The Illusion of Time: Gender Gaps in Job Search and Employment *

Preliminary Draft

Oriana Bandiera† Amen Jalal‡ Nina Roussille§

Click here for the most recent version

October 2024

Abstract

We study why women's educational gains are not reflected in labor market progress in the context of Pakistan. Leveraging surveys of ∼ 2,400 students fielded the month before their graduation from college, we first show that women's self-assessed likelihood of working six months later is similar to their male peers', at ~72%. By contrast, we find large employment gaps six months post graduation: only 37% of women were employed, vs. 64% of men. We find that the pipeline from education to employment only leaks at the very last stage: the decision to accept a job offer. Next, we uncover one key predictor of this decision for women: job search timing. Specifically, we find that applying promptly after graduation is highly predictive of women's future employment, while this is not true for men. To causally estimate the effect of timing, we experimentally shift the job applications of the next cohort of students closer to graduation. The intervention increases women's likelihood of working by 7.4ppt (22%) six months post graduation, and by 9.2 ppt (45%) when we restrict to full-time, firm employment. In contrast, the intervention shifts men's application earlier but has no effect on their employment outcomes. The experiment also provides suggestive evidence on the mechanism underlying women's overoptimism: their lack of foresight about competing forces from the marriage market.

*We are very grateful to Abi Adams, Natalia Emanuel, Camille Landais, Emily Nix, Anna Raute, Kate Vyborny and numerous seminar participants as well as the NBER SI Development Economics audience for comments and discussions. We thank Gabriela Diaz-Pardo, Jamie Emery, Emiliano Sandri, Sophia Scaglioni, Jiawei Tang, Viktor Veterinarov and Yifan Wang for outstanding research assistance. We pre-registered this experiment; registration ID: AEARCTR-0011298. We obtained IRB approval for this experiment; IRB number: 2304000976. We acknowledge financial support from the Melinda Gates Foundation.

†London School of Economics, o.bandiera@lse.ac.uk

‡London School of Economics, u.jalal@lse.ac.uk

§Massachusetts Institute of Technology, nroussil@mit.edu

1 Introduction

Limited access to education has been flagged as a major obstacle to women's labor market participation in low and middle income countries (Heath and Jayachandran 2018). Closing the gender education gap has therefore been a long-standing development goal, with numerous organizations dedicating vast resources to female educational initiatives. Such efforts have enabled women to achieve education levels that now match, if not surpass, those of men. Despite these advancements, women's labor force participation has plateaued, particularly in South Asia (Addati et al. 2018).

We study why women's educational gains are not reflected in labor market progress in the context of Pakistan, a country where the share of women with a college degree almost doubled in the past twenty years, from 4.2% in 1999 to 8.2% in 2018 (Pakistan Labor Force Survey). Meanwhile, college-educated female labor force participation has stagnated at 32.5%, amongst the lowest worldwide.

We field several rounds of surveys and an experiment with students graduating from college. The first series of surveys make up our diagnostic sample, and help identify where and why the female pipeline from education to employment leaks. It covers a panel of $\sim$ 1,000 students over three survey waves (baseline, two, then six months follow up) at a private university in 2022. In these surveys, we collect information on several important supply- and demand-side determinants of female labor force participation highlighted by the literature. This exercise reveals that timing of job search is uniquely predictive of women's employment. Using this insight from the diagnostic stage, we design a subsequent experiment which quantifies the casual effect of earlier job search on women's labor market outcomes. We study its impact through a second series of surveys, fielded over three waves with a panel of $\sim$ 1,400 students at a public university in 2023. In both settings, we first interview students a month before they graduate and then follow them over time to assess how their labor market expectations and outcomes evolve.

We begin our investigation of the pipeline from education to employment with the most straightforward supply-side question: do college-graduating women intend to work? At the time of college graduation, we find that men and women have strikingly similar beliefs about their likelihood of working: 71.8% of women believe they will work six months post graduation, compared to 77.0% of men. In addition, we find that women's stated intention to work is more than just talk: most women (70.2%) applied to at least one job in the six months following graduation, a share similar to men's.

Despite women's stated intentions to work and their corresponding search effort, we find that their realized employment falls short of expectations. Only 36.7% of women are working six months after graduation, which is 35.1 ppt lower than their baseline prediction. In contrast, 64.2% of men work six months after graduation, only 12.8 ppt lower than their original prediction. We also find that both women and men's baseline predictions about their peers' likelihood of working were much more accurate than predictions about their own: both women and men expected that $\sim 50\%$ of women and $\sim 65\%$ of men in their cohort would work. In short, women are largely aware of the aggregate labor force statistics but appear confident in their own ability to defy these odds.

Guided by the existing literature on gender employment gaps, we explore a wide range of traditional supply- and demand-side explanations. Controlling for various baseline supply-side characteristics (e.g. GPA, college major, industry of search, wage and non-wage preferences and expectations) reduces the 27.6 ppt gender gap in employment at six months by only 2.8 ppt. Further controlling for search effort has a modest effect, leaving the gap at 22.8 ppt. Typical demand-side factors also explain very little of the gap: controlling for the number of interviews and job offers leaves the gender gap in employment essentially unchanged, while further controlling for offered wages modestly shrinks the gap to 18.7 ppt. Instead, we find that the pipeline from education to employment mainly leaks at the very last stage: the decision to accept a job offer. The gender gap in acceptance is both statistically significant and economically meaningful: conditional on having received at least one job offer, women are 26.7 ppt (32.5%) less likely than men to be working six months after graduation. Even after we control for supply-side factors, including their expected and reservation wages, and demand-side factors, including the wage of the best offer they received, women are still 19.6 ppt (23.9%) less likely than men to be working six months after graduation.

We identify a key predictor of employment for college-graduating women: job search timing. Specifically, we find that, conditional on student characteristics and numbers of applications sent, the only factor that uniquely predicts women's probability of working, but not men's, is the time at which they start their job search. Women who started searching within two months of graduation are 17.0 ppt more likely to work six months post graduation than other women. This might be because women who apply earlier are positively selected or, alternatively, that applying earlier makes a difference. To disentangle the causal effect of timing from selection, we conduct an experiment that generates variations in timing that are orthogonal to individual traits. A month before graduation, a randomly selected group of students was promised $\approx$ USD 18 (PPP) if they applied to at least four relevant jobs by August 15th, about a month after graduation.[1] The control group received no financial incentive. We fielded the experiment at the largest public university in Pakistan, in June 2023, a year after the first round of surveys that created our diagnostic sample.

We find that our intervention has a large effect on the timing of applications: 81.7% of women in the treatment group applied to at least one job by September compared to 64.6% in the control group. The magnitude of this effect on application timing is similar for men. We follow our experimental sample six months after graduation (in January 2024) to test whether earlier applications have an effect on downstream labor market outcomes. We find large and significant ($p<0.01$) intent-to-treat effects on female labor force participation, defined as either working or having applied to at least one job in the last 30 days: 66.1% of women assigned to treatment are in the labor force six months after graduation, compared to 56.5% of control women (a 9.6 ppt difference). The intervention also increases women's likelihood of working by 7.4 ppt (22%) six months after graduation, and by 9.2 ppt (45%) when we restrict to full-time firm employment. In contrast, we find the treatment has no effect on the labor market outcomes of men, which is consistent with our descriptive evidence that earlier search is distinctively important for women.

We then turn to the mechanisms that underlie our results and first rule out a number of plausible ones. A natural explanation for our treatment effects on employment outcomes is the increase in search effort induced by our financial incentives. Indeed, we do find that six months post-graduation, treated women send 1.5 more applications on average than control women. However, we also find that they receive the same average number of job offers as control women. We further observe no treatment-control difference in working women's wages or occupations. As a result, women's work prospects are likely driven not by their increased search effort, but rather by their increased probability of accepting a job offer they received. This evidence also suggests that the demand side of the labor market is no different for control and treated women, and is unlikely to drive our findings.

Recognizing the pivotal role of search timing in our results, we develop a simple theoretical framework to model students' search timing choices, and explore the underlying mechanisms that make timing so influential. In this model, individuals derive utility from leisure and earn income from labor hours supplied. Various supply- and demand-side factors can create a wedge between the timing of application and eventual employment,

1. We consider that a job is relevant if it matches the skill set of the candidate. We manually check that this condition is met by reviewing screenshots of all four applications, which the student had to submit to receive their payment.

which is captured by a gap function that evolves over time. Drawing on the behavioral job search literature, our model allows individuals to form expectations over both the gap function and wages, which may be incorrect due to biases in their beliefs (Spinnewijn 2015a; Conlon et al. 2018; Mueller, Spinnewijn, and Topa 2021). The model reveals that, under biased beliefs, individuals who underestimate wages or the time gap between application and employment will delay their applications. Conversely, those expecting higher wages or a lower employment probability tend to apply earlier. From this setup, we derive two theoretical insights: first, for individuals with unbiased beliefs, our intervention will shift the timing of applications earlier but have no effect on their probability of work. Second, for individuals with high optimism about job prospects and/or pessimism about potential wages, incentives will lead to both earlier applications *and* an increased likelihood of employment. Moreover, biases that encourage delays in application timing may result in some individuals missing a critical threshold of opportunity beyond which job search will not yield employment, causing them to exit the labor force altogether. While we cannot directly measure biases empirically, we have measures of relative expectations. Relative employment expectations is the difference between a students' baseline beliefs about own likelihood of employment six months after graduation, minus that of their same-gender classmates. Relative wage expectations refers to the difference between their baseline beliefs about the wage offers they will personally get upon graduation, minus their beliefs about the offers their peers will get (holding constant the reference peer's gender, work schedule, job title, and cohort). We find that the treatment effects on women are driven by those who have high relative employment expectations, suggesting that for these women biased beliefs that encourage delays in application in the control group lead them to exit the labor force altogether.

We find substantial effects from a small nudge that alters women's job search and acceptance behavior in the months following graduation, while having no impact on men. Guided by the recent literature on how timing matters for marriage prospects, we turn to the dynamics of the marriage market as a potential explanation for why these months are so pivotal. Consistent with the finding in Adams and Andrew (2024) that families in South Asia believe that marriage-market prospects quickly deteriorate with age once a daughter is out of school, we document that marriage market activities start to unfold rapidly for women (but less so for men) shortly after graduation. We further find that the women who were most optimistic about their labor market prospects are also more likely to believe at baseline that they will marry later than the national average age of marriage for college-educated women (age 25). We show that these beliefs about their marriage timeline turn to be inaccurate. Indeed, when asked again six months later, the same women drastically revise their marriage expectation downwards, effectively converging towards age 25. Tying this result back with our experiment, we find that our treatment effects are concentrated among the women who had the highest expectations about their employment likelihood and had inaccurate beliefs about the marriage market at baseline.

Our evidence suggests that, as time passes post-graduation, the relative importance of the marriage market unexpectedly increases, at the expense of labor market opportunities. One plausible explanation for the relative change in priorities is parental intervention: parents may increasingly weigh the impact of employment on marriage prospects over time, possibly perceiving that working could limit marriage opportunities (as evidenced in other studies such as Subramanian (2024)). We are not able in this paper to provide definitive evidence on this explanation because we do not directly measure the beliefs of families or in-laws and the men we survey are not the right sample to infer the beliefs of future husbands, who are typically a few years older. We however can show that the daughters' self-reports reveal that family members play a significant role in the decision to accept job offers: 72.6% of women (and 48.3% of men) say they consult their family when making this decision.

This paper first contributes to a growing literature on behavioral job search, which highlights the role of biased beliefs about labor market prospects in the determination of employment and wages. Recent survey data reveals an optimistic bias among job seekers regarding their job finding rate (Spinnewijn 2015b; Mueller, Spinnewijn, and Topa 2021). Our paper most relates with two studies that further layer in a gender perspective. First, Kuziemko et al. (2018) show that, in the U.S., young women overestimate their chances of going back to work after childbirth. Second, Cortés et al. (2023) find that, among U.S. college graduates, women accept jobs earlier than men because they are more risk-averse and less over-optimistic than men. Our paper complements this study by describing the role of timing in a lower-income country, where college-graduating women's labor supply decisions are made more at the extensive margin (whether to work) than at the intensive margin (which job to choose). In this context, we find that the dynamics at play are very different: women are more over-optimistic than men about their job prospects and delaying job search decreases women's likelihood of employment.

To correct workers' misperceptions, recent research has honed in on information experiments (Jones and Santos 2022; Chakraborty, Negi, and Rao 2024; Aloud et al. 2020;

Jäger et al. 2024; Roussille 2024). Most related to the misperception we study, Alfonsi, Namubiru, and Spaziani (2024) find that debiasing overly optimistic job-seekers in Uganda increased their likelihood of working by 27% three months after the intervention. However, in our context, fixing misperception by providing information would most likely be to no avail. Indeed, women in our sample are well aware of aggregate labor force statistics: they nearly accurately predict the likelihood of their female peers to work six months after graduation. Rather, their misperception is about their own chances of working. This is why, rather than attempting to correct misperception with market-level information, we design an intervention that directly addresses the resulting job search distortions. In this way, our design relates to Banerjee and Sequeira (2023), where the intervention is a transportation subsidy that pushes job seekers to look for jobs beyond their local labor market. However, in Banerjee and Sequeira (2023), as jobs fail to materialize immediately, workers become increasingly impatient and redirect their search towards lower-paying jobs closer to home. In contrast, we find lasting positive effects of our intervention on employment and earnings.

We also contribute to a large literature that documents barriers to female labor force participation. On the labor supply side, social norms and intra-household constraints are the leading explanations (Goldin and Katz 1999; Bertrand, Goldin, and Katz 2010; Jayachandran 2021; Bursztyn, González, and Yanagizawa-Drott 2020; Agte and Bernhardt 2024; Kleven, Landais, and Leite-Mariante 2023). On the labor demand side, firm gender preferences have been shown to contribute to gender gaps in employment (Goldin and Rouse 2000; Kuhn and Shen 2023; Card, Colella, and Lalive 2021). In this literature, our findings most relate to Gentile et al. (2023). This paper leverages data from a job matching platform to describe the role of supply- and demand-side factors in the gender employment gap in Pakistan. The authors find that employers' gender restrictions place a larger constraint on women's job opportunities than supply-side decisions, these demand-side barriers relax at higher levels of education. Consistent with this, we find that firms make job offers to women at a rate similar to men, albeit at a lower wage. To the best of our knowledge, we are the first to uncover that the female pipeline from college education to employment leaks at the very last stage: the decision to accept a job offer.

Our paper additionally contributes to a growing literature on RCTs to remove barriers to women's access to the labor market. Many experiments in this field involve large investments in non-wage amenities and infrastructure. For example, Field and Vyborny (2022) increased women's job search in Pakistan by providing a women-only transport services to and from work. Ho, Jalota, and Karandikar (2023) offer women in India remote jobs with flexible schedules and see increased take-up of job offers by women. A range of studies reviewed in Halim, Perova, and Reynolds (2023) also show that improved access to childcare increases women's labor force participation. Such interventions hold tremendous promise but are costly to scale in lower-income countries. The experiment we design is distinct from previous studies in at least two ways. First, our intervention proves to be low-cost but effective: its positive effects on women's employment are almost immediate and persistent. Second, the design of our experiment is not based on a previously well-known cause of the gender employment gap, but rather on a cause we are the first to uncover in this context: the timing of job search. This relates to what we think is the fairly unique methodology we adopt in this paper. We follow the literature on randomized control trials in running an experiment to identify the causal effects of a specific factor. But the hypothesis we test, rather than coming from previous descriptive studies or experiments, is motivated by data we collect in a preceding round of descriptive surveys, with a population similar to that in our experiment. We think this sequencing of actions, in which the experiment design is informed by a year of descriptive work understanding the key local challenges, may, at least in part, explain the success of our intervention.

We finally relate to an extensive literature on the returns to education. This literature has long documented positive returns in the labor market (Becker 1975; Goldin 1992; Goldin and Katz 1999) and more recent evidence points to the positive returns in the marriage market as well (Lafortune 2013; Chiappori, Dias, and Meghir 2018; Attanasio and Kaufmann 2017; Ashraf et al. 2020). Particularly relevant to our setting is the finding in Adams and Andrew (2024) that perceived marriage market returns to education motivate investments in girls' schooling in India, a setting where gender norms are similar to Pakistan's. While most papers study marriage and labor market returns independently, the competition between these two markets has received less attention.[2] Yet, as Dhar (2021) documents in the Indian context, career-oriented women may face a penalty in the marriage market. Our paper adds to this literature by providing suggestive evidence that beliefs about the marriage market and labor market outcomes are interdependent.

2. Bursztyn, Fujiwara, and Pallais (2017) and Bertrand et al. (2021) are notable exceptions. Bursztyn, Fujiwara, and Pallais (2017) shows that unmarried women in an MBA program avoid career-enhancing actions when their actions are observed by unmarried male classmates. This is because women fear sending undesirable marriage market signals to potential marriage matches in their cohort. Bertrand et al. (2021) document a marriage/work trade-off across many developed countries which is especially binding for more educated women, and its intensity is driven by the strength of local gender norms. Both of these papers are set in higher income countries.

2 Background

2.1 Context

Gender gaps in college education in Pakistan have shown a remarkable decline over the years, as illustrated in Panel (a) of Figure 1. In 1999, 4.2% of women versus 8.0% of men (ages 18-35) held a college degree. By 2018, the share of women with a college degree had almost doubled (to 8.2%) and the gender education gap had nearly disappeared.

Despite these large strides in women's educational attainment, Panel (b) of Figure 1 shows that the share of college-educated women who are in the labor force has stagnated at 32.5% for the past twenty years. This share is not much higher than the average female labor force participation in the country (26%) and starkly contrasts with the labor force participation of college-educated men (81%).

As illustrated in Panel (a) of Figure A.1 for the United States, a standard pattern in higher income countries is that women's labor market entry rates are similar to men's but female labor force participation falls in childbearing years. In contrast, the life cycle of women's labor supply in Pakistan is flat (see Panel (b) of Figure A.1): rather than dropping out in childbearing ages, women in Pakistan rarely enter the labor market.[3]

The disconnect between a narrowing gender gap in education and a persistent gender gap in labor force participation is not unique to Pakistan: many countries across South Asia, North Africa, and the Middle East are facing similar challenges (see Jayachandran (2021) and Dinkelman and Ngai (2022) for comprehensive reviews). Still, Pakistan is a particularly interesting country to study: it stands out as one of the middle income countries where women's access to college has grown tremendously, yet women's involvement in the labor force remains one of the lowest globally, even when compared to countries with similar economic development (see Figure A.2).

We partnered with two of the most prominent universities in Pakistan to survey their students. One university is a large, mid-tier private university, and the other is the largest and oldest public university. The public university draws students from around the country, while students at the private university are mostly from Lahore, the second largest city in Pakistan. Both universities offer a comprehensive range of majors in the social sciences, physical sciences and humanities.

3. It is important to acknowledge that we are not accounting for cohort effects in Figure A.1. However, it is all the more striking that the labor force participation of young Pakistani women in 2018 is the same as that of older cohorts.

2.2 Two-Stage Strategy

This project can be separated into two distinct stages: diagnostic and experimentation. This section describes the reasons behind this "two-stage" approach. We illustrate the setting and timeline corresponding to each of the phases in Figure A.3.

Diagnostic Stage Because little is known about what determines the labor supply decisions of college-educated women in Pakistan, we run a series of descriptive surveys at the private university. Our goals for this descriptive exercise are threefold. First, we aim to determine whether women intend to work at the time they graduate from college. To do so, we run an in-person baseline survey in June 2022, one month before students' college graduation. We describe our results on labor market beliefs in Section 3.3. Our second goal is to diagnose whether the pipeline from labor market aspirations to entry breaks and, if so, where. To do so, we run follow-up phone surveys with the same students in September 2022 (two months after graduation) and in January 2023 (six months after graduation). The results comparing beliefs with realized labor market outcomes are described in Section 3.4. Last, we seek to provide evidence on the determinants of women's labor supply decisions. In particular, we test which characteristics are predictive of women's employment. We draw our list of potential predictors from the large literature on determinants of female labor force participation. We cover the variables we collect and construct, as well as their predictive power, in Section 4.

Experimentation Stage Armed with insights from the descriptive surveys, we identify a labor market behavior that is both highly predictive of women's employment, and plausibly receptive to change: the timing of women's applications. To test whether timing causally affects women's employment, we run an intervention that financially encourages a random subset of students to apply earlier. The experiment was implemented in June 2023, this time at Pakistan's largest public university. The main reason for switching to a new university is to verify that our descriptive insights generalize to a different context, particularly one with a more diverse student body representing various regions and economic backgrounds across the country. Here as well, we ran an in-person baseline survey a month before graduation, as well as 2- and 6-month follow up surveys over the phone. The design and results are detailed in Section 5.

3 Labor Market Outcomes: Aspirations and Reality

3.1 Recruitment and Balance

Recruitment We invited all 2,872 students (1,146 female and 1,726 male) at the private university to participate one month prior to their graduation date. Of these, 2,238 participated in our baseline survey (a response rate of 77.9%).[4] Since we are interested in labor market beliefs and outcomes, we exclude from our sample students who reported during the survey that they were already registered for graduate programs. This leaves us with a sample of 1,494 students in our baseline data.

Attrition and Balance Of the 1,494 students who are in our baseline sample, 1,080 responded to our 2-month survey and 1,029 responded to our 6-month survey. The 2-month and 6-month response rates are therefore 72.3% and 68.9%. These response rates are considerably higher than what is typical in the literature for phone surveys.[5] Table B.1 shows that the baseline, 2-month and 6-month samples have similar observables.[6]

Descriptive Sample Since our analysis systematically compares baseline beliefs with later realized outcomes, we define our descriptive sample as the 1,029 students who responded both to the baseline and 6-month follow-up surveys.

3.2 Descriptive Statistics

The baseline characteristics of our descriptive sample are shown in Table 1. The last column of the table reports the p-value associated with the test of equal means across

4. The response rate for our baseline survey is high compared to that of other surveys conducted in university settings. For instance, the response rate of Questrom graduating students in Cortés et al. (2024) was 20%, the response rate for Bertrand, Goldin, and Katz (2010)'s survey of University of Chicago MBA students was 31%, and the response rate was 10-12% across the 28 universities that participated in the recent Global COVID-19 Student Survey Jaeger et al. (2021). We think this high response rate was encouraged by the reward we offered for responding to the survey: a KFC meal. See Figure A.4 for a picture of our food stand.

5. We achieved this by calling students many (at least 3) times, systematically recording and varying the day/hour of the call to maximise our chances of response, and recording contact numbers of family members in case the student's own contact details change.

6. Even when the difference is statistically significant, it is not economically meaningful. For instance, the average GPA of non-attritors at 6-month is 3.09, while that of attritors is 3.03. This difference is statistically significant but economically small. We also note that there are no systematic patterns in attrition across waves; small but statistically significant differences that appear in the 2-months survey are eliminated during the 6-month survey, and vice versa.

gender. Women make up 42.7% of the sample, with 439 female and 590 male respondents. Men and women are both about 22 years-old, on average. Women's GPAs are about 6.7% higher than men's on average, and this difference is statistically significant. Men are more likely to major in Engineering and Computer Science (39.2% vs. 8.9%), whereas women are more likely to major in Life Sciences (21.6% vs. 5.1%) and Sciences (23.2% vs. 5.8%). Humanities, Languages and Education are moderately more popular among women than men (18.5% vs. 13.4%), while Social Sciences attract a higher proportion of men than women (36.6% vs. 27.8%). Finally, men and women come from similar familial backgrounds. On average, 41.0% of students have a college-educated mother, and 53.2% have a college-educated father, with no significant gender differences. Only 6.8% of women and 2.4% of men are married at the time they graduate from college, a similar fraction is engaged (7.5% of women and 5.9% of men). Figure A.5 also shows that these students have more liberal gender norms compared to national averages. Men in our sample are 18.3 ppt less likely to agree that being a housewife is just as fulfilling as being a working woman than the average male college student in Pakistan. Similarly, women are 26.5 ppt less like to agree that being a housewife is just as fulfilling as being a working woman than the average female college student in Pakistan. Both men and women in our sample are about 30 ppt less like to agree that when jobs are scarce, men have more right to a job than women. Thus, the barriers to female labor force participation that we document in this paper are plausibly more severe elsewhere in the country.

3.3 Baseline Beliefs about Labor Market Outcomes

Belief about Self: Elicitation In the baseline survey, we ask two main questions to measure students' beliefs about their future labor force participation. The first question is about their reservation wage for four work schedules (Full-time onsite, Part-time onsite, Full-time remote, Part-time remote), allowing them to leave the question unanswered if they would not be willing to work in a given schedule for any wage.[7] We consider that the

7. The exact wording of the question is: Imagine that you have graduated from your current degree and are offered a job with 4 possible schedules, which corresponds to [the job title they told us they'd prefer]. The four possible schedules are: Full-time (40 hours per week, 9am to 5pm, Monday to Friday) onsite, Part-time (25 hours per week, 9am to 2pm, Monday to Friday) onsite, Full-time remote, Part-time remote. There are no additional jobs currently available that are of interest to you so if you reject this job, you will be unemployed for the foreseeable future. What is the minimum monthly starting salary for which you would be willing to work for any of the following work schedules? Note: you may reject any or all schedules if you would not work on that schedule for any salary. Consider that in all options, the job and the employer are identical in all respects except the schedule, and the job is located in your preferred city. The job is a 20

student intends to be in the labor force at baseline if they provide a reservation wage for at least one of the schedules. The second question is probabilistic: "On a scale from 0 (very unlikely) to 100 (very likely), how likely is it that you will be working within 6 months of graduating? Work includes working for a private firm or running your own business."

Beliefs about Self: Results Across a wide range of measures, we find that, at the time they graduate, the majority of women intend to work. Virtually all women (and men) provide us with a reservation wage for at least one work schedule (and 95.0% of women provide a reservation wage for the Full-time, onsite schedule). Additionally, as illustrated in Figure A.6, the distribution of women's reservation wage lies to the left of men's, even after controlling for GPA, major and industry of search.[8] This suggests that women are not seeking unreasonably high wages. We then turn to the probabilistic question: The first two bars of Figure 2 show that women report a 71.8% likelihood of working within six months of graduation, only 5.2 ppt lower than men. These results are consistent with other related measures captured in our survey. Panel (a) of Figure A.5 shows that 60.9% of women in our sample report that being a housewife is not as fulfilling as being a working woman.[9] In addition, women's post-graduation actions support these intentions: panel (b) of A.7 shows 70.7% of women applied for at least one job within six months, mirroring 66.6% of men, who did the same.

