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

Oriana Bandiera Amen Jalal Nina Roussille

April 2025

Abstract

This paper asks how gender equality in education can coexist with large gender gaps in employment. We track the job search of 2,400 students graduating from two major universities in Pakistan — a country where two out of three college-educated women stay out of the labor force. We document that men and women begin with the same work aspirations, apply at the same time and at similar rates, and receive a comparable number of job offers. However, women are more likely to turn down these offers such that, six months post-graduation, the gender employment gap is 27 pp. Differences in the returns to application timing underpin this pattern: the likelihood of accepting an offer declines over time for women, not for men. This finding motivates our experiment, which financially incentivizes students to apply to jobs within a month of graduation. The intervention shifts applications earlier for both genders but raises employment for women only, reducing the gender employment gap by 34%. Treatment works on women who experience "the illusion of time", i.e. those who underestimate how quickly external constraints—particularly from families and the marriage market— arise. By inducing these women to search before constraints bind, our intervention increases their chances of accepting a job offer.

*Oriana Bandiera: London School of Economics, o.bandiera@lse.ac.uk. Amen Jalal: London School of Economics, u.jalal@lse.ac.uk. Nina Roussille, Massachusetts Institute of Technology, nroussil@mit.edu. We are grateful to Camille Landais, Peter Kuhn, Kate Vyborny, as well as to numerous participants in the 2024 sessions of NBER Summer Institute Development Economics, Stanford Institute for Theoretical Economics (Gender), Stanford Joint Applied Micro Seminar (JAMS), Oxford Research Jamboree (Gender at Work) and the Trans Pacific Labor Seminar conference (TPLS). We thank Gabriela Diaz-Pardo, Jamie Emery, Mars Leung, Emiliano Sandri, Sophia Scaglioni, 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.

1 Introduction

Large investments in women's education in low and middle income countries have successfully narrowed – and in many cases eliminated – the gender education gap. Yet, gender gaps in employment remain wide open (Addati et al. 2018). This is puzzling because limited access to education was believed to be a major obstacle to women's labor force participation (Heath and Jayachandran 2018).

Pakistan offers a compelling case study for this puzzle: the share of women with a college degree has nearly doubled over the past two decades, yet only one in three college-educated women is in the labor force (Figure A.1).[1] In contrast with Western economies where women work for several years before dropping out of the labor force, in Pakistan most women, even the ones with a college degree, never join. This paper asks why.

To provide answers, we run surveys and an experiment in partnership with two of the most prominent universities in Pakistan. We ensure representativeness of our results by picking universities with students from a wide range of socio-economic backgrounds and geographic origin: one is a large, mid-tier private university, and the other is the largest and oldest public university in the country.

Our research design has two phases: diagnostic and experimental. For the diagnostic phase, we track the full job search process of $\sim$ 1,000 students graduating in July 2022 from the private university. We interview students about their labor market aspirations a month before they graduate, then track their applications, offers and jobs two, six, and nine months post-graduation. This enables us to identify the critical stage in job search at which gender differences emerge, ultimately diverting women away from the workforce. Based on this insight, we design and test a policy that enables women to stay on the employment path. We test the policy with $\sim$ 1,400 students graduating in July 2023 from the public university. We track the effect of the experiment with a new series of surveys, similarly fielded over four waves, between June 2023 and September 2024.

Employment outcomes in our diagnostic sample are in line with the country-level statistics: six months after graduation, 36.9% of women are working compared to 64.2% of men. We set out to find where gender differences appear in the pipeline from education to employment. Guided by the literature on job search, we test for gender differences

1. As the world's 5th most populous nation, Pakistan is interesting to study in its own right. However, 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 face similar challenges (see Figure A.2 for cross-country evidence on female labor force participation, as well as Jayachandran (2021) and Dinkelman and Ngai (2022) for comprehensive reviews).

on both sides of the labor market, that is job seekers on the supply side and potential employers on the demand side.

On the supply side, we first explore gender gaps in labor market aspirations and find that, a month before graduating, women and men have similar beliefs about their likelihood of working six months post-graduation, averaging 77.0% for men and 71.8% for women. In line with these aspirations, we find that job search behavior in the six months following graduation is similar across genders: women are as likely as men to apply to jobs, and do so at the same time as them, such that controlling for these has no impact on the gender employment gap. Moreover, controlling for supply-side characteristics, such as GPA, major, industry of search, wage and non-wage preferences reduces the gender employment gap by only 1.9 pp (7.0%). That is because women's observables are on most dimensions similar, or better, than men's. For instance, women have on average higher GPAs, lower reservation wages, and similar preferences for work schedules and location, with the majority of both genders preferring full-time, onsite jobs. Turning to the demand side, we do not find evidence that a lack of firm demand for women explains the gender employment gap: women receive as many job interviews as men, and more job offers than them, albeit at lower wages. Taken together, the aforementioned supply- and demand-side factors, account for 11.4% of the employment gap, leaving 24.2 pp unexplained.[2] Given that women and men follow the same path all the way to the job offer stage, the gap must be due to women rejecting offers at a higher rate. Since this happens despite equality in personal traits and job search, the employment returns to these traits and to job search may differ across genders. We find that all returns are similar across genders, with one exception: women's chances of working are correlated with the timing of job applications, while men's are not. In particular, conditioning on the total number of applications sent, women who start applying within two months of graduation are 23.1 pp more likely to have accepted a job offer six months later than women who do not. In contrast, application timing does not correlate with men's employment.

The fact that application timing is strongly correlated with women's employment, and yet many women apply late, potentially adds new policy levers to foster female labor force participation. We design and implement an experiment to evaluate one such lever, specifically an incentive scheme that rewards early applications. The treatment pays a monetary reward to those who apply to jobs within one month of graduation. It has a

2. We replicate all the diagnostic sample evidence in the control group of the experimental sample and find the same descriptive patterns.

significant effect on the timing of applications: 57.1% of women in the treatment group apply to at least one job by August 15th, compared to 31.6% in the control group. We observe similar effects on application timing for men. The experiment allows us to estimate two parameters of interest, namely, the causal effect of the incentive policy on employment, and the causal effect of early applications on employment for the compliers.

The impact of the intervention on women's employment is large. The intent-to-treat estimates indicate that, by our six-month follow-up, 41.1% of women assigned to treatment are working, compared to 33.6% of control women (a 22.3% increase, p = 0.014). Fourteen months later, the gap between treated and control women remains substantial, at 6.9 pp (p=0.046). Since, in low-income countries, working on a self-employed basis indicates lower labor market attachment and earnings potential (Breza, Kaur, and Shamdasani 2021; Bandiera et al. 2023), we also consider the impact of our treatment on employment for a firm and find stronger effects: six months later, only 25.3% of control women work for a firm, compared to 35.5% of treated women (a 40.3% increase, p=0.001). This treatment effect also persists over time, at 9.5 pp (p=0.006) fourteen months later. In contrast, and consistent with our diagnostic evidence, we find no treatment effects on the labor market outcomes of men at either the six- or fourteen-month follow-up. As a result, the incentive policy leads to a 34% decrease in the gender employment gap in both follow-ups. The incentive policy can also be used to recover the local average treatment effect of early applications on employment. Using treatment assignment as an instrument for early applications, we find that applying within a month of graduation has a significant impact on women's employment outcomes, with 2SLS estimates indicating a 38.3 pp increase in female firm employment (vs. 23.6 pp with OLS), at the six-month follow-up. Similar effects are estimated at the fourteen-month follow-up. In contrast, the 2SLS estimates show that there is no impact of early applications on male compliers, confirming that timing matters only for women and more so for those who apply because of the policy.

Through what channel does the treatment operate? Studying the pipeline to employment, we find that the treatment increases the number of applications women have sent by the fourteen-month mark, but it has no downstream effect on the number of job offers that women in the treatment group receive, relative to the control group. Instead, the treatment operates through an increase in the likelihood that women accept one of the (earlier) job offers they have received. In doing so, it effectively makes the job search process closer to equitable between men and women.

It is important to note here that the reward was conditional on applying, not on working. Therefore, the employment increase in the treatment group suggests, by revealed preference, that women are better off working early, rather than cashing the reward and starting work later, or not at all (as they would have done in the absence of the policy). Further, employment is not an absorbing state: women can decide to revert to non-employment at any point in time. Therefore, the fact that employment stays higher in the treatment group more than a year after graduation indicates that women who start working early due to treatment prefer to keep working.

In light of these findings, we examine why women delay applying to jobs in the absence of the incentive - specifically, what gives rise to the illusion that they have more time than they actually do. We begin by noting that the large employment effect, despite only a modest reduction in application delay, points to the existence of a tipping point in job search that women do not anticipate. In other words, women must have inaccurate beliefs about *some* time-varying parameter of the job search process. Hence, we explore the role of factors that can unexpectedly change over time. We find no evidence that labor demand slows down over time. In fact, by endline, women in the control group have received as many job offers, and at the same wages, as women in the treatment group. We also find no evidence of changes in women's preferences for work: at the six-month follow-up, more than 80% of women state they want to work in the coming year, with no treatment effect on this answer. Moreover, women's reservation wages decline over time, suggesting that falling employment reflects binding constraints rather than declining interest in work. However, we find that the timing of women's primary outside option —marriage—closely mirrors labor market patterns. We provide three pieces of evidence on this front. The first builds on Adams and Andrew (2024) who find that parents in similar settings believe that a college degree improves marriage prospects, but that the quality of marriage offers still declines rapidly after graduation. In line with this, we find that marriage offers for women arrive quickly post-graduation, while job acceptance rates drop most sharply during the same period. Specifically, applying in September instead of June reduces employment by 30.4 pp, coinciding with a 29.8 pp rise in marriage offers. While this is purely correlational, the timing of the two variables is remarkably synchronized to be dismissed as entirely coincidental. For the marriage market demands to explain our results, it must be that women lack foresight about them. Supporting this, we find that, at baseline, many women expect to marry after age 25, but significantly revise their expectations downward by the six-month survey. In contrast, women who expected early marriage show no such updating. Further, treatment shifts employment precisely for the women who do not anticipate the speed of the marriage market. In particular, treatment effects on firm employment are larger for the group who initially expects a late marriage —17.1 pp-, versus 5.4 pp for those who expect an earlier marriage (p-value of the difference = 0.032). The fact that many women are caught off guard by the timing of marriage offers—and that treatment effects are concentrated among those who least expected them—suggests that external forces, rather than shifts in women's own preferences, play an unexpected role in shaping labor supply. In Pakistan, families not only arrange marriages but also exert significant control over women's work decisions: 79% of women report in a representative survey that the head of the household determines whether they are allowed to work. Motivated by this, we examine treatment effect heterogeneity by parental involvement in job decisions. We find that treatment effects on firm employment are nearly three times larger among women who consult their parents (16.3 pp) than among those who do not (6.5 pp), with a p-value for the difference of 0.069. This pattern suggests that the timing effects we document are driven by time-varying preferences of parents, particularly in the wake of marriage offers. For this explanation to hold, such preferences must also be unanticipated. Consistent with this, we find that most women do not anticipate family opposition at baseline: only 20% cite it as a potential barrier to their own employment, even as 91% recognize it as a barrier for others. This disconnect—between recognizing restrictive norms in general but believing they do not apply to oneself—mirrors the broader optimism discussed earlier. Together these findings suggest that some women operate under the illusion of time: they underestimate how quickly external constraints—particularly from families and the marriage market—can intensify. By inducing these women to search before constraints bind, our intervention increases their chances of accepting a job offer.

