence interval is [-1.48,-0.07], which is fully negative. We can thus rule out both null and positive results on employment.

Thus, our results show that agents do not act differently between overestimation and underestimation. Overestimators respond equally to information as underestimators.

7.2 Demographic and Educational Differences

Next, we analyze differences in information and information responsiveness by gender, race, and major.

Gender. Prior literature suggests that there may or may not be informational differences by gender. On the one hand, Gallen and Wasserman (2022) show that when students in a field experiment asked professionals for career advice, female students received different information than male students; Baker et al. (2019) also find that pay transparency at the university level reduces the gender wage gap in faculty salaries. On the other hand, Cortes et al. (2023) find no evidence on gender differences in information about population earnings between male and female students in their sample.

Similarly to Cortes et al. 2023, we find no differences in information by gender in our sample. In **Table 7** Column (1), we regress errors in log beliefs about the distribution on an indicator for female/non-binary students. We find no significant differences across gender in estimation of the center of the distribution (Panel A), or in the spread (Panel B). Coefficients are both insignificant and very close to 0. In **Table 8** Panel A, we analyze whether individuals respond differently to treatment across gender. We find no significant differences for either earnings or employment.

Race. Similar to gender gaps, racial gaps persist in earnings, even when controlling for education and major (Patten 2016). In the education literature, Hoxby and Turner (2015) show in a field experiment that high-achieving low-income students (who tend to be from minority backgrounds) are more likely to apply for college when given more information. This suggests that similar to gender, there may or may not be differences in information and information responsiveness.

We find very little significant differences in information by race in our sample. In **Table 7** Column (1), we regress errors in log beliefs about the distribution on an indicator for underrepresented minority (Black, Hispanic, and/or Native) students. We find no differences across underrepresented minority status in estimation of the center of the distribution (Panel

A), with a coefficient very close to 0. However, we do find that underrepresented minority students underestimate the spread (Panel B) of the distribution by 12.1% more, but this coefficient is not significant ($p = 0.144$). In **Table 7** Column (2), we regress errors in log beliefs about the distribution on an indicator for non-white (Asian American, Black, Hispanic, and/or Native) students. We find no differences across nonwhite status in estimation error on either the center (Panel A) or Spread (Panel B) of the distribution, with coefficients both very close to 0 and not statistically significant. In **Table 8** Panel B and C, we analyze whether individuals respond differently to treatment across race. Our only statistically significant result shows that increasing reservation wage is more likely to increase the amount of log bonus salary for underrepresented minority individuals, conditional on receiving a bonus.[39]

Major. Next, we examine whether students across different majors have information differences. Because high-paying majors have more room to negotiate (due to salaries being bounded below by 0), and because previous literature (Roussille 2023, Cortes et al. 2023) have particularly noted inefficient search behavior in computer science and economics (both of which are traditionally high-paying), it is more likely that high-paying majors exhibit higher error in estimating the availability of external wages. To classify high-paying majors, we split our sample by the median pay for the 75th percentile: majors with true 75th percentile pay above the median ($80,000) are considered high-paying majors, while majors below the median are considered low-paying majors.

We find that high-paying majors are more likely to underestimate both the center and spread of their major-specific distributions, but respond equally to the information treatment as low-paying majors. In **Table 6** Column (1),we regress errors in log beliefs about the distribution on an indicator for students in high-paying majors. We find that students in high-paying majors underestimate the center of the distribution (Panel A) by 18.7% more, significant at the 1% level. We also find that students in high-paying majors underestimate the spread of the distribution by 21.9% more, significant at the 10% level. However, regarding responsiveness to treatment (conditional on original estimation error), we find no differences in **Table 7** Panel D between high-paying and low-paying majors.

Across gender, race, and major, only major exhibits strong information differences, with students in high-paying majors more likely to underestimate both the spread and range of the

[39]We do, however, find a lot of selection into major by gender, consistent with results in Sloane et al. 2021, and by race, consistent with results in Bleemer and Mehta 2023. However, this is outside the scope of our experiment.

distribution. In responsiveness to the information treatment, our only statistically significant result shows that increasing reservation wage is more likely to increase the amount of log bonus salary for underrepresented minority individuals, conditional on already receiving a bonus.

8 Conclusion

In this paper, we conduct an information intervention experiment to examine the impact of correcting workers' beliefs about the external wage distribution prior to the start of their employment search. We find that consistently, the information intervention changes workers' reservation wages: students given the information treatment adjust their reservation wages based on how far they were from the accurate wage. This is robust to inclusions of prior behavioral explanations such as confidence and risk aversion. We find also that for actual job outcomes, students who increased their reservation wage as a result of treatment had higher actual wages, but a (non-statistically significant) decrease in likelihood of employment. On the other hand, students who decreased their reservation wage as a result of treatment had lower actual wages, but a (non-statistically significant) increase in likelihood of employment. This resulted in the wage distribution of treatment group students being more compressed than that of the control group.

The initial jobs that college graduates receive impact their life-long earnings. If higher information transparency does mitigate inefficiency from imprecise beliefs, this is a replicable and low-cost intervention that any university can implement for its students with high returns for efficiency.

9 References

Adams, S. (2015): "The College Majors with the Highest Starting Salaries," Forbes. Jul 2, 2015.

Agan, A. and S. Starr (2018): "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment," The Quarterly Journal of Economics, 191-235. doi:10.1093/qje/qjx028.

Avilova, T. and C. Goldin (2018): "What can UWE Do For Economics?" AEA Papers and Proceedings 108: 186-190.

Bertrand, M. and S. Mullainathan (2004): "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination," American Economic Review 94(4): 991-1013. https://www.jstor.org/stable/3592802

Blau, F.D. and L. Kahn (2017): "The Gender Wage Gap: Extent, Trends, and Explanations," Journal of Economic Literature 55 (3): 789–865. https://www.jstor.org/stable/27739928

"Career Fairs and Industry Showcases," Columbia University Center for Career Education. https://www.careereducation.columbia.edu/employers/recruit-campus. Accessed 2021 Jan 03.

Chakrabarti, R. and M. Jiang (2018): "Education's Role in Earnings, Employment, and Economic Mobility," Liberty Street Economics. https://libertystreeteconomics.newyorkfed.org/2018/09/educations-role-in-earnings-employmentand-economic-mobility.html

"College Scorecard," US Department of Education. https://collegescorecard.ed.gov/school/?190150-Columbia-University-in-the-City-of-New-York. Accessed 2021 Jan 03.

"Columbia University Tuition and Financial Aid," US News. https://www.usnews.com/best-colleges/columbia-university-2707/paying. Accessed 2021 Jan 03.

Comenetz, J. (2016): "Frequently Occurring Surnames in the 2010 Census." United States Census Bureau. https://www2.census.gov/topics/genealogy/2010surnames/surnames.

pdf. Accessed 2021 Jan 03.

Conlon, J., L. Pilossoph, M. Wiswall, and B. Zafar (2018): "Labor Market Search with Imperfect Information and Learning," NBER Working Paper No. w24988. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3246785

Cortes, P., J. Pan, E. Reuben, L. Pilossoph, and B. Zafar (2023): "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab." Working Paper.

Cox, J.C. and and Oaxaca R.L. (2000): "Good News and Bad News: Search from Unknown Wage Offer Distributions," Experimental Economics 2: 1997-225.

Delaney, J. and P.J. Devereux (2020): "Gender Differences in College Applications: Aspiration and Risk Management," Centre for Economic Policy Research Discussion Papers. https://cepr.org/active/publications/discussion_papers/dp.php?dpno=15589

"Frequently Occurring Surnames from the 2010 Census." United States Census Bureau. https://www.census.gov/topics/population/genealogy/data/2010_surnames.html

Gallen, Y. and M. Wasserman (2020): "Informed Choices: Gender Gaps in Career Advice," Working Paper.

Hernandez, M., D.R. Avery, S.D. Volpone, and C.R. Kaiser (2018): "Bargaining While Black: The Role of Race in Salary Negotiations," Journal of Applied Psychology 104(40): 581-592. https://doi.org/10.1037/apl0000363

Hoxby, C.M. and S. Turner (2013): "Expanding College Opportunities for High-Achieving, Low-Income Students," SIEPR Working Paper. https://siepr.stanford.edu/research/publications/expanding-college-opportunities-high-achieving-low-income-students

Hoxby, C.M. and S. Turner (2015): "What High-Achieving Low-Income Students Know About College," American Economic Review 105(5): 514-17. https://www.aeaweb.org/articles?id=10.1257/aer.p20151027

Huang, J. and C. Low (2018): "The Myth of the Male Negotiator: Gender's Effect on Negotiation Strategies and Outcomes," Working Paper. https://faculty.wharton.upenn.

edu/wp-content/uploads/2018/05/MythMaleNegotiator_WorkingPaper_JHuang-1.pdf

Jager, S. C. Roth, N. Roussille, and B. Schoefer (2021): "Worker Beliefs About Outside Options", Quarterly Journal of Economics, forthcoming.

Jansson, J. and B. Tyrefors (2020): "The Genius is a Male: Stereotypes and Same Sex Bias in Exam Grading in Economics at Stockholm University," IFN Working Paper.

Kahn, L. (2010): "The long-term labor market consequences of graduating from college in a bad economy," Labour Economics 17(2): 303-316. https://doi.org/10.1016/j.labeco.2009.09.002

Koc, E.W., J. Kahn, A.J. Koncz, A. Galbraith, and A. Longenberger. (2020): "NACE: Salary Survey Winter 2020," National Association of Colleges and Employers. https://www.naceweb.org/uploadedfiles/files/2020/publication/executive-summary/2020-nace-salary-surv
pdf

Leibbrandt, A. and J.A. List. (2015): "Do Women Avoid Salary Negotiations? Evidence from a Large-Scale Natural Field Experiment." Management Science 61:9: 2016-2024.

Low, C., J.B. Kessler, and C. Sullivan (2019): "Incentivized Resume Rating: Eliciting Employer Preferences without Deception," American Economic Review 109 (11): (11): 3713-44. https://www.aeaweb.org/articles?id=10.1257/aer.20181714

Maestas, N., K.J. Mullen, D. Powell, T. von Wachter, and J.B. Wenger (2018): "The Value of Working Conditions in the United States and Implications for the Structure of Wages," NBER Working Paper 25204. https://www.nber.org/papers/w25204

McCall, J.J. (1970): "Economics of Information and Job Search," Quarterly Journal of Economics 84(1): 113-126.

Mortensen, D.T. and C.A. Pissarides (1994): "Job Creation and Job Destruction in the Theory of Unemployment," The Review of Economic Studies 61(3): 397-415

National Association of Colleges and Employers. "Salary Survey," https://www.naceweb.org/store/subscription/salary-survey/. Accessed 2020 Dec 27.

Oreopoulos, P. (2011): "Why do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes," American Economic Journal: Economic Policy 3: 148-171.

Patten, E. (2016): "Racial, gender wage gaps persist in U.S. despite some progress," Pew Research Center. Jul. 1, 2016. https://www.pewresearch.org/fact-tank/2016/07/01/racial-gender-wage-gaps-persist-in-u-s-despite-some-progress/

Pedulla, D.S. and D. Pager (2019): "Race and Networks in the Job Search Process," American Sociological Review 84(6):983-1012. doi:10.1177/0003122419883255

Pissarides, C.A. (1985): "Short-run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages," American Economic Review 75(4): 676-690.

Reuben, E., M. Wiswall, and B. Zafar (2017): "Preferences and Biases in Educational Choices and Labor Market Expectations: Shrinking the Black Box of Gender," Economic Journal 127: 2153-2186.

Routt, A. "Getting and Cleaning Name Data from the U.S. Census Bureau and SSA for More (A)merican Random Sampling." http://rstudio-pubs-static.s3.amazonaws.com/24678_8234c801b4134f2cb32d8077f6cd75a2.html. Accessed 2021 Jan 03.

Sarsons, H. (2017): "Recognition for Group Work: Gender Differences in Academia," American Economic Review: Papers and Proceedings 107(5): 141-145

Sarsons H. (2017): "Interpreting Signals in the Labor Market: Evidence from Medical Referrals," Working Paper.

Shurchkov, O. and C.C. Eckel (2018): "Gender Differences in Behavioral Traits and Labor Market Outcomes," The Oxford Handbook of Women and the Economy, Oxford University Press, p. 481–512.

"The Labor Market for Recent College Graduates," Federal Reserve Bank of NY. https://www.newyorkfed.org/research/college-labor-market/college-labor-market_compare-majors.html. Accessed 2021 Jan 03.