Beliefs about Peers: Elicitation We asked students their beliefs about the future employment of their peers to test whether students' beliefs about themselves differ from their second order beliefs about their peers. Specifically, we asked: "Out of 100 randomly selected male students in your cohort at [the university], how many of them do you think would be employed within 6 months after graduating?".[10] We repeated the question replacing "male" with "female."

minute drive away from your house and is representative of other similar jobs in the industry in terms of career growth opportunities, non-wage benefits etc.

8. Industry of search fixed effects are constructed using students' preferred job title at baseline. For those who did not want to search for a job and were unable to report a relevant job title in case they *had* to search, we impute industry by asking them for the most common job title amongst students in their major.

9. We do not find any difference in women's responses to questions about their work preferences or adherence to traditional gender norms by the gender of the enumerator. This rules out concerns about demand effects due to the enumerator's gender, e.g. that women would be more inclined to abide by traditional gender norms when answering a man than when surveyed by a woman.

10. We made it clear we only asked about students that are not pursuing further education after graduating.

Beliefs about Peers: Results Even though women express strong labor market ambitions for themselves, both male and female students have much lower expectations about other women in their cohort. Figure 3 shows that men and women, respectively, estimate their female peers' chances of being employed six months post-graduation at 50.2% and 51.5%, while they believe their male peers' chances are higher (63.5% for men, 68.4% for women). Notably, women estimate their own likelihood of employment at 71.8%, similar to their belief about men's chances (68.5%). Thus, women see themselves as more likely to succeed than their female peers, and on par with men.

3.4 Beliefs Meet Reality

Beliefs About Self Meet Reality While men and women had similar beliefs about their employment likelihood, the third and fourth bars in Figure 2 reveal a large gender employment gap: 64.2% of men but only 36.7% of women were employed six months after graduation.[11] While both men and women overestimated their future likelihood of working, women did so to a much larger extent. At baseline, men estimated a 77.0% chance of working, with 64.2% actually employed six months later (overestimating by 12.8 ppt, or 15.8%), while women estimated a 71.8% chance of working, but only 36.7% end up employed (overestimating by 35.1 ppt, or 49.6%). Panel (b) of Figure 2 also shows that both genders had inaccurate beliefs, with relatively low correlations between baseline beliefs and outcomes (0.35 for men, 0.31 for women). However, the intercepts differ substantially: 14.4 ppt for women and 37.2 ppt for men, indicating that women overestimated across the distribution of baseline beliefs.

Beliefs About Others Meet Reality Both men and women accurately estimated their peers' employment outcomes. Men guessed 63.5% for other men, and women guessed 68.4%, close to the 64.2% actual employment rate for men. For female peers, men estimated 50.2% and women 51.6%. While these estimates are more optimistic than women's realized employment at 6 months (36.7%), they are much more accurate than women's belief about their own likelihood of working (which averaged 71.8%).

11. Illing, Schmieder, and Trenkle (2024) also find that men and women who are similar at baseline can experience different outcomes. This paper compares men and women who are displaced from similar jobs by applying an event study design combined with propensity score matching and reweighting to administrative data from Germany. After a mass layoff, women's earnings losses are about 35% higher than men's.

4 Diagnosing the Gender Gaps

Sections 3.3 and 3.4 reveal significant gender disparities in labor market outcomes six months post-graduation. This section provides evidence from a single context on various commonly suggested explanations. We first assess gender differences in relevant endowments and analyze if these differences explain the gender gaps, following the logic of a Oaxaca decomposition. We then test whether men and women receive different returns on their endowments, and whether potential differences in returns further explain the observed gaps. Before we proceed further, a quick note of caution: in this section we present diagnostic correlations without making causal claims.

Level Differences I: The Usual Supply-Side Suspects A leading explanation for low female labor force participation is that women are primarily responsible for household management Veerle (2011). This may imply a higher opportunity cost of working, and result in higher reservation wages or a preference for flexible or part-time work (Mas and Pallais 2017; Maestas et al. 2023). As a result, women may pursue fewer or rarer jobs than men, despite similar qualifications and aspirations. However, as discussed in Section 3.3, women's reservation wages are lower than men's, conditional or unconditional on covariates. Preferences over non-wage amenities are also similar across genders. Figure A.8 shows no gender differences in preferred work hours (averaging 6.4 per day for both), or preferences for remote work (about three-fourths of both genders prefer on-site jobs).[12] We also explore gender differences in GPA to assess if disparities in human capital contribute to these gaps. Figure A.7 Panel (a) illustrates that women have higher GPAs than men, and this holds even after controlling for their major.[13] Finally, another plausible explanation for the gender gap in employment is gender differences in job search (Cortés et al. 2023; Fluchtmann et al. 2024). For instance, if women apply to fewer jobs or apply later than men, they may be less likely to work. Panel (b) of Figure A.7 shows the cumulative distributions of the number of job applications are similar by gender and; when adjusted for GPA, major, and industry, the residualized gender gap in applications is null. Regarding application timing, we define "Applying early" as having sent at least one job application within two months of graduating. We find that women are as likely as men

12. At the 6-month follow-up, we also observe that working men and women's revealed preferences for non-wage amenities are similar: 72.5% of working women and 84.4% of working men work full time, 3.65% of working women and 3.07% of working men work remote.

13. The distribution of men and women across majors differ, as described in Table 1. However, we show in Figure 4 that gender differences in major choices do not drive the gender employment gap.

to apply early (77.1% vs. 80.0%). These findings suggest that college-graduating men and women are more alike in job preferences and search behavior than some literature suggests.

To formally examine how student characteristics affect the gender employment gap, we regress a six-month employment dummy on a female indicator, adding controls progressively (see Figure 4). The initial gap of 27.6 ppt decreases by 1.3-2.8 ppt as controls are added for education (GPA and major), industry of search, reservation wage, preferences for work hours and remote work, and baseline beliefs about chances of working in six months. Adding job search effort and work history in the final model reduces the gap by only 4.8 ppt (17.4%), suggesting that these factors do not explain the gender employment gap substantially.

Level Differences II: Demand-Side Factors The literature highlights a number of demand-side explanations that may alternatively explain gender gaps in employment, such as statistical, taste-based, and paternalistic gender discrimination (Goldin and Rouse 2000; Bertrand 2011; Kuhn and Shen 2013; Goldin 2014; Kline, Rose, and Walters 2022; Buchmann, Meyer, and Sullivan 2024). To investigate whether demand-side factors explain women's lower employment rates, we collect detailed information on the number of interviews as well as the number of job and salary offers received by both men and women. Figure A.9 draws the cumulative distributions of number of job interviews (Panel (a)) and job offers received (Panel (b)), separately for men and women. Strikingly, the same share of men and women get at least one interview and at least one job offer. We also find that there is no significant gender gap in the number of interviews students go on, and that women have a slightly higher (0.25) number of job offers than men. These conclusions are not altered when we further control for student baseline observables (i.e. GPA, major, and industry of search). Finally, we define "offered wage" as either the wage a person makes, if they are working, or the highest wage offer they have received (if they did not accept any job). Consistent with Brown (2022), we find that firms offer women lower wages, even after controlling for women's reservation and expected wages.

To formally examine whether demand-side factors explain lower female employment rates, we incrementally add students' number of interviews, number of job offers received, and the offered wage[14] as controls to the existing model in Figure 4, Rows 8-10. The gender gap

14. We construct the offered wage as equal to the current wage for individuals who are working, the

remains largely unaffected by these.[15]. The fact that even after holding job preferences, search behaviors, and the number of offers constant across men and women, women are still significantly less likely to be employed indicates that the gender gap emerges at an even later stage: job acceptance. Indeed, we find in Figure A.10 that upon receiving a job offer, women are 26.7 ppt less likely than men to be employed six months after graduation. Even after we control for supply-side factors and the wage offered by the firm, women are still 19.6 ppt less likely than men to be working six months post graduation.

Differences in Returns Controlling for student characteristics does not shrink gender gaps in a pooled regression, but a given characteristic may uniquely predict men or women's future employment. To test for this, we regress employment outcomes six months post-graduation separately for men and women on the main supply-side variables we used as controls in Figure 4. Figure 5 shows the results of this exercise in a pooled regression (all factors entered together) while Figure A.11 shows the results of this exercise in individual regressions (each factor entered separately). For all but two variables we find no meaningful difference in their ability to predict men and women's employment. For instance, in a bivariate model, a one-standard deviation increase in baseline belief about one's employment probability is associated with a 6.3 ppt increase in the likelihood of working 6 months later for men, and a 4.1 ppt increase for women. Once we control for all supply-side variables, the number of applications is not differentially predictive of employment for men and women. In contrast, we find that early applications uniquely matter for women: those applying within two months of graduation are, all else equal, 16.0 ppt more likely to be employed, while early job search has no effect on men. Additionally, internship experience increases employment probability for both men and women (by 11.8 ppt and 22.2 ppt, respectively). Moving forward, we narrow in on the effect of timing on the employment likelihood of women, as the focus of this paper is on job search rather than pre-graduation characteristics (e.g., doing an internship during college).

Why does timing matter for women? While we control for a rich set of factors in Figure 5, unobservable differences in the propensity to work might drive the observed correlation between the timing of applications and employment outcomes. To assess whether timing

highest rejected wage for individuals who are not working but have received a job offer in the past, and zero otherwise.

15. While none of these variables affects the gender employment gap, they do increase the adjusted R^2, confirming their expected relevance to students' labor market outcomes

has a causal effect on women's likelihood of working, we need to observe differences in timing that are independent of individual traits. We generate this exogenous variation by fielding an experiment. The experiment also allows us to shed light on the mechanisms underlying the causal effect of application timing on women's employment.

5 Experimental Evidence on the Timing of Job Search

5.1 Experimental Design

We field a survey experiment at the public university in Lahore in June 2023. The experiment was conducted with ≈ 2,000 students who were due to graduate in mid July. A randomly selected group was offered monetary incentives, which they would receive if they prove to us that they have applied to at least four relevant jobs by August 15th (about a month after their graduation). 40% of survey-takers were offered PKR 5,000 (≈$18 USD), 10% were offered PKR 20,000 (≈$72 USD), the remaining 50% constitute the control group and was offered no reward.[16] Assignment to treatment was presented to students as the result of a lottery.[17] Students had to submit proof of their applications to the research team to claim their reward. These submissions were made via a brief online questionnaire where students attached screenshots of their applications including the date of application and the title of the job. This allowed us to ensure that students indeed applied to real jobs that were relevant to their college degree, skill set as well as level of education and experience.

The goal of this experiment is to create variation in the timing of job applications that is orthogonal to students' traits. Given that the reward is a small and one-time incentive, it should not have any wealth effect on students' labor supply. Additionally, to the extent that our treatment is successful and increases the employment of women in the treatment group, there should be no spillovers (e.g. crowd-out or wage effects) on the control group since the treated group is infinitesimal relative to the broader labor market. In other words, our treatment group remains atomistic in this market.

We decide to re-survey the students twice after baseline. First, we re-survey them in early September 2023 (a couple of weeks after the treatment deadline) to measure

16. We pool the two treatment groups together in our main specification because they yield similar results.
17. Specifically, the wording was "You have now reached the last part of the survey which is experimental. At this stage, whether you are shown 2 modules or just 1 module will be randomly determined by a lottery." In the event students were not selected into treatment, they were told that "The lottery has decided that you will skip directly to the last module of the survey."

our treatment effect on the number of early applications. We also measure downstream outcomes (number of interviews, job offers and likelihood of working) but those are less likely to be different between the control and treatment groups since we survey students shortly after the treatment application deadline and there typically is a lag between the start of a job search and employment. That is why we re-survey the students in early January 2024, this time with the goal of measuring treatment effects on downstream outcomes, like employment.

5.2 Recruitment, Balance and Take-up

Recruitment Based on budget and power calculations, we were targeting a sample of about 2,000 students for our baseline survey. We stratified our sample by major and gender. Specifically, we over-sampled women (65%) since that is our population of interest for the experiment. For majors, we stratified to be representative of the full spectrum of majors at the university. The only exception is that we excluded majors with fewer than 25 students and we put a cap of 200 female and 100 male students per major to ensure broad representation across majors in the final sample. This cap was binding for a few majors, in which case sub-groups of students in the major were randomly and incrementally invited to participate until the target was reached. Within this sampling frame, each male and female student then had a 50% chance of being randomized into treatment. Since we are interested in labor market beliefs and outcomes, we exclude from our sample students who tell us during the survey that they already registered for graduate programs since these students had made prior commitments to not enter the labor force upon graduation. This leaves us with a sample of 1,931 students in the baseline survey.

Balance and Attrition The identification strategy for our RCT relies on the assumption that treatment and control students do not differ on average in all observable and unobservable characteristics. To support this hypothesis, we check for balance across treatment arms in each survey wave. Table B.2 Columns 1 to 4 show that the treatment and control groups are balanced on most of the key baseline variables, confirming the success of our randomization procedure. The remaining columns of Table B.2 show that balance on observables between treatment arms is maintained in the 2-month (Columns 5 to 8) and 6-month follow ups (Columns 9 to 12). Table B.3 further investigates whether there is selective attrition in the full study sample, that is whether attrition correlates with baseline observable characteristics or treatment status. Attrition is 15.4% and 26.0% at 2- and

6-months post-graduation, respectively. There is no differential attrition across treatment groups in the 6 month follow-up. There is a slight differential attrition of the treatment group in the 2 month follow-up (the share of treated among 2 months respondents is 50.79%, while it was 51.75% at baseline) but as noted previously balance is recovered in the 6 month follow-up, , which is the survey wave on which we compute all of our treatment effects. We also find no evidence of differential attrition on other attributes, with the exception of a modest deviation in the distribution of majors at the 2-month follow-up, and the share of women in the 6 month follow-up, where the latter is slightly lower at than at baseline (64.22 vs. 65.72).[18] However, as shown in Table B.2, Columns (6)-(12) these characteristics are not unbalanced by treatment status in the 6 month sample, and thus, their attrition across the survey waves was not systematically correlated with treatment. For this reason, we do not correct for attrition in our main regression specifications.

External Validity of Descriptive Findings Appendix D shows that the main descriptive findings, which are derived from a sample of private university students, replicate in the control group of the experimental sample, drawn from a large public university. This validates the relevance of our descriptive insights to a broad spectrum of college graduates.

Treatment Take-up 46% of women (and 44% of men) who were offered the treatment took it. Figure 6 summarizes the extent to which our financial incentives shifted the distribution of early applications. Specifically, Panel (a) shows, for women, the distribution of applications separately for the control (light red, right bar) and the treatment (dark red, left bar) in our 2-month follow up of early September, a couple of weeks after the treatment application deadline (August 15th). Panel (b) repeats the exercise for men. We observe that the treatment substantially increased the extensive and intensive margins of job applications. For women, only 24.1% of those offered the treatment had not applied to any jobs by September, compared to 42.1% in the control group. Additionally, the mean number of applications sent in the treated group was 28% higher than in the control one (6.4 vs 5.0 applications, respectively). The magnitudes (both at the extensive and intensive margin) are very similar for men in Panel (b). One may expect that our treatment would

18. Humanities, Languages and Education become less represented (27.0% of the baseline sample but 25.9% of the 2-month follow-up sample) and Sciences more represented (26.3% of the baseline sample but 27.7% of the 2-month follow-up sample). We do not think this deviation impacts the interpretation of our results since we find no selective attrition on any other dimension and, more importantly, this imbalance does not carry through to the 6-month follow-up, on which we perform virtually all of our analysis, reverts to the baseline distribution.

result in clustering at 4 job applications since that was the requirement to receive the financial incentive. However, Panel (a) and (b) illustrate that the entire distribution of students' applications shifted to the right, with more students in the treatment applying to five jobs and more than in the control group. Figure A.12 correlates the take-up dummy with students' baseline characteristics. It shows that both male and female who take up treatment are significantly more likely to be majoring in engineering or computer science and less likely to be majoring in Humanities. Second, women that are already engaged or married are, directionally, less likely to take up the treatment. Third, female compliance is positively correlated with baseline beliefs about employment at 6 months, internship experience and the amount of the financial incentive offered, while men's is not. Last, both for women and men, take-up is not correlated with gender norms, as measured by the World Value Survey questions, or with reservation wages.

5.3 Results

Empirical approach We run the following regression to estimate the impact of our intervention on students' labor market outcomes:

$$Y_i = \alpha_1 + \beta_1 T_i + \gamma_1 (T_i \times Male_i) + \lambda_1 Male_i + \epsilon_i \tag{1}$$

Where Y_i denotes the outcome of interest for individual i. The term α_1 represents the average outcome for women in the control group. The coefficient λ_1 captures the gender difference in outcomes within the control group, indicating the extent to which outcomes for male students differ from those of female students. The coefficient β_1 is the primary parameter of interest and measures the intent-to-treat effect of the intervention on women assigned to the treatment group. Finally, γ_1 represents the additional effect of the treatment on male students assigned to the treatment group. Our analysis is conducted separately for the 2-month and 6-month follow-up surveys. All standard errors are robust to heteroskedasticity.

Intent-to-Treat Effects on Labor Market Outcomes Table 3 shows results from our estimation of Equation 1. The effects for men and women are computed in the same pooled regression, using an interaction between treatment and gender. Panel A describes the effect for women, while Panel B shows the effects for men. In Column (1), we find a 9.6 ppt (17.0%) treatment effect on women's labor force participation six months post-graduation (p<0.01). In Column (2), we show that treated women are 7.4 ppt (22.0%) more likely to be employed than control women (p<0.05) six months post-graduation. In Column (3), our dependent variable is equal to one if the respondent found a full-time job with a positive wage. We find that the treatment has an even stronger effect on this outcome: treated women are 8.2 ppt (37.1%) more likely to be employed full-time for a wage than control women (p<0.01) six months post-graduation. Column 4 further examines the likelihood of working for a firm (as opposed to being self-employed,working in a household enterprise or being unemployed). We find that our treatment increased the probability of working for a firm by 9.2 ppt (45.3%) for women in the treatment group. We interpret these stronger effects on firm work as evidence that the treatment not only changed women's chances of working but also the composition of their work. This shift in composition follows directly from the fact that our intervention specifically induced women to apply for positions in the formal labor market, and thus does not have a channel through which it may influence their self employment prospects. Ensuring that women take up full-time firm work is a central contribution of our intervention since, in low income countries, working part-time, informally or on a self-employed basis is indicative of lower labor market attachment and earnings potential. For instance, Breza, Kaur, and Shamdasani (2021) show that a large fraction of self-employment in India is in fact "disguised" or "hidden" unemployment resulting from labor rationing. The World Bank's *World Development Report 2012* further argues that the disproportionate representation of women in part-time work or self-employment is in fact one of the the primary drivers of gender gaps in earnings and productivity, rather than factors such as human capital or worker and job characteristics (World Bank 2021).[19] In contrast, we detect no effect of the treatment on either men's labor force participation (Column 1 of Panel B), or their employment (Column 2 to 4 of Panel B). These results are consistent with our correlational findings, which found a positive association with early applications for women but not for men.

19. World Bank (2021) reviews extensive data and literature, showing that women in low-income countries are predominantly employed in low-productivity, low-wage jobs, such as small-scale farming, running small firms, and engaging in casual or piece-rate work. It also highlights that most female-headed enterprises are home-based and more often driven by "necessity" than by "opportunity." Reinforcing these findings, Ashraf, Delfino, and Glaeser (2022) reveals that female entrepreneurs typically earn less and are concentrated in low-return industries. Similarly, Our World in Data (2023) notes that most women in paid work in low and middle-income countries are employed within the informal economy. Together, these studies and meta-analyses emphasize how the types of economic activities women engage in contribute to persistent gender disparities in earnings and productivity. Therefore, the shift in the composition of work induced by our treatment holds substantial socioeconomic value for women.

2SLS Results Table 4 uses the experimental variation as an instrument for early applications, allowing us to estimate the causal effect of the treatment on compliers—those who respond to the financial incentive by applying earlier. We compare the magnitudes of our descriptive OLS and experimental 2SLS estimates to better understand the relative importance of selection into early applications versus the causal effect of our treatment. Column 1 presents the first stage, showing how the financial incentive influenced the likelihood of applying early. For women (Panel A), the treatment led to a significant 13.8 ppt increase in early applications. A similar pattern is observed for men (Panel B), where the intervention raised the likelihood of early applications by 9.5 ppt. For women, the OLS estimate in Column 2 suggests that early applications are associated with a 26.7 ppt increase in the probability of working. However, the 2SLS estimate in Column 3 indicates a much larger effect—a 63.9 ppt increase in the probability of working when women apply early. The OLS and 2SLS coefficients differ for several reasons. The OLS is upward biased because it captures the effect of selection, that is the fact that women who are more keen to work are more likely to apply earlier, but downward biased because of measurement error. In contrast, the IV does not suffer from selection bias, but it does measure the effect of the additional application on the compliers, that is, the women whose application behavior changes because of the incentives. The comparison of the two coefficients reveals that the latter effect prevails, that is, the compliers have a particularly strong effect of applying early, and the 2SLS coefficient is much larger than the OLS. For men (Panel B), neither the OLS nor the 2SLS estimates in Columns 2 and 3 are statistically significant, implying that the treatment did not have effect on their employment outcomes, even though it did increase their likelihood of applying early. This gender difference highlights that earlier job search is particularly important for women, who are more likely to benefit from applying early when incentivized.

Search Effort Columns (5) to (7) of Table 3 investigate the drivers of the positive treatment effects on female employment. While we find that six months post-graduation, treated women have sent 1.5 more applications on average than control women, we also show that treated women have received the same average number of job offers as control women. As a result, their work prospects are likely driven not by their increased search effort, but rather by their increased probability of accepting a job offer they received. This is confirmed in Column 7 which shows that, conditional on getting a job offer, treated women are 8.4ppt (18.1%) more likely to accept it. To test for the role of search effort, we also conduct a mediation analysis. This analysis looks at the extent to which total treatment effects are driven by search effort (i.e. the increase in total number of applications over six months) versus timing of search (i.e. the fact that treated women started searching earlier). We proxy timing of search here by the month of first job application. The empirical approach we adopt is described in detail in Appendix Section C. Figure 7 shows the results of this analysis. Panel (a) shows that the number of applications does not drive treatment effects: its average causal mediation effect is null. Meanwhile, Panel (b) shows that the total treatment effects operate predominantly through the month of first application. Together, these findings suggests that higher overall search effort is an unlikely explanation for our treatment effects. Rather, the timing of applications plays a key role.

6 Mechanisms

The results reveal that timing of job search matters distinctively and substantially for women. Why is this the case? In this section, we take insights from the job search literature to shed light on the mechanisms through which our treatment operates.

6.1 Framework

To interpret the findings in a coherent framework and to guide us to potential mechanisms underpinning them, we model the choice of application timing. After graduation, individuals choose whether to take time off before entering the labor market, trading off leisure and income. We model this choice in the standard labor supply framework, allowing for uncertainty on the wage and on the relationship between the timing of application and the probability to find a job.

Utility of individual i from leisure is $u(d_i)$, increasing and concave in the number of days before individual i starts applying ("days off"), d_i. Income is equal to $[T - g_i(d_i)]w_i$ where w_i is the daily wage rate individual i will receive once they work and T is the upper bound to the number of days without work - in our setting this is approximately six months as three out of four students plan to be working by then. In addition to labor income, individuals in the treatment group receive a one-off reward I if they apply before day $D < T$. The $g_i(.)$ function captures all the factors that create a gap between the time of application and the time of employment. These include factors that are time-invariant and specific to individual i, such as their preferences for job attributes. It also includes time-variant factors specific to individual i: the applicants' circumstances might change with time, restricting or broadening the set of jobs they are willing to accept. These finally further include factors that affect all graduates in the same cohort, like a change in labor demand over time. For instance, some employers might be concerned with large gaps in workers' C.V. so offers might dwindle.

Following the recent advances of the literature on behavioural job search we focus on expectations about wages and the probability of finding a job as key mechanisms (Spinnewijn 2015a; Krueger and Mueller 2016; Conlon et al. 2018; Mueller, Spinnewijn, and Topa 2021; Potter 2021). Denote $i's$ expectations by $g_i^e(.)$ and w_i^e. These might differ from the actual values of $g()$ and w that we denote by $g_i^a(.)$ and w_i^a.

Definition: *We define individual i to be optimistic about their employment prospects if $g_i^e(d) > g_i^a(d)$, unbiased if $g_i^e(d) = g_i^a(d)$, pessimistic if $g_i^e(d) < g_i^a(d)$. Similarly, we define individual i to be optimistic about their wage prospects if $w_i^e > w_i^a$, unbiased if $w_i^e = w_i^a$, pessimistic if $w_i^e < w_i^a$*

In what follows we show that the response to the application incentive can be used to infer whether these expectation are biased, and whether biased expectations can explain our results. Individuals choose d_i to maximise their payoff subject to $0 < d_i < T$ and $T - g_i(d_i) > 0$, where the latter constraint guarantees that income is not negative.

If individual i belongs to the control group their payoff is:

$$u(d_i) + [T - g_i^e(d_i)]w_i^e$$

If individual i belongs to the treatment group their payoff is:

$$\begin{cases} u(d_i) + [T - g_i^e(d_i)]w_i^e & \text{if } d_i > D \\ u(d_i) + [T - g_i^e(d_i)]w_i^e + I & \text{if } d_i \leq D \end{cases}$$

where I is the incentive payment that is awarded if individual i applies before day D.

Since the choice of d_i only affects incentive pay on the extensive margin, the first order condition is the same in treatment and control. The optimal application time for individual i, d_i^*, equalises the marginal benefit from one more day of leisure to the marginal cost in terms of forgone income, that is:[20]

$$u_d(d_i^*) - g_{id}^e(d_i^*)w_i^e = 0$$

Figure 8 illustrates the solution. The figure has three panels. The central panel depicts the first-order condition, which, as discussed above, is the same for the two groups. The top panel depicts timings of application and of work start date in the control group. It shows how biased expectations can lead people to drop out of the labour force. Beginning with the first order condition, the middle panel shows the marginal benefit of days off, which is decreasing and convex because of diminishing marginal utility. The optimal number of days of leisure before starting to search is determined by the intersection of the marginal benefit with the perceived marginal cost, here assumed to be constant. The intersection of the two gives the level d^*. Now, the timing of starting work depends on how quickly applications transform into a job offer and an acceptance. This is summarized by the $g()$ function. If expectations are correct, that is if $g^a() = g^e()$, we see that the individual applies at d^*. and starts working shortly thereafter at $g(d^*)$. If, however, expectations are biased, the relevant $g()$ function is not $g^e()$ but $g^a()$ and we see that drawing to the 45 degree line, the $g^a(d^*)$, is much to the right of $g^e(d^*)$. In the picture, we have depicted the case where it is even to the right of T, which means that individuals do not find a job in the time allocated for this purpose.