This paper 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 reveal an optimistic bias among job seekers regarding their job-finding rate (Spinnewijn 2015; Mueller, Spinnewijn, and Topa 2021; Abebe et al. 2024). In line with those papers, we show that biased beliefs affect job search in ways that cannot be undone, so that small differences between applicants at baseline place some women on different equilibrium paths leading to very different outcomes. Our paper complements 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 returning to work after childbirth. Similarly, we find that college-graduating women in Pakistan overestimate their near-future labor supply. 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. Our paper complements this study by describing the role of application timing in a context with low female labor force participation, where college-graduating women's labor supply decisions are shaped as much by extensive-margin decisions (whether to work) as by intensive-margin ones (which job to choose).

We also contribute to a large literature documenting barriers to female labor force participation (Goldin and Rouse 2000; Bertrand, Goldin, and Katz 2010; Bursztyn, González, and Yanagizawa-Drott 2020; Jayachandran 2021; Card, Colella, and Lalive 2021; Kleven, Landais, and Leite-Mariante 2023; Bursztyn et al. 2023; Kuhn and Shen 2023; Gentile et al. 2023; Agte and Bernhardt 2024). Much of this literature focuses on barriers created by gender roles in the household or by firm discrimination. But, in Pakistan, most women never enter the labor force in the first place, thus our focus is on the barriers that emerge immediately after graduation. Further, at least in our context, it seems like firm demand is not the primary barrier: women receive as many job offers as men. Our key finding, namely that women have a short window of time to apply and find jobs, might guide research and policy in the many settings where female labor force participation remains persistently low despite sustained policy efforts.

2 Labor Market: Beliefs vs. Outcomes

2.1 Our Diagnostic Sample

We study the labor supply decision of women at the time of their college graduation. This choice is motivated by the life cycle of women's labor supply in South Asia, which differs meaningfully from that in higher-income countries. Figure A.3 plots the labor force participation of men and women by age in 2018, for the United States in Panel (a) and for Pakistan in Panel (b). A standard pattern in higher-income countries such as the United States is that women's labor market entry rates are similar to men's but female labor force participation falls during the childbearing years. In contrast, the life cycle of women's labor supply in Pakistan is essentially flat: rather than dropping out in childbearing ages, women in Pakistan rarely enter the labor market in the first place. Similar patterns emerge when restricting the sample to college graduates (see Panel (c) and (d) of Figure A.3).[3] Since women do not enter the labor market from the onset of their adult life in Pakistan, understanding their labor supply decision right at the point of college graduation is essential.

Recruitment We invite all 2,872 graduating students (1,146 female and 1,726 male) at the private university to participate one month prior to their graduation. Of these, 2,238 participate 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 had already registered for a graduate program. This leaves us with 1,493 students in our baseline sample.

Attrition Of the 1,493 students in our baseline sample, 1,029 respond to our six-month survey, and 910 to our nine-month survey. The six-month and nine-month response rates are therefore 68.9% and 61.0%. These response rates are considerably higher than typical rates reported in the phone survey literature.[5] Table B.1 shows that the baseline, six-month, and nine-month samples have similar observables.[6]

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

3. Figure A.3 does not account for cohort effects. It is therefore all the more striking that the labor force participation of young Pakistani women in 2018 is the same as that of older cohorts.

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.5 for a picture of our food stand.

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

6. Among the many variables we test, a few observables display differences across waves. For instance, the average GPA of non-attritors at the six-month follow-up is 3.09, while that of attritors is 3.04. This difference is statistically significant but economically small. We also note that there are no systematic patterns in attrition across waves: most differences that appear in the six-month survey are eliminated at the nine-month survey, and vice versa. For readability, we exclude the two-month follow-up from the attrition table, as we ended up only using this survey for the definition of one variable in our analysis. The response rate at the two-month follow-up is higher than at the later follow-ups and attrition is no different either.

Descriptive Statistics The baseline characteristics of our diagnostic sample are shown in Table 1. The last column of the table reports the p-value from a test of equal means between genders. Women make up 42.7% of the sample, with 439 female and 590 male respondents. Men and women are both about 22.5 years old, on average. Women's GPAs are 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%). Only 6.8% of women and 2.4% of men are married at the time they graduate from college, with a similar fraction engaged (7.5% of women and 5.9% of men). Finally, men and women come from similar parental education 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.

2.2 Baseline Beliefs about Labor Market Outcomes

We begin our investigation of the education-to-employment pipeline by examining whether women, as they approach college graduation, expect that they will work in the near future. Indeed, it could be that college degrees serve purposes entirely outside of labor markets – notably, they could enhance marriage prospects (Adams and Andrew 2024). This raises the possibility that women want to pursue higher education primarily to improve marital outcomes rather than labor market positioning, potentially explaining the low labor force participation among college-educated women.

Belief: 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 their preferred job title across four work schedules (Full-time Onsite, Part-time Onsite, Full-time Remote, Part-time Remote). Importantly, for each schedule, there the option to declare that the respondent does not intend to work for any wage in this schedule.[7] The second question is probabilistic: "On a scale from 0 (very unlikely) to

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 [preferred job title]. 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

100 (very likely), how likely is it that you will be working within six months of graduating? Work includes working for a private firm or government, running your own business or your family business."

Beliefs: Results Across a wide range of measures, we find that, at the time they graduate, the majority of women expect to work. Virtually all women and men provide us with a reservation wage for at least one work schedule. Additionally, 95.0% of women provide a reservation wage for the full-time, onsite schedule. Importantly, as illustrated in Figure A.6, the distribution of women's reservation and expected wages lies to the left of men's, even after controlling for GPA, major, and industry of search. This shows that women are willing to work, even for wages lower than men's. Turning to the probabilistic question, the first two bars of Figure 1 Panel (a) show that women report a 71.8% likelihood of working within six months of graduation, only 5.2 pp lower than men. Beyond expecting to work, most women express a preference for working over being a housewife: 59.0% of women in our sample report that being a housewife is not as fulfilling as being a working woman. Finally, one may worry that stated beliefs at baseline may just be cheap talk (or the result of demand effects). Two pieces of evidence assuage this concern. First, 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., women may be more inclined to abide by traditional gender norms when answering a man than when surveyed by a woman). Second, women's post-graduation actions support their baseline expectations: Figure A.7 Panel (b) shows that 80.4% of women apply for at least one job within six months of graduating, a rate similar to that of men, at 78.5%.[8]

2.3 Beliefs Meet Reality

While men and women have similar beliefs about their employment likelihood, the third and fourth bars in Figure 1 Panel (a) reveal a large gender employment gap: 64.2% of men, but only 36.9% of women, were employed six months after graduation. Validating

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 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. We collect further evidence in the experimental sample by administering additional measures of baseline work intentions. Across all these measures, we find that nearly all women, irrespective of their family's wealth, want to work for at least two years after graduation. For more details, see Section C.

the representativeness of our sample, the share of working women closely aligns with the national average labor force participation rate for young college-educated women, which is 33.9% in 2018 as shown in Figure A.1 Panel (b). Comparing realized outcomes to the baseline beliefs represented in the first two bars of Figure 1 Panel (a), we find that on average men overestimate their likelihood of working by 12.8 pp (16.6%), while women overestimate it by 34.9 pp (48.6%).[9]

Figure 1 Panel (b) shows the relationship between baseline employment beliefs and realized employment six months post-graduation. The benchmark of accurate beliefs would manifest as observations on the 45-degree line.[10] The figure illustrates that, relative to this benchmark, both men and women have inaccurate beliefs about their future employment. Specifically, for both genders, the slope of the relationship between their baseline beliefs and realized outcomes is similar and closer to zero than to one (0.35 for men and 0.31 for women). However, the intercept is drastically different across genders: 14.7 pp for women, and 37.2 pp for men. Taken together, the slopes and intercepts imply that women overestimate their chances of working along most of the distribution of baseline beliefs, and do so significantly more than men. For example, among those who report an 80.0% chance of working six months post-graduation, only 32.6% of women, vs. 61.7% of men, end up working. In sum, despite the striking gender similarities in expected employment documented at the onset of job search, we uncover large gender gaps in employment rates six months post-graduation.

We also conduct a shorter survey nine months post-graduation to track students' employment outcomes. Figure A.8 shows that while women's employment levels rise substantially by 10.8 pp (from 26.1% to 36.9%) between the two- and six-month follow-up surveys, there is almost no change in employment levels for women between six-month and nine-month follow-ups. The nine-month employment rate is 38.8% for women, only 1.9 pp higher than the six-month employment rate. Additionally, the employment rates at both six and nine months closely align with the national average of 33.9% for college-educated women in Pakistan. Meanwhile, for men, the realized employment reaches 73.8% at the nine-month mark, which is precisely the baseline belief men held about their

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

10. Deviations from this benchmark take two forms. Observations above the 45-degree line correspond to students underestimating their chances of working. Conversely, observations below the 45-degree line imply students overestimate their chances.

employment at the six-month mark.