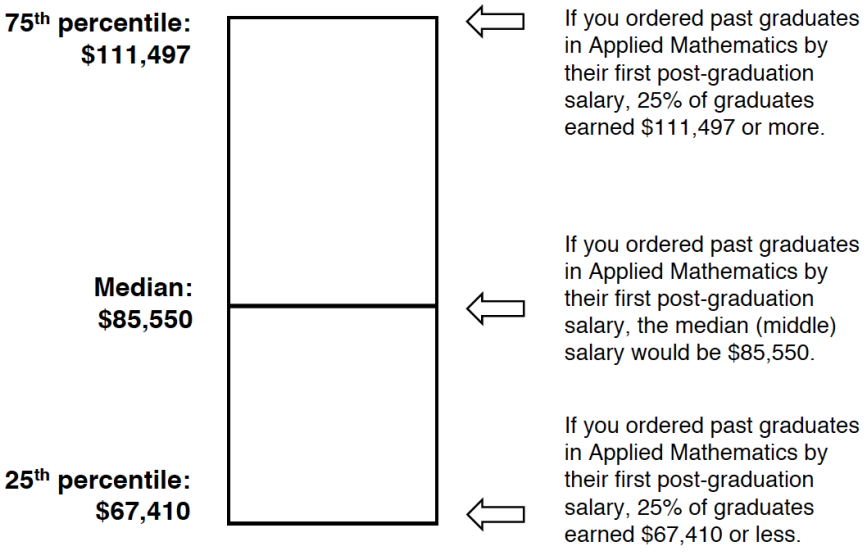
"Vital Statistics and Registry," State of New Jersey Department of Health. https://www.state.nj.us/health/vital/order-vital/online-requests/. Accessed 2021 Jan 02.

Whistle, W. (2019): "Department of Education Releases Earnings Data by Major," Forbes. Nov. 20, 2019. www.forbes.com/sites/wesleywhistle/2019/11/20/department-of-education-relea

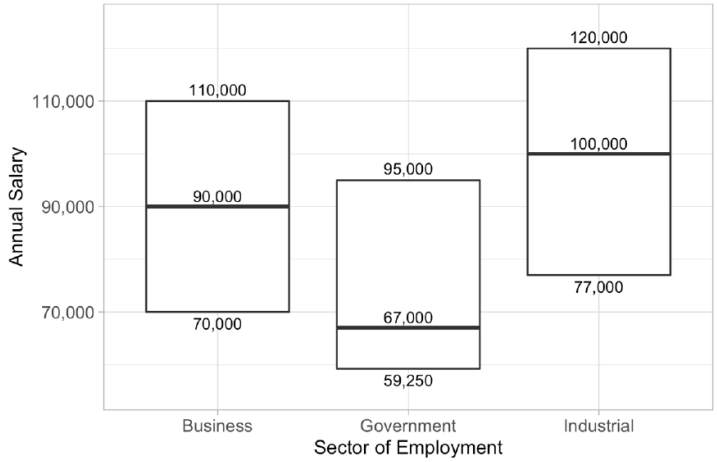
Wiswall, M. and Zafar (2018): "Preference for the Workplace, Investment in Human Capital, and Gender," The Quarterly Journal of Economics 133(1): 467-507, https://doi.org/10.1093/qje/qjx035

Figures and Tables

The **median annual starting salary** for **Applied Mathematics** majors in UC Berkeley's **College of Letters & Science** is **$85,550** in their first job after graduation.



Broken down by sector of employment, these diagrams of the 75th percentile, median, and 25th percentile starting salaries for **Applied Mathematics** majors in their first job after graduating are:



Data for majors are from the UC Berkeley Career Center First Destination Survey, using responses of UC Berkeley graduates in the years 2019-2021. Data for majors by sector are from the years 2019-2020. Sectors are self-reported by students, using the following options:

(1) Business (a for-profit organization that provides a service), e.g., financial consulting
(2) Industrial/Manufacturing (an organization that produces a product), e.g., computer manufacturer
(3) Nonprofit (other than education or government)
(4) Education
(5) Government

Figure 1: Sample of information treatment.

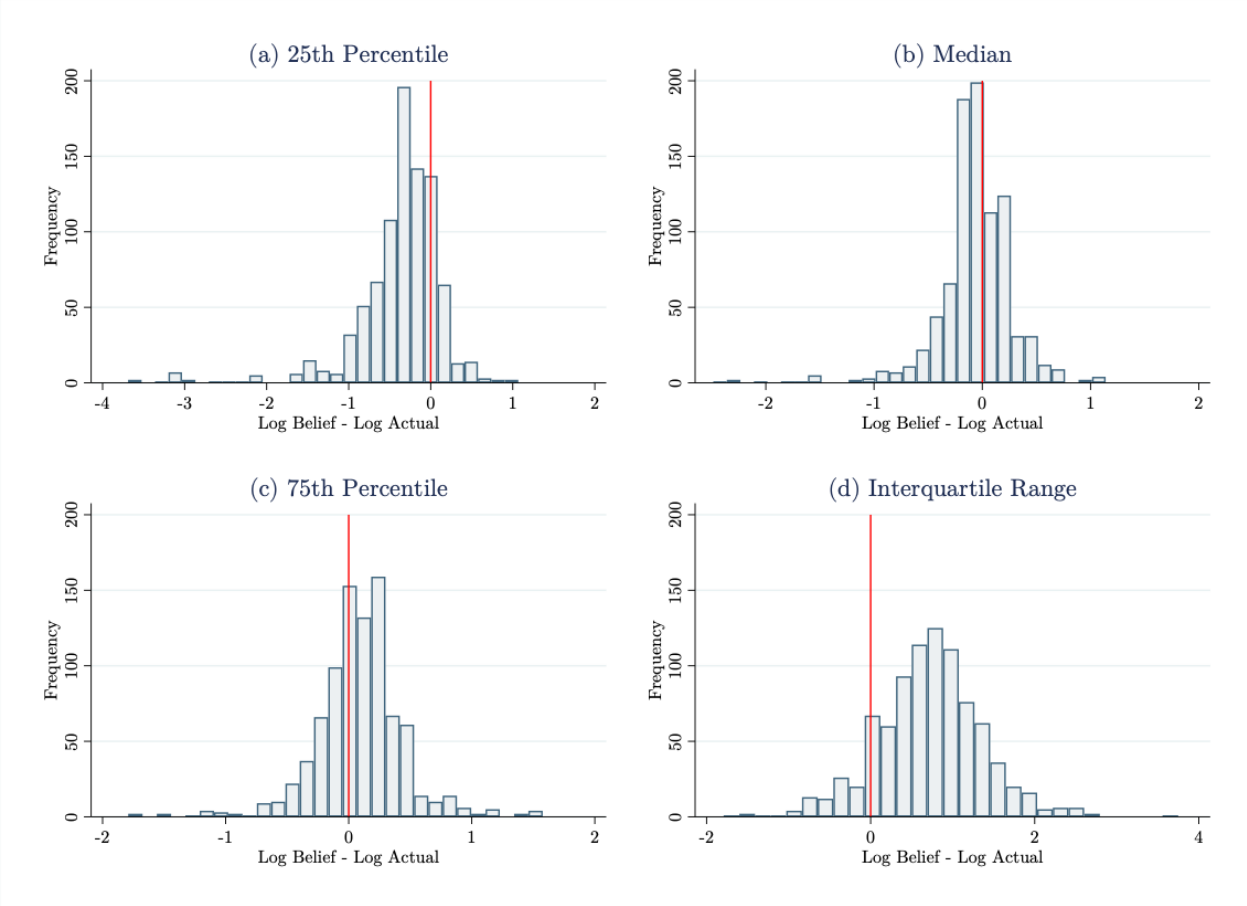


Figure 2: Log Deviation of Beliefs from True Salary (Within-Major).

Histogram of the error in students' pre-treatment beliefs about the major-specific distribution. Students were asked about the 25th percentile, average, and 75th percentile wage that students graduating within their major at UC Berkeley received last year. Graph (a) shows error in beliefs about the 25th percentile. Graph (b) shows beliefs about the average - the true median. Graph (c) shows error in beliefs about the 75th percentile. Graph (d) shows the range, calculated using 75th percentile - 25th percentile. The red line at 0 indicates no error in beliefs, i.e. the situation where studets have perfect information. Negative values in this histogram indicate underestimation, while positive values indicate overestimation.

Questions asked took the following format: "What do you believe was the 25th percentile (bottom 25%) annual salary for students graduating last year within your major at UC Berkeley? The 25th percentile is the salary of the bottom 25% of your major. Thus, if your major contains 100 people, and you rank them in order of salary from 1 (lowest) to 100 (highest), the 25th percentile salary would be that of the 25th person."

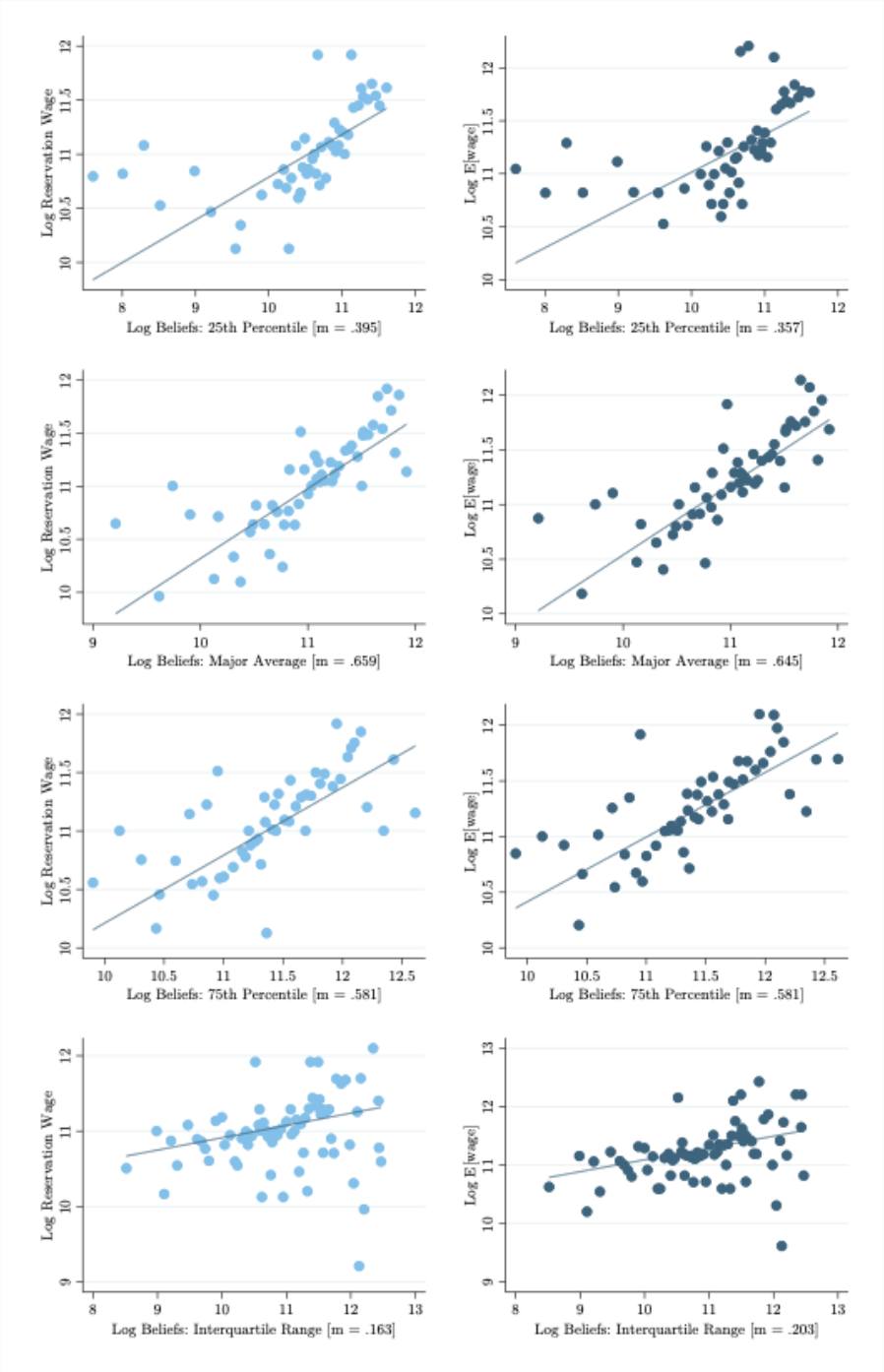


Figure 3: Log Personal Wage Expectations v. Log Beliefs about External Wage Distribution.

Scatter plots comparing log reservation wage (in light blue) versus log beliefs about the major-specific external wage distribution, as well as log average expected wage (in dark blue) versus log beliefs about the major-specific external wage distribution. Scatters are binned at the discrete level. In order of top row to bottom row, the independent variable is: (1) beliefs about the major-specific 25th percentile, (2) beliefs about the major-specific average, (3) beliefs about the major-specific 75th percentile, and (4) the interquartile range, computed via (3) - (1). Relationships are positively correlated, with slope m in the x-axis for each graph.