This leads to our first result:

Result 1 *If expectations are biased there is a threshold L such that all i's for whom $d_i^* > L$ will remain unemployed despite applying for jobs.*

Result 1 provides a mechanism that links application timing to the gender gap in employment. The incentive treatment is designed to exogenously shift applications earlier, thus it can tell us whether biased expectations lead women, but not men, to drop out. It is straightforward to show that incentives shift the application date and the job start date back for at least some of the applicants. Consider the choice faced by an individual j whose $d_j^* = D + \epsilon$. If they move the application date back by ϵ, by the envelope theorem their FOC would hold at D when ϵ is sufficiently small, while their payoff would increase by I. By the same logic, there must be a threshold v such that all i for whom $d_i^* < D + v$ will apply at $d = D$ instead of $d = d_i^*$. Since $g()$ is non decreasing in d, this implies

20. The second order condition is met by our assumptions on u and g

that they will start working earlier too. The bottom panel of Figure 8 shows the effect of the introduction of incentives in the treatment group. We first go back to the first-order condition and draw the line at D, which is the day by which applications must be sent in order to get the incentive payment. By definition, if D is to the left of d^*, it will not be optimal, and in particular the marginal benefit on additional day-off is larger than the marginal cost. However, once one takes into account the incentive payment, which is not marginal, it might as well be that the total utility is larger as explained above. Therefore, the first-order condition does not hold and we are at a boundary where individuals choose to start applications on day D. Bringing this to the timing graph, we see that the line hits the actual $g()$ at a much lower level than d^* does in the control group. This means that the $g(D)$ is much to the left of T. In other words, the incentive leads people to apply early, and by doing so, the difference between their perceived cost of effort and their actual cost of effort is much smaller. Therefore, they do not run the risk of applying too late and be left jobless. This is under the assumption that the payment I is not so large that people apply to cash the reward and decline all offers. The logic applies regardless of whether i's expectations are biased or not as summarised in the next result:

Result 2 *There is a threshold v_i such that all i's for whom $d_i^* < D + v_i$ will shift their application date back to D and begin working earlier.*

One implication of Result 2 is that compliers, that is those who apply earlier because of the incentive, are positively selected on expected wages because their optimal application date is earlier than the non-compliers. This is consistent with the take-up analysis in Figure A.12 and with the LATE estimates in Table 4. Result 2 also implies that if $d_i^* < D + v_i$ also $d_i^* - k < D + v_i$ for all $k > 0$. In line with this, Figure A.13 shows distributions of the months in 2023 when students sent their first applications and first offer, separated for female and male and by treatment status. This figure illustrates that treatment increases applications at or just before D and reduces them just after D. We also observe that the share of students first applying to jobs goes down in both treatment and control after the month of August, reaching levels below 5% of the sample in both groups by the end of our sample period.

Expectations come into play to determine the incentive effect on the probability that individual i is employed by day T. Since incentives shift the application date back, they can only increase the probability of employment. Since the probability is bound by 1 it can only increase if it is below 1 in the absence of incentives. This leads to our third and last result:

Result 3 *The incentive treatment can only increase the probability of working for individuals who have optimistic expectations about the employment probability or pessimistic expectations about the wage.*

Intuitively, incentives can only increase employment if some people remain jobless despite applying, which, by Result 1, requires biased expectations. However, having biased expectations does not guarantee that incentives increase employment. This only happens if $L < D + v_i$ that is, the size of the incentive is sufficient to attract individuals who would otherwise would apply too late.

6.2 Mechanisms: Theory

We have shown that the standard labor supply model with biased expectations provides a coherent framework for our findings; in this section we describe three mechanisms that emerge from the model and are consistent with the findings. We then propose empirical tests to gauge their practical relevance.

The first mechanism operates through gender differences in wage expectations. In particular, if women are more *pessimistic* they will delay applications more than men, and therefore they will be more likely to run out of time and end up jobless by T. The second mechanism operates through gender differences in employment expectations. In particular, if women are more *optimistic* they will delay applications more than men, and therefore they will be more likely to run out of time and end up jobless by T. Several papers align with these biases. A large body of literature suggests that women tend to be less confident about their expected wages. Similarly, there is evidence indicating that individuals are often overly optimistic about both their chances of securing a job and the time it will take. The third mechanism operates through gender differences in the cost of delay, that are unknown to the applicants before they start searching. In terms of the model, there is a systematic difference in $g()$ by gender, so that the gap between application and job start date is larger for women at any d and might be larger for larger d. The incentive treatment "reveals" $g()$, because each individual can see how long it takes for their applications to yield an offer they can accept.

Separating these mechanisms, especially the third from the first two, is important because the long-term consequences can be quite different. In the first two cases, the gender difference in expectations may lead to a delay in employment, but women eventually catch up. Of course, this comes at the cost of foregone income, but there will be convergence nevertheless. In the third case, however, the difference becomes bigger over time, so there is no convergence of sorts, and tiny differences at the point of graduation result in big differences down the line.

6.3 Mechanisms: Tests

Gender Differences in Relative Expectations: Elicitation Measuring bias (in any direction, for any outcome) requires collecting expected and potential outcomes for each individual. Expected outcomes can be elicited via surveys but potential outcomes cannot be observed because expectations affect realized outcomes. To test the relevance of mechanisms 1 and 2 we rely on the randomized allocation of treatment and test whether treatment effects are heterogeneous by expectations. The sign of the interaction between treatment effect and expected outcomes is informative of the sign of the bias under the assumption that control individuals are a valid counterfactual.

In addition to employment beliefs, which we already collected in the diagnostic sample, we also elicited in the experiment sample individuals' expectations about their wages and the wages of their fellow graduates from the same university, year of graduation, same gender and same field.[21] To measure employment expectations, we use students' baseline beliefs about their employment likelihood six-month post-graduation, as well as their beliefs about the employment likelihood of their same-gender peers six-month post-graduation. Both of these variables are defined in Section 3.3. We finally construct a measure of "relative expectations", defined as the difference between what an individual believes they will achieve, either in terms of wages or employment probabilities, and what they believe their similar peers will achieve. For wages, we express it in percentage of the same-gender peers wage. The reason we subtract individuals' beliefs about their fellow students is to incorporate their understanding of general market conditions so that we extract their expectations about themselves net of their beliefs about the labor market in

21. The two exact questions are: 1) Imagine a firm wants to hire you in a full-time on-site job for the job [job_title]. How much do you think you would be offered in monthly starting salary for the job? 2) Think of the typical [gender_student] student graduating from Punjab University looking for a job as a [job_title]. What do you think that student's monthly starting salary for a full-time on-site job (40 hours per week, 9am to 5pm, Monday to Friday) will be? The inputs in those questions are: [job_title] is the job title the student previously told the enumerator they would be most interested in, [gender_student] is the gender of the respondent.

general.

Gender Differences in Relative Expectations: Results Figure A.14 plots the distributions of the two measures, relative wage expectations and relative employment expectations, for men and women separately. Panel (a) shows that wage expectations are similar between genders, the median is $0pp$ for both women and men. In contrast, Panel (a) shows that both women and men have high expectations about employment prospects and women more so than men: the median woman believes to be 30pp more likely than her peers to find a job within six months, the median men 10pp.

Table 6 re-estimates our baseline specification interacting all treatment indicators with the relative expectation measure described below, centered at its mean. We find that women in the control group who have higher expected wages are more likely to work, that is, labor supply is positively sloped. The interaction between treatment and wage expectations for women is negative and of a similar magnitude which implies that the effect of the treatment is equivalent to eliminating differences in expected wages in this sample. However, as shown in Table 7, the effect of wage expectations does not go through the timing of applications as neither the expected wage nor its interaction with treatment correlates with application timing. Overall, our analysis of wage expectations does not lend support to the first mechanism but it does provide a useful quantification of the treatment effect.

Next we investigate the second mechanism, namely whether the effect of treatment goes through employment expectations. Column 2 of 6 show that women who believe are more likely than their peers to work are indeed more likely to be working at the 6 months mark and the effect is more than twice as strong for the treatment group. Again the evidence on application timing (Table 7), is not consistent with the mechanism as employment expectations do not predict the time at which women start applying for jobs.

Taken together the evidence in this section suggest that there are differences in expectations that determine employment outcomes but, in contrast to the logic of mechanisms 1 and 2, do not determine application timings. In the model notation we do not detect differences in the expected $g()$ that can be corrected by treatment. The effect of treatment must therefore be channeled through the actual $g()$ as described in mechanism 3. That is, the incentive encourages early applications and upon applying women discover what it takes to go from application to job offer to working. The next sections investigate the stages of mechanism 3.

Gender Differences in the Cost of Delay -Demand The cost of delaying applications might be due to an unexpected slowdown in labor demand over time. This could only be specific to "women's jobs" since we know already that treatment shifted men's job search earlier but had no effect on their employment, which suggests that men's vacancy creation and offer arrival rates do not decay with time. We test this by considering the top 3 occupations women most frequently work in. Table B.5 shows that the shares of women in a given occupation are similar across treatment and control groups. For example, teaching is the most common job amongst women, and plausibly sensitive to seasonal changes in demand. Thus, we might expect that earlier applications led treated women to disproportionately sort into teaching in the months when demand for teachers is highest. However, treated and control women are similarly likely to work as teachers (35.9% vs. 34.6%). This suggests that the positive employment effects of applying and receiving job offers earlier are not driven by women sorting into a different category of more seasonal jobs. Another piece of evidence here is that while women in the treatment group sent more application, their number of job offers is the same as in the control group, making a relative demand slowdown for the control group unlikely. This can be seen from Column 6 of Table 3. Finally, Table B.6 shows the effects of the experiment on wages for men and women. 58% of the sample is not working so their wages are coded as 0. The table then assesses whether the treatment effects the likelihood that an individual is pushed above the 60th, 75th or 90th (gender specific) percentiles of the wage distribution. Column 1 shows a clear extensive margin effect: treatment increases women's but not men's likelihood of being in the 60th percentile, which effectively means it increases the likelihood that their wage is non-zero. But across the other thresholds, we see no strong evidence for men or women that treatment induces them to earn more than the control group. Further, conditional on working, Figure A.16 shows that the distribution of wages for control and treated women are similar, and the average gap between the two is essentially zero. This also goes against the idea that control women are facing a different demand curve than treated ones.

Gender Differences in the Cost of Delay–Supply Slow Down If temporal changes in labor demand do not explain our results, what changes on the supply side over time? To understand this better, we test whether earlier applications matter for women because they enable women to get ahead of any unexpected changes in their labor supply. To measure unexpected changes in women's labor supply we examine, in the control group, the likelihood that women reject job offers with offered salaries above their baseline reservation wage. This outcome captures changes in labor supply that women did not anticipate at baseline. We also collect, in the experimental sample, more granular information on timing. For instance, we ask respondents to tell us the date of their first application and their first job offer. This allows us to look at the effect of monthly delays in application and job offer, as opposed to relying on whether a respondent has applied by the two month survey to construct a measure of "early" applications, as we do in the diagnostic phase. Table 5 shows that women are 9.2 ppt more likely than men to reject a wage offer above their baseline reservation wage. We find that timing of the job offer is a key determinant of that decision: women who receive a job offer at graduation are no more likely than men to reject it. However, with each additional month since graduation, this likelihood to turn down a sufficiently paid job offer rises by 3.2 ppt for women. This effect is stable despite controlling for students' industry of search, GPA, baseline employment beliefs, and college major. By contrast, there is no effect of timing on men's likelihood of rejecting job offers above their baseline reservation wages. This suggests that timing matters for women mainly because it allows them to receive – and accept – job offers before their reservation wages rise and labor market attachment weakens.

How can incorrect beliefs prevail in equilibrium? One critical insight our paper offers is that, on average, women only have modestly incorrect beliefs about their female peers' chances of overcoming gender barriers. In fact, most women understand that familial and marital barriers will prevent many other women from working, but believe that they personally will overcome these barriers. In support of this, Figure A.17 illustrates that at baseline, 91% of women acknowledge that other women may struggle to work after graduation due to various "family reasons," such as not receiving approval from their families or being pressured to prioritize finding a marriage match. In contrast, only 20% express that they themselves may face similar obstacles. Therefore, the main misconception is the belief that their own labor market outcomes will be substantially better than their peers'. As a result, women do not learn from the experiences of the women they observe around them, and it is expected in that context that their biased beliefs persist in equilibrium. Relatedly, the observation that women are not meaningfully misinformed at the overall labor market outcomes for women motivates our choice of a financial intervention rather than any type of information intervention about their peers.

6.4 Relationship with the Marriage Market

We find substantial effects from a small incentive that alters women's job search and acceptance behavior in the months following graduation, while having no impact on men. This suggests that this period plays a particularly critical role in shaping women's labor market prospects. To understand why these months are so pivotal for women, we turn to the dynamics of the marriage market as a potentially natural explanation.

Marriage Market Context Significant gender disparities characterize not only the labor but also the marriage market in Pakistan. Data from Pakistan's Demographic Health Survey 2017-18 (DHS) reveal that women marry significantly earlier than men: 61% of women versus only 24% of men (ages 25-49) were married by age 22. This trend persists even among more educated groups, where women with higher education marry at a median age of 25, 3.3 years earlier than men with similar educational attainment. Since women's marriage prospects are older men, it's important to note that the men in our sample are not the right cohort to consider when thinking about marriage matches for our female respondents. Given that decisions about both labor and marriage markets are often made in the short period of time following graduation, the timing of marriage could be crucial in understanding why this post-graduation period is so pivotal for women.

In line with DHS data, we find in Panel (a) of Figure A.15 that once the students in our sample graduate, women expect their marriage to happen sooner than men: 52% of women expect to be married within 2 years of graduation and 75% expect to be married within 3 years of graduation (compared to 37% and 57% of men, respectively). The modal woman in our sample is 22 years of age which means that, consistent with DHS data, most women expect to be married by age 24-25. While marriage may be a few years away, the search for a suitable match starts right after graduation: Panel (b) of Figure A.15 shows that the share of women who received at least one marriage offer post-graduation grows dramatically from 3.8% at graduation to 30% two months, and 39% six months after graduation (compared to 2.6%, 18% and 25%, respectively, for men). For the women who have not yet received a marriage offer by six months post-graduation, the pressures from the marriage market may be similar, if not accentuated by the scarcity of marriage offers. Thus, women's potential entry into the labor and marriage markets is determined soon after graduation, often simultaneously.

Marriage Market Beliefs: Elicitation To understand women's beliefs about marriage timing, we asked them at baseline: 'When do you expect to get married? Please answer in months or years from now.' Adding this belief to their current age we can assess the age at which they expect to get married. Using this information, we construct a relative measure of women's marriage expectations, as we did with their employment expectations. However, unlike with employment expectations, we use the ground truth about marriage age in Pakistan, rather than women's second order beliefs about the marriage timelines of other women, as we did not collect the latter. In practice, this means our measure of relative expectations about marriage timeline is an indicator equal to one if a woman believes they will marry later than age 25, the national median age of marriage for college-educated women.

Incorrect Beliefs or Differences in Type? We characterize women who expect to marry later as being "relatively optimistic." However, these women do not necessarily hold incorrect beliefs. Instead, they may simply be different from the average woman in a way that will lead them to actually marry later. If these women are indeed correct about their marriage timelines, we should observe that their beliefs remain stable after graduation, when more precise information about the marriage market is acquired. To test this, we asked women about their marriage expectations again six months after graduation, allowing us to examine whether women revised these expectations since baseline. Figure 9 illustrates the relationship between women's initial expectations (x-axis) and their updated beliefs about their age at marriage six months post-graduation (y-axis), separately for control and treated women. The 45 degree line represents the benchmark of time-invariant beliefs. While there is a positive slope (of ~ 0.6), it is far from 1. This means women who initially expected to marry after age 25 are much more likely to revise downward their expectations. For instance, women who believed they would get married at age 26 have now converged to believing they will get married a year earlier (matching the national median). In contrast, women who initially anticipated marrying sooner (e.g., by age 24) have near-stable expectations. The fact that those who expected to marry later are the ones who update their beliefs the most suggests that they held incorrect beliefs at first, and were exposed to an earlier onset of marriage-related activities than they had expected. It is also worth noting that the treatment has no effect on these marriage timeline beliefs.

Marriage Market Belief: Heterogeneous Treatment Effects We find that our treatment effects are stronger for women who are relatively optimistic about their marriage timeline. Table 8 shows that these women are 11.8 ppt (20.0%) more likely to be in the labor force, 10.3 ppt (29.4%) more likely to work, 11.0 ppt (46.8%) more likely to work full-time, and 14.3 ppt (70.8%) more likely to work full-time for a firm. Additionally, they submitted 2.3 more job applications (a 26.3% increase), did not receive more offers than the control group, but were 14.8 ppt (33.6%) more likely to accept an offer they received. In contrast, for women who are relatively pessimistic about their marriage timeline, the treatment effects are much smaller and not statistically significant across all labor market outcomes.

Underestimating the speed at which the marriage market unfolds may have led optimistic women to delay their job search. Our findings provide suggestive evidence of this phenomenon: Column 1 of Table 9 demonstrates that greater relative expectations about employment prospects at baseline is positively associated with longer expected time until marriage. Specifically, a 10 ppt increase in relative employment expectations is associated with a 1.5 ppt ($p<0.05$) increase in the chances of expecting to be married after age 25. Column 3 shows that a 10 ppt increase in relative employment expectations is further associated with about an additional one month reduction in the expected age of marriage by the six-month follow-up. This indicates that women who were more optimistic about their employment prospects at baseline are subsequently also more likely to lower their anticipated marriage age. While Column 2 shows no significant relationship between relative wage expectations and the likelihood of expecting to marry after age 25, we find that higher relative wage expectations is associated with further downward updating of marriage beliefs.

Identifying Women of a Different Type If the role of timing in labor market outcomes is at least partly due to competing forces in the marriage market, then our treatment should have little impact on the women for whom these two markets are independent. While identifying this sub-population of women is challenging, we can proxy for it by looking at women from families with a working mother or working sister(s). The idea is that the families of these women may be more supportive of women's work, regardless of the marriage market. If treatment operates by reducing overlap between launch of the marriage and labor markets, women from such families should be less affected by it. Table B.4 tests this hypothesis by separately looking at our treatment effects for women with and without a working mother or sister(s). Panel A focuses on women with a working mother or sister(s). The impact of treatment on this subgroup is generally small and statistically insignificant across all labor market outcomes. For example, in Column 1 (Labor Force Participation), the treatment effect for women with a working mother or sister is just 0.049 ($p > 0.1$), and in Column 2 (Working), the treatment effect is 0.003 ($p > 0.1$). In contrast, Panel B shows the results for women without a working mother or elder sister, where the treatment effects are larger and statistically significant across most labor market outcomes. In Column 1, the treatment increases labor force participation by 10.5 ppt ($p < 0.01$), while in Column 2, it increases the likelihood of being employed by 10.4 ppt ($p < 0.01$). Overall, the results suggest that while the treatment has a broadly positive effect on women's labor market outcomes, it makes little difference for the women who are from working women families.

Treatment Effects on Marriage Outcomes A related question that arises is whether our intervention not only influenced the timing of women's entry into the labor market but also impacted their engagement with the marriage market. This is relevant as encouraging women to enter the workforce earlier may inadvertently lead to a delay in their entry into the marriage market or affect the nature of the marriage offers they receive. Marriage is still a few years out and therefore not an outcome we are able to track, but we can see whether treatment impacted the number or quality of early marriage offers, which are a key signal of a woman's prospects on the marriage market. We test this in Table 10. In Column 1, we find that the treatment had no effect on whether women received any marriage offers after graduation. Similarly, Column 2 shows that there was no intensive margin effect on the number of marriage offers received. Column 3 shows that there is also no effect on the quality of marriage offers received by women, as proxied by an indicator for whether the highest education level among all the received marriage offers is a Master's degree or higher.[22] Overall, the results suggest that the treatment had no significant impact on the quantity or quality of marriage market offers for either gender. This sheds further light on the idea that our treatment works not by delaying the onset of the marriage market, but by inducing the labor market to unfold earlier.

Discussion In previous sections we found that our treatment effects were driven by women who, at baseline, had high relative optimism about their future employment prospects. In this section, we show that, at baseline, these women are also more likely to expect to marry

22. This question was only asked to women, and not men.

later than the average college educated woman in Pakistan. However, they seem to hold incorrect beliefs about their marriage timeline because they exhibit a higher tendency to adjust their marriage expectations downwards six months after graduation. This update in their future projections indicates that they had underestimated how soon they would enter the marriage market. We find that the group of women who had jointly overestimated their future labor supply and underestimated the immediacy of their exposure to the marriage market stand to gain the most from our treatment. This suggests that, as time passes post-graduation, the relative importance of the marriage market unexpectedly increases, at the expense of labor market opportunities. One plausible explanation for the relative change in priorities is parental intervention: parents may increasingly weigh the impact of employment on marriage prospects over time, possibly perceiving that working could limit marriage opportunities (as evidenced in other studies such as Subramanian (2024)). We are not able in this study to provide definitive evidence on this explanation because we do not directly measure the beliefs of families or in-laws and the men we survey are not the right sample to infer the beliefs of future husbands, who are typically a few years older. We can however provide evidence that parents are indeed very involved in students' job search, especially for women. Indeed, we asked both men and women in our descriptive sample, six months after they graduated: "Whom do you (or would you) seek advice from when deciding whether to accept a job offer?". Figure A.18 shows the distribution of the responses for both men and women. The daughters' self-reports reveal that family members play a significant role in the decision to accept job offers: 72.2% of women (and 48.3% of men) say they consult their family when making this decision.

7 Conclusion

This study presents novel findings establishing that college-educated women in Pakistan overwhelmingly expect to work at the time of their college graduation. However, most of them are not employed or in the labor force six months post-graduation. Traditional demand and supply-side explanations do not explain the resulting gender employment gap: women have more human capital (higher GPA conditional on choice of major), lower reservation wages, and no differential preference for non-wage amenities relative to their male peers. Women are just as likely as men to receive interviews and job offers. However, conditional on receiving a job offer, they are more likely than men to reject it and remain unemployed.

In this context, our study identifies a novel predictor of women's employment: the timing of their job search. Women who had applied to jobs within two months of graduation were 19.1 ppt more likely to work compared to women who had not. By contrast, there are no returns to early applications for men. To differentiate between selection and the causal effects of early job search, we designed and implemented an experiment that offered a small financial incentive to a randomly selected group of students to encourage earlier job applications. This incentive induced both men and women to expedite their job search and led to a significant increase in employment among women, while men's employment remained unaffected. Treatment effects were highest for women with high relative expectations (measured by their beliefs about their own employment prospects relative to their female peers). We show that our intervention worked by helping women overcome an unexpected decline in their labor supply over time. We rule out alternative mechanisms: increased search effort, and overcoming an unexpected slowdown of labor demand. Additionally, we provide suggestive evidence that the concurrent timing of the marriage market may partly explain the observed decline in women's labor market attachment over time.

Our study offers valuable insights for shaping cost-effective and scalable policies aimed at bolstering highly educated women's employment. One key takeaway for future policy design is that intervening early, ideally while women are still in college, is crucial. At this time, they hold strong beliefs that they will work, and have weaker marriage market exposure. College counseling programs or role model interventions that induce women to apply earlier are therefore promising, and so would be interventions that encourage early labor market exposures, such as internship programs.

References

Adams, Abi, and Alison Andrew. 2024. "Revealed Beliefs and the Marriage Market Return to Education." *IFS Working Paper.*

Addati, Laura, Umberto Cattaneo, Valeria Esquivel, and Isabel Valarino. 2018. *Care Work and Care Jobs for the Future of Decent Work.* International Labour Organisation (ILO).

Agte, Patrick, and Arielle Bernhardt. 2024. "The Economics of Caste Norms: Purity, Status, and Women's Work in India." *Working Paper.*

Alfonsi, Livia, Mary Namubiru, and Sara Spaziani. 2024. "Meet Your Future: Experimental Evidence on the Labor Market Eects of Mentors." *Working Paper.*

Aloud, Monira Essa, Sara Al-Rashood, Ina Ganguli, and Basit Zafar. 2020. "Information and Social Norms: Experimental Evidence on the Labor Market Aspirations of Saudi Women." *NBER Working Paper,* no. 26693.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena. 2020. "Bride Price and Female Education." *Journal of Political Economy* 128 (2): 591–641.

Ashraf, Nava, Alexia Delfino, and Edward Glaeser. 2022. "Female Entrepreneurship and Trust in the Market."

Attanasio, Orazio P, and Katja M Kaufmann. 2017. "Education Choices and Returns on the Labor and Marriage Markets: Evidence from Data on Subjective Expectations." *Journal of Economic Behavior & Organization* 140:35–55.

Banerjee, Abhijit, and Sandra Sequeira. 2023. "Learning by searching: Spatial mismatches and imperfect information in Southern labor markets." *Journal of Development Economics* 164:103111.

Becker, Gary S. 1975. "Front matter, Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education." In *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, Second Edition,* 7–36. NBER.

Bertrand, Marianne. 2011. "New Perspectives on Gender." In *Handbook of Labor Economics,* 4:1543–1590. Elsevier.

Bertrand, Marianne, Patricia Cortes, Claudia Olivetti, and Jessica Pan. 2021. "Social Norms, Labour Market Opportunities, and the Marriage Gap Between Skilled and Unskilled Women." *The Review of Economic Studies* 88 (4): 1936–1978.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2 (3): 228–255.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani. 2021. "Labor rationing." *American Economic Review* 111 (10): 3184–3224.

Brown, Christina. 2022. "Understanding Gender Discrimination by Managers." *Working Paper.*

Buchmann, Nina, Carl Meyer, and Colin D. Sullivan. 2024. "Paternalistic Discrimination." *Working Paper.*

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais. 2017. "'Acting Wife': Marriage Market Incentives and Labor Market Investments." *American Economic Review* 107 (11): 3288–3319.

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott. 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review* 110 (10): 2997–3029.

Card, David, Fabrizio Colella, and Rafael Lalive. 2021. *Gender Preferences in Job Vacancies and Workplace Gender Diversity.* Working Paper, Working Paper Series 29350. National Bureau of Economic Research.

Chakraborty, Aranya, Digvijay Negi, and Rahul Rao. 2024. "Inefficiency in Agricultural Production: Do Information Frictions Matter?" *Working Paper.*

Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir. 2018. "The Marriage Market, Labor Supply, and Education Choice." *Journal of Political Economy* 126 (S1): S26–S72.