3 Diagnosing the Gender Gaps

Sections 2.2 and 2.3 reveal significant gender disparities in labor market outcomes six months post-graduation. We now examine several commonly suggested explanations to account for these disparities. We first assess gender differences in relevant endowments (e.g., GPA, major, or search effort) and analyze whether 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 (e.g., whether labor market returns to a higher GPA differ by gender), 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 and/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 expectations. However, as discussed in Section 2.2, women's reservation wages are lower than men's, conditional or unconditional on covariates. Turning to non-wage amenities, Figure A.9 shows no gender differences in preferred work hours (averaging 6.4 hours per day for both men and women), or preferences for remote work (about three-fourths of both genders prefer on-site jobs). The similarities in preferences for non-wage amenities show that, absent marriage and children, there is no intrinsic gender gap. 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.[11] 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. Figure A.7 Panel (b) presents the cumulative distributions of the number of job applications by gender. It shows that, at the extensive margin, women are as likely as men to apply to any job.

11. The distribution of men and women across majors differ, as described in Table 1. However, we later show in Figure 2 that gender differences in major choices do not drive the gender employment gap.

Additionally, the median number of applications, 4, is the same for men and women. While, on average, women apply to 2.8 fewer jobs than men, when adjusted for other observables, such as GPA, major and industry of search, the residualized gender gap in applications is null. Regarding application timing, we define "applying early" as sending at least one job application within two months of graduation and find that women are as likely as men to apply early. Overall, college-graduating men and women are more alike in job preferences and search behavior than we anticipated based on the existing literature.

To formally examine how student characteristics affect the gender employment gap, we regress a six-month employment indicator on a female indicator, adding controls progressively in Figure 2. The initial gap of 27.3 pp decreases by between 1.1 and 2.8 pp as we sequentially add controls for GPA and major (Row 2), industry of search (Row 3), reservation and expected wages (Row 4), preferences regarding work hours and remote work (Row 5), as well as baseline beliefs about one's chances of working six months post-graduation (Row 6). In the final model of supply-side factors (Row 7), we add controls for search effort and past internship experience, which only moderately decrease the gap to 22.8 pp. These results suggest that supply-side factors do not play a substantial role in the gender employment gap.

Level Differences II: Demand-Side Factors The literature highlights a number of demand-side factors that may alternatively explain gender gaps in employment, including statistical, taste-based, and paternalistic forms of 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, the number of job offers, and the salary offers received by both men and women. Figure A.10 shows the cumulative distributions of the number of job interviews in Panel (a), and job offers in Panel (b), presented separately for men and women. Strikingly, a similar share of men and women receive at least one interview and one job offer. We also find that there is no significant gender gap in the number of interviews attended, or in the number of job offers they receive. When we further control for student baseline observables (e.g., GPA, major, industry of search, as well as wage and non-wage preferences), women get 0.7 more job offers than men. Finally, consistent with Brown (2022), we find in Figure A.10 Panel (c) that firms offer women lower wages, even after controlling for cumulative GPA, major, and industry of search.[12]

To formally examine whether demand-side factors explain lower female employment rates, we extend the existing model in Figure 2 by incrementally adding controls for students' number of interviews (Row 8) and number of job offers received (Row 9). The gender disparities in employment remain largely unaffected by these additional controls.[13] Finally, in Row 10, we control for offered wage. After controlling for all these supply- and demand-side variables, the unexplained gender gap in employment remains very large, at 24.2 pp. Even after holding constant job preferences, search behaviors, and the number of offers across men and women, women remain significantly less likely to be employed, indicating that the gender gap emerges at an even later stage: job acceptance. Indeed, we find in Figure A.11 that upon receiving a job offer, women are 27.7 pp less likely than men to be employed six months after graduation. Even after controlling for the supply-side factors detailed above, the number of interviews and offers received, and the wage offered by the firm, women remain 18.4 pp less likely than men to have accepted an offer six months post-graduation.

Differences in Returns Controlling for student characteristics does not shrink gender gaps in a pooled regression, but a given individual characteristic may uniquely predict men or women's future employment. To test for this, we regress employment outcomes at six months post-graduation separately for men and women on the same supply-side variables used as controls in Figure 2. Figure 3 presents the multivariate results, while Figure A.12 shows the bivariate models.[14] Only one variable, early application, is differentially predictive of men's and women's employment. None of the other variables, including baseline beliefs about employment and number of job applications, exhibit differential predictive power by gender. For example, a one-standard deviation increase in baseline belief about employment probability is associated with a 7.5 pp increase in men's likelihood of working six months later, compared to a 4.3 pp increase for women (p-value = 0.305). Similarly, after controlling for all supply-side variables, the number of

12. We define the offered wage as the highest wage offer received by a student for a job (regardless of whether they have accepted it). In the regression on Row 10 of Figure 2 we impute a zero offered wage for all individuals who are not working and have not received a job offer in the past.

13. 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

14. Corresponding means and standard deviations for the independent variables shown in the figure are provided in Table B.10.

applications also has no differential impact across genders. In contrast, early application timing uniquely matters for women: those applying within two months of graduation are, all else equal, 23.1 pp more likely to be employed six months later, while early job applications have no effect on men's employment. Finally, internship experience increases employment probability for both men and women.

4 Experimental Evidence on the Timing of Job Search

4.1 Experimental Design and Implementation

Motivation The finding that early labor market engagement is strongly correlated with women's employment potentially suggests new policy levers to foster female labor force participation. We design and implement an experiment to evaluate one such lever, specifically an incentive scheme that rewards early applications.[15] We chose to intervene on the supply side, rather than the demand side because in our sample the gender gap opens due to differences among the applicants, rather than the employers. In doing so, we complement a growing experimental literature that focuses on testing firm-level interventions aimed at reducing barriers to women's labor market access (Field and Vyborny 2022; Cheema et al. 2022; Ho, Jalota, and Karandikar 2023). We also considered an information campaign about the low employment rate of women (and its correlation with application timing), following recent research that has run information experiments to correct workers' misperceptions (Aloud et al. 2020; Jäger et al. 2024; Roussille 2024). For such a campaign to be effective, applicants would need to hold inaccurate beliefs about the employment rate of college-graduating women. We test for this in the diagnostic sample by collecting students' beliefs about the employment rate of their female (and male) peers.[16]. We display students' responses to this question in Figure 4. We find that women estimate that only 51.6% of their graduating female peers will be employed within six months of graduation. In contrast, women estimate their own likelihood of employment at 71.8%, which is not only much higher than their estimate for other women, but also comparable to their belief about men's chances (68.5%). In sum, even though women have high labor market expectations for themselves, they have much lower expectations about

15. We incentivize applications for jobs rather than for internships, which was also predictive of employment for women, because the former is directly linked to employment and logistically much simpler.

16. Specifically, we ask: "Think of women in your cohort at [the university] who are not pursuing further education after graduating. Out of 100 randomly selected female students in this sample, how many of them do you think would be employed within six months after graduating?" We repeat the question replacing "female" with "male."

other women in their cohort. Therefore, providing aggregate statistics about women may not result in any behavioral change if female respondents do not internalize these statistics as relevant to them.

The Test Our null hypothesis is that applying early has no impact on subsequent employment. There are two scenarios under which the null holds. First, there is no causal link between timing and employment, and the observed correlation reflects individual heterogeneity that affects both variables independently. Second, there is a causal link between timing and employment, but it is fully internalized so the incentive does not alter the optimal timing of job search. If we reject the null, it must be that there is a causal link between timing and employment, and women have inaccurate beliefs about *some* factor whose influence on employment changes over time, so that the same woman, applying for jobs at time t and $t + 1$, faces different probabilities of accepting an offer. The fact that job seekers have biased beliefs that can lead to unemployment is at the core of the recent literature on behavioral job search (Spinnewijn 2015; Mueller, Spinnewijn, and Topa 2021; Abebe et al. 2024) and motivates our test and interpretation of the findings.

Design To incentivize early applications we randomly assign students to treatment and control groups. The treatment consists of a small monetary reward paid to those who apply to jobs within one month of graduation. The experiment allows us to estimate two parameters of interest, that is the causal effect of the incentive policy on employment, β_{ITT}, and the causal effect of early applications on employment for the compliers, β_{LATE}. [17] To identify β_{ITT}, two assumptions need to hold: (i) that treatment is orthogonal to individual characteristics correlated with employment and (ii) the stable unit treatment value assumption (SUTVA). A successful randomization guarantees assumption (i) holds, and Table B.2 Columns 1 to 4 confirm that the treatment and control groups are balanced on all key variables measured at baseline. Violations of SUTVA are unlikely because students from both the treatment and the control group apply to jobs in the broader labor market and compete with other graduates from all over Lahore, the capital and largest city of the Pakistani province of Punjab. This implies that our treatment group can be considered as atomistic in this labor market. To identify β_{LATE}, we require the additional

17. We note that our experiment does not identify the average causal effect of timing on employment. To do that we would need to randomize applicants into different application timings, which would make some apply later than they would have in the absence of the experiment. This would be unethical as we would jeopardize their chance of ever finding a job. More importantly, this experiment would be irrelevant for the design of policy.

assumption that treatment affects employment solely through the timing of applications. This is the exclusion restriction necessary for treatment to serve as a valid instrument for early applications. It is likely to hold because the financial reward itself is too small to generate meaningful income effects on labor supply.

Implementation We field the experiment at the public university in Lahore in June 2023. The experimental sample consists of 1,947 students scheduled to graduate in mid-July 2023. A randomly selected 50% of the sample is offered a monetary reward on the condition that they apply to at least four relevant jobs by August 15th, approximately one month after graduation.[18] A job is relevant if it matches their skill set and, to claim their reward, they have to provide proof of their applications to the research team through a brief online questionnaire, including screenshots showing the application date and job title.[19] The incentive amount is PKR 5,000 ($\sim$ 18 USD or 89 USD at Purchasing Power Parity), which is equivalent to 2 days of pay at the median monthly salary in our diagnostic sample at the six-month follow-up. To insure against the risk of low take-up, we offer a much higher reward—PKR 20,000—to 10% of the sample.[20] Ex-post, this risk did not materialize and the two treatments have similar effects on applications and employment, thus we pool them in the analysis that follows.