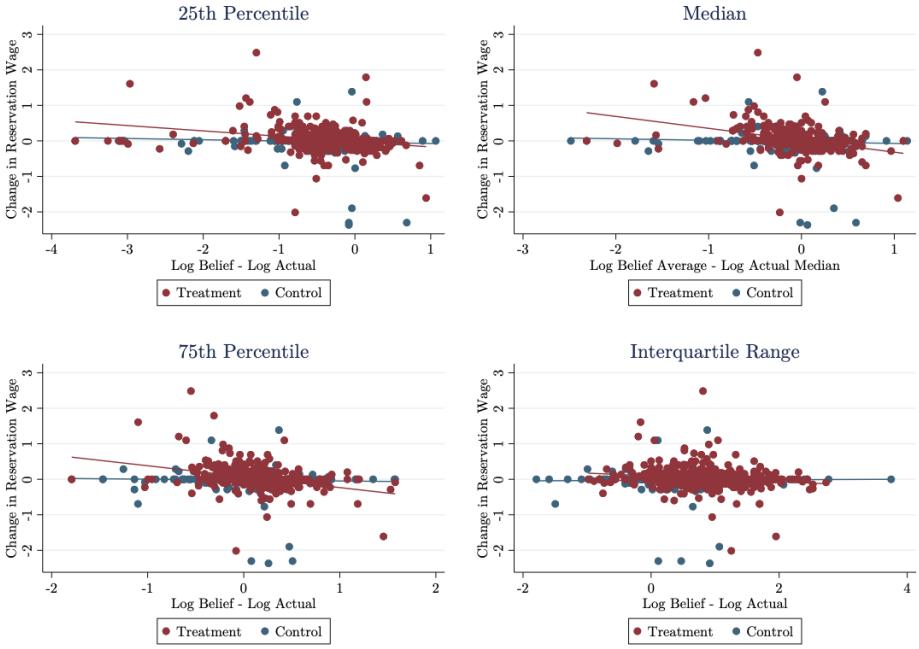


Figure 4: Post-Treatment Change in Log Reservation Wage.

We take two measures of the reservation wage - one post-treatment and one pre-treatment (the control group is asked at the same survey page timing as the treatment group). Reservation wages are asked as "What is the minimum annual salary you would require in order to accept a job?" Figure 4 plots the change in log reservation wage over the error in log beliefs (log belief - log actual). Red dots indicate the treatment group, while blue dots indicate the control group. A linear relationship is graphed for both - as expected, the treatment group has a negative relationship, while the slope for the control group is flat.

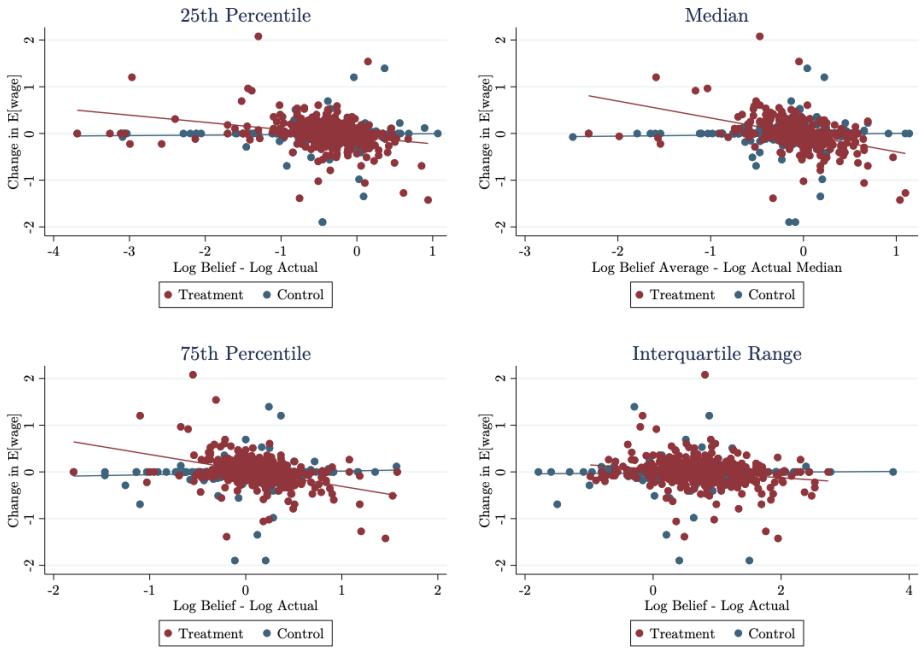


Figure 5: Post-Treatment Change in Log Reservation Wage.

We take two measures of the average expected wage - one post-treatment and one pre-treatment (the control group is asked at the same survey page timing as the treatment group). Average expected wage is asked as "What is the annual salary you expect to accept?" Figure 5 plots the change in log reservation wage over the error in log beliefs (log belief - log actual). Red dots indicate the treatment group, while blue dots indicate the control group. A linear relationship is graphed for both - as expected, the treatment group has a negative relationship, while the slope for the control group is flat.

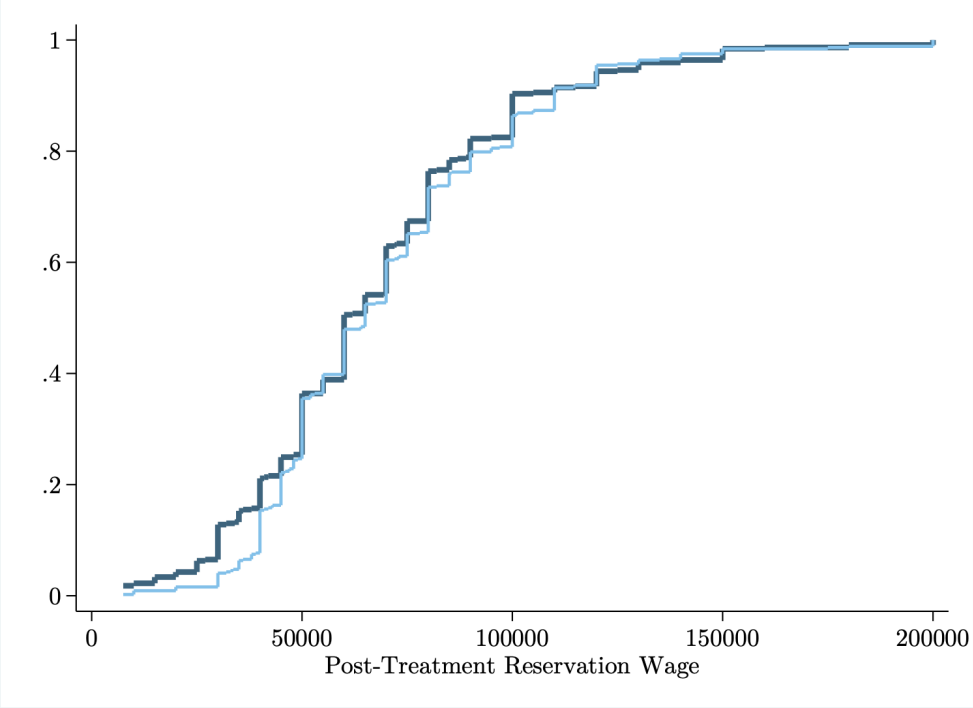


Figure 6: CDF of reservation wages, treatment v. control.

Figure 6 plots the CDF for reservation wages. Light blue indicates treatment, while dark blue indicates control. A Kolmogorov-Smirnov test establishes that treatment first order stochastically dominates control ($p = 0.053$).

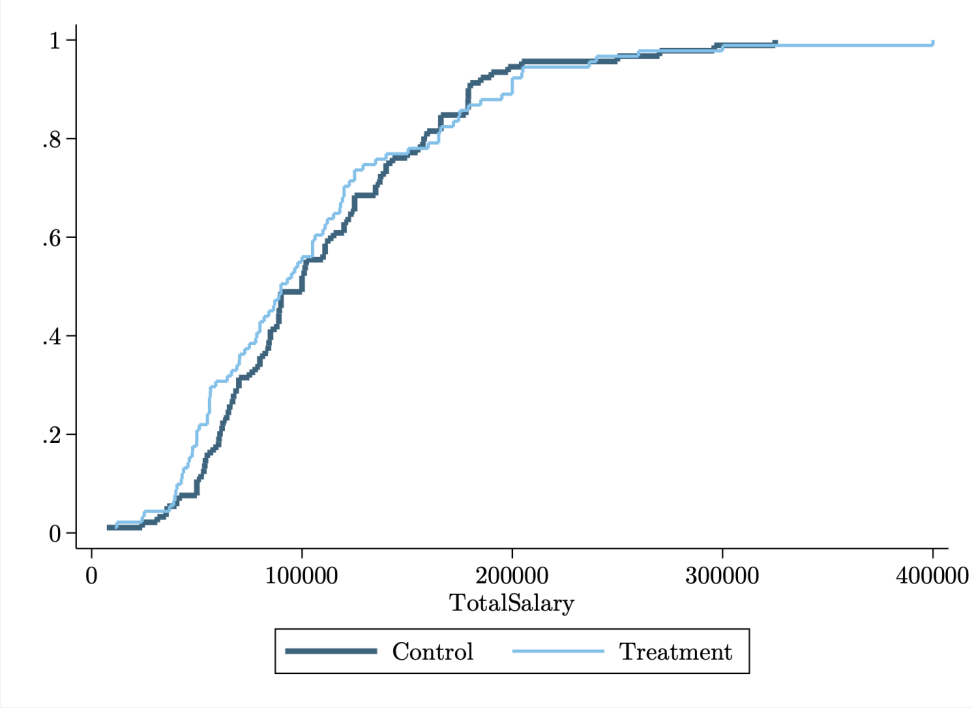


Figure 7: CDF of accepted wages, treatment v. control.

Figure 7 plots the CDF for accepted wages. Light blue indicates treatment, while dark blue indicates control.

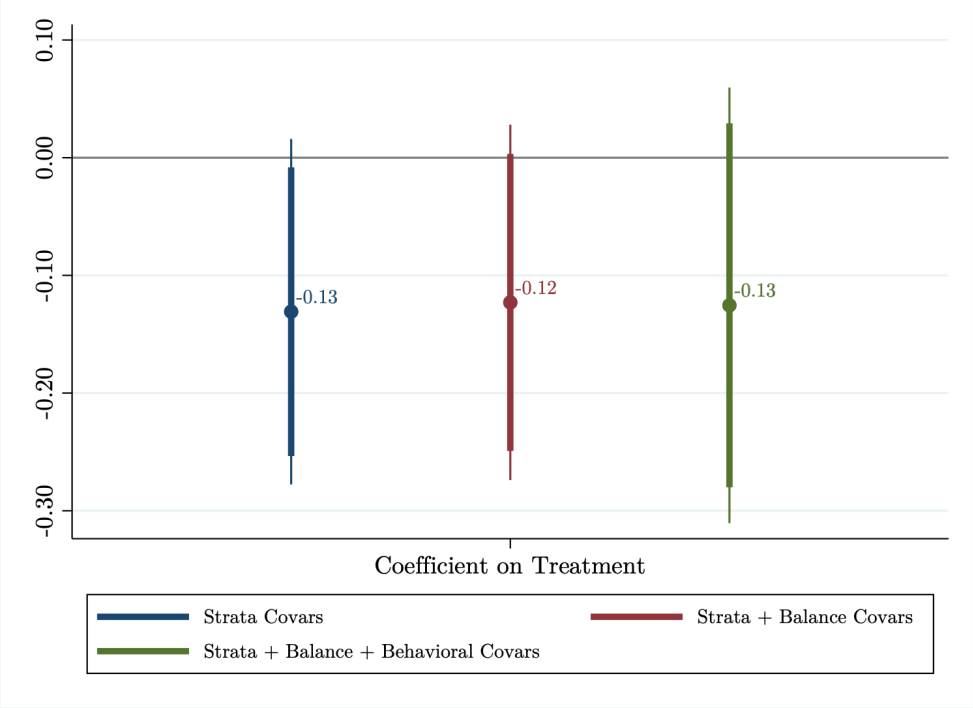


Figure 8: Coefficient on impact of treatment on mean-min wage raio.

Figure 9 plots coefficients for a regression of mean-min wage ratio on treatment.

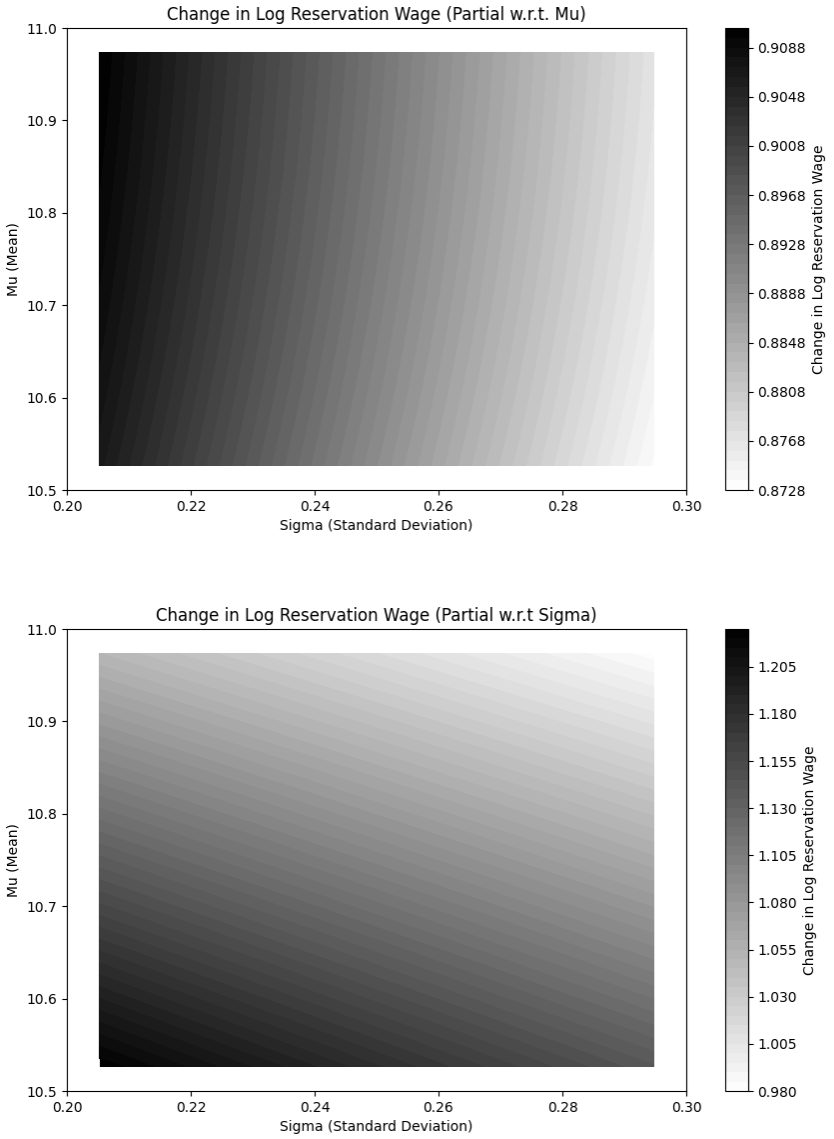


Figure 9: Simulation of $\frac{\partial \log w_R}{\partial \mu}$ and $\frac{\partial \log w_R}{\partial \sigma}$.