Conlon, John, Laura Pilossoph, Matthew Wiswall, and Basit Zafar. 2018. "Labor Market Search With Imperfect Information and Learning." *NBER Working Paper.*

Cortés, Patricia, Gizem Koşar, Jessica Pan, and Basit Zafar. 2024. "Should Mothers Work? How Perceptions of the Social Norm Affect Individual Attitudes Toward Work in the US." *Review of Economics and Statistics,* Forthcoming.

Cortés, Patricia, Jessica Pan, Laura Pilossoph, Ernesto Reuben, and Basit Zafar. 2023. "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab." *The Quarterly Journal of Economics* 138 (4): 2069–2126.

Dhar, Diva. 2021. "Indian Matchmaking: Are Working Women Penalized in the Marriage Market in India?" *SSRN Working Paper.*

Dinkelman, Taryn, and L Rachel Ngai. 2022. "Time Use and Gender in Africa in Times of Structural Transformation." *Journal of Economic Perspectives* 36 (1): 57–80.

Field, Erica, and Kate Vyborny. 2022. "Women's Mobility, and Labor Supply: Experimental Evidence from Pakistan." *Asian Development Bank Economics Working Paper Series,* no. 655.

Fluchtmann, Jonas, Anita M. Glenny, Nikolaj A. Harmon, and Jonas Maibom. 2024. "The Gender Application Gap: Do Men and Women Apply for the Same Jobs?" *American Economic Journal: Economic Policy* 16:182–219.

Gentile, Elisabetta, Nikita Kohli, Nivedhitha Subramanian, Zunia Tirmazee, and Kate Vyborny. 2023. "Barriers to entry: Decomposing the gender gap in job search in urban Pakistan." *Asian Development Bank Economics Working Paper Series,* no. 707.

Goldin, Claudia. 1992. "The Meaning of College in the Lives of American Women: The Past One-Hundred Years." *NBER Working Paper,* no. 4099.

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

Goldin, Claudia, and Lawrence F Katz. 1999. "The Returns to Skill in the United States across the Twentieth Century." *NBER Working Paper,* no. 7126.

Goldin, Claudia, and Cecilia Rouse. 2000. "Orchestrating Impartiality: The Impact Of 'Blind' Auditions On Female Musicians." *American Economic Review* 90 (4): 715–741.

Halim, Daniel, Elizaveta Perova, and Sarah Reynolds. 2023. "Childcare and Mothers' Labor Market Outcomes in Lower-and Middle-Income Countries." *The World Bank Research Observer* 38 (1): 73–114.

Heath, Rachel, and Seema Jayachandran. 2018. "The Causes and Consequences of Increased Female Education and Labor Force Participation in Developing Countries." In *The Oxford Handbook of Women and the Economy.* Oxford University Press.

Ho, Lisa, Suhani Jalota, and Anahita Karandikar. 2023. "Bringing Work Home: Flexible Arrangements as Gateway Jobs for Women in West Bengal." *Working Paper.*

Illing, Hannah, Johannes Schmieder, and Simon Trenkle. 2024. "The Gender Gap in Earnings Losses After Job Displacement." *Journal of the European Economic Association.*

Jaeger, David A, Jaime Arellano-Bover, Krzysztof Karbownik, Marta Martınez-Matute, John M Nunley, Alan Seals, Mackenzie Alston, Sascha O Becker, Pilar Beneito, René Böheim, et al. 2021. "The Global Covid-19 Student Survey: First Wave Results." *IZA Discussion Paper,* no. 14419.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer. 2024. "Worker Beliefs About Outside Options." *The Quarterly Journal of Economics* 139 (3): 1505–1556.

Jayachandran, Seema. 2021. "Social Norms as a Barrier to Women's Employment in Developing Countries." *IMF Economic Review* 69 (3): 576–595.

Jones, Sam, and Ricardo Santos. 2022. "Can information correct optimistic wage expectations? Evidence from Mozambican job-seekers." *Journal of Development Economics* 159:102987.

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante. 2023. *The Child Penalty Atlas.* Working Paper, Working Paper Series 31649. National Bureau of Economic Research.

Kline, Patrick, Evan K Rose, and Christopher R Walters. 2022. "Systemic Discrimination Among Large U.S. Employers." *The Quarterly Journal of Economics* 137 (4): 1963–2036.

Krueger, Alan B., and Andreas I. Mueller. 2016. "A Contribution to the Empirics of Reservation Wages." *American Economic Journal: Economic Policy* 8, no. 1 (February): 142–79.

Kuhn, Peter, and Kailing Shen. 2013. "Gender Discrimination in Job Ads: Theory and Evidence." *The Quarterly Journal of Economics* 128 (1):287–336.

———. 2023. "What Happens When Employers Can No Longer Discriminate in Job Ads?" *American Economic Review* 113, no. 4 (April): 1013–48.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington. 2018. "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?" *NBER Working Paper,* no. 24740.

Lafortune, Jeanne. 2013. "Making Yourself Attractive: Pre-marital Investments and the Returns to Education in the Marriage Market." *American Economic Journal: Applied Economics* 5:151–78.

Maestas, Nicole, Kathleen J. Mullen, David Powell, Till von Wachter, and Jeffrey B. Wenger. 2023. "The Value of Working Conditions in the United States and the Implications for the Structure of Wages." *American Economic Review* 113 (7): 2007–47.

Mas, Alexandre, and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 107 (12): 3722–59.

Mueller, Andreas I., Johannes Spinnewijn, and Giorgio Topa. 2021. "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias." *American Economic Review* 111, no. 1 (January): 324–63.

Our World in Data. 2023. *Female Labor Supply.* Accessed: 2024-10-01. https://ourworldindata.org/female-labor-supply.

Potter, Tristan. 2021. "Learning and job search dynamics during the Great Recession." *Journal of Monetary Economics* 117:706–722.

Roussille, Nina. 2024. "The Role of the Ask Gap in Gender Pay Inequality." *The Quarterly Journal of Economics* 139 (3): 1557–1610.

Spinnewijn, Johannes. 2015a. "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs." *Journal of the European Economic Association* 13 (1): 130–167.

———. 2015b. "Unemployed But Optimistic: Optimal Insurance Design With Biased Beliefs." *Journal of the European Economic Association* 13 (1): 130–167.

Subramanian, Nivedhitha. 2024. "Workplace Attributes and Women's Labor Supply Decisions: Evidence from a Randomized Experiment." *Working Paper.*

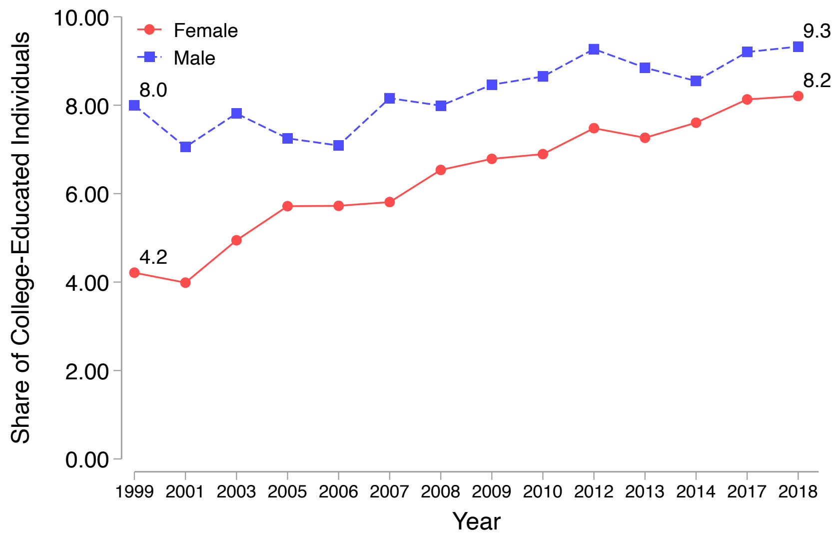
Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele, and Kosuke Imai. 2014. "mediation: R Package for Causal Mediation Analysis." *Journal of Statistical Software* 59 (5): 1–38.

Veerle, Miranda. 2011. "Cooking, Caring and Volunteering: Unpaid Work Around the World." *OECD Working Paper Series,* no. 116.

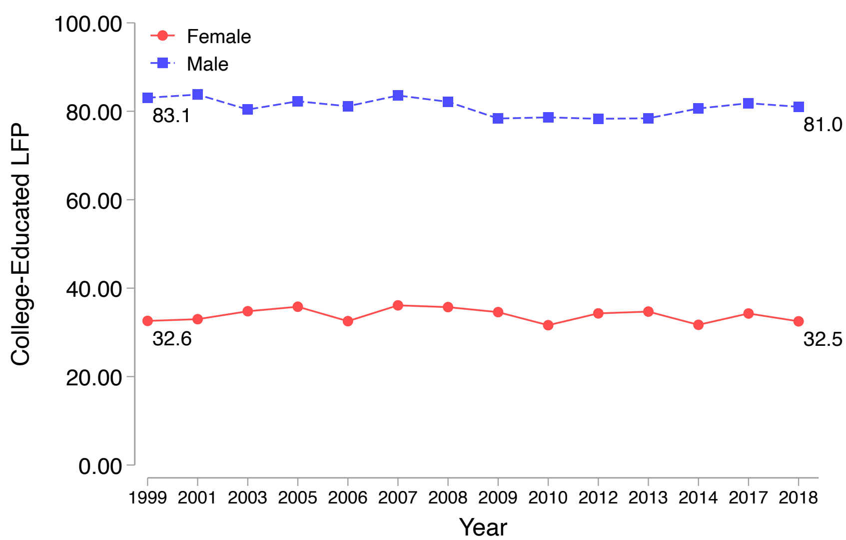
World Bank. 2021. *Women, Business and the Law 2021.* Accessed: 2024-10-01. World Bank. https://documents1.worldbank.org/curated/en/492221468136792185/pdf/Main-report.pdf.

Figures

Figure 1: College Education and Labor Force Participation in Pakistan



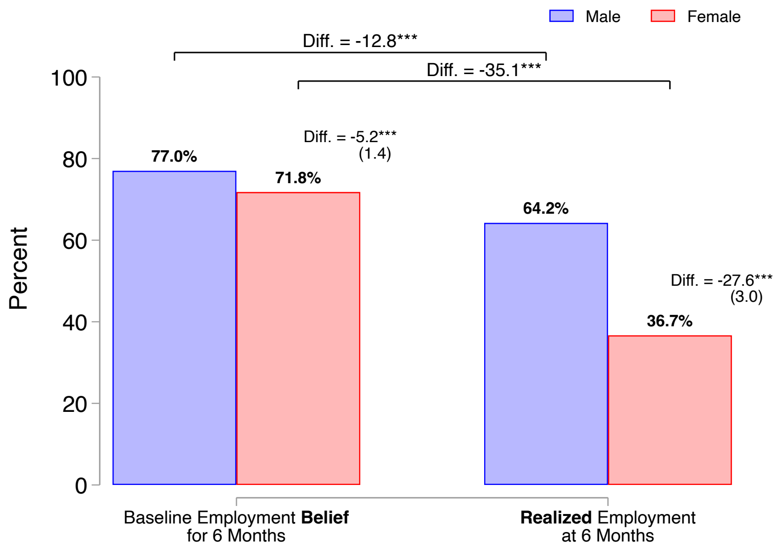
(a) Gender Gaps in College Education



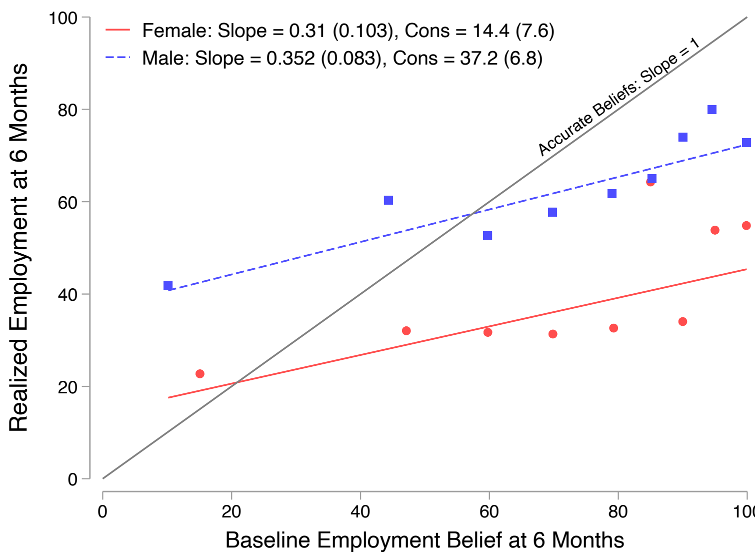
(b) Gender Gaps in Labor Force Participation of College-Educated Individuals

Notes: The figures show trends in college education and labor force participation between 1999-2018 in Pakistan. Panel (a) shows the shares of men and women between age 18-35 who are college-educated. Panel (b) shows the labor participation rates of college-educated men and women between age 18-35. Data is obtained from the Pakistan Labor Force Surveys.

Figure 2: Baseline Employment Beliefs vs. Realized Employment Outcomes



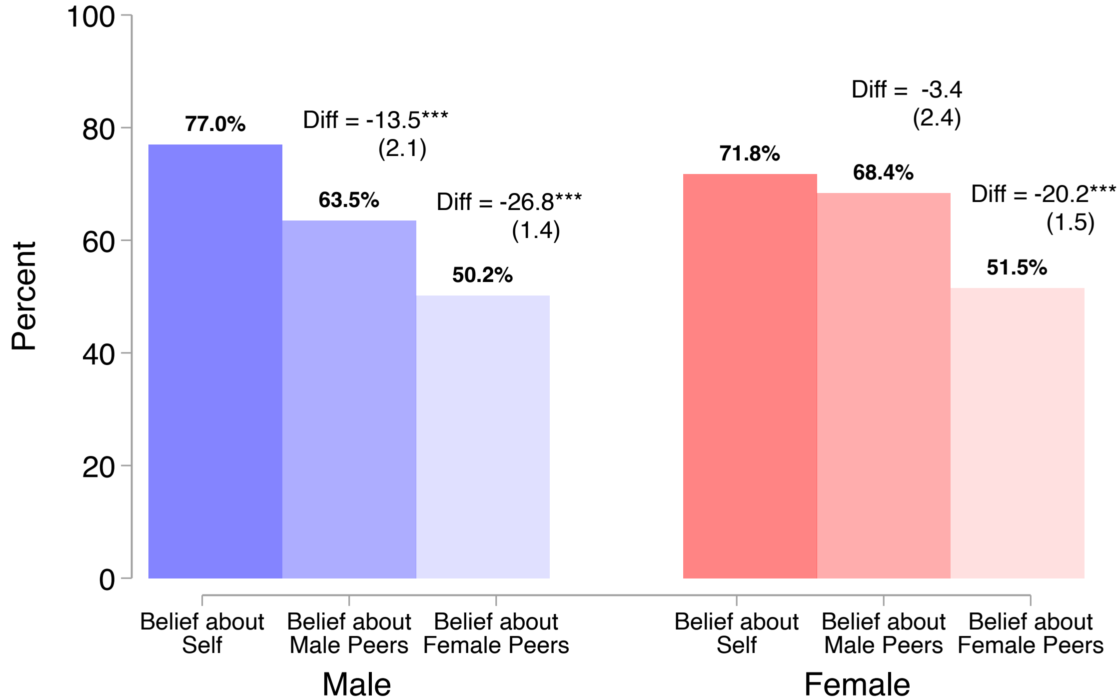
(a) Mean Levels: Intended vs. Realized Employment



(b) Binscatter: Intended vs. Realized Employment

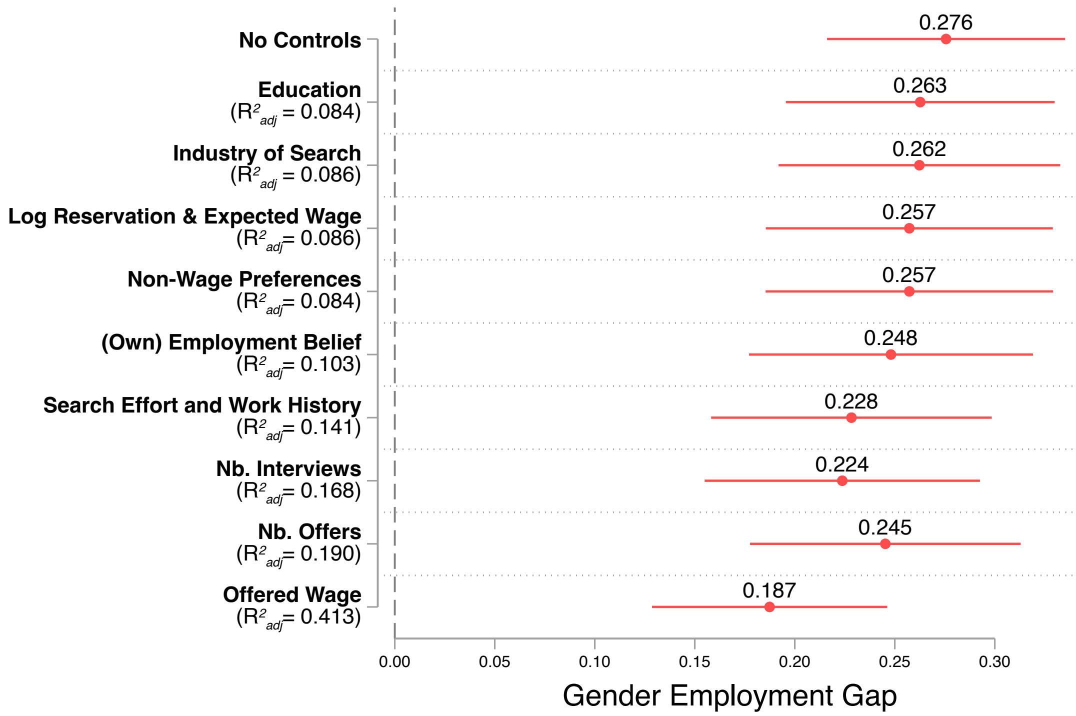
Notes: The figure compares students' baseline beliefs about employment with realized employment. Panel (a) plots the average self-assessed likelihood of employment among men and women 6 months after graduation (left) vs. the percentages who are actually employed by then (right). It also estimates and tests for the differences between men and women (immediately above the bars) and the differences between belief and realized employment within gender (above the brackets). Panel (b) is a binscatter plot of realized employment vs. baseline beliefs by gender. The 45-degree line corresponds to accurate beliefs. Students above the line overestimate their chances, while those below underestimate them. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure 3: Employment Beliefs about Self vs. Employment Beliefs about Peers



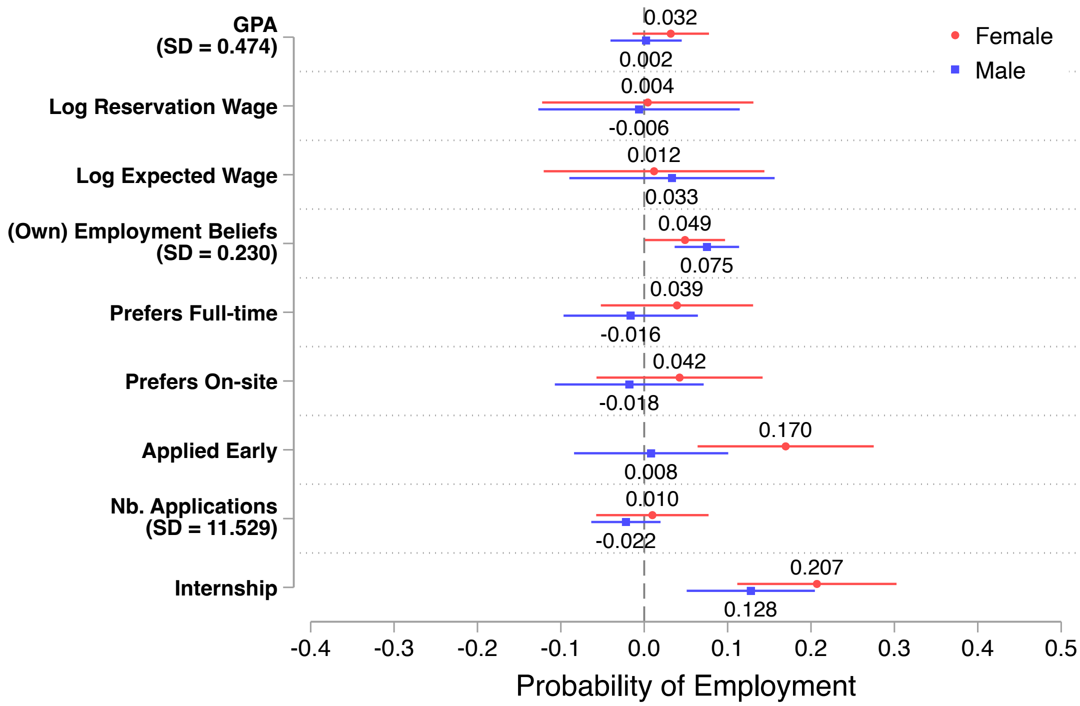
Notes: The figure shows men's (left) and women's (right) average baseline beliefs about the employment likelihood of different groups. Within gender, each bar from left to right reflects average baseline beliefs about oneself, one's male peers, and one's female peers. Within gender, numbers floating above the middle bar estimate and test for the difference between "Belief about Male Peers" and "Belief about Self", and numbers floating above the right bar estimate and test for the difference between "Belief about Female Peers" and "Belief about Self". * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure 4: Explaining the Gender Employment Gap 6 Months after Graduation



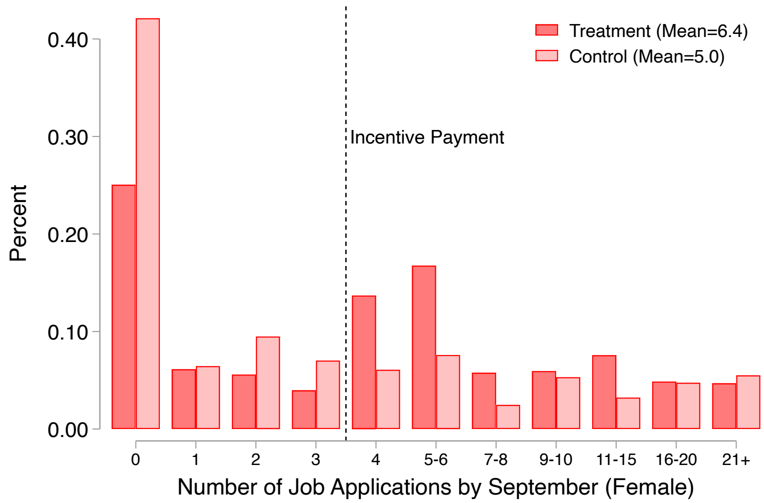
Notes: This figure reports coefficients (re-scaled to be positive) from regressing an employment indicator, measured at six months after graduation, on the female dummy while incrementally adding controls. The goal is to assess to what extent the observed gender employment gap can be explained by observable characteristics. Each row shows the coefficient after controlling for variables specified in that row and all rows above. Horizontal bars indicate 95% confidence intervals. The **Education** control includes cumulative GPA and major fixed effects. The **Industry of Search** control includes SOC sub-major group fixed effects based on semantic occupation classification. The **Reservation and Expected Wage** control adjusts for respondents' wage expectation at baseline. The **Non-Wage Preferences** control adjusts for preference for onsite vs. remote work and the number of preferred hours per day. The **Employment Belief** control includes baseline belief about one's probability of employment at 6 months post graduation. The **Search Effort and Work History** control includes the total number of job applications sent by the 6-month follow-up, a dummy for applying early, and a dummy for having internship experience. The **Nb. Interviews** control adjusts for the number of interviews by the 6-month follow-up. The **Nb. Offers** control includes the number of job offers by the 6-month follow-up. The **Offered Wage** control is one's current wage if the respondent is working, or the highest rejected wage offer if a respondent does not work. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure 5: Timing Distinctively Predicts Women's Employment (Pooled Estimates)

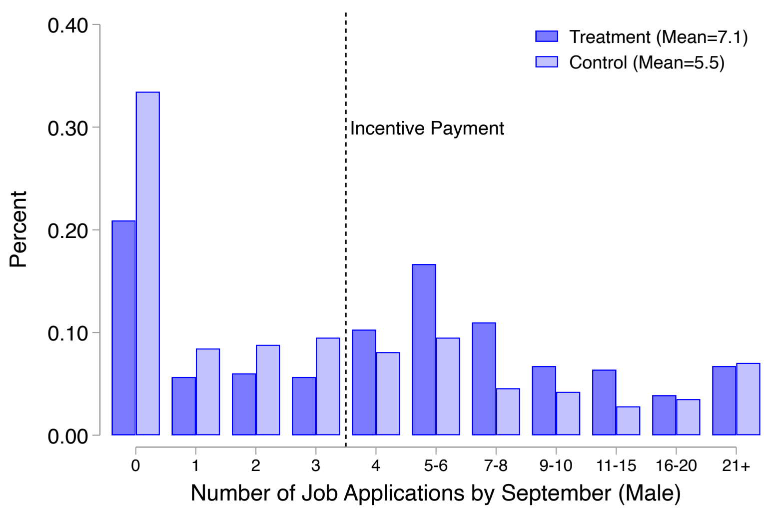


Notes: The figure reports coefficients from a pooled regression where we regress the probability of employment simultaneously on a set of potential employment predictors. The regression is run separately for men and women. Horizontal bars indicate 95% confidence intervals. The predictors include a student's GPA, reservation wage, expected wage, baseline self-reported likelihood of employment six months post-graduation (on a 0-1 scale), preference for full-time vs. part-time work, preference for on-site vs. remote work, whether the student applied to jobs early, the number of job applications submitted (measured in standard deviations), and prior internship experience. Figure A.11 shows the coefficients of individual, rather than pooled, regressions using the same set of predictors. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure 6: Distribution of Number of Job Applications (2-Month Follow-Up)



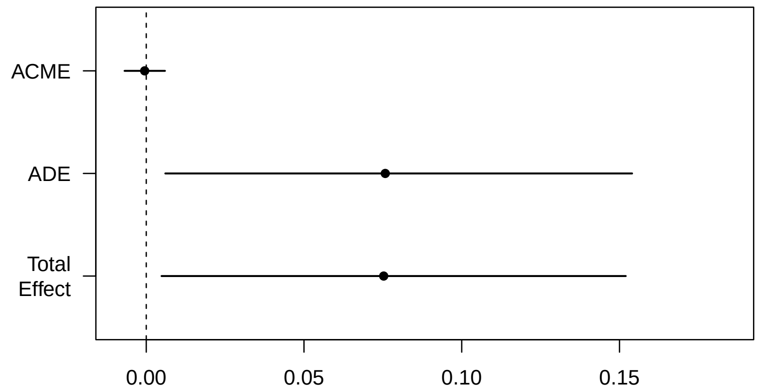
(a) Female: Distribution of Job Applications



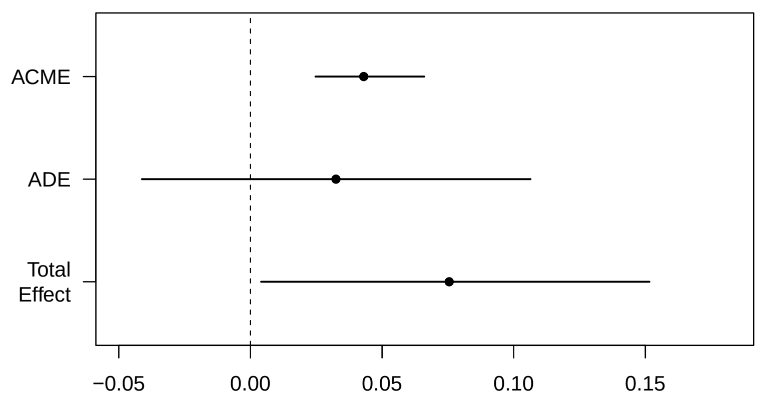
(b) Male: Distribution of Job Applications

Notes: The figures show distributions of the numbers of job applications students sent between baseline and the first follow-up survey in September 2023. Panel (a) shows a histogram of women's numbers of applications by treatment status, and Panel (b) shows a histogram of men's numbers of applications by treatment status. The dotted line locates our experimental incentive: students who sent 4 or more applications receives a payment. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Figure 7: Mediation Analysis: Search Effort vs. Timing



(a) Search Effort



(b) Timing (Month of Application)

Notes: The figures show results of a mediation analysis that assesses whether treatment effects operate through timing of search rather than search effort. It does so by decomposing the total effect into Average Causal Mediation Effect (ACME), and the Average Direct Effect (ADE). Panel (a) tests whether the number of job applications contributes to treatment effects. Panel (b) tests whether the total treatment effects are driven by the month of first application. Controls include GPA, college major fixed effect, belief about employment in six months, preferences for full-time and on-site work, and baseline reservation wage. The confidence intervals are calculated at 95% levels, using heteroskedasticity-consistent standard errors of type 1 (HC1). The sample consists of female students in the experimental sample at the public university (see Section 5.1 for details).