For transparency and to ensure that students do not perceive the assignment to treatment as a signal of their labor market prospects, we tell them explicitly that assignment is done by lottery. Specifically, students are told "You have now reached the last part of the survey which is experimental. At this stage, whether you are shown two modules or just one module will be randomly determined by a lottery." Next, a wheel spins in front of them on the tablet with its outcome jointly observed by the student and the enumerator (see Figure A.13 for a visual of the wheel). If the student is not selected into treatment, they are told that "The lottery has decided that you will skip directly to the last module of the survey." We run three follow-up surveys. First, we re-survey students in early September 2023 (a couple of weeks after the deadline to receive the treatment incentive) to measure whether we have a first stage, that is a treatment effect on the number of early ap-

18. We condition the reward on students submitting at least four applications, as four was the median number of applications early appliers had sent by the two-month follow-up in the diagnostic sample.

19. Our research team reviewed all the screenshots to ensure students applied to real jobs that are relevant to their skill set and found that all students who took up treatment had complied with the terms of the reward.

20. We obtain PPP estimates from the World Bank: https://data.worldbank.org/indicator/PA.NUS.PPP?year=2023..

plications. We then re-survey students in early January 2024 to measure treatment effects on our outcome of interest: employment. We re-survey students a final time in September 2024, fourteen months after graduation to measure the persistence of the employment effects. The timeline is illustrated in Figure A.4.

4.2 Our Experimental Sample

Recruitment Based on budget and power calculations, we target a sample of about 2,000 students for our baseline survey. We stratify our sample by major and gender. Specifically, we over-sample women (65%), as they are our population of interest for the experiment. For majors, we stratify the sample to be representative of the full spectrum of majors at the university. The only exceptions are majors with fewer than 25 students, which we exclude, and a cap of 200 female and 100 male students per major, imposed to ensure broad representation in the final sample. This cap is binding for a few majors, in which case sub-groups of students in the major are randomly and incrementally invited to participate until the target is reached. Within this sampling frame, each male and female student has a 50% chance of being randomized into treatment. We receive 2,468 responses at baseline. Since we are interested in labor market beliefs and outcomes, we exclude 223 (9.0%) students who are enrolled in graduate school, as well as 299 (12.1%) students who have already secured a job[21]. After these adjustments, we obtain a baseline sample of 1,947 students.

Balance and Attrition We test for balance across treatment arms on a wide range of observable characteristics.[22] Table B.2 Columns 1 to 4 show that, at baseline, the treatment and control groups are balanced on all key baseline variables, confirming the success of our randomization procedure. Specifically, no significant differences are observed in gender composition, GPA, the distribution of majors, marital or familial backgrounds, and key outcomes between treatment and control groups. The remaining columns of Table B.2 show that balance on observables between treatment arms is maintained in the six-month

21. The share of students that secured a job before graduation (12.1%) is much lower than in some higher-income countries. For instance, in the United States, many college students interview and receive job offers in the year preceding their graduation (Cortés et al. 2023). This low rate of pre-graduation job acceptance informs the design of our experiment: offering financial incentives to apply the month preceding graduation can meaningfully impact students' application timeline as the vast majority has yet to start their job search.

22. For readability, we exclude the two-month follow-up from the attrition and balance tables since this follow-up is not used in the analysis, beyond showing descriptives about the shift in applications right after the application deadline. The two-month follow-up had a higher response rate than the later ones, and the balance test results are consistent with those from the six- and fourteen-month surveys.

(Columns 5 to 8) and fourteen-month follow-ups (Columns 9 to 12). Table B.3 further investigates whether attrition is systematically correlated with treatment status or baseline characteristics. Attrition is 25.9% and 37.4% at six and fourteen months post-graduation, with sample sizes of 1,442 and 1,218 students, respectively. There is no differential attrition by treatment status in the six- or fourteen-month follow-ups, as shown in Columns 5 to 12 of Table B.3. Similarly, attrition in either wave is not driven by specific majors, type of marital or familial profile, or differences in baseline values of key outcomes. An exception is that women are slightly overrepresented among those lost to attrition. However, Table B.2 confirms that the share of women remains balanced across treatment arms in both waves, indicating that this pattern is unrelated to treatment assignment. GPA is the only other characteristic that differs systematically between attritors and non-attritors. On average, individuals who remain in the panel have a GPA that is 0.05 points higher than those who leave. However, this difference is economically negligible (1.5%) and unlikely to meaningfully affect our results.

Treatment Take-up Take-up, defined as the share of respondents in the treatment group that claim and receive the financial award, is 53.5% for women and 48.5% for men. Figure A.14 correlates an indicator for take-up with students' baseline characteristics. It shows that both male and female students who take up treatment are more likely to major in engineering or computer science and less likely to major in Humanities. Second, women who are already engaged or married are, directionally, less likely to take up the treatment. Third, both women's and men's compliance are positively correlated with baseline beliefs about their own employment at six months, as well as baseline beliefs about the employment of same-gender peers at six months. Last, for both women and men, take-up is not correlated with gender norms, as measured by the World Value Survey questions. We also want to test for the "relevance" assumption, namely that the financial incentive leads some women to start applying earlier. To do so, we ask respondents, at the six- and fourteen-month follow-ups, the date of their first job application. As shown in Table 3, Column 1, we find that 57.1% of women in the treatment group have sent at least one job application by August 15th (the application deadline), compared to 31.6% in the control group. We observe similar effects on application timing for men: 54.7% of men in the treatment group have sent at least one early application, compared to 37.0% in the control group.

Descriptive Statistics Table 2 presents descriptive statistics for the six-month experimental sample, i.e., students who respond to both the baseline and six-month follow-up surveys. The last column of the table reports the p-value from a test of equal means between genders. Women make up 64.2% of the sample, with 926 female and 516 male respondents. On average, graduating students are about 23 years-old. As in the diagnostic sample, women's GPAs are slightly higher ($\sim$ 5%) than men's on average, and this difference is statistically significant. Unlike the diagnostic sample, where men and women tend to major in different fields, the experimental sample shows no systematic gender differences in the distribution of majors. Men and women also share similar backgrounds in terms of parental education, parental employment, and indicators of familial wealth. Only around 4% of both men and women are married at the time they graduate from college, and a similar fraction is engaged (4.3% of women and 3.1% of men).

External Validity of Diagnostic Findings Appendix C shows that the main diagnostic findings, derived from private university students, replicate using the control group of the experimental sample, drawn from a large public university. This validates the relevance of our diagnostic insights to a broad spectrum of college graduates.

4.3 Results

Intent-to-Treat: Specification To estimate the effect of our intervention on students' labor market outcomes, we run the following intent-to-treat analysis:

$$Y_{it} = \alpha_0 + \alpha_1 Male_i + \alpha_2 T_i + \alpha_3 (T_i \times Male_i) + \alpha_4 X_i + \epsilon_{it} \tag{1}$$

where Y_{it} denotes the outcome of interest for individual i at time t. The coefficient α_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 α_2 is the primary parameter of interest and measures the intent-to-treat effect of the intervention on women assigned to the treatment group ($T_i = 1$). α_3 represents the additional effect of the treatment on male students assigned to the treatment group. For interpretational ease, in our results table we show the treatment effect on male students as $\alpha_1 + \alpha_3$, instead of α_3. In our main specification, X_i is a vector of individual covariates measured at baseline and selected on the basis of their ability to predict the primary outcomes to improve statistical power (McKenzie 2012).[23] We report treatment effects separately for two periods t: six months and fourteen months post-graduation. All standard errors are robust to heteroskedasticity.

Intent-to-Treat: Results Table 3 shows results from our estimation of Equation 1. Panel A describes the treatment effect for women, while Panel B shows the treatment effects for men. In Panel A Column 2, we observe that treated women are 7.5 pp (22.3%, p=0.014) more likely to be employed six months post-graduation than control women. This increase in women's employment is equivalent to a ~35% decrease in the gender employment gap, which amounts to 21.5 pp in the control group. As shown in Column 4, treatment effects on employment persist in the fourteen-month follow-up, where treatment increases the likelihood of work for women by 6.9 pp (p=0.046) and has no effect on men's employment. Similar to the six-month results, the gender employment gap is 34% smaller in the treatment group relative to the control group, indicating that the effect remains stable over time. We further examine the likelihood of working for a positive wage at a firm (as opposed to being self-employed, working in a household enterprise, working for no wage, or being unemployed). Whether women take up firm work is an important outcome for two reasons. First, firm work is, at baseline, the preferred mode of working for the vast majority of women in our sample.[24] Second, in low-income countries, working informally for the family firm or on a self-employed basis is associated with lower labor force 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 self-employment is one of the primary drivers of gender gaps in earnings (World Bank 2021).[25] In Panel A

23. We adapt the post-double-selection approach set forth by Belloni, Chernozhukov, and Hansen (2014). We show that our results are robust to the exclusion of these controls.

24. We ask women who told us they planned to work at baseline: "What type of work would you prefer to do after you graduate?". The options were "Work for a private firm (or government)", "Work in my family-owned business", "Be self-employed" or "Be an intern at a firm". Strikingly, 90.7% of our female respondents said they prefer to work for a private firm.

25. 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

Column 3, we show that our treatment increases women's probability of working for a firm six months after graduation by 10.2 pp (40.3%, p=0.001). Column 5 shows that this effect persists fourteen months after graduation: while control women have a 41.6% likelihood of working for a firm, the likelihood for treated women is 9.5 pp higher (p=0.006). We interpret these stronger effects on firm work as evidence that the treatment not only changes women's chances of working but also the composition of their work. This shift in composition is consistent with the fact that our intervention induces women to apply for positions in the formal labor market, and thus does not have a channel through which it may influence self-employment prospects. The persistence of our effects more than a year after the incentive is awarded indicates that the intervention does not just prompt women who were already likely to work to start a job sooner. Rather, it sustainably increases levels of employment for women. Turning to men in Panel B, we detect no treatment effect on either men's likelihood of working overall, or of working for a firm. These results are consistent with our correlational findings, which show a positive association between early applications and employment for women, but not for men.[26]

Why A Few Months Have a Significant Impact Figure 5 provides further insights into the role that the few months around graduation play in women's longer-term employment, and how our treatment changes women's employment path. Since we ask women in the experimental sample about the exact date of their first job application, we can unpack the employment rate of women as a function of the month they first applied. Panel (a) shows that the share of working women drops dramatically between those who apply around graduation and those who wait just a couple of months. For instance, in the control group, 70.4% of women who apply in June 2023 are working for a firm by September 2024. This share goes to 54.5% for women applying in July 2023 and to only 41.7% for the women applying in August 2023. This illustrates a very steep decline in the employment profile of women with just a few months difference in their application behavior. To the extent that these patterns are not solely driven by selection (i.e. more employable women apply earlier), there is large room for improving women's employment by moving their

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.