Using a grid of $\mu \in [10.5, 11]$ and $\sigma \in [0.2, 0.3]$, we simulate reservation wages using Equation (1) from Section 2.2:

$$\Big(1 - \beta(1-\lambda)\Big)u(w_R) - u(b) = \beta\Big(\lambda u(E[w|\mu_i, \sigma_i]) - u(b)\Big) + \beta\lambda \int_0^{w_R} u'(w)F(w|\mu_i, \sigma_i)dw$$

and solve for the slope of the reservation wage with respect to μ, $\frac{\partial \log w_R}{\partial \mu}$, and the slope of the reservation wage with respect to σ, $\frac{\partial \log w_R}{\partial \sigma}$. We assume CRRA utility $u(w) = \frac{w^{1-\eta}}{1-\eta}$, with risk aversion parameter $\eta = 1.5$. We also set a discount factor $\beta = 0.95$, an offer arrival rate $\lambda = 0.3$, and unemployment earnings $b = 10000$.

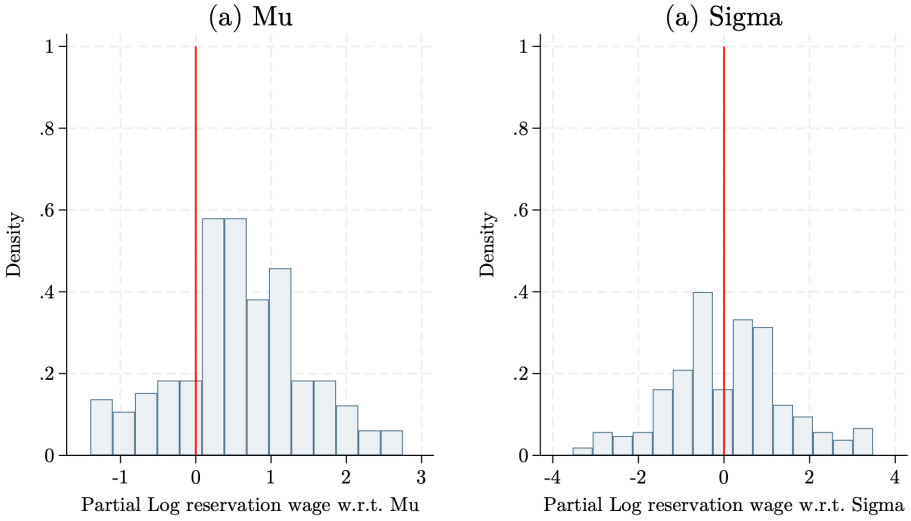


Figure 10: Empirical Estimates of $\frac{\partial \log w_R}{\partial \mu}$ and $\frac{\partial \log w_R}{\partial \sigma}$.

Figure 10 solves for the slope of the reservation wage with respect to μ, $\frac{\partial \log w_R}{\partial \mu}$, and the slope of the reservation wage with respect to σ, $\frac{\partial \log w_R}{\partial \sigma}$, using treatment students' beliefs about the 25th percentile and 75th percentile earnings compared to true 25th percentile and 75th percentile earnings in the denominator, and the difference in post-treatment versus pre-treatment reservation wages in the numerator. We take the assumption that to have a relevant slope, students must have responded the information in some way, so we drop in this exercise any student who had a reservation wage change in 0. In the opposite direction, we trim the slopes at the 1% level to avoid outliers of students who greatly changed their reservation wages. For visual reasons, this graph trims values at the 10%, but the same graph at the 1% level is available in the Appendix B: Additional Figures and Tables.

Table 1: Descriptive Statistics and Balance

	Control			Treatment			Difference	
	Mean	SD	N	Mean	SD	N	Coef.	p-value
Panel A: Demographics								
Gender: Female or Non-binary	0.62	0.48	445	0.60	0.49	442	-0.02	[0.44]
Gender: Male	0.38	0.48	445	0.40	0.49	442	0.02	[0.44]
Race: Asian American	0.57	0.50	445	0.57	0.50	442	-0.00	[0.98]
Race: Black, Hispanic, Native	0.20	0.40	445	0.18	0.38	442	-0.02	[0.42]
Race: White	0.19	0.39	445	0.19	0.40	442	0.00	[0.89]
Citizenship: USA	0.88	0.32	445	0.88	0.33	442	-0.00	[0.89]
Family Income Group (1-5)	3.26	1.36	377	3.18	1.43	368	-0.08	[0.40]
Parent's Highest Educational Level Group (1-6)	3.74	1.61	438	4.04	1.51	436	0.30***	[0.01]
Panel B: Educational background								
Major: STEM	0.74	0.44	445	0.76	0.43	442	0.01	[0.66]
Major: High-paying	0.55	0.50	445	0.52	0.50	442	-0.03	[0.26]
Transfer student	0.25	0.43	445	0.26	0.44	442	0.01	[0.81]
Receiving financial aid	0.47	0.50	445	0.42	0.49	442	-0.05	[0.12]
GPA	3.57	0.37	440	3.59	0.38	442	0.02	[0.39]
Panel C: Behavioral characteristics								
Risk attitude: Daily (1-7)	4.90	1.29	414	4.83	1.31	394	-0.07	[0.42]
Risk attitude: Financial (1-7)	3.54	1.56	386	3.61	1.49	405	0.07	[0.45]
Confidence (-1 to 1)	0.07	0.24	445	0.06	0.22	440	-0.01	[0.50]
Panel D: Beliefs about wages, external								
External Beliefs: Avg. Earnings of Major	68140.00	25056.37	445	70416.29	26432.51	442	2276.29	[0.12]
External Beliefs: P25 Earnings of Major	46044.94	21733.60	445	46149.32	21242.69	442	104.38	[0.93]
External Beliefs: P75 Earnings of Major	94073.03	38381.62	445	98984.62	43226.11	442	4911.58*	[0.06]
External Beliefs: Job Finding Rate of Major	56.63	20.10	443	55.66	20.39	440	-0.97	[0.41]
Panel E: Beliefs about wages, self								
Reservation Wage (Baseline)	68712.58	33436.78	445	69172.17	31508.16	442	459.59	[0.81]
Expected Wage (Baseline)	83961.80	42705.25	445	84293.89	40594.39	442	332.09	[0.90]

We compare the means, standard deviation, and number of observations between treatment and control for a variety of pre-treatment variables. In the last two columns, we regress the demographic of interest on an indicator for treatment, and report the coefficient on treatment. Standard errors are clustered by major. The p-value for the coefficient is in brackets. Stars indicate: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2: Post-Attrition: Balance and Descriptive Statistics

	Control			Treatment			Difference	
	Mean	SD	N	Mean	SD	N	Coef.	p-value
Panel A: Demographics								
Gender: Female or Non-binary	0.57	0.50	217	0.59	0.49	219	0.02	[0.58]
Gender: Male	0.43	0.50	217	0.41	0.49	219	-0.02	[0.58]
Race: Asian American	0.59	0.49	217	0.59	0.49	219	0.00	[0.91]
Race: Black, Hispanic, Native	0.16	0.36	217	0.17	0.38	219	0.01	[0.70]
Race: White	0.21	0.41	217	0.20	0.40	219	-0.01	[0.78]
Citizenship: USA	0.90	0.30	217	0.89	0.31	219	-0.01	[0.75]
Family Income Group (1-5)	3.39	1.32	183	3.24	1.48	181	-0.15	[0.27]
Parent's Highest Educational Level Group (1-6)	3.98	1.60	216	4.17	1.44	214	0.20	[0.23]
Panel B: Educational background								
Major: STEM	0.81	0.39	217	0.81	0.39	219	-0.00	[0.93]
Major: High-paying	0.63	0.48	217	0.60	0.49	219	-0.03	[0.46]
Transfer student	0.19	0.40	217	0.22	0.41	219	0.03	[0.50]
Receiving financial aid	0.46	0.50	217	0.36	0.48	219	-0.10**	[0.01]
GPA	3.61	0.35	215	3.64	0.34	219	0.03	[0.29]
Panel C: Behavioral characteristics								
Risk attitude: Daily (1-7)	4.95	1.24	197	4.77	1.32	198	-0.19	[0.16]
Risk attitude: Financial (1-7)	3.49	1.56	189	3.58	1.46	200	0.08	[0.58]
Confidence (-1 to 1)	0.06	0.21	217	0.05	0.22	218	-0.01	[0.52]
Panel D: Beliefs about wages, external								
External Beliefs: Avg. Earnings of Major	73571.43	26408.57	217	75155.25	28286.24	219	1583.82	[0.44]
External Beliefs: P25 Earnings of Major	50751.15	22644.25	217	48675.80	22256.43	219	-2075.35	[0.15]
External Beliefs: P75 Earnings of Major	99472.35	36539.73	217	103991.78	45245.81	219	4519.43	[0.16]
External Beliefs: Job Finding Rate of Major	59.26	20.04	216	56.96	20.14	218	-2.30	[0.28]
Panel E: Beliefs about wages, self								
Reservation Wage (Baseline)	72776.96	34776.41	217	72102.74	30627.01	219	-674.22	[0.83]
Expected Wage (Baseline)	88543.78	42183.67	217	88363.01	41426.98	219	-180.77	[0.96]

We compare the means, standard deviation, and number of observations between treatment and control for a variety of pre-treatment variables. In the last two columns, we regress the demographic of interest on an indicator for treatment, and report the coefficient on treatment. Standard errors are clustered by major. The p-value for the coefficient is in brackets. Stars indicate: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 3: Regressions: Personal Wage Expectations on Treatment (First-Stage)

	(1) Log Change in Reservation Wage	(2) Log Change in Average Expected Wage
Treatment	0.087***	0.026
	(0.031)	(0.025)
Treatment x Δ Log P50	-0.231**	-0.369***
	(0.096)	(0.084)
Treatment x Δ Log IQR	-0.044*	-0.031
	(0.025)	(0.026)
Δ Log P50	-0.080*	0.019
	(0.045)	(0.017)
Δ Log IQR	0.024	0.004
	(0.023)	(0.017)
Constant	0.052	0.027
	(0.039)	(0.032)
Control Group Mean	-0.024	-0.018
R-squared	0.115	0.160
Number of observations	717	717

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We regress change in reservation wage and change in average expected wage on an indicator for treatment, treatment interacted with centrality measurement error (based on the median), treatment interacted with spread measurement error (based on the interquartile range), centrality measurement error, spread measurement error, and controls. Centrality measurement error, labeled Δ Log $P50$, is measured as log(Believed Average of the external wage distribution) - log(True Median of the external wage distribution), as we do not have data on the true average. Spread measurement error, labeled Δ Log IQR, is measured as log(Believed Interquartile Range) - log(True Interquartile Range). Controls include indicators for strata variables (male, white, STEM), balance variables (parental education level and financial aid status), and behavioral covariates (risk aversion and confidence). Standard errors are clustered at the major level.

Table 4: 2SLS: Job Outcomes

	(1) Log Total Salary	(2) Log Base Salary	(3) Received Bonus Salary	(4) Log Bonus Salary	(5) Employed	(6) Attending Graduate School
Log Post-Treatment Reservation Wage	1.283**	1.342*	1.262*	2.000	-0.564	-0.256
	(0.634)	(0.755)	(0.712)	(3.974)	(0.438)	(0.269)
Constant	-3.077	-3.850	-13.243*	-13.113	6.533	3.001
	(6.876)	(8.207)	(7.777)	(43.178)	(4.704)	(2.941)
Control Group Mean	11.462	11.346	0.576	9.739	0.535	0.207
First Stage: Anderson-Rubin Test (χ^2 statistic)	9.86	7.28	20.35	6.23	4.85	11.91
First Stage: Anderson-Rubin Test (pval)	.02	.06	0	.1	.18	.01
Number of observations	147	146	147	84	280	351
Anderson-Rubin 95% CI	[.304316, 3.76696]	[-1.61892, 4.30387]	[.613089, 4.05398]	[-13.5797, 17.5791]	[-2.28023, .112679]	[-.926775,-.564695]

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first stage regresses reservation wage on an indicator for treatment, treatment interacted with centrality measurement error (based on the median), treatment interacted with spread measurement error (based on the interquartile range), centrality measurement error, spread measurement error, and controls. The second stage regresses job outcome on predicted reservation wage, measurement errors, and controls. Centrality measurement error is measured as log(Believed Average of the external wage distribution) - log(True Median of the external wage distribution), as we do not have data on the true average. Spread measurement error is measured as log(Believed Interquartile Range) - log(True Interquartile Range). Controls include indicators for strata variables (male, white, STEM), balance variables (parental education level and financial aid status), and behavioral covariates (risk aversion and confidence). Standard errors are clustered at the major level. At the bottom of the table, we report (in order) control group mean, the first stage Anderson-Rubin χ^2 test for weak instruments, the correspondent p-value for the Anderson-Rubin test, number of observations, and the 95% Anderson-Rubin confidence interval.