Figure 8: Theoretical Framework

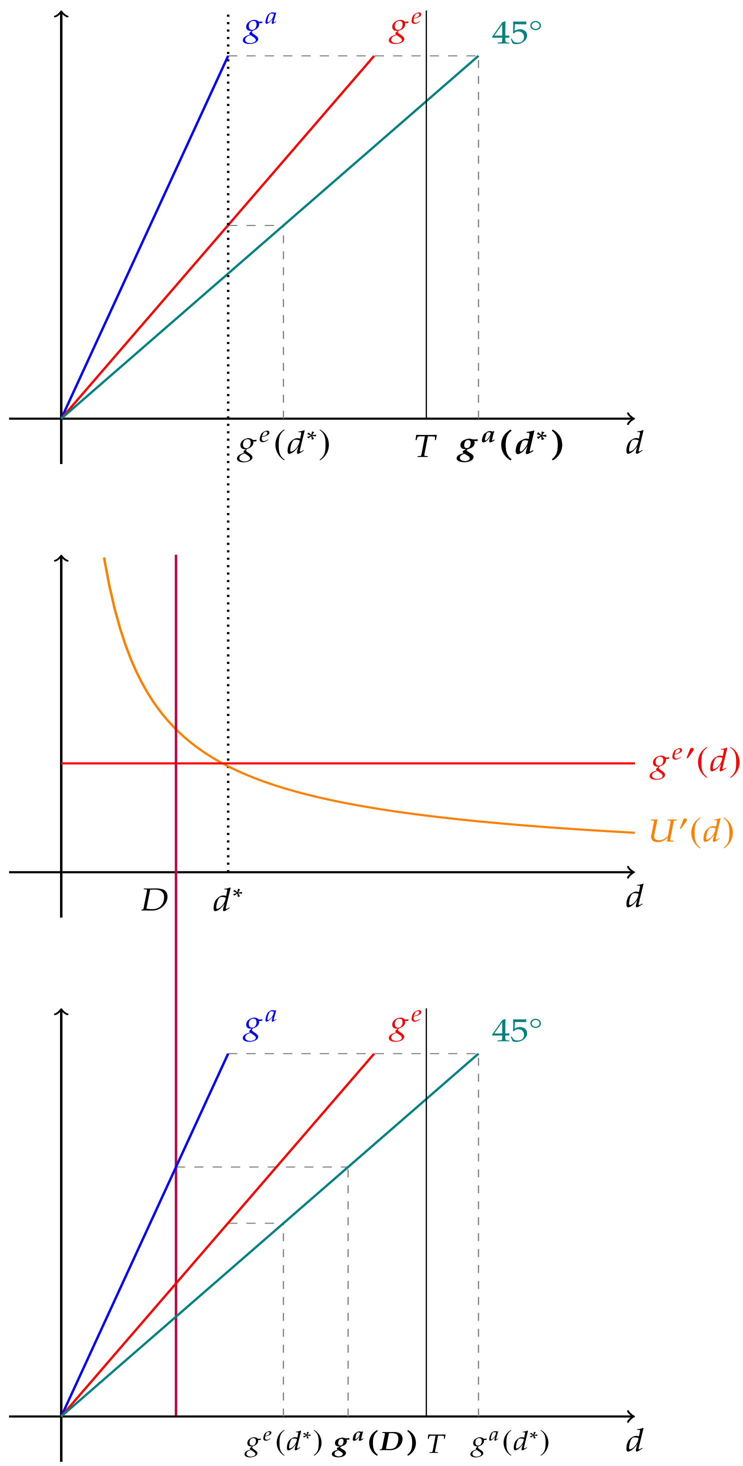
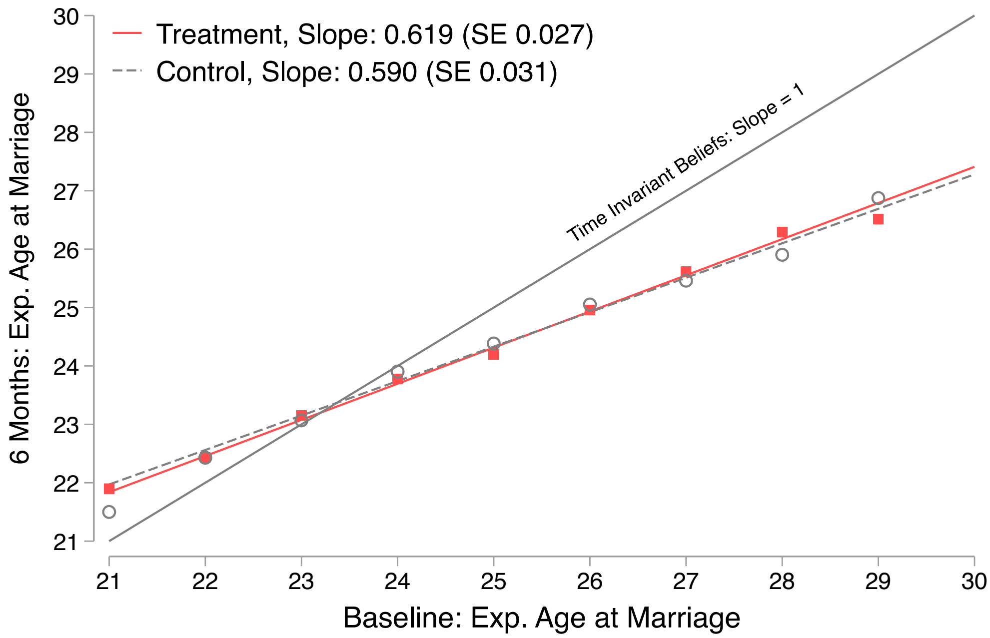


Figure 9: Women's Belief Updating About Expected Age at Marriage



Notes: The figure shows the relationship between women's expected age of marriage at baseline and their updated beliefs reported after 6 months, split by treatment status. The 45-degree line corresponds time-invariant beliefs, meaning respondents provided the same estimates for the age at marriage at both the baseline and the 6-month follow-up. The sample consists of female students in the experimental sample at the public university (see Section 5.1 for details).

Tables

Table 1: Descriptive Statistics for the Descriptive Sample

	All	Male	Female	Diff.	p-value
Nb. Observations	1,029	590	439		
Age	22.5	22.7	22.2	0.5	0.00
GPA	3.1	3.0	3.2	-0.3	0.00
Married	4.3	2.4	6.8	-4.5	0.00
Engaged	6.6	5.9	7.5	-1.6	0.32
Majors:					
Engineering / Computer Science	26.2	39.2	8.9	30.3	0.00
Humanities, Languages and Education	15.5	13.4	18.5	-5.1	0.03
Life Sciences / Pharmacy	12.1	5.1	21.6	-16.6	0.00
Sciences	13.2	5.8	23.2	-17.5	0.00
Social Sciences (inc. Business and Law Degrees)	32.8	36.6	27.8	8.8	0.00
Parental Background:					
College-Educated Mother	41.0	40.7	41.5	-0.8	0.80
College-Educated Father	53.2	52.0	54.7	-2.6	0.40

Notes: This table presents descriptive statistics for respondents who answered the baseline and six-month surveys in the descriptive sample at the private university (see Section 3.1 for details). The last two columns compute the difference between gender and report the p-value for testing equality of means between genders.

Table 2: Descriptive Statistics for the Experimental Sample

	All	Male	Female	Diff.	p-value
Nb. Observations	1,442	516	926		
Age	22.8	23.3	22.5	0.8	0.00
GPA	3.3	3.2	3.4	-0.2	0.00
Married	4.4	4.3	4.4	-0.2	0.88
Engaged	3.9	3.1	4.3	-1.2	0.23
Majors:					
Engineering / Computer Science	7.1	6.4	7.5	-1.1	0.44
Humanities, Languages and Education	26.3	27.9	25.4	2.5	0.30
Life Sciences / Pharmacy	12.6	10.1	14.0	-4.0	0.02
Sciences	27.0	29.7	25.6	4.1	0.10
Social Sciences (inc. Business and Law Degrees)	27.0	26.0	27.5	-1.6	0.52
Parental Background:					
College-Educated Mother	27.9	18.8	32.9	-14.1	0.00
College-Educated Father	42.6	36.4	46.1	-9.7	0.00
Working Mother	6.9	5.6	7.6	-1.9	0.15
Working Father	86.0	84.9	86.6	-1.7	0.43
Family Owns Car	49.3	44.4	51.9	-7.4	0.02
Family Owns Motorbike	93.3	93.0	93.5	-0.5	0.73
Family has Internet	86.7	80.3	90.1	-9.8	0.00
Family has Laptop	84.9	84.7	85.0	-0.3	0.88
Family has Smartphone	99.6	99.5	99.6	-0.1	0.81

Notes: This table presents descriptive statistics for respondents who answered the baseline and six-month surveys in the experimental sample at the public university (see Section 5.1 for details). The last two columns compute the difference between gender and report the p-value for testing equality of means between genders.

Table 3: Labor Market Effects of the Experiment

	LFP	Working	Working FT	Working Firm FT	Nb. Apps	Nb. Offers	Accepted Offer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Female							
Treatment	0.096***	0.074**	0.082***	0.092***	1.469*	0.144	0.084**
	(0.032)	(0.032)	(0.029)	(0.028)	(0.766)	(0.172)	(0.041)
Female Control Mean	0.565	0.336	0.221	0.203	8.163	2.007	0.464
Panel B: Male							
Treatment	-0.008	0.014	-0.016	-0.002	0.524	0.186	-0.022
	(0.037)	(0.044)	(0.043)	(0.041)	(1.104)	(0.213)	(0.050)
Male Control Mean	0.776	0.551	0.413	0.307	9.530	1.909	0.710
Nb. observations	1,442	1,442	1,442	1,442	1,433	1,433	927

Notes: The table presents estimated effects of the treatment on labor market outcomes. Panel A presents results for women (926 observations), comparing the treatment's effects (treated-control differences) against control means, while Panel B presents results for men (516 observations). Column (1) reports the treatment's effects on labor force participation. Column (2) reports effects on employment. Column (3) reports effects on working full time for a wage. Column (4) reports effects on working full time for a firm and a wage. Column (5) reports effects on the number of job applications sent. Column (6) reports effects on the number of job offers received. The numbers of observations in Columns (5) and (6) are slightly lower because a few respondents did not report the number of applications or offers. Column (7) reports effects on the likelihood of accepting and offer, among those who received at least one job offer after graduation (or are currently working). Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Table 4: OLS and IV for Employment by Early Application Indicator

	Applied Early	OLS Working	2SLS Working
	(1)	(2)	(3)
Panel A: Female			
Treatment	0.138***		
	(0.030)		
Applied Early		0.267***	0.639***
		(0.032)	(0.243)
Female Control Mean	0.664	0.184	0.184
Panel B: Male			
Treatment	0.095***		
	(0.036)		
Applied Early		0.064	0.090
		(0.055)	(0.469)
Male Control Mean	0.744	0.510	0.510
Nb. observations	1,384	1,384	1,384

Notes: The table presents OLS (Ordinary Least Squares) and 2SLS (Two-Stage Least Squares) estimates of the effects of the treatment and applying early. Panel A reports results for women, while Panel B presents results for men. Column 1 reports the treatment's effects on whether one applied early. Column 2 reports the OLS estimates from regressing employment at 6 months on whether one applied early. Column 3 uses the exogenous treatment as an instrumental variable and reports the 2SLS estimates of the effect of applying early on employment at 6 months. The last rows in each panel show the means of the outcome variables in the control group. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Table 5: Unexpected Changes in Women's Labor Supply

	Rej. Offer above RW		
	(1)	(2)	(3)
Female	0.091***	0.002	0.001
	(0.025)	(0.050)	(0.052)
First Job Offer Month		-0.006	-0.009
		(0.010)	(0.011)
First Job Offer Month*Female		0.032**	0.034**
		(0.016)	(0.016)
Male Control Mean	0.093	0.093	0.093
Supply-Side Controls			X
Adjusted R-squared	0.014	0.018	0.019
Nb. observations	774	774	774

Notes: The table presents evidence that job offers' timing shapes women's likelihood of rejecting an offer above their reported baseline reservation wages. Column (1) reports the raw difference in rejection likelihood between women and men. Column (2) reports regression coefficients controlling for the month of first job offer interacted with a female dummy. Column (3) adds further supply-side controls, including GPA, baseline employment belief, and college major and industry of search fixed effects. Reservation wage is normalized to full-time for part-time work. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details). There are fewer observations in this table than in Table 3 because 668 respondents either had zero offer or did not provide the month of the job offer.

Table 6: Employment Outcome by Relative Wage and Employment Expectations

	Outcome: Working at 6 Months					
	Female			**Male**		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.071**	0.069**	0.064**	0.006	0.032	0.024
	(0.032)	(0.032)	(0.032)	(0.044)	(0.045)	(0.046)
Relative Wage Expectation	0.120*		0.110*	-0.045		-0.040
	(0.063)		(0.063)	(0.083)		(0.081)
Treatment*Relative Wage Expectation	-0.175*		-0.194**	0.181		0.172
	(0.091)		(0.089)	(0.119)		(0.118)
Relative Emp. Expectation		0.161*	0.144*		-0.116	-0.112
		(0.086)	(0.086)		(0.124)	(0.125)
Treatment*Relative Emp. Expectation		0.221*	0.248**		0.235	0.224
		(0.119)	(0.119)		(0.164)	(0.165)
Control Mean	0.336	0.336	0.336	0.551	0.551	0.551
Nb. observations	1,442	1,442	1,442	1,442	1,442	1,442

Notes: The table presents relationships between work status at 6 months and relative wage and employment expectations, and whether these relationships differ between treated and control students. Columns (1)-(3) report regression estimates for females, while columns (4)-(6) report regression estimates for males. Within gender, the first two columns estimate relationships between work and each of the two variables separately, interacted with treatment. The third column includes both relative wage and relative employment expectations in the model. Relative expectations are the difference between a student's self-assessed likelihood of employment or expected wage (in percentage units of peer wage) vs. what they expect for same-gender peers, within 6 months of graduation. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Table 7: Application Timing by Relative Wage and Employment Expectations

	Outcome: Months Between Graduation and First Application					
	Female			Male		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-1.543***	-1.564***	-1.574***	-0.702**	-0.739**	-0.666**
	(0.218)	(0.221)	(0.223)	(0.291)	(0.309)	(0.307)
Relative Wage Expectation	0.411		0.446	0.933		0.937
	(0.483)		(0.487)	(0.634)		(0.634)
Treatment*Relative Wage Expectation	-0.209		-0.245	-1.082		-1.099
	(0.631)		(0.636)	(0.846)		(0.846)
Relative Emp. Expectation		-0.434	-0.502		0.012	-0.086
		(0.693)	(0.691)		(0.780)	(0.782)
Treatment*Relative Emp. Expectation		0.468	0.511		0.353	0.461
		(0.890)	(0.892)		(1.048)	(1.050)
Control Mean	3.931	3.931	3.931	3.583	3.583	3.583
Nb. observations	1,442	1,442	1,442	1,442	1,442	1,442

Notes: The table presents relationships between application timing and relative wage and employment expectations, and whether these relationships differ between treated and control students. Application timing is measured as the number of months between graduation and first application and ranges between 0 and 6 months. We code people who did not submit any application by our 6-month followup as "7 months" and use a tobit model to address censoring. Columns (1)-(3) report regression estimates for females, while columns (4)-(6) report regression estimates for males. Within gender, the first two columns estimate relationships between application timing and each of the two variables separately, interacted with treatment. The third column includes both relative wage and relative employment expectations in the model. Relative expectations are the difference between a student's self-assessed likelihood of employment or expected wage (in percentage units of peer wage) vs. what they expect for same-gender peers, within 6 months of graduation. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Table 8: Labor Market Effects of the Experiment with Heterogeneity by Marriage Expectations

	LFP	**Working**	**Working FT**	**Working Firm FT**	**Nb. Apps**	**Nb. Offers**	**Accepted Offer**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Exp. Marriage > 25							
Treatment	0.128***	0.114**	0.119***	0.153***	2.704**	0.264	0.149**
	(0.048)	(0.050)	(0.046)	(0.045)	(1.205)	(0.274)	(0.062)
> 25 Control Mean	0.585	0.347	0.233	0.199	8.301	2.029	0.446
Panel B: Exp. Marriage ≤ 25							
Treatment	0.065	0.012	0.022	0.023	0.758	0.056	-0.015
	(0.046)	(0.045)	(0.040)	(0.040)	(1.095)	(0.246)	(0.059)
≤ 25 Control Mean	0.559	0.342	0.234	0.225	8.509	2.091	0.493
Nb. observations	845	845	845	845	837	837	551

Notes: The table presents evidence of heterogeneous treatment effects on women's labor market outcomes between those who expect to marry later vs. those who expect to marry sooner. We define "later" and "sooner" based on baseline responses relative to the age 25 — the national median age at marriage for college-educated women. Panel A (455 observations) presents results for women who expect to marry later, comparing the treatment's effects (treated-control differences) against control means, while Panel B (389 observations) presents results for women who expect to marry sooner. Column (1) reports the treatment's effects on labor force participation. Column (2) reports effects on employment. Column (3) reports effects on working full time and for a wage. Column (4) reports effects on working full time for a firm and a wage. Column (5) reports effects on the number of job applications sent. Column (6) reports effects on the number of job offers received. The numbers of observations in Columns (5) and (6) are slightly lower because a few respondents did not report the number of applications or offers. Column (7) reports effects on the likelihood of accepting a job offer, among respondents who received at least one job offer after graduation or are currently working. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of female students in the experimental sample at the public university (see Section 5.1 for details), excluding those who are already married or engaged at baseline. As a result, the sample size is smaller than in Table 3.

Table 9: Relationship between Employment, Wage and Marriage Expectations

	Exp. Age at Marriage >25		**Update in Exp. Age at Marriage**	
	(1)	(2)	(3)	(4)
Relative Emp. Expectations	0.156**		-0.508***	
	(0.070)		(0.190)	
Relative Wage Expectations		0.025		-0.243*
		(0.050)		(0.137)
Constant	0.417	0.458	-0.650	-0.771
Nb. observations	845	845	791	791

Notes: The table presents the relationship between women's beliefs about their marriage timelines and the labor market. Column (1) reports the coefficient from regressing an indicator for expecting to marry later than at age 25 on relative employment expectations, both measured at baseline. Column (2) reports the coefficient from regressing the same marriage indicator on relative wage expectations, both measured at baseline. Column (3) reports the coefficient from regressing an update measure — the number of years by which a woman updates her expected age at marriage by the 6-month follow-up — on baseline relative employment expectations. Column (4) reports the coefficient from regressing the same update measure on baseline relative wage expectations. Robust standard errors are reported in parentheses. Relative expectations are the difference between a student's self-assessed likelihood of employment or expected wage (in percentage units of peer wage) vs. what they expect for same-gender peers, within 6 months of graduation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of female students in the experimental sample at the public university (see Section 5.1 for details), excluding those who are already married or engaged.

Table 10: Treatment Effects of the Experiment on Marriage Outcomes

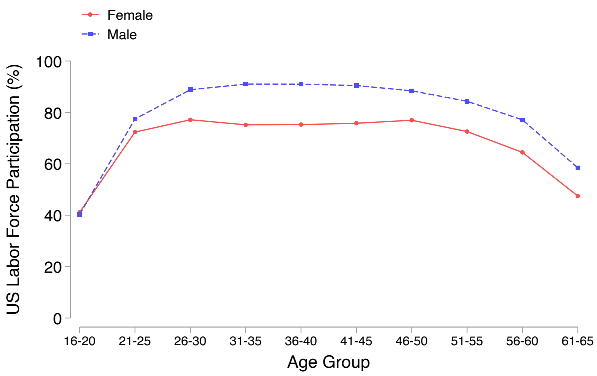
	Any Marriage Offers	Nb. Marriage Offers	Highest Ed. > Master's
	(1)	(2)	(3)
Panel A: Female			
Treatment	0.050	0.082	0.004
	(0.036)	(0.159)	(0.033)
Female Control Mean	0.461	1.526	0.303
Panel B: Male			
Treatment	0.044	0.119	
	(0.044)	(0.160)	
Male Control Mean	0.300	0.843	
Nb. observations	1,233	1,233	805

Notes: The table presents the estimated effects of the treatment on marriage outcomes. Panel A presents results for women, comparing the treatment's effects (treated-control differences) against control means, while Panel B presents results for men. Column (1) reports the treatment's effects on receiving at least one marriage offer after graduation (extensive margin). Column (2) reports effects on the number of marriage offers received after graduation (intensive margin). Column (3) reports effects (for women only) on the "quality of marriage offers", measured by an indicator for whether the highest education level among all received offers is a Master's degree or higher. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details), excluding those who are already married or engaged.

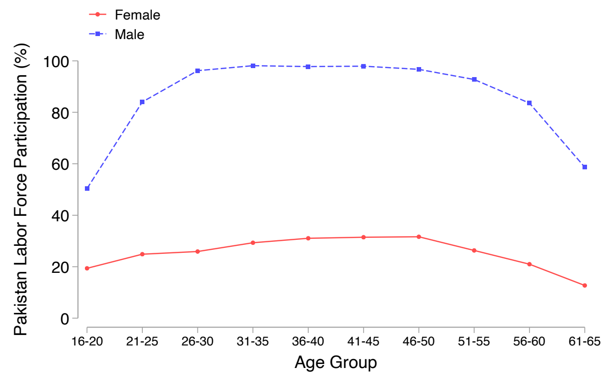
Appendix

A Appendix Figures

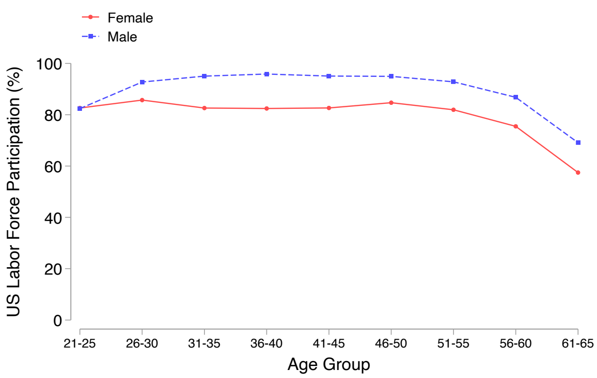
Figure A.1: Labor Force Participation by Age and Gender, US vs. Pakistan (2018)



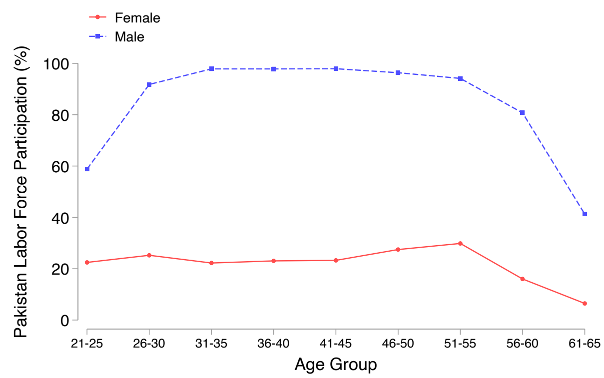
(a) U.S. Overall



(b) Pakistan Overall



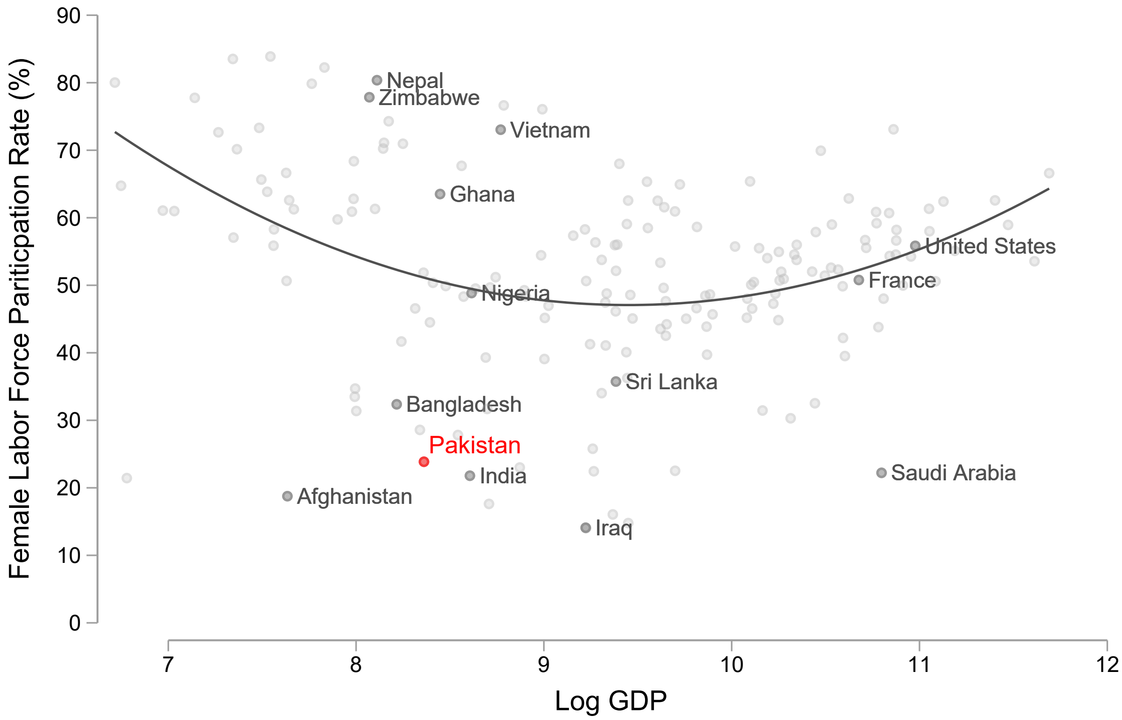
(c) U.S. College Graduates



(d) Pakistan College Graduates

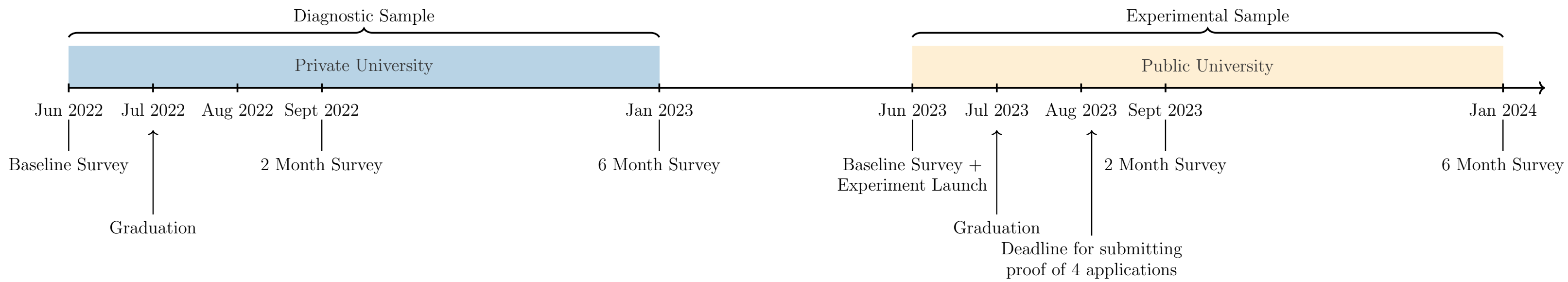
Notes: The figures contrast the relationships between age and labor force participation between the U.S. and Pakistan, separately for men and women. Panels (a) and (b) show the overall age-labor force participation profiles between age 16-65 in both countries. Panels (c) and (d) show the age-labor force participation profiles among college graduates between age 21-65 in both countries. Data is obtained from the Current Population Survey (2018) and the Pakistan Labor Force Surveys (2018).