26. Appendix Table B.4 shows results from our estimation of Equation 1 without the controls selected by the post-double-selection procedure. The treatment effects for women are of similar magnitude as in Table 3 and remain statistically significant throughout. At 6 months, effects on employment are significant at the 5% and effects on firm employment at the 1% level; at 14 months, they are statistically significant at the 10% and 5% levels, respectively. We continue to find no effects for men.

application date by only a month or two. The application and employment patterns in the treatment group confirm that such large room exists. Indeed, Panel (b) illustrates that our treatment triples the share of women having applied by end of July 2023 (from 9.2% in control to 24.4% in treatment). If selection was at play, then the compliers to our treatment, i.e. the many women who shifted their applications earlier, should have a different (lower) employment rate, such that the average share of working women among early appliers would be lower in the treatment than in the control group. Contrary to this narrative, Panel (a) shows that the share of women that are working at a firm in December 2024, conditional on being early applicants (i.e., applying by July 2023), is the same for the treatment and control groups. Taken together, the large shift to early application, the steep gradient in employment rates by application month and the absence of selection explain the large magnitude of our treatment effects.

Dynamics of Female Labor Supply Figure 6 examines the treatment effects on firm employment over time. Specifically, Panel (a) presents the treatment effect for women on having ever worked at a firm by a given date, Panel (b) repeats the exercise for men.[27] By May 2023, just before the treatment announcement, there is no difference in employment between the control and treatment group. For women, treatment effects progressively rise in the months following the announcement of the intervention, become marginally significant by the incentive deadline in mid-August 2023, and stabilize around 10 pp by January 2024 (about six months post-graduation). These treatment effects remain stable for the subsequent eight months. In contrast, the intervention has no effect on men's employment at any point in time. Since the results remain stable after six months post-graduation, we do not pursue further data collection after the fourteen-months follow-up. We also ask working women, at six months, how long they expect to continue in their *current* job: the average answer is 3.6 years (similar in both treatment and control). Given this, we consider the fourteen-month mark a relevant stopping point to study the post-graduation employment effects of the treatment. However, women's labor supply decisions may still change after marriage (which on average occurs at age 25 for college-educated women, that is about three years after college graduation). As a result, we cautiously consider this study as testing for the effect of early applications on women's pre-marriage labor supply. We consider early labor market experience to be an important

27. To get a start date on the first job that a respondent ever held, at the fourteen-month follow-up, we ask respondents about the start date of their first job. We only ask those who tell us they have worked at some point post-graduation.

outcome in its own right, as various studies show that it shapes women's lifelong labor supply. For instance, Jensen (2012) shows that entering the workforce early increases young Indian women's aspirations to long-term, post-marriage careers. Working early also facilitates women's return to work post-child, as past co-worker networks play an important role in informing the job search of mothers (Henke, Schmieder, and Berge 2021). In the context of our experiment, we directly ask working women about their long-term labor supply and find that the vast majority of them intend to pursue employment in the long-run. Specifically, we ask women in the six-month survey: "Do you intend to work (for a firm or in your own business) after getting married?" 85.0% of working women answer "Yes" to this question. While these responses may reflect women's inaccurate beliefs about their future labor supply, they provide evidence that women intend to pursue long-term employment.

Treatment on the Treated We show in Table B.5 the effect of treatment on the treated (the students who were offered the treatment and took it up) and on the non-treated (students who were offered the treatment but did not take it up), relative to the control group. We find that, for women (Panel A), the employment of the treated group is 13.6 pp higher (p=0.000) than the control group's at 6 months, and 11.6 pp higher (p=0.003) at 14 months. By contrast, the employment of the non-treated is only marginally higher than in the control group, reflecting limited selection patterns into treatment take-up. For men (Panel B), the treatment effect on the treated is also positive (7.9-10.3 pp). However, the treatment effects on the non-treated men are of similar magnitude as the effects on the treated men, but of the opposite sign. This reflects selection patterns into treatment take-up and helps explain why, in aggregate, the male intent-to-treat estimates are null.

Local Average Treatment Effects Beyond the intent-to-treat effects of our intervention, we are also interested in the effect of applying early itself. Not everyone selected by our lottery ended up sending early applications and some students in the control group applied early, even absent the financial incentive. Therefore, we use a standard instrumental-variable approach with lottery selection as an instrument for applying early to estimate a local average treatment effect. Formally, we estimate the following model with 2SLS:

$$Early_i = \beta_0 + \beta_1 Male_i + \beta_2 T_i + \beta_3 T_i \times Male_i + \epsilon_i \tag{2: 1st Stage}$$

$$Y_{it} = \delta_0 + \delta_1 Male_i + \delta_2 \widehat{Early_i} + \nu_{it}. \tag{3: 2nd Stage}$$

where $Early_i$ denotes whether the student has sent an early application. In our first stage, the coefficients β_2 and β_3 capture the effect of the financial incentive (treatment) on the likelihood of applying early for women and men, respectively. Our second stage estimates the effects of applying early, as instrumented by the treatment indicator interacted with the student's gender, on labor market outcomes (e.g., employment). Empirically, we define applying early as a dummy equal to one if one has sent their first application before August 15th (the application deadline to receive our financial incentive).[28]

Local Average Treatment Effects: Results In Table 4, 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 effect of our treatment. For women, the OLS estimate in Column 1 suggests that early applications are associated with a 23.6 pp increase in the probability of working for a firm. The 2SLS estimate in Column 2 is higher than the OLS one, at 38.3 pp. These effects persist at the fourteen-month follow-up, with the 2SLS estimate at 30.8 pp. The OLS and 2SLS coefficients differ for several reasons. The OLS is upward-biased because it captures the effect of selection—individuals who are more eager to work apply earlier—but downward-biased because of measurement error. In contrast, the 2SLS estimator measures the causal effect of early application on the compliers, that is, individuals whose application behavior changes because of the financial incentives. For women, the 2SLS estimate of the causal effect of applying early is larger than the OLS. In contrast, for men (Panel B), the point estimates in the 2SLS estimation (in Columns 2 and 4) are small and statistically insignificant (ranging from -0.03 pp to 0.05 pp). This implies that early applications have no effect on the likelihood of employment for male compliers.

Lack of Anticipation about the Role of Timing As explained in Section 4.1, the fact that we find a positive treatment effect on women's employment not only demonstrates that there is policy space to reduce the gender employment gap, but also implies that women hold inaccurate beliefs about *some* time-varying parameter of the job search process. However, we find direct empirical evidence that women in our treatment group do not anticipate the causal effect of timing. At baseline, after randomizing respondents into

28. This definition is more granular than in the diagnostic part of the paper, where applied early was defined as whether one had sent an application by the second follow-up. In the experimental surveys, we added a question in each wave asking respondents about the exact date at which they sent their first job application. Therefore, we can define applying early precisely as the date of the incentive deadline (August 15).

being offered treatment, we ask them about their job search timing and employment expectations. On search timing, we ask them both when they think they will send their first job application, and how many applications they think they will have sent by August 15th (the deadline to apply to four relevant jobs in order to get the reward we offer). On employment, we ask them to report their perceived likelihood of working six months post-graduation. The latter question is worded the same way as for the diagnostic sample (see Section 2.2). Table 5 shows the intent-to-treat estimates for these variables. We find that treatment significantly shifts women's beliefs about when they will start applying earlier (Column 2, p=0.000). Being assigned to treatment also increases the number of applications women expect to make by August 15th (Column 1, p=0.067). However, we find no treatment effect on women's beliefs about their likelihood of working (Column 3, p=0.391). This provides direct evidence that treated women did not anticipate the role of timing in employment.

Discussion We find that a small financial incentive to send a few applications right after graduation has a large and persisting effect on women's employment. In the next section, we dive into the mechanisms driving our results, which will help make sense of the magnitudes. As a preamble to this analysis, we note that the few months around graduation are a critical juncture in the personal and professional lives of the women we study. Most of them have no work experience so their post-graduation job search is not informed by past labor market exposure. During their college years, women spend their days at school and may live away from their parents. As Table 1 and Table 2 highlight, very few of them are married or even engaged during college. But, at the time they graduate, women suddenly return to their parents' home and/or spend significantly more time at home as school has ended. At around the same time, parents ramp up the search for a suitable marriage match (Calvi, Farooqi, and Kandpal 2024), which we discuss in more detail in Section 5.2. At this critical juncture, even a small behavioral change can plausibly generate large, path-dependent effects on women's labor market outcomes. The next section explores the channels through which these large effects operate.

5 Why and How Does Timing Matter for Employment?

As explained in Section 4.1, the fact that we find a positive treatment effect on women's employment not only demonstrates that there is policy space to reduce the gender employment gap, but also implies that women hold inaccurate beliefs about *some* time-varying parameter of the job search process. This section explores *which* time-varying parameter these inaccurate beliefs are about. To do so, we first identify demand and supply-side determinants of labor market outcomes that can plausibly change over time. We then provide empirical evidence on which of these determinants drives our experiment results, and the channels through which they do so.

5.1 Changes in Labor Demand

One possible explanation for our treatment effects is that labor demand slows down over time, and women do not anticipate it. This hypothesis is motivated by a well-established literature documenting that longer unemployment spells result in worse labor market outcomes (Schmieder, Wachter, and Bender 2016), as well as a growing literature showing that workers do not anticipate this (Mueller, Spinnewijn, and Topa 2021). We observe three facts that, together, rule out this hypothesis in our context. The first is that there is no slowdown in labor demand for men. As shown in the previous section, our treatment has a similar effect on application timing for both genders but only affects employment outcomes for women, implying that men's vacancy creation and offer arrival rates do not decay with time. Thus, our treatment effects are unlikely to be driven by diminishing labor demand over time for the occupations that employ both men and women. Additionally, Table 6 shows treatment effects on intermediate labor market outcomes that ultimately result in employment: number of applications and number of job offers. While treated women send more applications than control women at the six and fourteen months marks (Column 1 and 3), the number of job offers received is the same for treated and control women at the six- and fourteen-month marks (Column 2 and 4). These results have two implications. First, treatment-induced changes in search effort are unlikely to explain our results since they do not translate into a higher number of job offers. Second, since the group that applies later (control) receives the same number of offers as the group applying earlier (treatment), the lower employment rate in the control group has to be driven by lower job acceptance rates, rather than by a slowdown in demand. In Appendix Table B.6, we also show results on applications and job offers at the extensive margin (i.e. whether or not a student sends at least one job application and receives at least one job offer). Reflecting the results on number of applications and offers, we find that the treatment increases the likelihood of sending at least one application even by the fourteen-month mark but has no effect on the likelihood of getting at least one offer by the fourteen-month follow-up. Overall, the evidence does not support a labor demand slowdown as the reason for our treatment effects on employment.