Table 5: Overestimators vs. Underestimators: Change in Reservation Wage on Treatment x Estimation Error (First-Stage)

	(1) Overestimators	(2) Underestimators
Treatment	0.029	0.097**
	(0.030)	(0.037)
Treatment x Δ Log P50	-0.237	-0.267*
	(0.277)	(0.143)
Treatment x Δ Log IQR	0.001	-0.082**
	(0.034)	(0.038)
Δ Log P50	-0.127	-0.046
	(0.160)	(0.036)
Δ Log IQR	0.009	0.026
	(0.022)	(0.016)
Constant	-0.002	0.040
	(0.046)	(0.037)
Control Group Mean	-0.034	-0.016
R-squared	0.063	0.145
Number of observations	383	484

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We regress change in reservation wage and change in average expected wage on an indicator for treatment, treatment interacted with centrality measurement error (based on the median), treatment interacted with spread measurement error (based on the interquartile range), centrality measurement error, spread measurement error, and controls. Centrality measurement error, labeled Δ Log $P50$, is measured as log(Believed Average of the external wage distribution) - log(True Median of the external wage distribution), as we do not have data on the true average. Spread measurement error, labeled Δ Log IQR, is measured as log(Believed Interquartile Range) - log(True Interquartile Range). Controls include indicators for strata variables (male, white, STEM) and balance variables (parental education level and financial aid status). Behavioral covariates (risk aversion and confidence) were omitted due to lack of power. Standard errors are clustered at the major level.

Table 6: 2SLS: Job Outcomes - Split First Stage between Overestimators and Underestimators

	(1) Log Total Salary	(2) Log Base Salary	(3) Received Bonus Salary	(4) Log Bonus Salary	(5) Employed	(6) Attending Graduate School
Log Post-Treatment Reservation Wage	1.065**	1.174***	0.824***	0.136	-0.313	-0.089
	(0.458)	(0.448)	(0.317)	(1.367)	(0.320)	(0.219)
Constant	-0.851	-2.124	-8.727**	7.318	3.834	1.242
	(4.945)	(4.803)	(3.474)	(14.390)	(3.463)	(2.404)
Control Group Mean	11.462	11.346	0.576	9.739	0.535	0.207
First Stage: Anderson-Rubin Test (χ^2 statistic)	21.81	22.8	23.23	8.050000000000001	16.6	17.17
First Stage: Anderson-Rubin Test (pval)	0	0	0	.23	.02	.02
Number of observations	147	146	147	84	280	351
Anderson-Rubin 95% CI	[.357475, 2.86146]	[.37572, 2.9296]	[.385554, 2.06578]	[-5.22108, 1.70513]	[-1.49297,-.071752]	null set

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first stage regresses reservation wage on an indicator for treatment interacted with an indicator for overestimation of the center, treatment interacted with an indicator for overestimation of the range, treatment, treatment interacted with overestimation of the center interacted with the absolute value of centrality measurement error (based on the median), treatment interacted with underestimation of the center interacted with the absolute value of centrality measurement error (based on the median), treatment interacted with overestimation of the range interacted with the absolute value of spread measurement error (based on the interquartile range), treatment interacted with underestimation of the range interacted with the absolute value of spread measurement error (based on the interquartile range), an indicator for overestimation of the center, an indicator for overestimation of the range, centrality measurement error, spread measurement error, and controls. The second stage regresses job outcome on predicted reservation wage, measurement errors, and controls. Centrality measurement error is measured as log(Believed Average of the external wage distribution) - log(True Median of the external wage distribution), as we do not have data on the true average. Spread measurement error is measured as log(Believed Interquartile Range) - log(True Interquartile Range). Controls include indicators for strata variables (male, white, STEM), balance variables (parental education level and financial aid status), and behavioral covariates (risk aversion and confidence). Standard errors are clustered at the major level. At the bottom of the table, we report (in order) control group mean, the first stage Anderson-Rubin χ^2 test for weak instruments, the correspondent p-value for the Anderson-Rubin test, number of observations, and the 95% Anderson-Rubin confidence interval.

Table 7: Regression: Estimation Error on Demographic or Educational Background

Panel A: Estimation Error - Centrality (Median)	(1) Error: Log P50	(2) Error: Log P50	(3) Error: Log P50	(4) Error: Log P50
Gender: Female or Non-binary	0.025			
	(0.028)			
Race: Black, Hispanic, Native		-0.030		
		(0.043)		
Race: Non-white			0.006	
			(0.032)	
Major: High-paying				-0.187***
				(0.036)
Constant	-0.182***	-0.156**	-0.163**	-0.078
	(0.066)	(0.061)	(0.068)	(0.059)
R-squared	0.011	0.012	0.011	0.071
Number of observations	724.000	724.000	724.000	724.000
Panel B: Estimation Error - Spread (Interquartile Range)	(1) Error: Log IQR	(2) Error: Log IQR	(3) Error: Log IQR	(4) Error: Log IQR
Gender: Female or Non-binary	-0.034			
	(0.057)			
Race: Black, Hispanic, Native		-0.121		
		(0.082)		
Race: Non-white			0.013	
			(0.064)	
Major: High-paying				-0.219*
				(0.123)
Constant	0.774***	0.746***	0.726***	0.948***
	(0.160)	(0.154)	(0.157)	(0.110)
R-squared	0.036	0.041	0.036	0.057
Number of observations	717.000	717.000	717.000	717.000

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Panel A, we regress errors in log beliefs about the center of the distribution (calculated as log(Believed Average) - log(True Median)) on an indicator for demographic / education group. Column (1) shows the results for gender, columns (2) and (3) for race, and column (4) for major choice. Controls include indicators for strata variables (male, white, STEM - omitting the category of interest (e.g. male for the gender regression) to avoid collinearity), balance variables (parental education level and financial aid status), and behavioral covariates (risk aversion and confidence). Standard errors are clustered at the major level.

Table 8: 2SLS: Job Outcomes (Interacted with demographic/education variables)

Panel A: Gender	(1)	(2)	(3)	(4)	(5)	(6)
	Log Total Salary	Log Base Salary	Received Bonus Salary	Log Bonus Salary	Employed	Attending Graduate School
Log Post-Treatment Reservation Wage x Female or Non-binary	0.560	0.513	1.232	-2.314	0.446	0.449
	(0.882)	(0.843)	(1.160)	(2.666)	(0.840)	(0.576)
Log Post-Treatment Reservation Wage	1.038**	1.039*	0.534	4.055	-0.688	-0.525
	(0.526)	(0.535)	(0.500)	(2.670)	(0.617)	(0.429)
Female or Non-binary	-6.477	-5.845	-13.999	25.146	-5.051	-4.991
	(9.849)	(9.420)	(12.892)	(29.999)	(9.275)	(6.387)
Constant	-0.214	-0.448	-4.806	-35.517	7.941	6.115
	(5.862)	(5.914)	(5.606)	(30.083)	(6.846)	(4.828)
Control Group Mean	11.462	11.346	0.576	9.739	0.535	0.207
First-Stage χ^2-stat	8.15	8.07	8.15	5.13	6.12	6.04
First-Stage χ^2-stat (pval)	.148	.153	.148	.4	.294	.302
Number of observations	147	146	147	84	280	351
Panel B: URM (Black/Hispanic/Native)	(1)	(2)	(3)	(4)	(5)	(6)
	Log Total Salary	Log Base Salary	Received Bonus Salary	Log Bonus Salary	Employed	Attending Graduate School
Log Post-Treatment Reservation Wage x URM	-1.092	-0.240	-0.673	2.753*	0.420	0.111
	(2.049)	(1.757)	(1.213)	(1.652)	(0.802)	(0.451)
Log Post-Treatment Reservation Wage	0.983	0.730	0.225	-0.574	-0.590	-0.290
	(1.026)	(0.950)	(0.682)	(1.373)	(0.626)	(0.346)
Black, Hispanic, Native (URM)	12.209	2.698	7.703	-30.251	-4.445	-1.290
	(22.770)	(19.453)	(13.436)	(18.566)	(8.762)	(4.916)
Constant	-0.064	2.667	-2.234	13.832	6.609	3.440
	(10.951)	(10.230)	(7.306)	(14.286)	(6.743)	(3.792)
Control Group Mean	11.462	11.346	0.576	9.739	0.535	0.207
First-Stage χ^2-stat	1.67	1.64	1.67	3.57	5.06	6.13
First-Stage χ^2-stat (pval)	.893	.897	.893	.614	.408	.293
Number of observations	147	146	147	84	280	351
Panel C: Non-white	(1)	(2)	(3)	(4)	(5)	(6)
	Log Total Salary	Log Base Salary	Received Bonus Salary	Log Bonus Salary	Employed	Attending Graduate School
Log Post-Treatment Reservation Wage x Non-white	-0.140	-0.302	-0.137	1.796	-0.778	-0.221
	(0.907)	(0.906)	(1.145)	(1.288)	(0.514)	(0.373)
Log Post-Treatment Reservation Wage	1.136*	1.258*	1.045	-0.093	0.137	-0.158
	(0.646)	(0.763)	(1.027)	(0.348)	(0.360)	(0.237)
Non-white	1.630	3.350	1.335	-18.759	8.543	2.382
	(10.107)	(10.103)	(12.817)	(14.602)	(5.669)	(4.111)
Constant	-1.728	-3.121	-10.771	8.239**	-1.389	1.949
	(7.169)	(8.416)	(11.516)	(3.835)	(3.783)	(2.587)
Control Group Mean	11.462	11.346	0.576	9.739	0.535	0.207
First-Stage χ^2-stat	2.4	2.26	2.4	4.44	9.77	9.87
First-Stage χ^2-stat (pval)	.791	.812	.791	.488	.082	.079
Number of observations	147	146	147	84	280	351
Panel D: High-paying major	(1)	(2)	(3)	(4)	(5)	(6)
	Log Total Salary	Log Base Salary	Received Bonus Salary	Log Bonus Salary	Employed	Attending Graduate School
Log Post-Treatment Reservation Wage x High-paying major	-1.411	-0.817	0.597	16.233	-1.077	-0.184
	(1.605)	(1.220)	(1.150)	(24.721)	(1.046)	(0.749)
Log Post-Treatment Reservation Wage	1.201	0.909	-0.673	-13.765	0.124	-0.114
	(0.796)	(0.644)	(0.510)	(24.374)	(0.405)	(0.706)
Major: High-paying	15.755	9.179	-6.024	-175.513	12.204	2.005
	(17.574)	(13.348)	(12.664)	(267.262)	(11.557)	(8.054)
Constant	-1.965	1.120	7.608	157.679	-0.968	1.483
	(8.591)	(6.947)	(5.544)	(263.883)	(4.310)	(7.531)
Control Group Mean	11.462	11.346	0.576	9.739	0.535	0.207
First-Stage χ^2-stat	4.25	3.96	4.25	.55	5.68	7.31
First-Stage χ^2-stat (pval)	.514	.555	.514	.758	.339	.199
Number of observations	147	146	147	84	280	351

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In this specification, we use two endogenous variables: Log Reservation Wage and Log Reservation Wage interacted with an indicator for the relevant demographic variable. Instruments are the same as before. Controls include indicators for strata variables (male, white, STEM), balance variables (parental education level and financial aid status), and behavioral covariates (risk aversion and confidence). Standard errors are clustered at the major level. At the bottom of the table, we report (in order) control group mean, the first-stage Sanderson-Windmeijer χ^2 test for weak instruments and correspondent p-value, and number of observations.

10 Appendix A: Model Derivation and Proofs

In Appendix A, we solve for reservation wages as a function of the perceived wage distribution. We derive the comparative statics and propositions discussed in Section 2.

10.1 Model Derivation

Briefly re-summarizing the model, workers in each period t receive an employment offer with probability λ. An offer pays wage w, randomly drawn from wage distribution $F(w)$. Workers maximize expected utility:

$$\max_{\{x_t\}} \sum_{t=0}^{\infty} \beta^t u(x_t), \beta \in (0,1) \tag{1}$$

where:

$$x(t) = \begin{cases} w, & \text{accept job, receive wage } w \\ b, & \text{reject job, receive unemployment} \end{cases}$$

A worker chooses between (1) accepting the job and receiving wage w for the rest of time t, or (2) rejecting the job, receiving b, and continuing the search next period.