Figure A.2: Female Labor Force Participation vs. (Log) GDP per capita across Countries (2015)



Notes: Using World Bank (2015) data, the figure shows a scatterplot of the relationship between the female labor force participation (FLFP) rate and log GDP per capita across countries, which exhibits a U shape (shown in a solid black line). Pakistan is highlighted in red, while several countries with similar levels of GDP per capita (e.g., Nepal) or similar FLFP rates (e.g., Saudi Arabia) are identified in dark grey. We also identify a few countries that lie exactly on the U-shape (Nigeria, France, and the United States). Despite the overall U shape, the figure shows considerable heterogeneity: Nepal's FLFP rate is more than 3 times Pakistan's despite a similar GDP per capita, while Saudi Arabia's FLFP rate matches Pakistan's even though its GDP per capita is as high as France's.

Figure A.3: Research Timeline

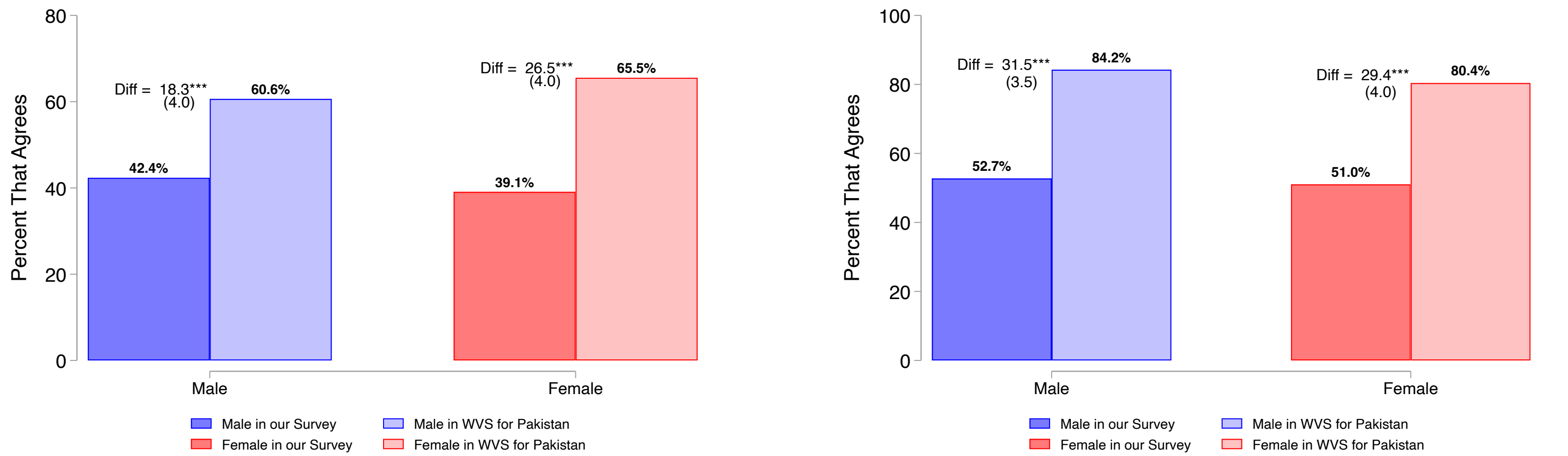


Notes: This figure illustrates the timing of our data collection waves relative to students' graduation timeline. We first surveyed students at the private university (highlighted in blue) starting in June 2022, one month prior to the end of their academic semester, which we refer to as "graduation" (the convocation ceremonies were scheduled later, at different times by different departments). These surveys gave us our descriptive sample described in Section 3.1. We then followed this cohort with additional surveys 2 and 6 months post graduation, in September 2022 and January 2023, respectively. Insights from these follow-ups informed the intervention implemented one year later at the public university (highlighted in yellow and described in Section 5.1). The baseline survey and experiment were fielded in June 2023 at this university, one month prior to end of the academic term in July 2023. The deadline given to the treatment group to show proof of 4 applications was August 15, 2023. The follow-ups were conducted in September 2023 and January 2024, 2 and 6 months post graduation, respectively.

Figure A.4: Survey Incentives

Notes: This photo was taken on June 9, 2022 at the private university in Lahore, Pakistan. It shows one of our food stands being set up during the early days of data collection at the private university. All students who completed the survey were given vouchers to collect their KFC meals and a bakery item from the food stand.

Figure A.5: Social Norms in Our Sample Relative to National Average

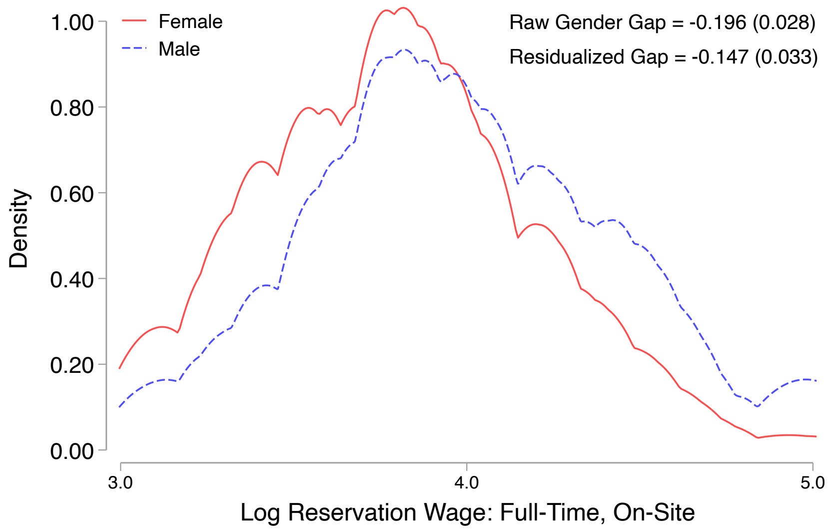


(a) Being a Housewife is Just as Fulfilling as Being a Working Woman

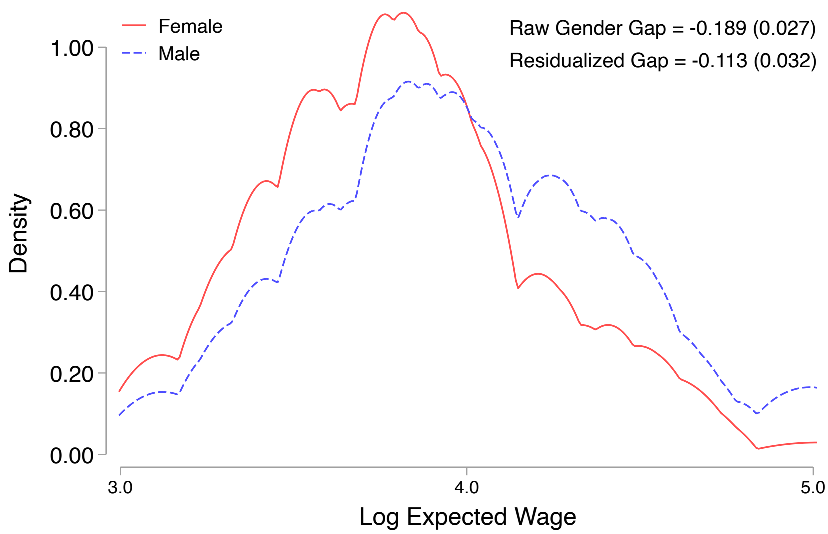
(b) When Jobs are Scarce, Men Have More Right to a Job than Women

Notes: The figure shows baseline average responses to two questions from the World Values Survey (WVS), which we replicated on our descriptive sample, benchmarked against the national average response to each question in the WVS. In both panels, men's responses are colored in blue and women's in red. Darker shades denote students in the descriptive sample at the private university (see Section 3.1 for details), while lighter shades denote the nationally representative sample from the WVS (Wave 7: 2017-21), restricted to students under age 29. Panel (a) shows the percentage of respondents agreeing with the statement: "Being a housewife is just as fulfilling as working for pay." Panel (b) shows the percentage of respondents agreeing with the statement: "When jobs are scarce, men should have more right to a job than women." Within gender in each panel, numbers floating above the bars estimate and test for the differences between the levels in our descriptive sample and the national average. Overall, our descriptive sample expresses more progressive attitudes than the national average. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.6: Supply-Side Factors I: Reservation Wage and Expected Wage



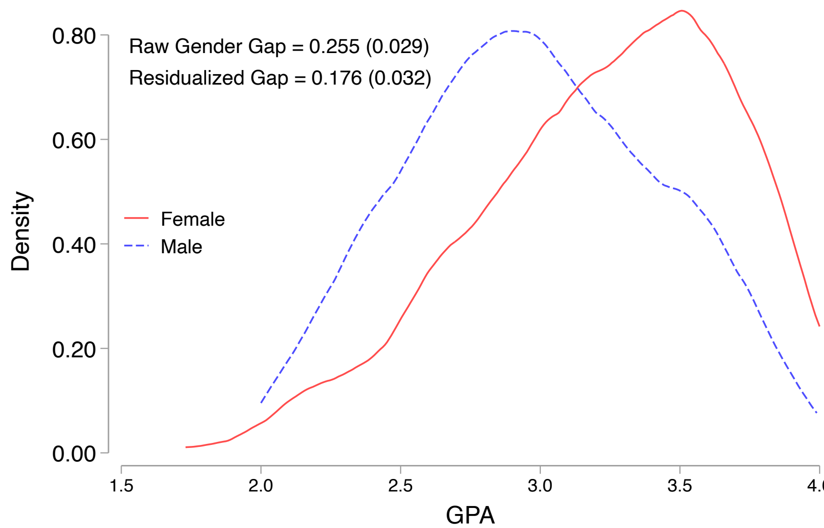
(a) Kernel Density: Log Reservation Wage



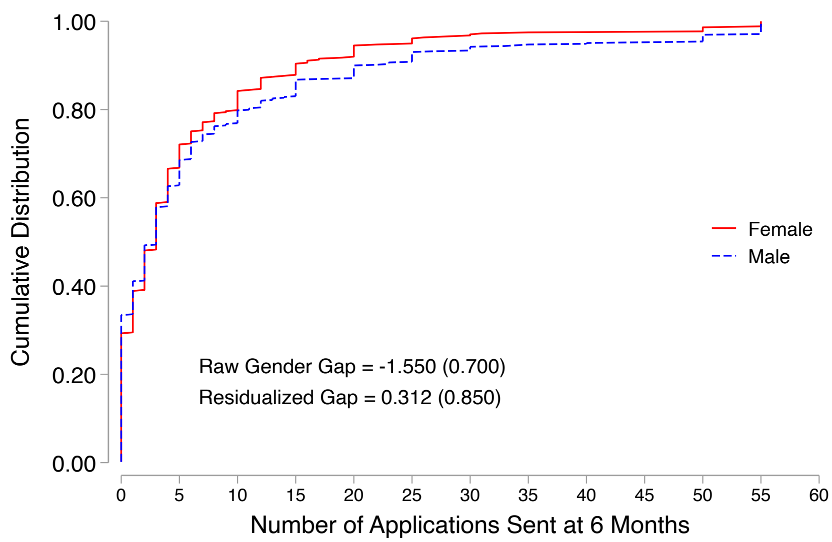
(b) Kernel Density: Log Expected Wage

Notes: The figures show gender differences in labor supply factors that may drive the gender employment gap. Panel (a) shows kernel density of log reservation wage for a full-time on-site job. Panel (b) shows kernel density of log expected wage for a full-time onsite job for respondents' preferred job title. Each also displays estimates and standard errors for raw and residualized (controlling for GPA, major, and industry of search) gender gaps in the mean. For the few women (5.1%) and men (3.4%) who report that they would not work in a full-time onsite job at any wage, we report their reservation wages for a full-time remote job. If the student prefers part-time work, we use twice their reservation wage for a part-time onsite job, if available, or a part-time remote job. Virtually everyone provides at least one reservation wage. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure A.7: Supply-Side Factors II: GPA and Search Effort



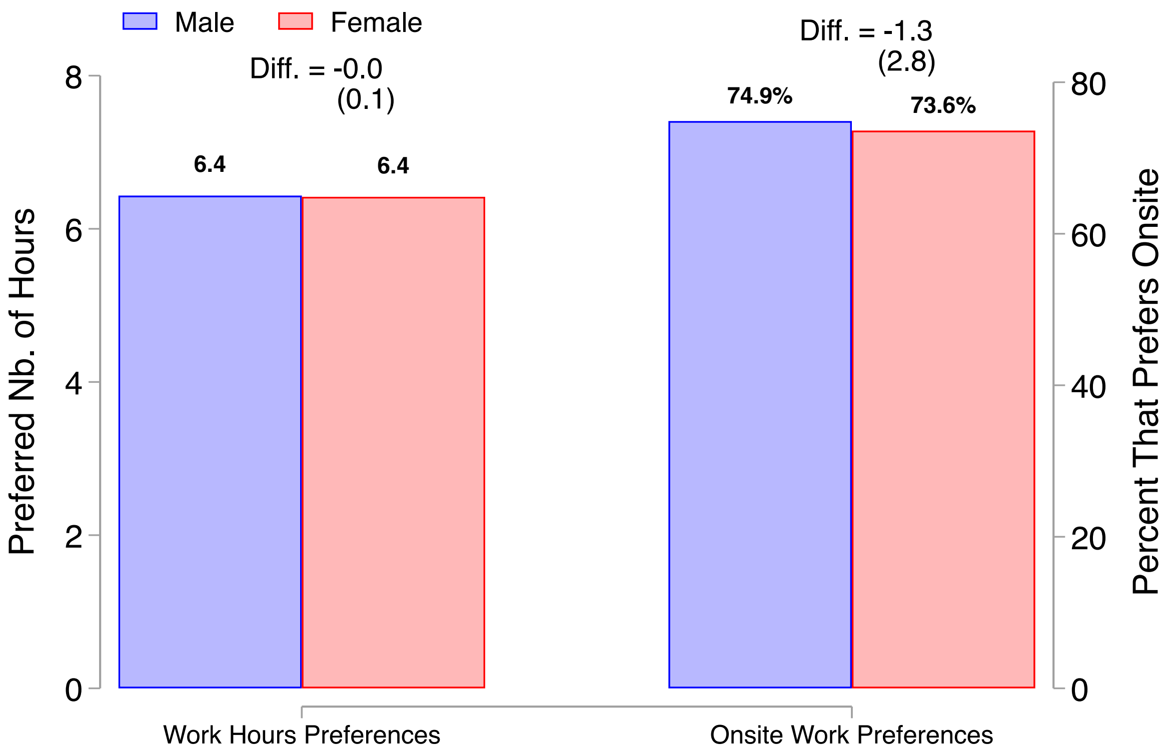
(a) Kernel Density: GPA



(b) CDF: Job Applications at 6 Months

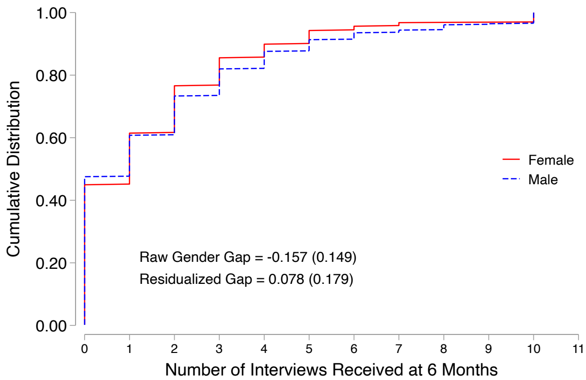
Notes: The figures show gender differences in further labor supply factors that may drive the gender employment gap. Panel (a) shows kernel density of GPA, and estimates and standard errors for raw and residualized (controlling for major) gender gaps in mean GPA. Panel (b) shows the cumulative distribution of the numbers of applications sent at 6 months, and estimates and standard errors for raw and residualized (controlling for GPA, major, and industry of search) gender gaps in the mean number of applications. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure A.8: Supply-Side Factors III: Preferred Work Arrangements

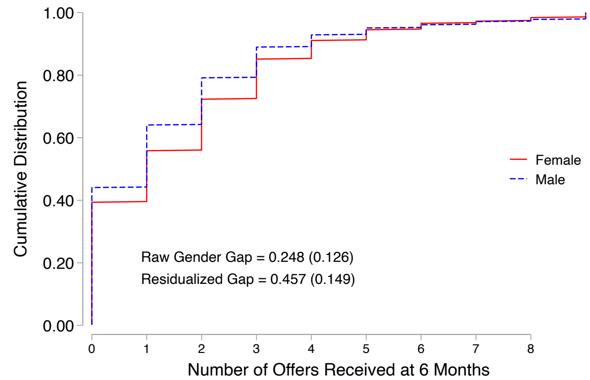


Notes: The figure shows gender differences in baseline preferences over work arrangements that may drive the gender employment gap. The left panel shows the average numbers of work hours preferred by men and women. The right panel shows the percentage of men and women that prefer onsite to remote work. Within each panel, numbers floating above the bars estimate and test for the difference between men and women. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

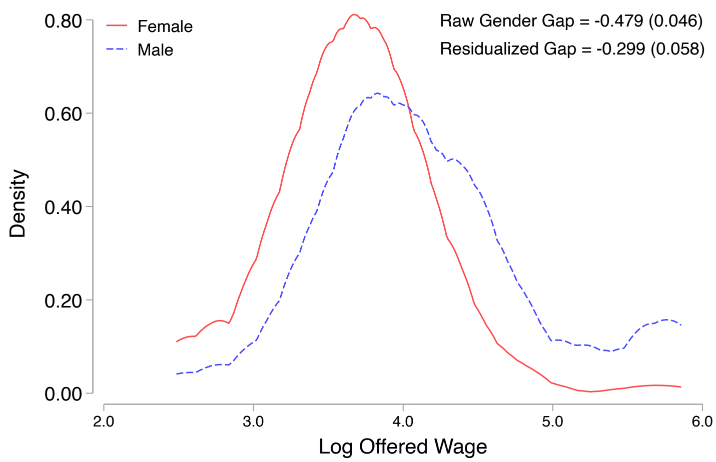
Figure A.9: Demand Factors: Interviews, Offers, and Offered Wages



(a) CDF: Job Interviews at 6 Months



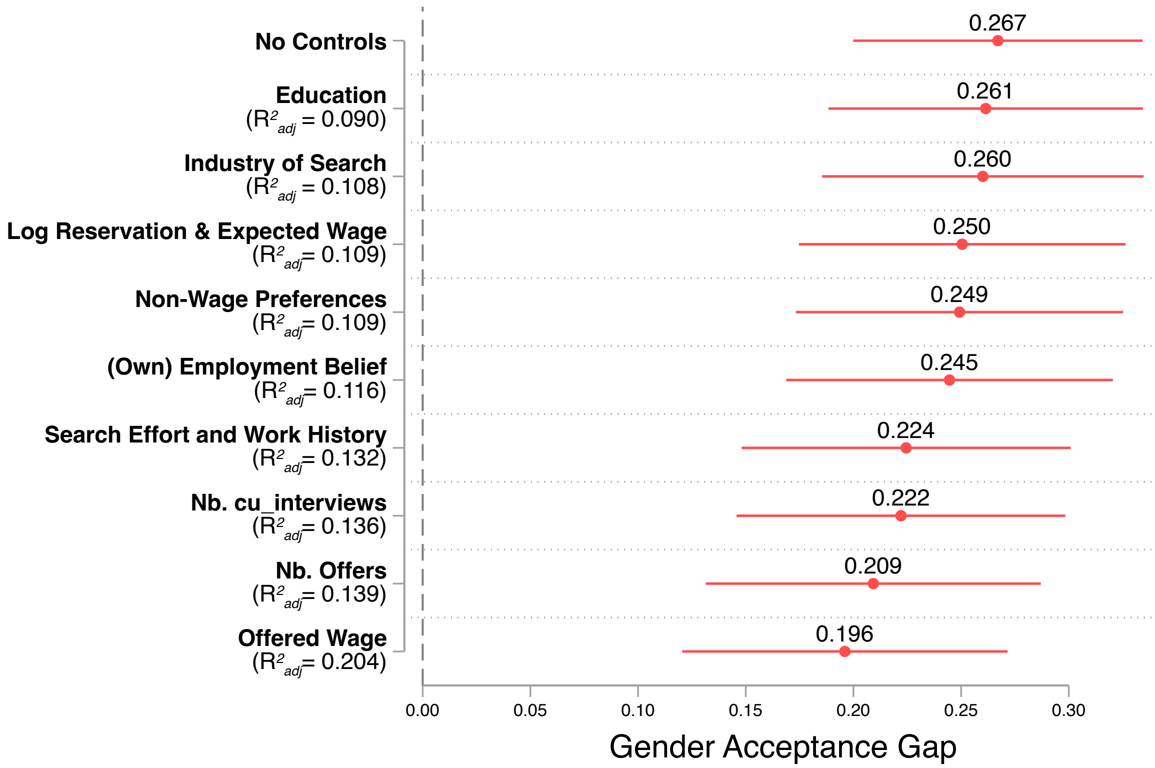
(b) CDF: Job Offers at 6 Months



(c) Kernel Density: Log Offered Wages at 6 Months

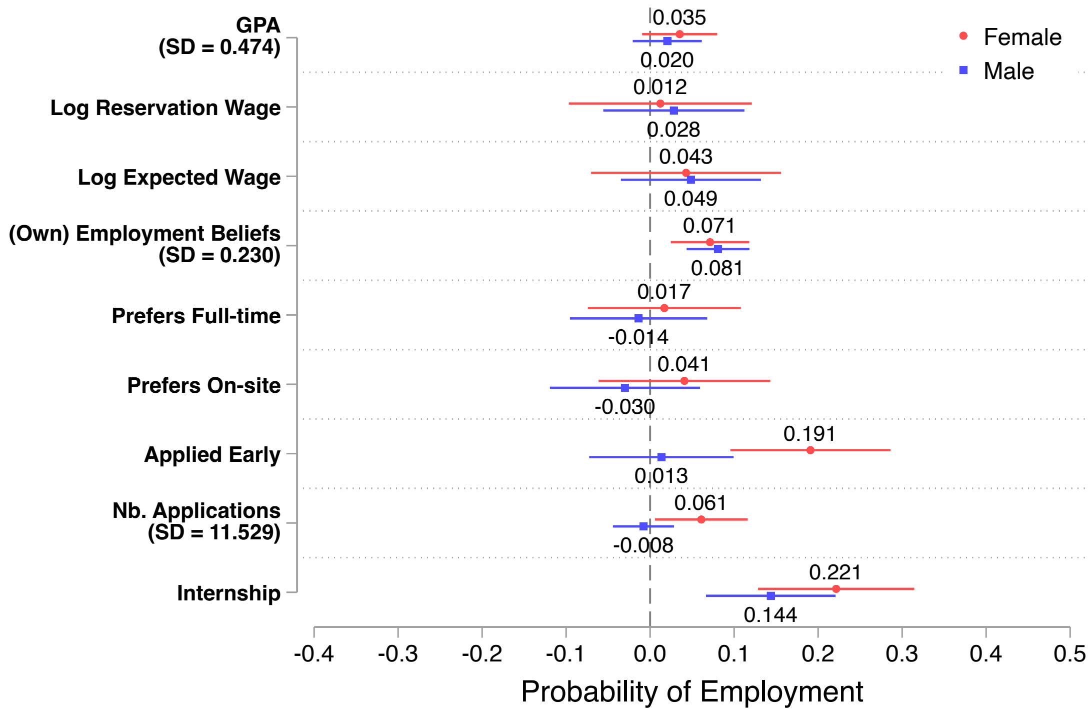
Notes: The figure shows gender differences in labor demand factors that may drive the gender employment gap. Panel (a) shows the cumulative distribution of the numbers of interviews received at 6 months. Panel (b) shows the cumulative distribution of the numbers of offers received at 6 months. Panel (c) shows kernel density of log offered wage at 6 months. Each also displays estimates and standard errors for raw and residualized (controlling for GPA, major, industry of search, intern experience, preference for full- vs. part-time work, preference for onsite vs. remote work, and, for log offered wages only, baseline log expected wage) gender gaps in the mean. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure A.10: Explaining the Gender Acceptance Gap 6 Months after graduation