One last dimension we study is the effect of our treatment on the composition of jobs (and corresponding wages). Indeed, it could be that the control group is as likely as the treatment group to secure a job offer, but control group members may receive lower wage offers or end up in different occupations. Considering the top 3 occupations women most frequently work in, Panel A of Table B.7 shows that the share of women in a given occupation is similar across treatment and control groups at the fourteen-month follow-up. For example, teaching is the most common job among women, and plausibly sensitive to seasonal changes in demand. If the peak hiring period in teaching coincides with our deadline given to the treatment group for applying early, we may expect that earlier applications lead treated women to disproportionately sort into teaching. However, treated and control women are equally likely to work as teachers—23.8% of women in the treatment group become teachers relative to 24.1% of women in the control group (p-value of the difference = 0.916). This suggests that the positive employment effects of applying earlier and subsequently receiving job offers are not driven by women sorting into a different category of jobs. Further, Figure A.15 presents the distribution of current wages in Panel (a) and offered wages in Panel (b) for control and treated women at the six and fourteen-month follow-ups. The distributions for both current and offered wages are similar across treatment groups, as indicated by absolute z-scores below one. This finding contradicts the idea that control women face a different demand curve than treated women.

5.2 Changes in Labor Supply

We already discussed in the previous section that while there is indeed a higher total number of applications sent by treated relative control women, these cannot explain our results as the number of offers remain the same. In this section, we further investigate whether women's preferences for work change over time and if so, whether women anticipate these changes.

Change in Preferences Just as in the diagnostic sample, at baseline women in the experimental sample perceive their likelihood of employment as high. For instance, as shown in Figure C.1, control women estimate a 79.8% chance of working six months post-graduation. However, only 33.6% of them are actually employed by the six-month mark. This stark gap between expectations and reality points to two supply-side factors as potential explanations for our treatment effects: either women unexpectedly revise their preferences for work downward over time, or they encounter unanticipated constraints that limit their ability to work. To disentangle these two explanations, we field a direct question about women's preference for working in the near future at the six-month follow-up: "Do you want to work in the coming 12 months?"[29] We find that 87.6% of all women, and 83.4% of unemployed women, answer "yes" to this question, with no treatment effects on the response.

As an additional test of this hypothesis, we examine whether women's reservation wages rise over time, which would be consistent with a declining interest in work among women. Table B.8 reports estimated time-trends and treatment effects on log reservation wages. Among control group women, reservation wages *fall* significantly over time: by 14.5 log points at the six-month follow-up and by 7.1 log points at the fourteen-month follow-up, both relative to baseline. There is no evidence that treatment affects this trajectory. Taken together, the results suggest that women continue to express a desire to work and appear increasingly willing to do so at lower wages—supporting the interpretation that binding constraints, rather than shifting preferences, are driving the decline in employment.

The Role of Beliefs about Labor Markets There is a longstanding literature on how biased beliefs about wages and job offer arrival rates influence the labor market outcomes of job seekers (Spinnewijn 2015; Mueller, Spinnewijn, and Topa 2021; Abebe et al. 2024). This literature also engages with time-varying factors: job seekers with optimistic bias tend to turn down early offers in the hope of getting better offers later on. While these workers adjust their wage expectations over time, they often adjust too slowly and end up in long-term unemployment. Motivated by this literature, we test whether treatment effects are heterogeneous across baseline labor market beliefs.[30] In Figure 7 Panel (a), we show that treatment effects on employment are the same for individuals who, at baseline, have above median vs. below median reservation wages, expected wage offers and beliefs about market wages. We also find that these effects are not driven by beliefs about broader market conditions, as proxied by second-order beliefs regarding other women's job prospects. Figure A.16 Panel (a) also examines heterogeneity in treatment effects on early application behavior using the same baseline wage beliefs. Again, we find no

29. We purposefully made the wording about preference for work, rather than expectations to work, to ensure that we are isolating women's desire to work, outside of external factors that could impact their ability to work.

30. To maximize power, we pool data from the six- and fourteen-month survey waves, with wave fixed effects, and cluster errors at the individual level.

evidence of differential effects. Taken together, these results suggest that our treatment effects do not operate through differences in beliefs or expectations about wages and job prospects.

5.3 Changes in the Outside Option

The Marriage Market While men and women share similar preferences for work and pursue a similar search strategy, their outside options differ sharply: women may specialize in housework after marriage, whereas men do not. Therefore, to explain the labor force participation decision of women, it is important to understand the marriage market. In Pakistan, college-educated women get married, on average, three to four years after graduation. However, the marriage market—comprising the process of receiving and evaluating proposals—tends to unfold quickly for them in the immediate aftermath of graduation. Figure 8 illustrates this dynamic. The solid red line shows the cumulative share of women in the control group who have received at least one marriage offer by a given date after graduation. While most women marry several years later, a striking one-third receive a marriage offer within just two months of graduating. The figure also overlays the female employment share in the control group, plotted by the month of first job application. What becomes evident is that the steepest decline in job acceptance rates—occurring in the initial months post-graduation—closely coincides with the sharpest rise in marriage offer arrivals. Applying in September instead of June lowers the employment rate by 30.4 pp at the same time marriage offers increase by 29.8 pp. While this is purely correlational, the timing of the two variables is too closely synchronized to be dismissed as entirely coincidental. This pattern is consistent with a "race" between the marriage and the labor market: for women who have not searched for a job by the time marriage market activities ramp up, the chances of accepting a job offer are lower (more on potential explanations below). We combine this observation with the fact that inaccurate beliefs about a time-varying parameter of search must underlie our treatment effects to formulate the following hypothesis: some women may underestimate the speed of the marriage market onset such that, absent treatment, they suffer from "the illusion of time": they delay applications until after marriage market activities ramp up, at which point starting work is harder. We provide evidence on this phenomenon below.

Change in Beliefs about Marriage First, we test whether women change their expectations about their age of marriage over time. Figure 9 illustrates the relationship between women's baseline beliefs about age at marriage (x-axis) and women's updated beliefs, six months post-graduation (y-axis), separately for control and treatment. The 45-degree line represents the benchmark of time-invariant beliefs. While there is a positive slope (of ∼ 0.6), it is significantly below 1, suggesting that women who initially expected to marry later are likely to revise their expectations downward. For instance, women who believe at baseline that they would get married at age 26 have converged, by the six-month mark, to believing they will get married a year earlier, at age 25. In contrast, women who initially anticipated marrying sooner (e.g., by age 24) have near-stable expectations. The fact that women who expected to marry later are the ones who update their beliefs the most, suggests that they held inaccurate beliefs at baseline and were exposed to an earlier onset of marriage-related activities than they had expected. Consistent with a convergence of beliefs towards the truth, the beliefs of women who initially thought they would marry late update downwards to age 25, the national median age at marriage for educated women. Importantly, the treatment has no effect on these beliefs about the marriage timeline.

Having established that women who initially thought they would marry late lack foresight about the speed of the marriage market, we turn to testing whether these same women are driving our treatment effects on employment. Panel (b) of Figure 7 tests for heterogeneity in treatment effects based on baseline marriage beliefs, separating women who initially think they will marry earlier than age 25 (63.3% of our sample) vs. later (37.0% of our sample). We find that our treatment effects on working for a firm are much stronger for women who expect to marry late (past age 25), at 17.1 percentage points versus only 5.4 percentage points for those expecting early marriage (p-value of the difference = 0.032).

Change in Family Preferences Our last test sheds light on family involvement, following a large and growing literature emphasizing the role of families in shaping women's labor supply, but not men's (Dean and Jayachandran 2019; Bursztyn, González, and Yanagizawa-Drott 2020; Jayachandran 2021; Lowe and McKelway 2024).[31]

31. Families are particularly relevant in Pakistan. In a representative sample of urban households, 79% of women report that the head of the family decides (unilaterally or in consultation with them) whether they should work (Junaid et al. 2021). The rapid unfolding of marriage offers is 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. Several other papers point to the role of families in restricting women's labor supply in Pakistan. In an experiment on women in Lahore with similar age, marital, and educational profiles to our sample, Subramanian (2024) finds that increasing the salience of family discussions about job search led to a 60% decrease in job applications to positions with male supervisors, compared to a control group in which family opinions were not highlighted. Field and Vyborny (2016) also document that 40% of non-working women in Pakistan cite lack of permission from

To test for heterogeneous treatment effects by family involvement in job search, we ask all respondents in our baseline survey: "Whom do you consult when making decisions on a job?" We consider that parents are involved in women's job search if their response is either that they consult their mother, their father, or both.[32] We find that 60.0% of women say that they consult their parents (vs. 38.8% of men). We treat heterogeneity by this variable as a test of the hypothesis that time-varying preferences of parents can explain our result. If timing only matters in situations where women consult their parents about job decisions, this suggests that families, rather than the women themselves, are the ones whose preferences shift over time, especially given our earlier finding that women's own preferences did not change. Figure 7 Panel (b) presents treatment heterogeneity results by this variable. While treatment effects on early applications are similar across women who do and do not rely on their parents to make job decisions (as shown in Figure A.16), treatment effects on firm employment differ. In families where daughters do not consult their parents about job offers, the treatment effect on firm employment is smaller than average, at 6.5 pp. In contrast, in families where parents play a role in women's job search decisions, the treatment effect is larger, at 16.3 pp (p-value of the difference = 0.069)— close to a 3-fold increase compared to women who do not involve their parents. The fact that timing matters more for women who involve their parents in their job search points to time-varying preferences among parents about their daughters' work. For these time-varying preferences to explain our treatment effects, they must be unanticipated. In support of this, the left panel of Figure A.17 shows that only 20% of women report that they may struggle to work after graduation due to "family reasons," such as not receiving approval from their families or being pressured to prioritize finding a marriage match. In contrast, 91% of women recognize that *other* women will face those obstacles. Further, confirming women's lack of foresight about parents' opposition, we find no relationship between women reporting parental involvement in their job search and their recognition that family may pose a barrier to one's own employment. In sum, while women acknowledge that parents generally play a significant role in restricting daughters from working, they do not view their own families as a factor limiting their employment. This contrast between beliefs about the self and about others is in line with the tension we find between women's high expectations of own work, but low expectations

their husbands or fathers as the main reason for not working.