Using value functions, an employed worker receives:

$$W(w) = u(w) + \beta W(w) \implies W(w) = \frac{u(w)}{1-\beta} \tag{2}$$

And an unemployed worker receives:

$$U(\mu_i, \sigma_i) = u(b) + \beta\lambda \int_w \max\{W_{t+1}(w|\mu_i, \sigma_i), U_{t+1}(\mu_i, \sigma_i)\} dF(w|\mu_i, \sigma_i) + \beta(1-\lambda) U_{t+1}(\mu_i, \sigma_i) \tag{3}$$

A worker is indifferent between accepting wage w and rejecting wage w at the reservation wage w_R. Thus, at the reservation wage w_R,

$$W(w_R|\mu_i, \sigma_i) = \frac{u(w_R)}{1-\beta} = U(\mu_i, \sigma_i) \tag{4}$$

Then, we know:

$$\frac{u(w_R)}{1-\beta} = u(b) + \beta\lambda \int_w \max\{W_{t+1}(w|\mu_i, \sigma_i), U_{t+1}(\mu_i, \sigma_i)\} dF(w|\mu_i, \sigma_i) + \beta(1-\lambda) U_{t+1}(\mu_i, \sigma_i) \tag{5}$$

For wages $w \leq w_R$, $U_{t+1}(\mu_i, \sigma_i) \geq W_{t+1}(w|\mu_i, \sigma_i)$ and the worker rejects the job. For $w \geq w_R$, $U_{t+1}(\mu_i, \sigma_i) \leq W_{t+1}(w|\mu_i, \sigma_i)$ and the worker accepts the job. Then, we can simplify equation (5):

$$\frac{u(w_R)}{1-\beta} = u(b) + \beta \int_0^{w_R} U_{t+1}(\mu_i, \sigma_i) dF(w|\mu_i, \sigma_i) + \beta\lambda \int_{w_R}^{\infty} W_{t+1}(w|\mu_i, \sigma_i) dF(w|\mu_i, \sigma_i) \tag{6}$$

Substituting (2) and (4) into (6), we get:

$$\frac{u(w_R)}{1-\beta} = u(b) + \beta \int_0^{w_R} \frac{u(w_R)}{1-\beta}(\mu_i, \sigma_i) dF(w|\mu_i, \sigma_i) + \beta\lambda \int_{w_R}^{\infty} \frac{u(w)}{1-\beta} dF(w|\mu_i, \sigma_i) \tag{7}$$

We can then rewrite the left-hand side and simplify:

$$\frac{u(w_R)}{1-\beta}\int_0^{w_R} dF(w|\mu_i,\sigma_i) + \frac{u(w_R)}{1-\beta}\int_{w_R}^{\infty} dF(w|\mu_i,\sigma_i) =$$

$$u(b) + \beta\int_0^{w_R}\frac{u(w_R)}{1-\beta}dF(w|\mu_i,\sigma_i) + \beta\lambda\int_{w_R}^{\infty}\frac{u(w)}{1-\beta}dF(w|\mu_i,\sigma_i)$$

$$u(w_R)\int_0^{w_R} dF(w|\mu_i,\sigma_i) + \frac{u(w_R)}{1-\beta}\int_{w_R}^{\infty} dF(w|\mu_i,\sigma_i) = u(b) + \beta\lambda\int_{w_R}^{\infty}\frac{u(w)}{1-\beta}dF(w|\mu_i,\sigma_i)$$

$$u(w_R)\int_0^{w_R} dF(w|\mu_i,\sigma_i) - u(b) = u(b) + \int_{w_R}^{\infty}\frac{\beta\lambda u(w) - u(w_R)}{1-\beta}dF(w|\mu_i,\sigma_i)$$

$$u(w_R)\int_0^{w_R} dF(w|\mu_i,\sigma_i) - u(b) + u(w_R)\int_{w_R}^{\infty} dF(w|\mu_i,\sigma_i) = u(b) + \int_{w_R}^{\infty}\frac{\beta\lambda u(w) - u(w_R) + (1-\beta)u(w_R)}{1-\beta}dF(w|\mu_i,\sigma_i)$$

$$u(w_R) - u(b) = \int_{w_R}^{\infty}\frac{\beta\lambda u(w) - \beta u(w_R)}{1-\beta}dF(w|\mu_i,\sigma_i)$$

$$u(w_R) - u(b) = \frac{\beta}{1-\beta}\int_{w_R}^{\infty}\big(\lambda u(w) - u(w_R)\big)dF(w|\mu_i,\sigma_i) \tag{8}$$

From equation (8), we know that the reservation wage is the wage at which the expected cost of searching again (rejecting w_R, left hand side) equals the expected benefit of searching again (right hand side). We can define the right hand side of the equation as $g(w_R) = \frac{\beta}{1-\beta}\int_{w_R}^{\infty}\big(\lambda u(w) - u(w_R)\big)dF(w|\mu_i,\sigma_i)$. Then, the general $g(\cdot)$ function would be:

$$g(s) = \frac{\beta}{1-\beta}\int_{s}^{\infty}\big(\lambda u(w) - u(s)\big)dF(w|\mu_i,\sigma_i) \tag{9}$$

Differentiating $g(s)$, we get:

$$g'(s) = -\frac{\beta}{1-\beta}\big(1-F(s)\big)u'(s) < 0 \tag{10}$$

This tells us that equation (8) has a unique solution. Differentiating once more (assuming our wage distribution $F(s)$ has density $f(s)$):

$$g''(s) = -\frac{\beta}{1-\beta}\big(1-F(s)\big)u''(s) + \frac{\beta}{1-\beta}f(s)u'(s) \geq 0 \tag{11}$$

So $g(s)$ is convex.

From equation (8), we can also re-formulate the expression to more easily derive proofs of Proposition 1A, 1B, and 2. To do so, we add and subtract $\frac{\beta}{1-\beta}\int_0^{w_R}\big(\lambda u(w) - u(w_R)\big)dF(w)$ to the right-hand side of (8):

$$u(w_R) - u(b) = \frac{\beta}{1-\beta}\int_0^{\infty}\big(\lambda u(w) - u(w_R)\big)dF(w|\mu_i,\sigma_i) - \frac{\beta}{1-\beta}\int_0^{w_R}\big(\lambda u(w) - u(w_R)\big)dF(w|\mu_i,\sigma_i)$$

$$u(w_R) - u(b) = \frac{\beta\lambda u(E[w|\mu_i,\sigma_i])}{1-\beta} - \frac{\beta u(w_R)}{1-\beta} - \frac{\beta}{1-\beta}\int_0^{w_R}\big(\lambda u(w) - u(w_R)\big)dF(w|\mu_i,\sigma_i)$$

$$(1-\beta)u(w_R) - (1-\beta)u(b) = \beta\lambda u(E[w|\mu_i,\sigma_i]) - \beta u(w_R) - \beta\int_0^{w_R}\big(\lambda u(w) - u(w_R)\big)dF(w|\mu_i,\sigma_i)$$

$$u(w_R) - u(b) = \beta\Big(\lambda u(E[w|\mu_i,\sigma_i]) - u(b)\Big) - \beta\int_0^{w_R}\big(\lambda u(w) - u(w_R)\big)dF(w|\mu_i,\sigma_i) \tag{12}$$

Using integration by parts, we know $\int_0^{w_R} u(w)dF(w|\mu_i,\sigma_i) = u(w)F(w)|_0^{w_R} - \int_0^{w_R} u'(w)F(w)dw$, which

simplifies to $\int_0^{w_R} u(w)dF(w|\mu_i,\sigma_i) = u(w_R)F(w_R) - \int_0^{w_R} u'(w)F(w)dw$. Then, equation (12) becomes:

$$u(w_R) - u(b) = \beta\Big(\lambda u(E[w|\mu_i,\sigma_i]) - u(b)\Big) + \beta\lambda \int_0^{w_R} u'(w)F(w|\mu_i,\sigma_i)dw + \beta(1-\lambda)u(w_R)F(w_R) \quad (13)$$

Rearranging, we have:

$$\Big(1-\beta(1-\lambda)\Big)u(w_R) - u(b) = \beta\Big(\lambda u(E[w|\mu_i,\sigma_i]) - u(b)\Big) + \beta\lambda \int_0^{w_R} u'(w)F(w|\mu_i,\sigma_i)dw \quad (14)$$

We can use equation 14 to prove our propositions.

10.2 Impact of Information

There are two ways in which information, defined as different expectations of the wage distribution, can impact this model:

1. First order stochastic dominance (FOSD): Differences in beliefs that occur via a shift of the entire wage distribution, i.e. the wage distribution has the same shape, but one person believes it's shifted to the right of the true distribution (overestimation) and one person believes it's shifted to the left of the true distribution (underestimation).
2. Second order stochastic dominance (SOSD): A shift from F to $\tilde{F}$ that is a mean-preserving spread, where $\int_0^{w_R} \tilde{F(w)}dw > \int_0^{w_R} F(w)dw$.

Our proofs characterize both.

10.3 Proofs of Propositions

Proposition (1A). *Holding all else constant, reservation wages $w_R(\mu_i,\sigma_i)$ are increasing in $\mu_{\epsilon i}$, that is $\frac{\partial w_R}{\partial \mu_{\epsilon i}} > 0$.*

We can implicitly differentiate equation (14) to find $\frac{\partial w_R}{\partial \mu_{\epsilon i}}$.

$$\Big(1-\beta(1-\lambda)\Big)\frac{\partial u(w_R)}{\partial w_R}\frac{\partial w_R}{\partial \mu_{\epsilon i}} = \beta\lambda\frac{\partial u(E[w|\mu_i,\sigma_i])}{\partial E[w|\mu_i,\sigma_i]}\frac{\partial E[w|\mu_i,\sigma_i]}{\partial \mu_{\epsilon i}} + \beta\lambda F(w_R)\frac{\partial u(w_R)}{\partial w_R}\frac{\partial w_R}{\partial \mu_{\epsilon i}}$$

Rearranging, we get:

$$(1-\beta)\frac{\partial u(w_R)}{\partial w_R}\frac{\partial w_R}{\partial \mu_{\epsilon i}} + \beta\lambda[1-F(w_R)]\frac{\partial u(w_R)}{\partial w_R}\frac{\partial w_R}{\partial \mu_{\epsilon i}} = \beta\lambda\frac{\partial u(E[w|\mu_i,\sigma_i])}{\partial E[w|\mu_i,\sigma_i]}\frac{\partial E[w|\mu_i,\sigma_i]}{\partial \mu_{\epsilon i}}$$

which simplifies to:

$$\Big((1-\beta)+\beta\lambda[1-F(w_R)]\Big)\frac{\partial u(w_R)}{\partial w_R}\frac{\partial w_R}{\partial \mu_{\epsilon i}} = \beta\lambda\frac{\partial u(E[w|\mu_i,\sigma_i])}{\partial E[w|\mu_i,\sigma_i]}\frac{\partial E[w|\mu_i,\sigma_i]}{\partial \mu_{\epsilon i}}$$

On the left-hand side, we know $(1-\beta) > 0$, $\beta > 0$, $\lambda > 0$, $[1-F(w_R)] > 0$, and $\frac{\partial u(w_R)}{\partial w_R} > 0$. Thus, $\Big((1-\beta)+\beta\lambda[1-F(w_R)]\Big)\frac{\partial u(w_R)}{\partial w_R} > 0$. On the right-hand side, we know $\beta > 0$, $\lambda > 0$, $\frac{\partial u(E[w|\mu_i,\sigma_i])}{\partial E[w|\mu_i,\sigma_i]} > 0$,

and since under log normality $E[w|\mu_i, \sigma_i] = \exp[\mu_w + \mu_{\epsilon i} + \frac{\sigma_w^2 + \sigma_{\epsilon_i}^2}{2}]$, $\frac{\partial E[w|\mu_i,\sigma_i]}{\partial \mu_{\epsilon i}} > 0$. Thus, $\frac{\partial w_R}{\partial \mu_{\epsilon i}} > 0$. $\square$

Proposition (1B). *Holding all else constant, reservation wages $w_R(\mu, \sigma)$ are increasing in $\sigma_{\epsilon i}^2$, that is $\frac{\partial w_R}{\partial \sigma_{\epsilon i}^2} > 0$.* Similar to Proposition (1A), we can implicitly differentiate equation (14) to find $\frac{\partial w_R}{\partial \sigma_{\epsilon i}^2}$.

$$\Big(1 - \beta(1-\lambda)\Big)\frac{\partial u(w_R)}{\partial w_R}\frac{\partial w_R}{\partial \sigma_{\epsilon i}^2} = \beta\lambda\frac{\partial u(E[w|\mu_i,\sigma_i])}{\partial E[w|\mu_i,\sigma_i]}\frac{\partial E[w|\mu_i,\sigma_i]}{\partial \sigma_{\epsilon i}^2} + \beta\lambda F(w_R)\frac{\partial u(w_R)}{\partial w_R}\frac{\partial w_R}{\partial \sigma_{\epsilon i}^2}$$