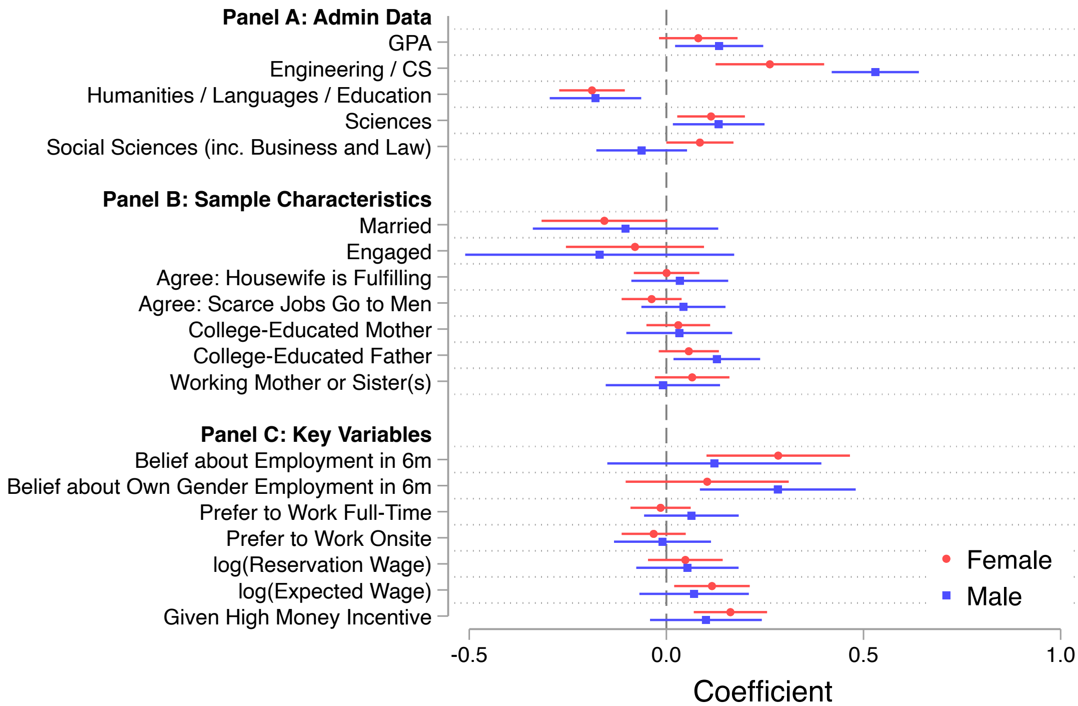
Notes: This figure reports coefficients (re-scaled to be positive) from regressing an indicator for accepting a job offer, conditional on getting at least one six months after graduation, on the female dummy while incrementally adding controls. The goal is to assess to what extent the observed gender acceptance gap can be explained by observable characteristics. Each row shows the coefficient after controlling for variables specified in that row and all rows above. Horizontal bars indicate 95% confidence intervals. The **Education** control includes cumulative GPA and major fixed effects. The **Industry of Search** control includes SOC sub-major group fixed effects based on semantic occupation classification. The **Reservation and Expected Wage** control adjusts for respondents' wage expectations at baseline. The **Non-Wage Preferences** control adjusts for preference for onsite vs remote work and the number of preferred hours per day. The **Employment Belief** control includes baseline belief about one's probability of employment at 6 months post graduation. The **Search Effort and Work History** control includes the total number of job applications sent by the 6-month follow-up, a dummy for applying early, and a dummy for having internship experience. The **Nb. Interviews** control adjusts for the number of interviews by the 6-month follow-up. The **Nb. Offers** control includes the number of job offers by the 6-month follow-up. The **Offered Wage** control is one's current wage if the respondent is working, or the highest rejected wage offer if a respondent does not work. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure A.11: Timing Distinctively Predicts Women's Employment (Bivariate Regression Estimates)



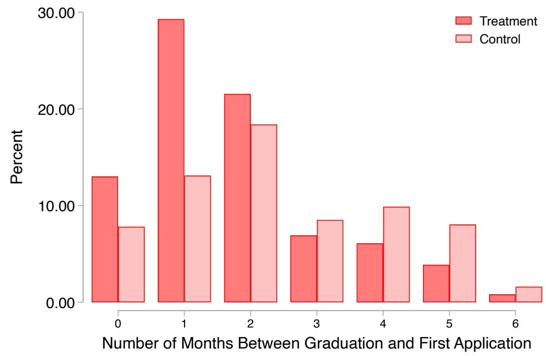
Notes: The figure shows coefficients from multiple individual regressions, each of which regresses an employment indicator on a potential predictor. The regressions are run separately for men and women. Horizontal bars indicate 95% confidence intervals. The predictors include a student's GPA, reservation wage, expected wage, baseline self-reported likelihood of employment six months post-graduation (on a 0-1 scale), preference for full-time vs. part-time work, preference for onsite vs. remote work, whether the student applied to jobs early, the number of job applications submitted (measured in standard deviations), and prior internship experience. Figure 5 shows the coefficients of pooled, rather than individual, regressions using the same set of predictors. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

Figure A.12: Characteristics of the Treated

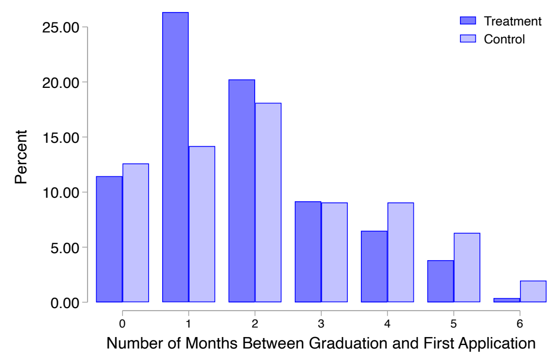


Notes: The figure examines whether there are systematic differences in observable characteristics between students who were offered and took up treatment and those who were offered but did not take up. The coefficients come from individually regressing the take-up indicator on each of the listed characteristics, estimated separately for men and women. Horizontal bars indicate 95% confidence intervals. Panel A contains on administrative academic data, including GPA and fields of study. Panel B contains students' demographic background, including relationship status, agreement with two statements that reflect gender values, parental education, female family members' work status. Panel C contains key determinants of labor force participation, including employment beliefs about oneself and others, work arrangement preferences, reservation and expected wages, and whether the student belongs to the high-money incentive category. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

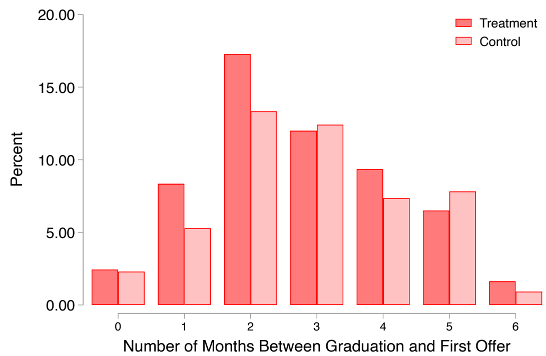
Figure A.13: Timing of First Job Applications and Offers After Graduation



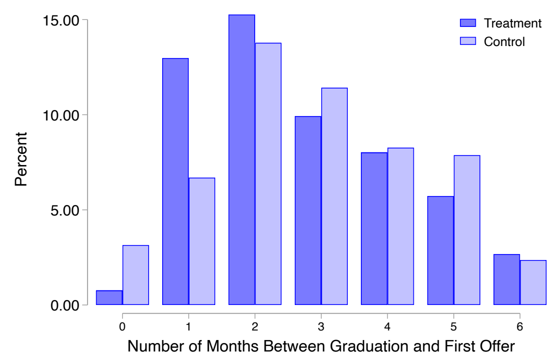
(a) Applications: Female



(b) Applications: Male



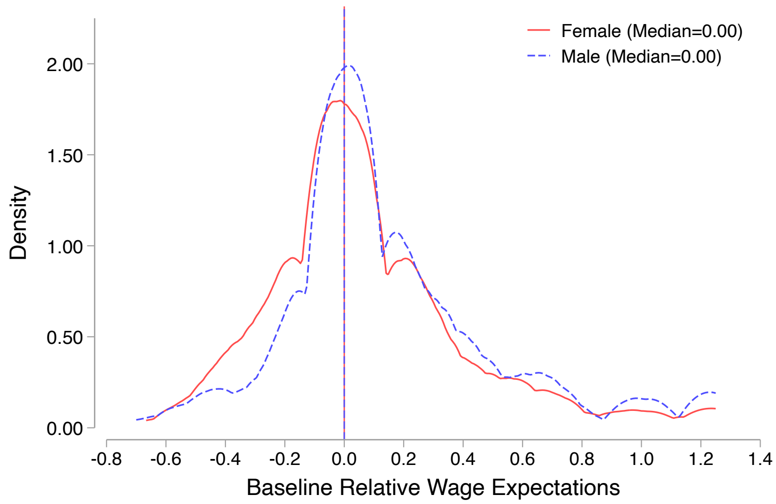
(c) Offers: Female



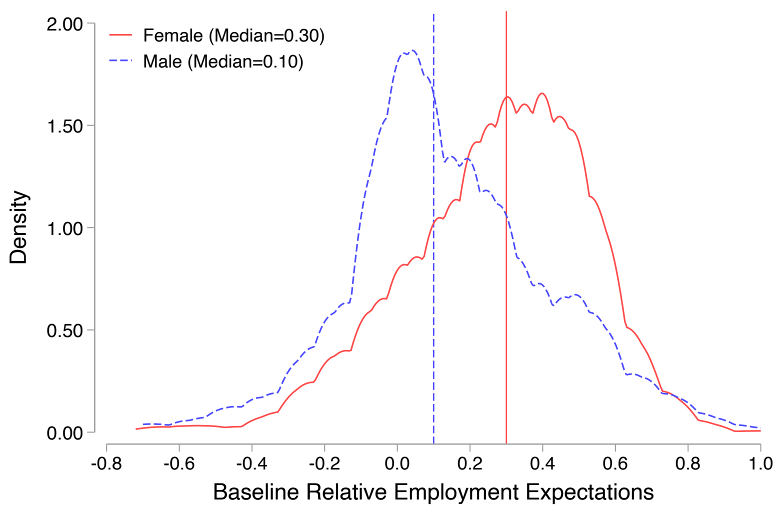
(d) Offers: Male

Notes: The figures show the distributions of the number of months between graduation and when students sent their first application and received their first offer, split by treatment status. Panels (a) and (c) show histograms for women, and panels (b) and (d) show histograms for men. Students without any application or offer are included in the denominator but ommited from these figures. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Figure A.14: Gender Differences in Relative Expectations



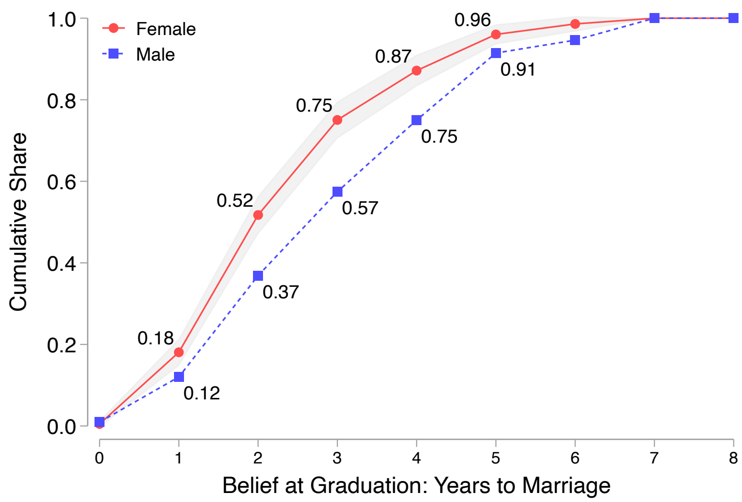
(a) Kernel Density: Baseline Relative Wage Expectations



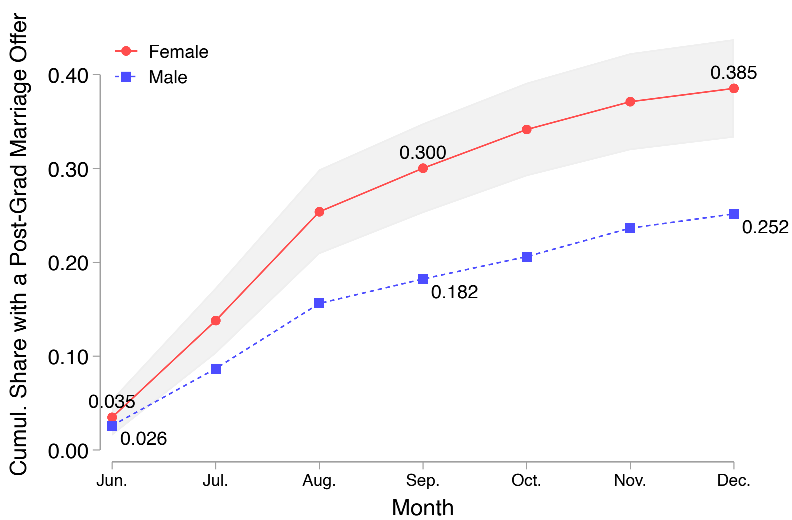
(b) Kernel Density: Baseline Relative Employment Expectations

Notes: The figures show gender differences in relative expectations at baseline. Panel (a) shows kernel density of baseline relative wage expectations. Panel (b) shows kernel density of baseline relative employment expectations. A relative expectation is defined as the difference between what an individual believes they will achieve, either in terms of wages or employment probabilities, and what they believe their similar peers will achieve. For wages, we express it in percentage of similar peers' wage. We define similar peers as students from the same university, year of graduation and same gender. For wages, we also restrict it to the same job title. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Figure A.15: Unfolding of the Marriage Market Post Graduation



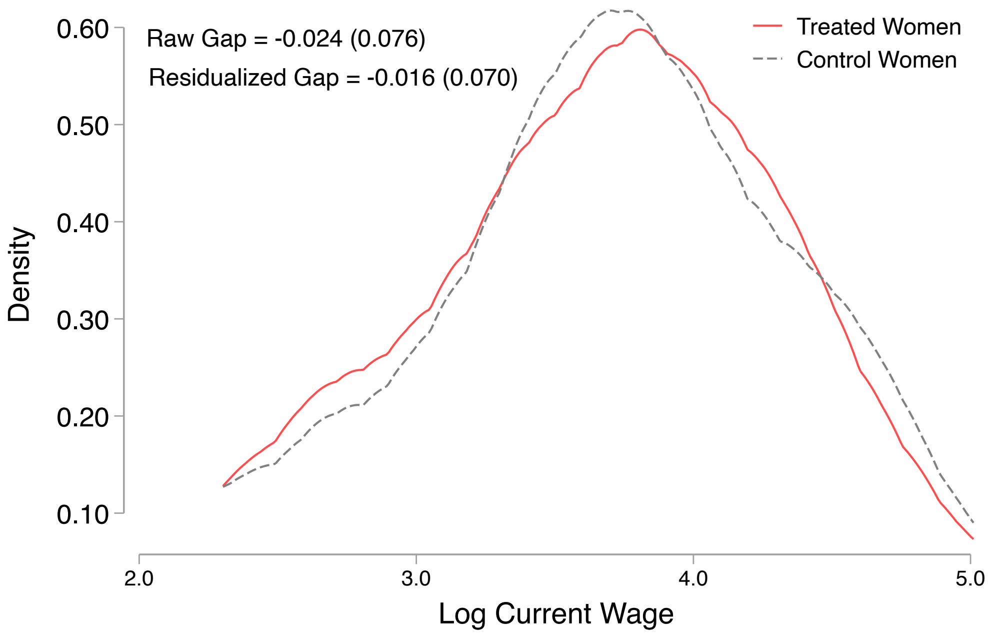
(a) CDF: Baseline Belief on Number of Years to Marriage



(b) CDF: Share with Marriage Offer over Time

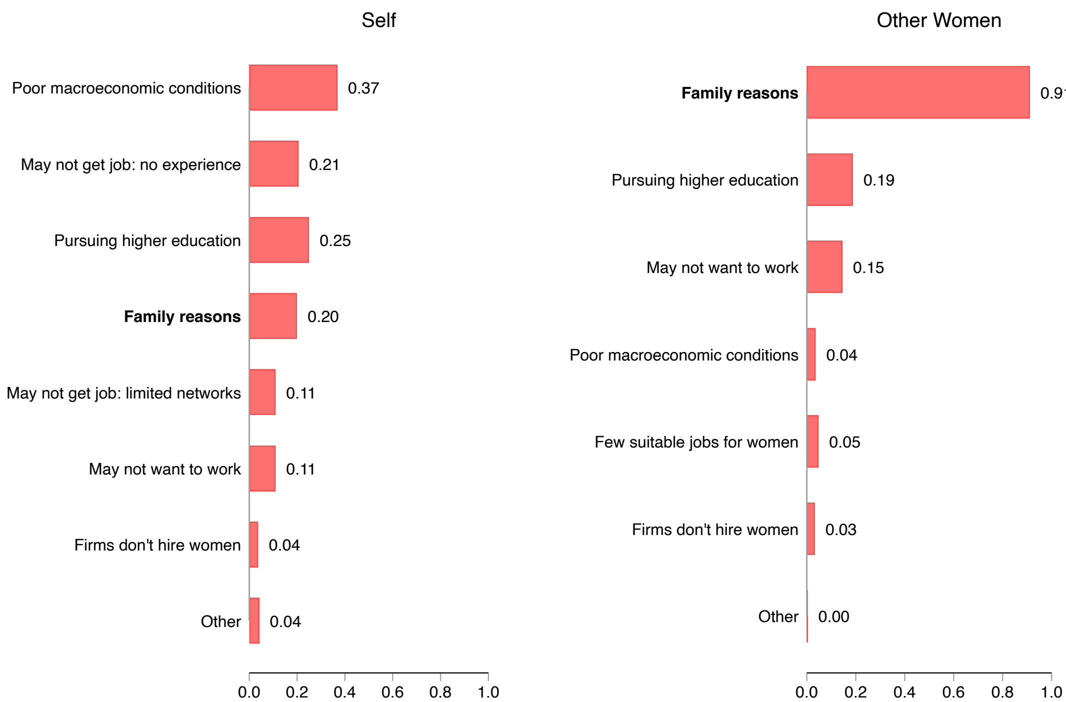
Notes: The figures shows the distribution of respondents' beliefs about their marriage timelines at graduation and their timelines of receiving marriage offers after graduation by gender. Panel (a) shows the cumulative shares of respondents who at graduation believe they will be married by the given number of years. Panel (b) shows the cumulative shares of respondents who have received at least one marriage offer by given number of months after graduation. Grey shades denote areas formed by the 95% confidence intervals among women. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Figure A.16: Wages Distributions for Treated and Control Women



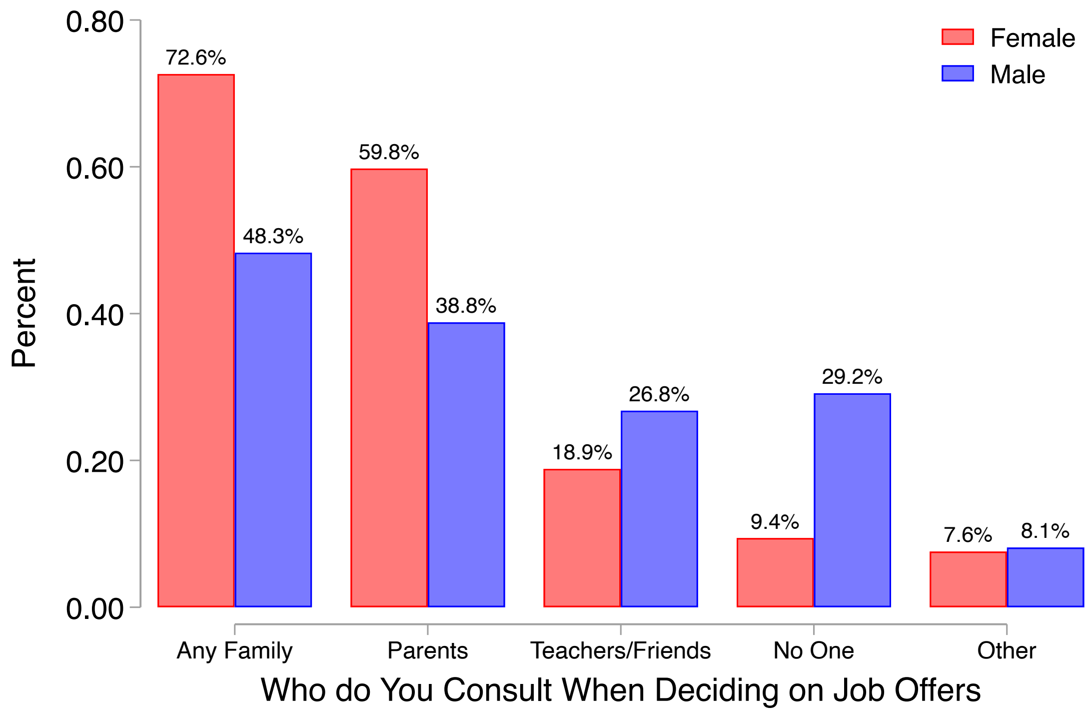
Notes: The figure shows kernel density of log offered wages for both treated and control women. It also provides estimates and standard errors for raw and residualized (controlling for GPA, major, preferences for full- vs. part-time work and for onsite vs. remote work, and baseline log expected wage) gaps in mean log current wages between treatment and control. Industry of search and internship experience controls are not included since there were not fielded on the experimental sample. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

Figure A.17: Potential Reasons for Remaining Unemployed: Self vs. Peers



Notes: The figure shows the most common reasons women provided at baseline for why there is a chance they may not work in the future (titled "Self"), and their reasons for why other women in their class may not work (titled "Other women"). The left panel comes from responses to the question: "Why do you think there is a chance you may not work?" The question was asked to all women who reported less than a 100% likelihood of being employed in 6 months from graduation. The right panel comes from responses to the question: "Out of the remaining XX women, why do you think they are not working?" This question was asked to all respondents who reported that they expect a non-zero share of women in their cohort to not work. After respondents provided open-ended answers, enumerators then coded them into various categories, or entered responses in the "Other" category if none is applicable. Enumerators could select multiple categories for a single answer. The sample consists of all female students in the control group at the public university (see Section 5.1 for details).

Figure A.18: Women Lean More Heavily on Family When Deciding on Job Offers



Notes: The figure shows the percentages of women and men who responded with each category of people when asked “Who do you consult when deciding on job offers". The “Any Family Member” category includes parents, brothers, sisters, cousins, and husband or fiancee, or other unspecified family members. The sample consists of students in the descriptive sample at the private university (see Section 3.1 for details).

B Appendix Tables

Table B.1: Attrition in the Descriptive Sample: Baseline, 2-month and 6-month Follow-Ups

		2m Follow-Up				6m Follow-Up			
	Baseline	Non-Attritors	Attritors	Diff.	P-value	Non-Attritors	Attritors	Diff.	P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nb. Observations	1,494	1,080	414			1,029	465		
Panel A: Administrative Data									
Female	43.98	44.72	42.03	2.69	0.35	42.66	46.88	-4.22	0.13
GPA	3.07	3.06	3.08	-0.02	0.44	3.09	3.03	0.06	0.02
Age	22.47	22.53	22.32	0.21	0.00	22.50	22.41	0.09	0.24
Majors:									
Engineering / Computer Science	25.10	26.11	22.46	3.65	0.06	26.24	22.58	3.66	0.29
Humanities / Languages / Education	13.32	11.30	18.60	-7.30	0.00	15.55	8.39	7.16	0.00
Life Sciences / Pharmacy	14.59	16.39	9.90	6.49	0.00	12.15	20.00	-7.85	0.00
Sciences	13.32	14.26	10.87	3.39	0.10	13.22	13.55	-0.33	0.86
Social Sciences (inc. Business and Law)	33.67	31.94	38.16	-6.22	0.03	32.85	35.48	-2.64	0.25
Panel B: Survey Responses (Sample Characteristics)									
Married	4.48	5.09	2.90	2.19	0.05	4.28	4.95	-0.67	0.67
Engaged	6.89	6.94	6.76	0.18	0.96	6.61	7.53	-0.92	0.59
College-Educated Mother	40.76	39.44	44.20	-4.76	0.10	41.01	40.22	0.80	0.77
College-Educated Father	53.75	52.31	57.49	-5.17	0.07	53.16	55.05	-1.90	0.52
Panel C: Survey Responses (Key Variables)									
Belief About Own Employment in 6m	74.21	74.02	74.71	-0.69	0.67	74.78	72.95	1.82	0.21
Belief About Women's Employment in 6m	50.86	50.44	51.96	-1.51	0.26	50.78	51.04	-0.26	0.90
Prefer to Work Full-Time	41.58	42.21	39.95	2.26	0.51	39.92	45.26	-5.34	0.08
Prefer to Work Onsite	72.44	72.49	72.30	0.18	0.92	74.33	68.26	6.07	0.02
Reservation Wage	54.77	54.43	55.64	-1.21	0.56	55.37	53.43	1.95	0.30
Expected Wage	54.30	53.29	56.93	-3.64	0.03	55.08	52.58	2.50	0.15

Notes: The table compares observable characteristics between attritors and non-attritors in the descriptive sample from the private university (see Section 3.1 for details). Panel A contains administrative data such as GPA, age, and college major. Panel B contains demographic characteristics from survey responses, including relationship status and parental education. Panel C contains survey responses that are key determinants of labor force participation, including employment beliefs about oneself and others, work arrangement preferences, and reservation and expected wages. Column (1) reports these characteristics in the baseline sample. Columns (2) and (3) report these characteristics for respondents who answered our 2-month follow-up survey ("non-attritors") and those who did not ("attritors"), respectively. Column (4) computes the differences between non-attritors and attitors, and column (5) reports the p-values for testing the differences after controlling for gender. Columns (6) through (9) repeat this process for the 6-month follow-up survey. Overall, minimal differences between non-attritors and attritors in both follow-up waves suggest that attrition did not meaningfully impact our surveys.