32. The full list of options is Friends, Mother, Father, Brother(s), Sister(s), Cousin(s), Husband or fiancé, Rest of family, Teachers (or career office, career councilor, university website), Classmates enrolled in my course / seniors, No one, Other.

about other women's chances of employment. Together, these observations draw a picture in which women are aware of the barriers the "typical" woman faces, but consider these barriers do not apply to them.

In sum, consistent with our main hypothesis that timing matters because of an unexpected shift in attention to the marriage market, our treatment effects on employment are driven by women who lack foresight about the imminence of the marriage market. Together with the evidence on parental involvement, this result provides insight into how the treatment may operate: in the treatment group, women start searching for jobs earlier and, as a result, consult their parents on job offers at a time where the marriage market has yet to unfold, such that parents may be more inclined to let their daughters start a job.

Treatment Effects on the Marriage Market We just showed that marriage and labor market activities are negatively correlated in timing. Since our treatment persistently shifts women into work, one open question is whether the treatment impacts the marriage market outcomes of women in our sample. While marriage itself is still many years out for most women and therefore cannot be measured by our data, we collect early marriage market signals: number of marriage offers and the quality of these offers, proxied by the highest education achieved among the potential grooms. Columns 1 and 4 of Table B.9 show that women in the treatment group are no less likely to get married or engaged than control women in the first six and fourteen months post-graduation. Columns 2 and 5 also show that women in the treatment group receive the same number of marriage proposals as women in the control group at both the six- and fourteen-month marks. Finally, Column 3 shows that the quality of marriage offers—proxied by the highest education level among received proposals—does not differ by treatment status. In sum, our treatment works not by delaying the launch of the marriage market, but by inducing the labor market to unfold before the marriage market.

This result suggests that, in our context, women induced to work early by the treatment do not face a marriage penalty. However, a penalty for starting work later may still exist. One way to reconcile these two possibilities is if marriage offers are endogenous to a woman's work status: women who begin working early may attract more 'progressive' men who are supportive of female employment, while those who delay entering the labor force may receive offers from men less accepting of working wives. This dynamic could create a self-fulfilling prophecy—women who postpone work risk losing existing marriage offers if they start working later, and therefore choose not to work at all. Meanwhile, working women receive offers from more progressive men and are able to continue working. The difference in views among men might be innate or shaped by the breadwinner norm–that is, men may be indifferent to their wives' employment status but suffer a reputation loss if this is perceived as a necessity (started working after the marriage offer) rather than a choice (started work beforehand).

6 Conclusion

This study provides new evidence that most college-educated women in Pakistan expect to work at graduation, yet most remain unemployed six months later. Traditional supply-side explanations fail to account for this gender employment gap: women have higher GPAs (conditional on major choice), lower reservation wages, and similar preferences for non-wage amenities as their male peers. Demand from firms is also not the limiting factor, as women are just as likely as men to receive interviews and job offers—though at lower wages. Instead, the gap is primarily driven by women rejecting job offers at significantly higher rates than men.

This raises the question: why do women's employment outcomes fall short of their expectations? Our analysis reveals that the timing of job search is the strongest predictor of gender differences in employment. Women who apply for jobs within two months of graduation are 23.1 percentage points more likely to be employed than those who apply later, while men experience no such returns to earlier applications. Motivated by this finding, we design an experiment offering a small financial incentive to randomly selected students to encourage earlier job search. While the incentive prompts both men and women to apply sooner, it significantly increases employment only among women. These findings suggest that women hold inaccurate beliefs about some time-varying aspect of job search, leading them to delay applications more than they should.

We show that preferences do not change over time and beliefs about standard labor market outcomes do not predict the strength of the treatment effect. In contrast, the expected age at marriage falls over time and women who, mistakenly, expected to marry later than the national average age of 25, are the most likely to suffer from"the illusion of time".

These findings imply that policies that lead to marginal changes on one dimension—here the timing of applications—can have large impacts if selection is driven by a second variable—in our case expected age at marriage. The findings also imply that policies designed to achieve labor market outcomes must take into account the effect on the outside option, whatever that may be.

References

Abebe, Girum, Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, Simon Quinn, and Forhad Shilpi. 2024. "Matching Frictions and Distorted Beliefs: Evidence from a Job Fair Experiment." *Economic Journal,* Forthcoming.

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.*

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, Alexia Delfino, and Edward Glaeser. 2022. "Female Entrepreneurship and Trust in the Market." *Journal of Political Economy,* Forthcoming.

Bandiera, Oriana, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali. 2023. "The search for good jobs: evidence from a six-year field experiment in Uganda." *Journal of Labor Economics.*

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives* 28, no. 2 (May): 29–50.

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

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, Alexander W Cappelen, Bertil Tungodden, Alessandra Voena, and David H Yanagizawa-Drott. 2023. "How are gender norms perceived?" *NBER Working Paper,* no. 31049.

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.

Calvi, Rossella, Hira Farooqi, and Eeshani Kandpal. 2024. *Perceived Marital Returns to Education and the Demand for Schooling.* Working Paper 709. Center for Global Development.

Card, David, Fabrizio Colella, and Rafael Lalive. 2021. "Gender Preferences in Job Vacancies and Workplace Gender Diversity." *NBER Working Paper,* no. 29350.

Cheema, Ali, Asim I. Khwaja, M. Farooq Naseer, and Jacob N. Shapiro. 2022. "Glass Walls: Experimental Evidence on Access Constraints Faced by Women." *G^2LM | LIC Working Paper,* no. 65.

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.

Dean, Joshua T., and Seema Jayachandran. 2019. "Changing Family Attitudes to Promote Female Employment." *AEA Papers and Proceedings* 109:138–42.

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. 2016. "Female labor force participation: Pakistan country study for the Asian Development Bank." *Policy brief, Asian Development Bank* 4.

———. 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 (2): 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. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–1119.

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

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.

Henke, Alexander, Johannes F. Schmieder, and Philipp vom Berge. 2021. "Mothers' Job Search after Childbirth." *IZA Discussion Paper,* no. 14593.

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* 22 (5): 2108–2147.

Jaeger, David A., Jaime Arellano-Bover, Krzysztof Karbownik, Marta Martínez-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.

Jensen, Robert. 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India." *The Quarterly Journal of Economics* 127 (2): 753–792.

Junaid, Novaira, Ume Farwa, Asima Ihsan, and Shahzada M. Naeem Nawaz. 2021. *Female Labour Force Participation Rate in Punjab: Challenges and Way Forward.* Punjab Economic Research Institute.

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante. 2023. "The Child Penalty Atlas." *NBER Working Paper,* no. 31649.

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.

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 (4): 1013–1048.

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.

Lowe, Matt, and Madeline McKelway. 2024. "Coupling Labor Supply Decisions: An Experiment in India." *Working Paper.*

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): 2007–2047.

Mas, Alexandre, and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 107 (12): 3722–3759.

McKenzie, David. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99 (2): 210–221.

Mueller, Andreas I., Johannes Spinnewijn, and Giorgio Topa. 2021. "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias." *American Economic Review* 111 (1): 324–363.

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

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

Schmieder, Johannes F., Till von Wachter, and Stefan Bender. 2016. "The Effect of Unemployment Benefits and Nonemployment Durations on Wages." *American Economic Review* 106 (3): 739–777.

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

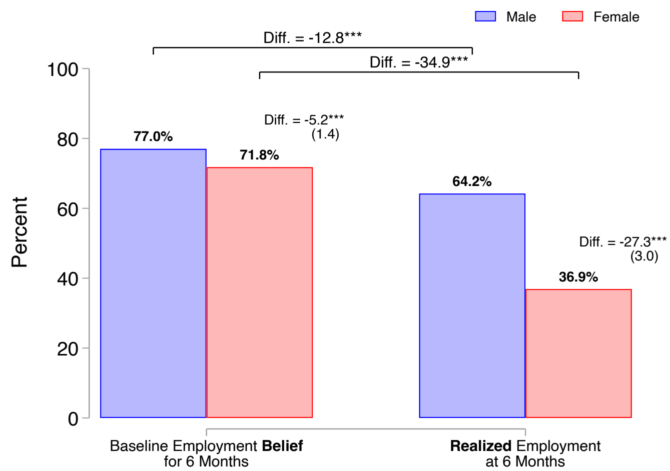
Subramanian, Nivedhitha. 2024. "Workplace Attributes and Women's Labor Supply Decisions: Evidence from a Randomized Experiment." *Working Paper.*

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

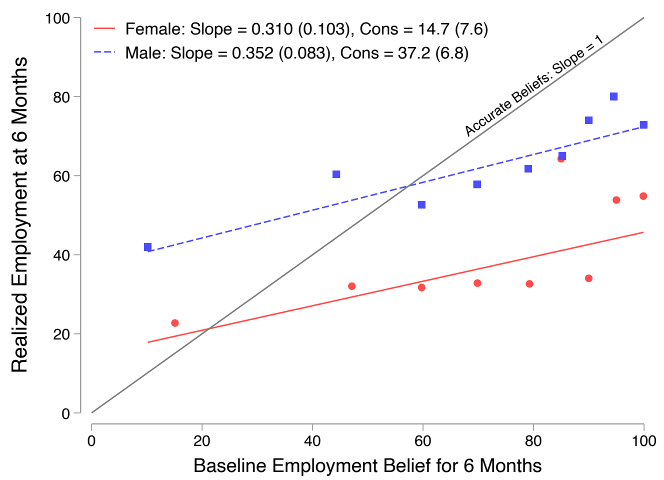
World Bank. 2021. *Women, Business and the Law 2021.* Accessed: 2024-10-01. World Bank. https://documents1.worldbank.org/curated/en/492221468136792185/pdf/Main-report.pdf.