Rearranging and simplifying, we get:

$$\Big((1-\beta) + \beta\lambda[1 - F(w_R)]\Big)\frac{\partial u(w_R)}{\partial w_R}\frac{\partial w_R}{\partial \sigma_{\epsilon i}^2} = \beta\lambda\frac{\partial u(E[w|\mu_i,\sigma_i])}{\partial E[w|\mu_i,\sigma_i]}\frac{\partial E[w|\mu_i,\sigma_i]}{\partial \sigma_{\epsilon i}^2}$$

On the left-hand side, we know $(1-\beta) > 0$, $\beta > 0$, $\lambda > 0$, $[1 - F(w_R)] > 0$, and $\frac{\partial u(w_R)}{\partial w_R} > 0$. Thus, $\Big((1-\beta) + \beta\lambda[1 - F(w_R)]\Big)\frac{\partial u(w_R)}{\partial w_R} > 0$. On the right-hand side, we know $\beta > 0$, $\lambda > 0$, $\frac{\partial u(E[w|\mu_i,\sigma_i])}{\partial E[w|\mu_i,\sigma_i]} > 0$, and since under log normality $E[w|\mu_i, \sigma_i] = \exp[\mu_w + \mu_{\epsilon i} + \frac{\sigma_w^2 + \sigma_{\epsilon_i}^2}{2}]$, $\frac{\partial E[w|\mu_i,\sigma_i]}{\partial \sigma_{\epsilon i}^2} > 0$. Thus, $\frac{\partial w_R}{\partial \mu_{\epsilon i}} > 0$. $\square$

Proposition (2). *The dispersion of reservation wages $w_R(\mu, \sigma)$ across agents is higher in a model with subjective beliefs than in a model where every worker has perfect information and derives a reservation wage $w_R(\mu^*, \sigma^*)$.*

Proof. Recall w is drawn from a distribution with mean $\mu_w + \mu_{\epsilon i}$ and variance $\sigma_w^2 + \sigma_{\epsilon i}^2$. We can thus simplify $w_R(\mu, \sigma) = w_R(\epsilon)$. Under a compound probability distribution,

$$p_H(w_R) = \int p_F(w_R|\epsilon)p_g(\epsilon)$$

A compound distribution's second moment is given by:

$$\begin{aligned} Var_H(w_R) &= E_G[Var_F(w_R|\epsilon)] + Var_G(E_F[w_R|\epsilon]) \\ Var_H(w_R) &= Var_F(w_R|\epsilon = 0)] + Var_G(\epsilon) \\ Var_H(w_R) &= Var_F(w_R|\mu^*, \sigma^*) + \sigma^2 \end{aligned}$$

This results in a wider variance. $\square$

Appendix B: Additional Figures and Tables

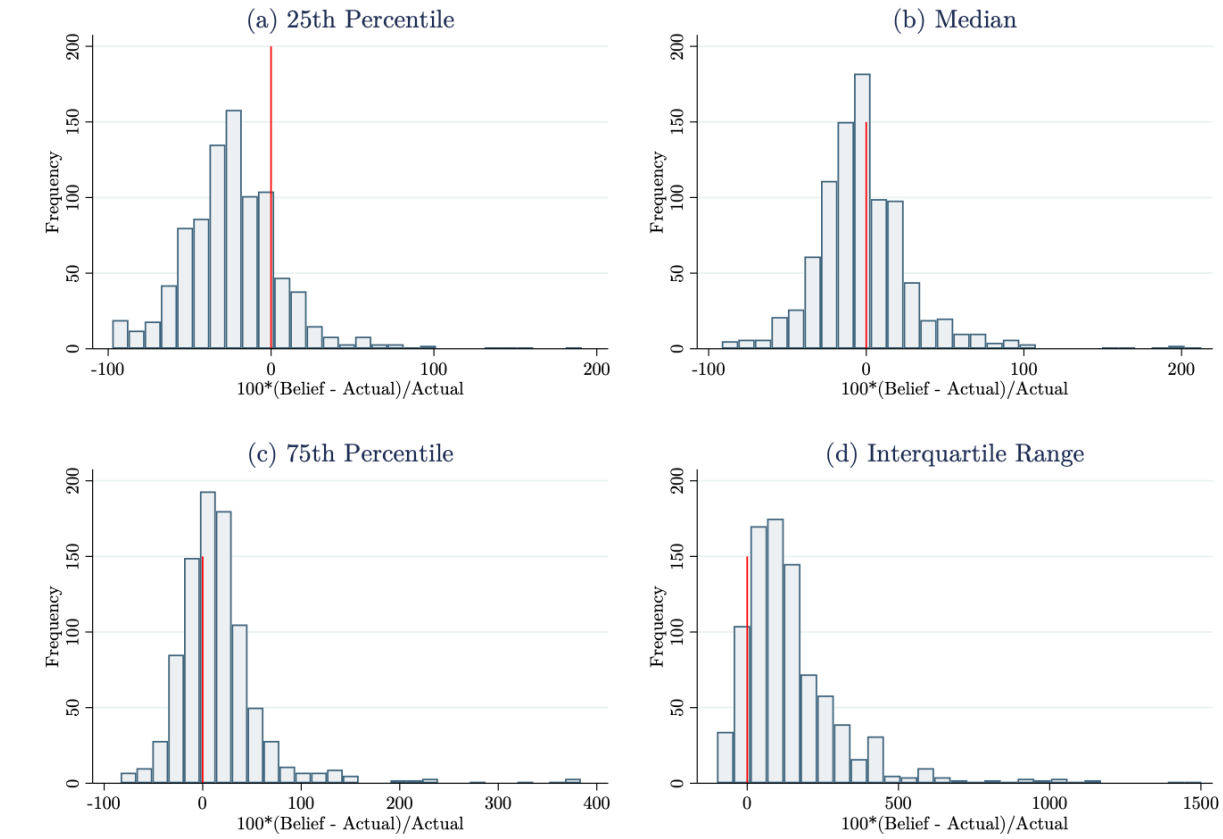


Figure B1: Percent Deviation of Beliefs from True Salary (Within-Major).

Histogram of the error in students' pre-treatment beliefs about the major-specific distribution. Students were asked about the 25th percentile, average, and 75th percentile wage that students graduating within their major at UC Berkeley received last year. Graph (a) shows error in beliefs about the 25th percentile. Graph (b) shows beliefs about the average - the true median. Graph (c) shows error in beliefs about the 75th percentile. Graph (d) shows the range, calculated using 75th percentile - 25th percentile. The red line at 0 indicates no error in beliefs, i.e. the situation where studets have perfect information. Negative values in this histogram indicate underestimation, while positive values indicate overestimation.

Questions asked took the following format: "What do you believe was the 25th percentile (bottom 25%) annual salary for students graduating last year within your major at UC Berkeley? The 25th percentile is the salary of the bottom 25% of your major. Thus, if your major contains 100 people, and you rank them in order of salary from 1 (lowest) to 100 (highest), the 25th percentile salary would be that of the 25th person."

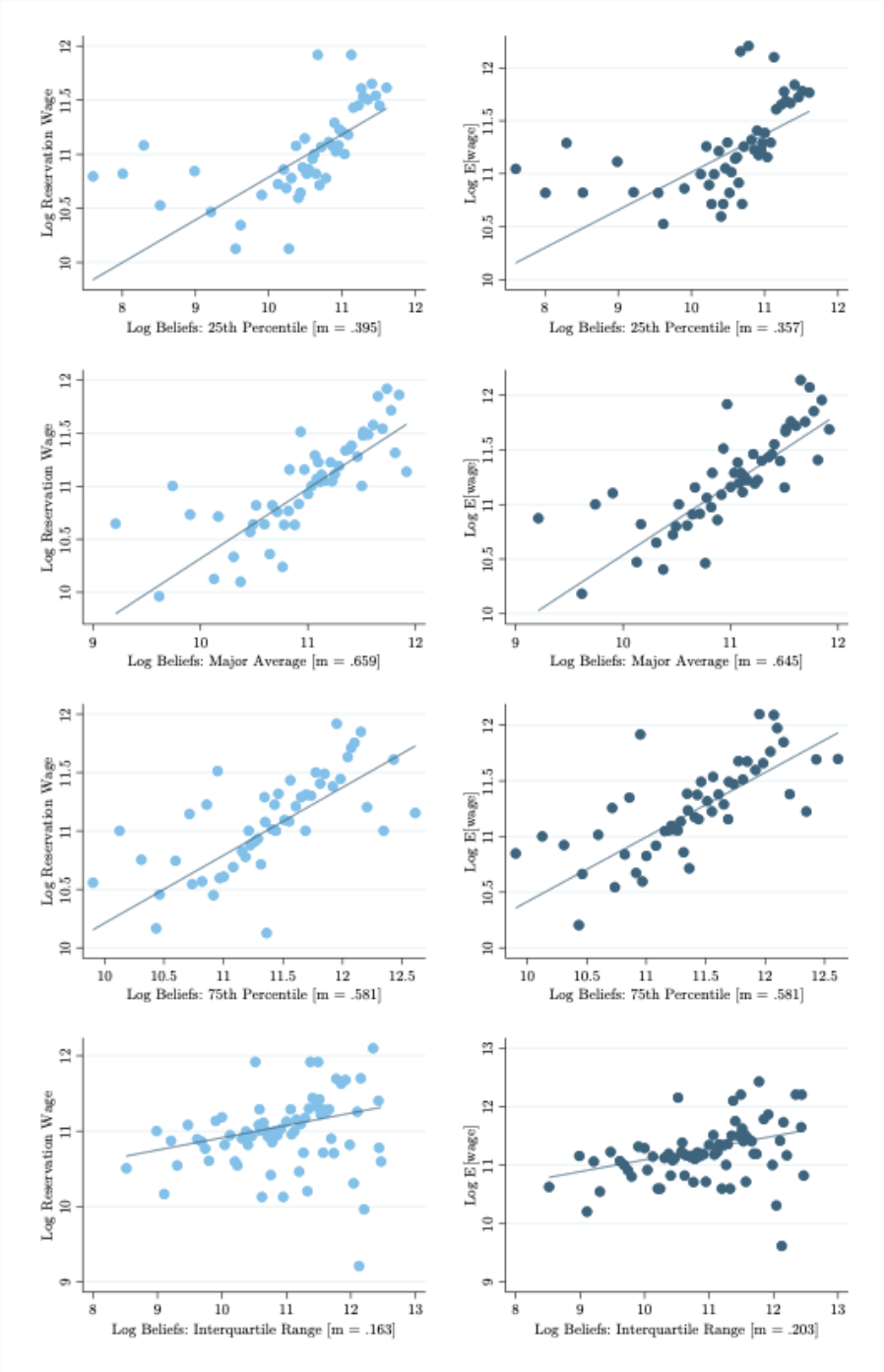


Figure B2: Personal Wage Expectations v. Beliefs about External Wage Distribution (In Levels)

Scatter plots comparing reservation wage (in light blue) versus log beliefs about the major-specific external wage distribution, as well as average expected wage (in dark blue) versus log beliefs about the major-specific external wage distribution. Scatters are binned at the discrete level. In order of top row to bottom row, the independent variable is: (1) beliefs about the major-specific 25th percentile, (2) beliefs about the major-specific average, (3) beliefs about the major-specific 75th percentile, and (4) the interquartile range, computed via (3) - (1). Relationships are positively correlated, with slope m in the x-axis for each graph.

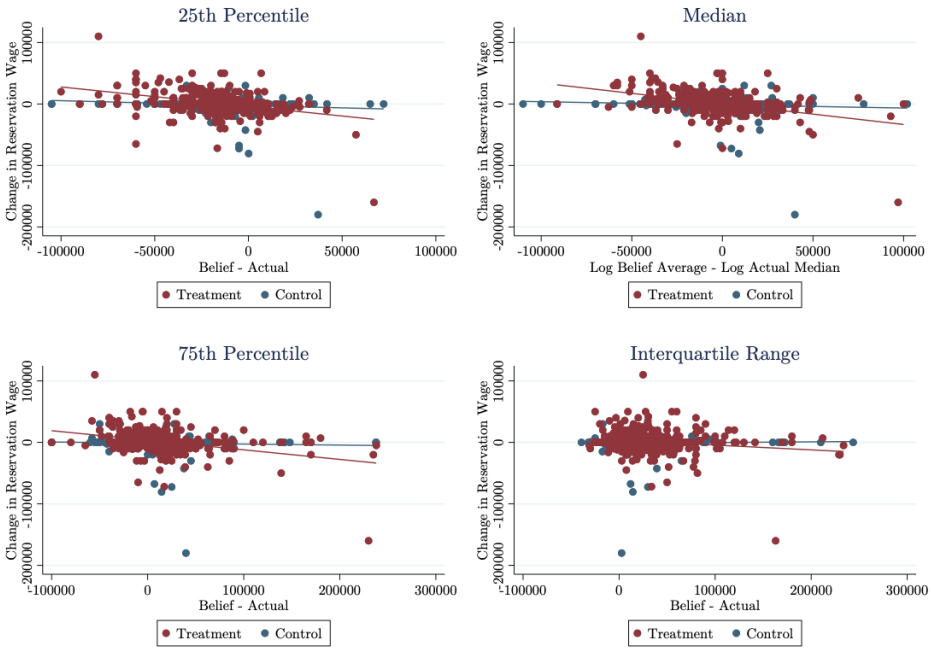


Figure B3: Post-Treatment Change in Reservation Wage (In Levels).