Table B.2: Experimental Sample Treatment Assignment Balance: Baseline, 2-month and 6-month Follow-Ups

	Baseline				**2m Follow-Up**				**6m Follow-Up**			
	Control (1)	Treatment (2)	Diff. (3)	P-value (4)	Control (5)	Treatment (6)	Diff. (7)	P-value (8)	Control (9)	Treatment (10)	Diff. (11)	P-value (12)
Nb. Observations	939	1,007			811	837			688	754		
Panel A: Administrative Data												
Female	65.07	66.34	-1.27	0.56	64.98	66.31	-1.33	0.57	63.08	65.25	-2.17	0.39
GPA	3.31	3.31	0.00	0.89	3.32	3.32	0.00	0.80	3.32	3.33	-0.01	0.88
Majors:												
Engineering / Computer Science	8.20	6.36	1.84	0.12	7.64	6.45	1.19	0.35	7.56	6.63	0.93	0.48
Humanities / Languages / Education	28.01	26.12	1.89	0.35	26.26	25.45	0.82	0.71	28.49	24.27	4.22	0.07
Life Sciences / Pharmacy	12.78	11.82	0.96	0.49	13.19	11.95	1.25	0.42	13.08	12.20	0.88	0.58
Sciences	25.13	27.41	-2.28	0.25	26.51	28.79	-2.28	0.29	25.15	28.78	-3.63	0.11
Social Sciences (inc. Business and Law)	25.88	28.30	-2.42	0.23	26.39	27.36	-0.97	0.66	25.73	28.12	-2.39	0.31
Panel B: Survey Responses (Sample Characteristics)												
Married	4.47	5.26	-0.79	0.42	4.19	5.14	-0.95	0.37	3.92	4.77	-0.85	0.43
Engaged	5.01	3.87	1.13	0.21	4.56	3.46	1.10	0.25	4.65	3.18	1.47	0.14
College-Educated Mother	28.43	28.60	-0.17	1.00	28.36	29.15	-0.79	0.78	27.62	28.12	-0.50	0.93
College-Educated Father	44.09	42.50	1.59	0.45	44.02	43.13	0.89	0.68	43.17	42.18	0.99	0.64
Panel C: Survey Responses (Key Variables)												
Belief About Own Employment in 6m	79.96	79.57	0.39	0.71	79.56	79.48	0.08	0.98	80.47	80.14	0.33	0.82
Belief About Women's Employment in 6m	51.22	50.71	0.50	0.55	50.91	50.37	0.54	0.55	51.15	50.08	1.07	0.28
Prefer to Work Full-Time	59.74	61.17	-1.43	0.43	58.82	62.01	-3.19	0.14	59.74	61.14	-1.40	0.45
Prefer to Work Onsite	69.18	67.46	1.72	0.46	70.35	68.16	2.18	0.37	69.80	67.81	1.99	0.49
Reservation Wage	53.49	52.52	0.97	0.43	52.94	52.46	0.48	0.76	53.01	52.72	0.29	0.93
Expected Wage	62.06	60.50	1.55	0.25	61.24	60.93	0.31	0.91	62.42	60.85	1.57	0.35

Notes: The table compares observable characteristics between treatment and control groups in each wave in the experimental sample from the public university (see Section 5.1 for details). Panel A contains administrative data such as GPA, age, and college major. Panel B contains demographic characteristics from survey responses, including relationship status and parental education. Panel C contains survey responses that are key determinants of labor force participation, including employment beliefs about oneself and others, work arrangement preferences, and reservation and expected wages. Column (1) reports these characteristics for the control group at baseline. Columns (2) reports these characteristics for the treatment group at baseline. Column (3) computes the differences between baseline control and treatment groups, and column (4) reports the p-values for testing the differences after controlling for gender. Columns (5) through (8) repeat this process for the 2-month follow-up survey. Columns (9) through (12) repeat this process for the 6-month follow-up survey. Overall, almost no p-value is significant, indicating no detectable differences between control and treatment groups in each wave and ensuring the balance of our experiment.

Table B.3: Attrition in the Experimental Sample: Baseline, 2-month and 6-month Follow-Ups

		2m Follow-Up				6m Follow-Up			
	Baseline	Non-Attritors	Attritors	Diff.	P-value	Non-Attritors	Attritors	Diff.	P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nb. Observations	1,946	1,648	298			1,442	504		
Percentage Treated	51.75	50.79	57.05	-6.26	0.05	52.29	50.20	2.09	0.40
Panel A: Administrative data									
Female	65.72	65.66	66.11	-0.45	0.88	64.22	70.04	-5.82	0.02
GPA	3.31	3.32	3.26	0.06	0.02	3.33	3.28	0.05	0.01
Age	22.80	22.79	22.84	-0.05	0.63	22.82	22.73	0.09	0.54
Majors:									
Engineering / Computer Science	7.25	7.04	8.39	-1.35	0.43	7.07	7.74	-0.66	0.62
Humanities / Languages / Education	27.03	25.85	33.56	-7.71	0.01	26.28	29.17	-2.88	0.21
Life Sciences / Pharmacy	12.28	12.56	10.74	1.82	0.35	12.62	11.31	1.31	0.34
Sciences	26.31	27.67	18.79	8.88	0.00	27.05	24.21	2.84	0.23
Social Sciences (inc. Business and Law)	27.13	26.88	28.52	-1.64	0.56	26.98	27.58	-0.60	0.78
Panel B: Survey Responses (Sample Characteristics)									
Married	4.88	4.67	6.04	-1.37	0.36	4.37	6.35	-1.98	0.11
Engaged	4.42	4.00	6.71	-2.71	0.08	3.88	5.95	-2.07	0.10
College-Educated Mother	28.52	28.76	27.18	1.58	0.55	27.88	30.36	-2.48	0.46
College-Educated Father	43.27	43.57	41.61	1.96	0.52	42.65	45.04	-2.39	0.45
Panel C: Survey Responses (Key Variables)									
Belief About Own Employment in 6m	79.76	79.52	81.09	-1.58	0.17	80.30	78.20	2.10	0.08
Belief About Women's Employment in 6m	50.96	50.64	52.73	-2.09	0.10	50.59	52.01	-1.42	0.22
Prefer to Work Full-Time	60.48	60.44	60.74	-0.30	0.89	60.47	60.52	-0.04	0.59
Prefer to Work Onsite	68.29	69.24	63.07	6.17	0.04	68.76	66.94	1.82	0.61
Reservation Wage	52.99	52.70	54.59	-1.89	0.20	52.86	53.38	-0.52	0.40
Expected Wage	61.25	61.08	62.20	-1.11	0.49	61.60	60.28	1.32	0.65

Notes: The table compares observable characteristics for attritors and non-attritors in the experimental sample from the public university (see Section 5.1 for details). Panel A contains administrative data such as GPA, age, and college major. Panel B contains demographic characteristics from survey responses, including relationship status and parental education. Panel C contains survey responses that are key determinants of labor force participation, including employment beliefs about oneself and others, work arrangement preferences, and reservation and expected wages. Column (1) reports these characteristics in the baseline sample. Columns (2) and (3) report these characteristics for respondents who answered our 2-month follow-up survey ("non-attritors") and those who did not ("attritors"), respectively. Column (4) computes the differences between non-attritors and attitors, and column (5) reports the p-values for testing the differences after controlling for gender. Columns (6) through (9) repeat this process but for the 6-month follow-up survey. Overall, minimal differences between non-attritors and attritors in both follow-up waves suggest that attrition did not meaningfully impact our experiment.

Table B.4: Labor Market Effects of Experiment on Women with Heterogeneity by Working Mother or Sister

	LFP	**Working**	**Working FT**	**Working Firm FT**	**Nb. Apps**	**Nb. Offers**	**Accepted Offer**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Working Mother or Sister(s)							
Treatment	0.049	0.003	0.045	0.061	0.160	0.292	-0.029
	(0.069)	(0.073)	(0.068)	(0.067)	(2.015)	(0.395)	(0.087)
Control Mean	0.650	0.475	0.300	0.275	9.092	2.038	0.647
Panel B: No Working Mother or Sister(s)							
Treatment	0.106***	0.106***	0.113***	0.124***	1.851**	0.045	0.119**
	(0.040)	(0.039)	(0.036)	(0.035)	(0.883)	(0.208)	(0.051)
Control Mean	0.549	0.311	0.203	0.185	11.063	2.018	0.442
Nb. observations	786	786	786	786	778	778	506

Notes: The table presents evidence of heterogeneous treatment effects on women's labor market outcomes, between those with working mother or sister(s) and those without. Panel A (193 observations) presents results for women who have working mother or sister(s), comparing the treatment's effects (treated-control differences) against control means, while Panel B (592 observations) presents results for women who do not have working mother or sister(s). Column (1) reports the treatment's effects on labor force participation. Column (2) reports effects on employment. Column (3) reports effects on working full time and for a wage. Column (4) reports effects on working full time for a firm and a wage. Column (5) reports effects on the number of job applications sent. Column (6) reports effects on the number of job offers received. The numbers of observations in Columns (5) and (6) are slightly lower because a few respondents did not report the number of applications or offers. Column (7) reports effects on the likelihood of accepting a job offer among respondents who received at least one job offer after graduation or are currently working. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of female students in the experimental sample at the public university (see Section 5.1 for details). Because not everyone reports if they have working mother or sister(s) at baseline, the sample size is smaller than in Table 3.

Table B.5: Distribution of Workers across Top 3 Occupations by Gender and Treatment

	Treatment (1)	Control (2)	Diff (3)	p-value (4)
Panel A: Women				
Teachers	35.90	34.56	1.34	0.80
Sales and Marketing	10.77	12.50	-1.73	0.63
Software Developers	6.15	8.09	-1.93	0.51
Panel B: Men				
Teachers	14.38	17.78	-3.39	0.44
Sales and Marketing	12.33	12.59	-0.26	0.95
Administration Professionals	10.27	11.11	-0.84	0.82

Notes: This table presents the 3 most common occupations, by treatment status, for women and men who are employed by the 6-month follow-up survey. Panel A presents results for women (330 observations) and Panel B for men (281 observations). Columns (1) and (2) report top occupation shares for treatment and control groups, respectively. Column (3) computes the differences between treatment and control groups, and Column (4) reports the p-values for testing the differences. The sample consists of female students in the experimental sample at the public university (see Section 5.1 for details).

Table B.6: Wage Effects of the Experiment

	Wage > 60th pct	Wage > 75th pct	Wage > 90th pct
	(1)	(2)	(3)
Panel A: Female			
Treatment	0.086***	0.052*	0.011
	(0.032)	(0.028)	(0.019)
Female Control Mean	0.304	0.207	0.080
Panel B: Male			
Treatment	-0.038	-0.043	-0.023
	(0.044)	(0.039)	(0.027)
Male Control Mean	0.420	0.272	0.111
Nb. observations	1,382	1,382	1,382

Notes: The table presents the estimated effects of the treatment on the likelihood of earning above various wage percentiles. Panel A presents results for women, comparing the treatment's effects (treated-control differences) against control means, while Panel B presents results for men. Wage is set to 0 for individuals who are not working, which makes up 58% of our sample. Column (1) reports the treatment's effects on the likelihood of earning above the 60th percentile. Column (2) reports effects on the likelihood of earning above the 75th percentile, while Column (3) reports effects on the likelihood of earning above the 90th percentile. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The sample consists of students in the experimental sample at the public university (see Section 5.1 for details).

C Mediation Analysis

We conduct a mediation analysis to estimate the extent to which treatment effects are mediated by search effort versus timing of search. Let $M_i(t)$ denote the potential value of the mediator of interest for unit i with treatment status T_i. For each individual we observe outcome $Y_i(T_i, M_i(T_i))$. We then obtain the total treatment effect as follows:

$$\tau_i = Y_i(1, M_i(1)) - Y_i(0, M_i(0))$$

We decompose this total effect into two components, one at a time: number of applications sent and month of first application. The causal mediation effects for each mediator M_i are represented by:

$$D_i(t) = Y_i(t, M_i(1)) - Y_i(t, M_i(0))$$

for each treatment status $t \in \{0, 1\}$. All other mechanisms that may drive treatment effects can be represented by:

$$\phi_i(t) = Y_i(1, M_i(t)) - Y_i(0, M_i(t))$$

Together, they sum up to the total treatment effect:

$$\tau_i = D_i(t) + \phi_i(1 - t)$$

Then we consider $\bar{D}$ to be the average causal mediation effect (ACME) for mediator M_i, and $\bar{\phi}$ to be the average direct effect (ADE). If $\bar{D}$ is small and insignificant while $\bar{\phi}$ is large and significant, we conclude that most of the treatment affects the outcome through channels other than the M_i of interest. If $\bar{D}$ is large and significant while $\bar{\phi}$ is small and insignificant, then we conclude that most of the treatment affects the outcome through the mediator.
Mediation analysis relies on two assumptions:

$$Y_i(t, m), M_i(t) \perp T_i \mid X_i = x \tag{2}$$

$$Y_i(t, m) \perp M_i(t) \mid T_i = t, X_i = x \tag{3}$$

The first is satisfied by our randomization process. For the second, which requires that conditional on pre-treatment covariates and treatment, the mediator is orthogonal to the outcome, we control for GPA, baseline belief about employment prospects in six months, preferences for full-time and on-site work, and baseline reservation wage.
ACMEs are estimated using 1,000 quasi-Bayesian Monte Carlo simulations. To this end, we use the R package developed and described by Tingley et al. (2014).

D External Validity of Descriptive Results

In this section we establish the external validity of our descriptive findings from the private university by replicating them at Pakistan's oldest and largest public university. This new setting is ideal for testing the external validity of our descriptive insights as it attracts students from diverse socioeconomic backgrounds and regions nationwide. To ensure comparability of labor market expectations and outcomes between the descriptive and experimental samples, we test whether the descriptive insights from Figures 2-4 hold when reproduced on the control group in our experimental sample.

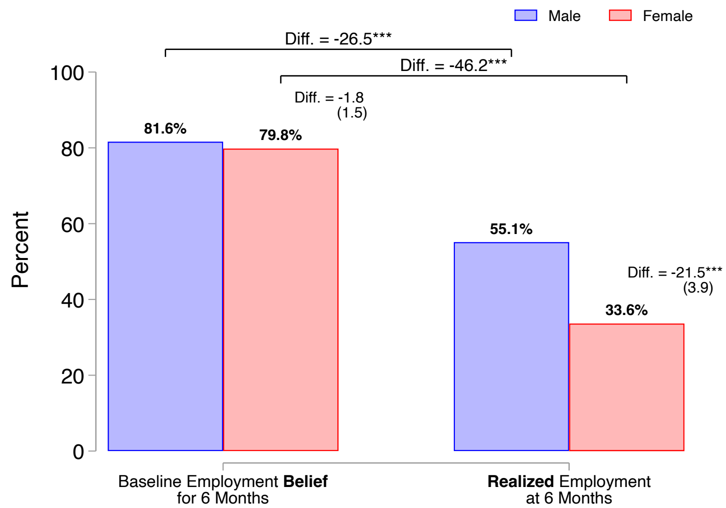
Panel (a) of Figure D.1 compares baseline employment expectations with realized outcomes by gender in the control group of the experimental sample, replicating the analysis from Panel (a) of Figure 2 for the descriptive sample. Consistent with the patterns observed earlier, men and women in the control experimental sample exhibit high and similar expectations for future employment at baseline: 80.4% of women and 82.7% of men expect to be employed within six months of graduation. The modest gender gap in expectations (2.3 ppt) is comparable to the gap in the descriptive sample (5.2 ppt), where employment expectations were similarly high (71.8% for women, 77.0% for men). Additionally, the realized outcomes for women diverge sharply from their initial expectations in both samples. In the experimental sample, only 33.6% of women in the control group are employed six months after graduation—46.8 ppt below their expectations, and 21.6 ppt lower than men's realized employment rate. This overestimation of future employment among women mirrors the descriptive sample, where 36.7% of women are employed six months post-graduation, 35.1 ppt below their average baseline expectations and 28.9 ppt lower than men's realized employment rate.

Panel (b) of Figure D.1 shows the correlation between baseline employment beliefs and realized outcomes in the experimental sample, mirroring the analysis in Panel (b) of Figure 2 done on the descriptive sample. Correlations are stronger in the experimental sample than the descriptive sample (0.48 for women and 0.65 for men in the experimental sample, versus 0.32 for women and 0.31 for men in the descriptive sample). However, they are also less precise especially for men due to a smaller sample size, which results from restricting our analysis to the control group, and the fact that we over-sampled women for this phase of the study. Nonetheless, the take-away is consistent across samples: both genders have inaccurate beliefs across the full distribution, and markedly different intercepts (-4.7 for women and 1.6 for men).

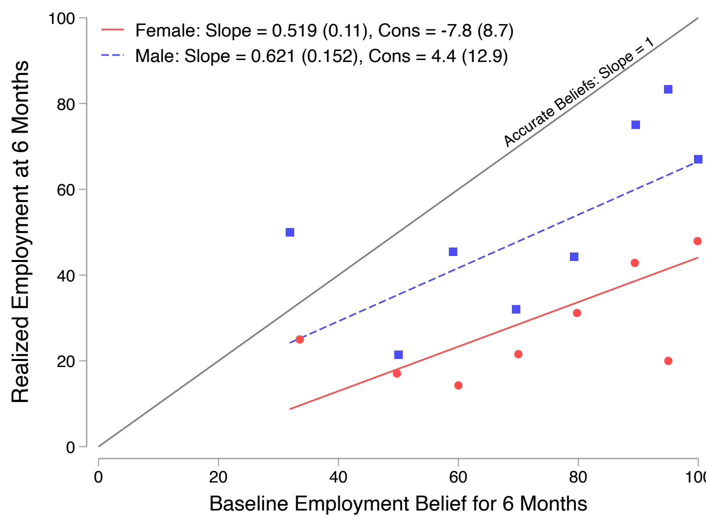
Figure D.2 compares the baseline first and second order employment beliefs of men and women in the experimental sample, replicating insights from Figure 3 produced on the descriptive sample. In the descriptive sample, we saw that even though men and women have inaccurate beliefs about own future employment, they predict their peers' future labor supply more accurately. Specifically, both men and women correctly predict that their male peers' chances of employment are relatively high, estimated at 63.5% by men, and 68.5% by women in the descriptive sample. This matches closely the true employment rate for men of 64.2%. Similar responses were provided by the control group of the experimental sample, where men estimated other men's employment prospects at 68.1% while women estimated them at 73.0%, both of which are comparable to men's actual employment rate of 55.1%. Similarly, both genders correctly assess that women have relatively lower chances of working six months later, estimated in the control experimental sample at 51.0% by men and 51.2% by women. This is closer than women's first order beliefs to the true female employment rate of 33.6%, and remarkably close the second order beliefs reported in the descriptive sample (50.2% by men and 51.6% by women).

Figure D.3 follows Figure 4 in analyzing whether common demand- and supply-side barriers to female employment explain the gender employment gap observed six months after graduation. In the descriptive sample (illustrated in Figure 4), the raw gender gap in employment was 28.9 ppt. In the control group of the experimental sample, the gap is smaller but still substantiallyy large at 21.6 ppt. As in Figure 4, gradually adding controls for education (GPA and major), job search industry, reservation wage, preferences for work hours and remote work, and baseline beliefs about employment prospects reduces the gap modestly by 2.9 ppt. Adding controls for job search effort and work history further narrows the gap by just 0.9 ppt. Finally, accounting for demand-side factors (number of interviews, job offers, and offered wages) reduces the gap by another 1.5 ppt. Taken together, these controls reduce the raw gap from 21.6 ppt to 16.3 ppt. This is similar to the descriptive sample, where the gap shrinks only modestly from 28.9 ppt to 20.1 ppt. The key takeaway remains: even after controlling for student characteristics, job preferences, and demand-side factors (especially the number of job offers), the gender gap persists. This suggests that the gap likely emerges at a later stage, when women decide whether to accept the job offers they receive.

Figure D.1: Baseline Employment Beliefs vs. Realized Employment Outcomes: Experiment Control Sample



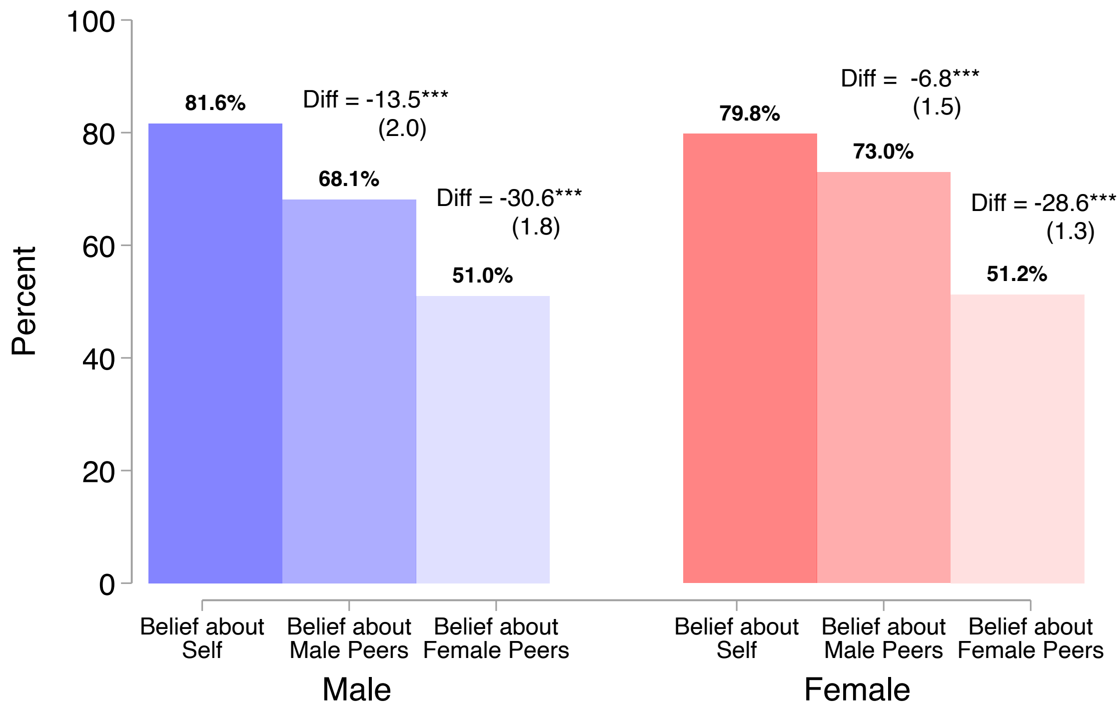
(a) Mean Levels: Intended and Realized Employment



(b) Binscatter: Intended and Realized Employment

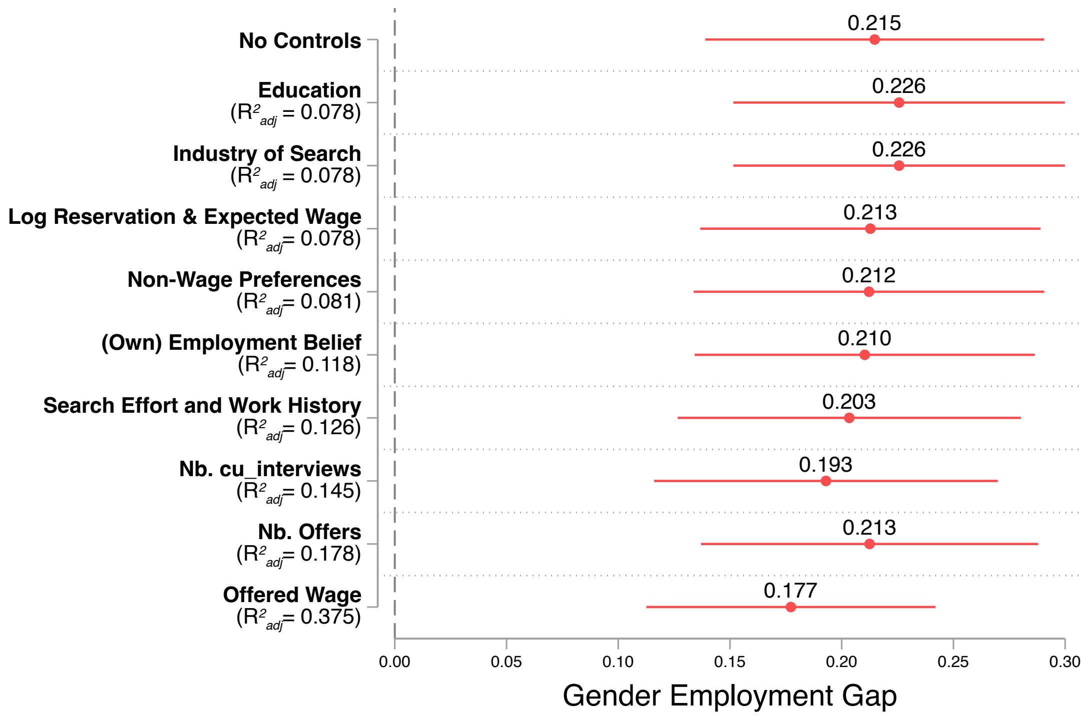
Notes: This figure replicates Figure 2 using data from the control group of the experimental sample (see Section 5.1 for details). Panel (a) plots the percentages of men and women who believe they will be employed 6 months after graduation (left) vs. the percentages who are actually employed by then (right). Panel (b) is a binscatter plot of realized employment vs. baseline beliefs by gender. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure D.2: Employment Beliefs about Self vs. Employment Beliefs about Peers: Experiment Control Sample



Notes: The above figure replicates Figure 3 using data from the control group of the experimental sample (see Section 5.1 for details). The figure shows men's (left) and women's (right) average baseline beliefs about the employment likelihood of different groups. Within gender, numbers floating above the middle bar estimate and test for the difference between "Belief about Male Peers" and "Belief about Self", and numbers floating above the right bar estimate and test for the difference between "Belief about Female Peers" and "Belief about Self". * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure D.3: Explaining Gender Employment Gap at 6 Months: Experiment Control Sample



Notes: The above figure replicates Figure 4 using data from the control group of the experimental sample (see Section 5.1 for details). The figure reports coefficients (re-scaled to be positive) from regressing an employment indicator, measured at six months after graduation, on the female dummy while incrementally adding controls. The goal is to assess to what extent the observed gender employment gap can be explained by observable characteristics. Each row shows the coefficient after controlling for variables specified in that row and all rows above. Horizontal bars indicate 95% confidence intervals. See notes under Figure 4 for detailed descriptions of the controls.