We take two measures of the reservation wage - one post-treatment and one pre-treatment (the control group is simply asked twice for their reservation wages). Reservation wages are asked as "What is the minimum annual salary you would require in order to accept a job?" The figure plots the change in reservation wage over the error in beliefs (belief - actual, in levels). Red dots indicate the treatment group, while blue dots indicate the control group. A linear relationship is graphed for both - as expected, the treatment group has a negative relationship, while the slope for the control group is flat.

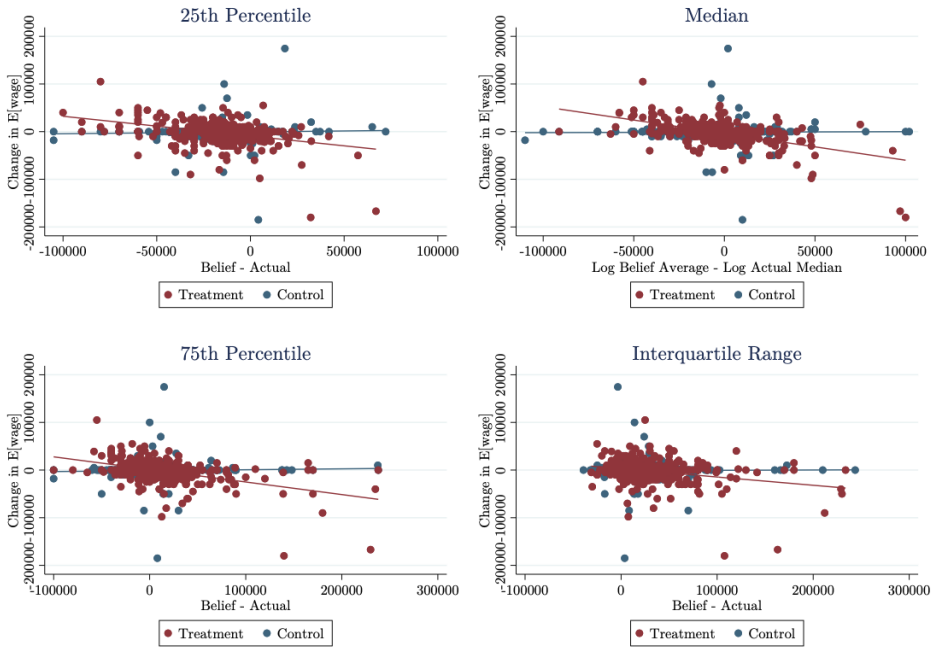


Figure B4: Post-Treatment Change in Average Expected Wage (In Levels).

We take two measures of the average expected wage - one post-treatment and one pre-treatment (the control group is asked at the same survey page timing as the treatment group). Average expected wage is asked as "What is the annual salary you expect to accept?" The figure plots the change in log reservation wage over the error in log beliefs (log belief - log actual). Red dots indicate the treatment group, while blue dots indicate the control group. A linear relationship is graphed for both - as expected, the treatment group has a negative relationship, while the slope for the control group is flat.

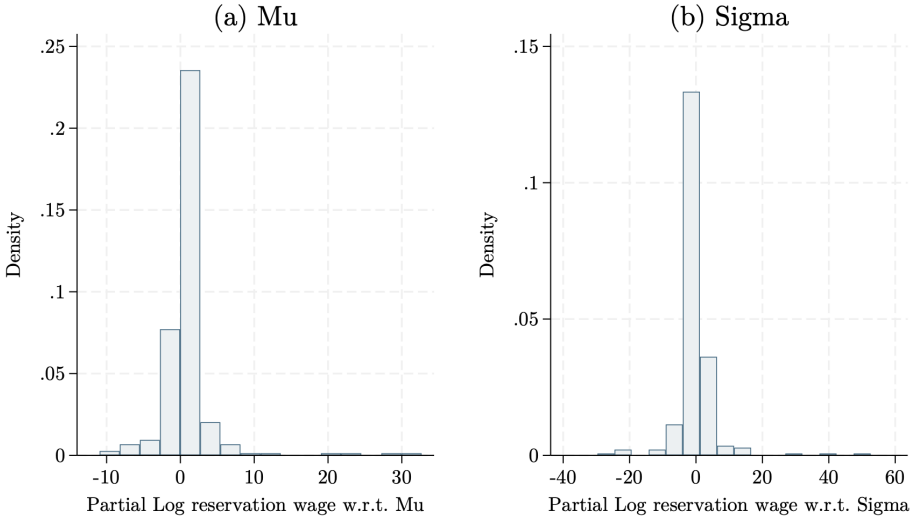


Figure B5: Empirical Estimates of $\frac{\partial \log w_R}{\partial \mu}$ and $\frac{\partial \log w_R}{\partial \sigma}$.

Version of Figure 10, using the full data. Figure B5 solves for the slope of the reservation wage with respect to μ, $\frac{\partial \log w_R}{\partial \mu}$, and the slope of the reservation wage with respect to σ, $\frac{\partial \log w_R}{\partial \sigma}$, using treatment students' beliefs about the 25th percentile and 75th percentile earnings compared to true 25th percentile and 75th percentile earnings in the denominator, and the difference in post-treatment versus pre-treatment reservation wages in the numerator. We take the assumption that to have a relevant slope, students must have responded the information in some way, so we drop in this exercise any student who had a reservation wage change in 0. In the opposite direction, we trim the slopes at the 1% level to avoid outliers of students who greatly changed their reservation wages.

Table B1: Regressions: Personal Wage Expectations on Treatment (First-Stage), Alternative Specification

	Log Change in Reservation Wage		Log Change in Average Expected Wage	
	(1)	(2)	(3)	(4)
Treatment	0.096***	0.086**	0.036	0.010
	(0.031)	(0.034)	(0.025)	(0.028)
Treatment x Δ Log P25	-0.091*		-0.163***	
	(0.051)		(0.044)	
Treatment x Δ Log P75		-0.250**		-0.470***
		(0.111)		(0.097)
Treatment x Δ Log IQR	-0.088***	0.012	-0.098***	0.091**
	(0.023)	(0.046)	(0.029)	(0.043)
Δ Log P25	-0.046		0.012	
	(0.028)		(0.017)	
Δ Log P75		-0.108*		0.048
		(0.058)		(0.043)
Δ Log IQR	0.009	0.055	0.006	-0.012
	(0.019)	(0.038)	(0.016)	(0.030)
Constant	0.047	0.057	0.027	0.045
	(0.036)	(0.037)	(0.038)	(0.038)
Control Group Mean		-0.024		-0.018
R-squared	0.086	0.100	0.107	0.144
Number of observations	717	717	717	717

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This is a robustness version of Table 3. We regress change in reservation wage and change in average expected wage on an indicator for treatment, treatment interacted with centrality measurement error (based on the 25th percentile in columns (1) and (3) and the 75th percentile in columns (2) and (4)), treatment interacted with spread measurement error (based on the interquartile range), centrality measurement error, spread measurement error, and controls. Centrality measurement error is measured as log(Believed Percentile) - log(True Percentile). Spread measurement error, labeled Δ Log IQR, is measured as log(Believed Interquartile Range) - log(True Interquartile Range). Controls include indicators for strata variables (male, white, STEM), balance variables (parental education level and financial aid status), and behavioral covariates (risk aversion and confidence). Standard errors are clustered at the major level.

Table B2: Reduced form: Job Outcomes

	(1) Log Post-Treatment Reservation Wage	(2) Log Total Salary	(3) Log Base Salary	(4) Received Bonus Salary	(5) Log Bonus Salary	(6) Employed	(7) Attending Graduate School
Treatment	0.089**	-0.159	-0.137	-0.137	0.357	-0.038	-0.152**
	(0.039)	(0.136)	(0.110)	(0.124)	(0.238)	(0.110)	(0.059)
Treatment x Δ Log P50	-0.209*	-0.601*	-0.630**	-0.695***	0.508	0.321*	-0.214
	(0.108)	(0.327)	(0.283)	(0.175)	(0.605)	(0.189)	(0.132)
Treatment x Δ Log IQR	-0.039	0.035	0.058	0.149	-0.429*	0.003	0.177***
	(0.054)	(0.106)	(0.088)	(0.106)	(0.208)	(0.128)	(0.061)
Δ Log P50	0.152	-0.036	0.038	0.026	-0.348	-0.031	0.123
	(0.108)	(0.147)	(0.100)	(0.123)	(0.303)	(0.097)	(0.081)
Δ Log IQR	0.039	0.053	0.038	0.025	0.216	0.055	-0.105**
	(0.072)	(0.064)	(0.052)	(0.077)	(0.189)	(0.069)	(0.046)
Constant	10.807***	10.898***	10.769***	0.467**	8.477***	0.480***	0.295**
	(0.095)	(0.249)	(0.229)	(0.178)	(0.613)	(0.162)	(0.118)
Control Group Mean	10.991	11.462	11.346	0.576	9.739	0.535	0.207
R-squared	0.150	0.424	0.447	0.275	0.376	0.119	0.105
Number of observations	717	147	146	147	84	280	351

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Reduced form impacts on job outcomes. We regress job outcome on treatment, treatment interacted with centrality measurement error (based on the median), treatment interacted with spread measurement error (based on the interquartile range), centrality measurement error, spread measurement error, and controls. Centrality measurement error, labeled Δ Log $P50$, is measured as log(Believed Average of the external wage distribution) - log(True Median of the external wage distribution), as we do not have data on the true average. Spread measurement error, labeled Δ Log IQR, is measured as log(Believed Interquartile Range) - log(True Interquartile Range). Controls include indicators for strata variables (male, white, STEM), balance variables (parental education level and financial aid status), and behavioral covariates (risk aversion and confidence). Standard errors are clustered at the major level. A negative coefficient on Treatment x Estimation Error indicates that those who overestimated will have decreased job outcome Y, while those who underestimated will have increased job outcome Y.

Table B3: 2SLS: Job Outcomes (Robustness)

Panel A: Centrality using 25th Percentile	(1)	(2)	(3)	(4)	(5)	(6)
	Log Total Salary	Log Base Salary	Received Bonus Salary	Log Bonus Salary	Employed	Attending Graduate School
Log Post-Treatment Reservation Wage	0.228	0.044	-1.163	1.329*	-0.477	-0.350
	(0.762)	(0.722)	(1.452)	(0.765)	(0.507)	(0.222)
Constant	8.315	10.227	12.992	-6.005	5.640	4.037*
	(8.301)	(7.888)	(15.811)	(8.279)	(5.433)	(2.451)
Control Group Mean	11.462	11.346	0.576	9.739	0.535	0.207
First Stage: Anderson-Rubin Test (χ^2 statistic)	4.91	3.3	5.58	6.100000000000001	2.36	7.06
First Stage: Anderson-Rubin Test (pval)	.18	.35	.13	.11	.5	.07
Number of observations	147	146	147	84	280	351
Anderson-Rubin 95% CI	[-2.75975, 3.21631]	[-2.78675, .815971] U [1.38783, 2.87467]	[-6.85612, .044466] U [1.88462, 4.52985]	[-.338183, 4.32917]	[-2.46426, .345679]	[-1.22061, .046002]
Panel B: Centrality using 75th Percentile	(1)	(2)	(3)	(4)	(5)	(6)
	Log Total Salary	Log Base Salary	Received Bonus Salary	Log Bonus Salary	Employed	Attending Graduate School
Log Post-Treatment Reservation Wage	1.605	1.790	1.065	2.197	-0.570	-0.262
	(1.107)	(1.497)	(0.953)	(1.555)	(0.465)	(0.198)
Constant	-6.584	-8.688	-11.125	-15.330	6.582	3.086
	(11.959)	(16.186)	(10.313)	(16.655)	(4.973)	(2.176)
Control Group Mean	11.462	11.346	0.576	9.739	0.535	0.207
First Stage: Anderson-Rubin Test (χ^2 statistic)	11.71	13.07	13.86	9.18	5.3	6.96
First Stage: Anderson-Rubin Test (pval)	.01	0	0	.03	.15	.07
Number of observations	147	146	147	84	280	351
Anderson-Rubin 95% CI	[-2.73409,-.805663] U [.684484, 5.94383]	[-4.0785,-.996216] U [.900571, 7.65788]	[-2.67033,-.708384] U [1.1781, 4.80014]	[.289011, 8.29142]	[-2.39448, .11132]	[-.897802, .059334]

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first stage regresses reservation wage on an indicator for treatment, treatment interacted with centrality measurement error (based on the 25th percentile in Panel A and the 75th percentile in Panel B), treatment interacted with spread measurement error (based on the interquartile range), centrality measurement error, spread measurement error, and controls. The second stage regresses job outcome on predicted reservation wage, measurement errors, and controls. Centrality measurement error is measured as log(Believed Percentile) - log(True Percentile). Spread measurement error is measured as log(Believed Interquartile Range) - log(True Interquartile Range). Controls include indicators for strata variables (male, white, STEM), balance variables (parental education level and financial aid status), and behavioral covariates (risk aversion and confidence). Standard errors are clustered at the major level. At the bottom of the table, we report (in order) control group mean, the first stage Anderson-Rubin χ^2 test for weak instruments, the correspondent p-value for the Anderson-Rubin test, number of observations, and the 95% Anderson-Rubin confidence interval.