Figures

Figure 1: Baseline Employment Beliefs vs. Realized Employment Outcomes



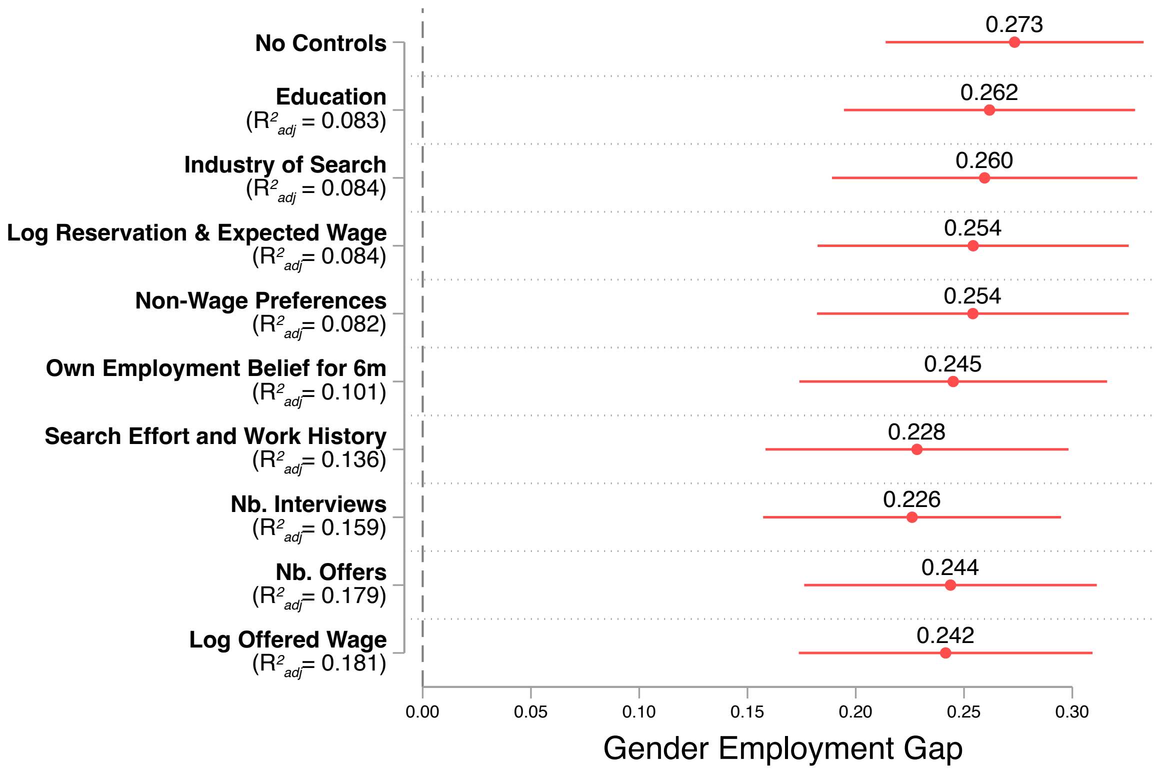
(a) Mean Levels: Expected vs. Realized Employment



(b) Binned Scatter: Expected vs. Realized Employment

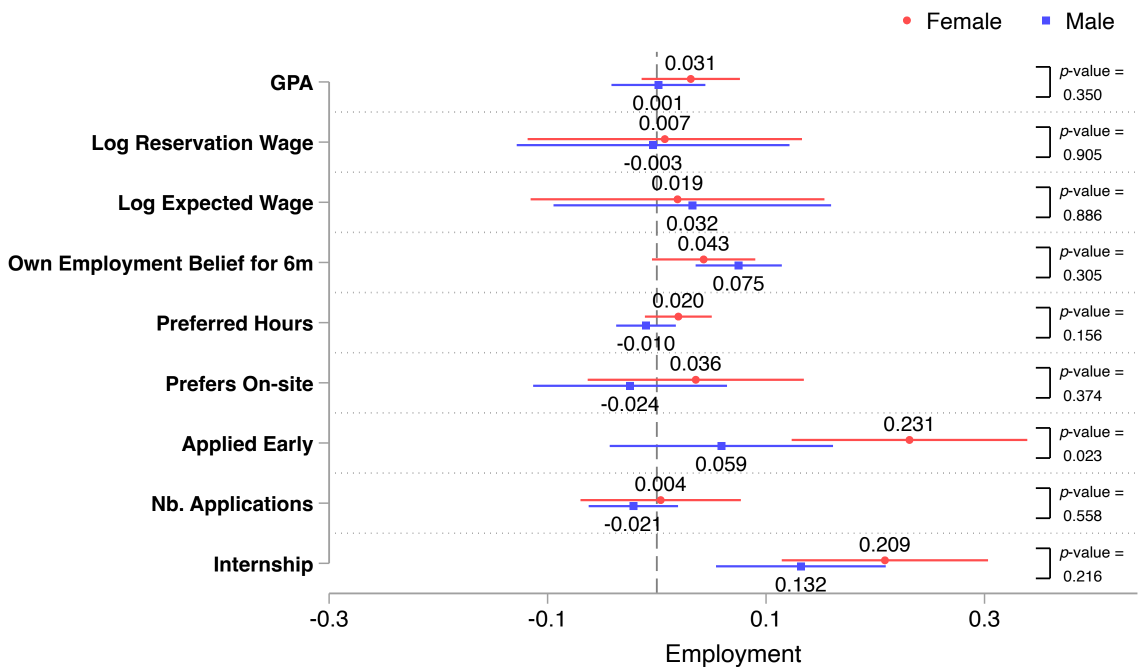
Notes: This figure compares baseline beliefs about one's future employment with realized employment. The sample consists of students in the diagnostic sample (see Section 2.1 for details). Panel (a) contrasts the average baseline belief of students about their own chances of working within six months of graduation (left cluster of two bars) with the same students' realized employment at the six-month mark (right cluster of two bars), separately for men (bars 1 and 3) and women (bars 2 and 4). The gender gap in responses is shown directly above the female bar. The average within-gender difference between baseline beliefs and realized employment is shown above the horizontal brackets. Panel (b) shows a binned scatter plot of baseline employment beliefs against realized employment, separately for men (dashed blue line with square markers) and women (solid red line with circular markers). The 45-degree line represents accurate beliefs, points above (below) the line indicate underestimation (overestimation) of employment chances. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 2: Explaining the Gender Employment Gap



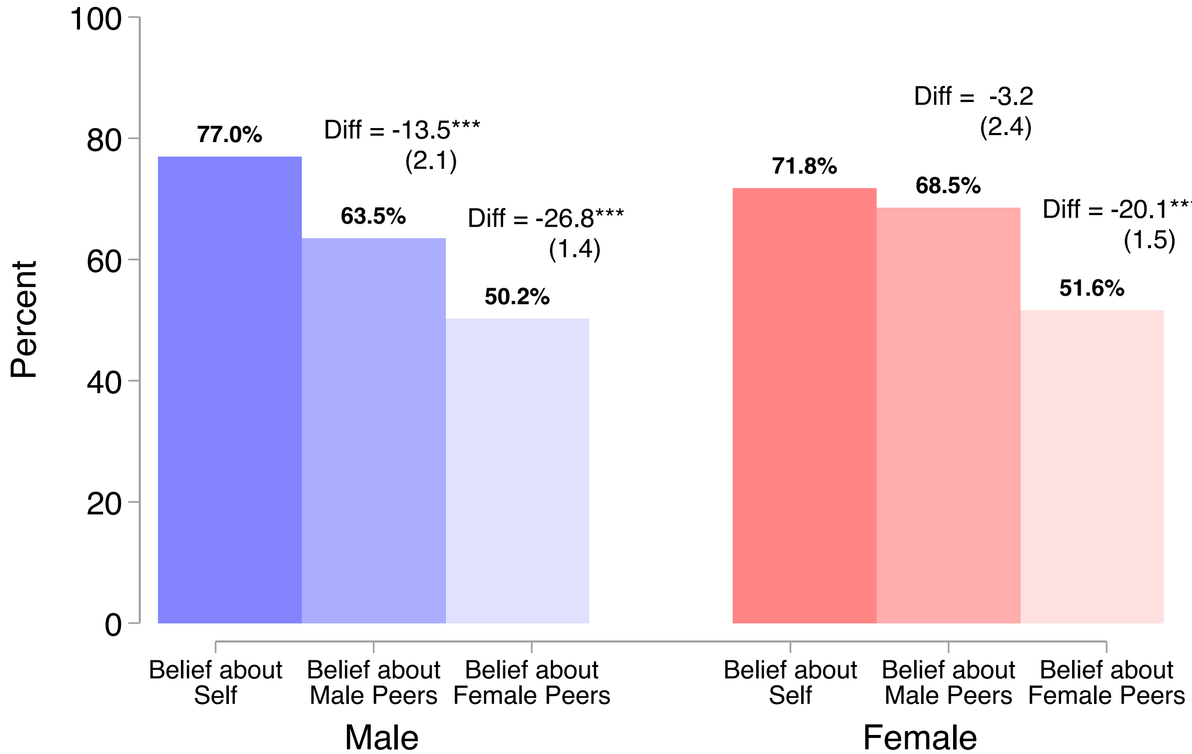
Notes: This figure presents the difference between male and female employment six months post-graduation, with observable baseline characteristics added incrementally as controls. The sample consists of students in the diagnostic sample (see Section 2.1 for details). The Education controls include cumulative GPA and major fixed effects. The Industry of Search controls are fixed effects based on the SOC (Standard Occupational Classification) of a respondent's text-entry baseline responses about what job they are searching for. The Non-Wage Preferences controls include baseline preferences regarding onsite vs. remote work and preferred number of daily work hours. The Own Employment Belief for 6 Months control includes baseline belief about one's own probability of employment six months later. The Search Effort and Work History controls include the total number of job applications sent by the six-month follow-up, an indicator for applying early (that is for having sent at least one application by the two-month follow-up), and an indicator for internship experience. The Nb. Interviews (Offers) control includes the number of interviews (offers) received by the six-month follow-up. The Offered Wage control includes the highest wage offer a student has received for a job (regardless of whether they have accepted it).To address missing values (e.g., if a student did not receive a job offer), we assign a constant (999) to missing entries and include a binary indicator (e.g., no_offered_wage) in the regression. All unbounded continuous variables are winsorized at the 2% level. Each coefficient is from a multivariate regression that includes all preceding rows' variables as controls. Adjusted R-squared values are shown for each row. The horizontal bars show 95% confidence intervals.

Figure 3: The Determinants of Gender-Specific Labor Supply



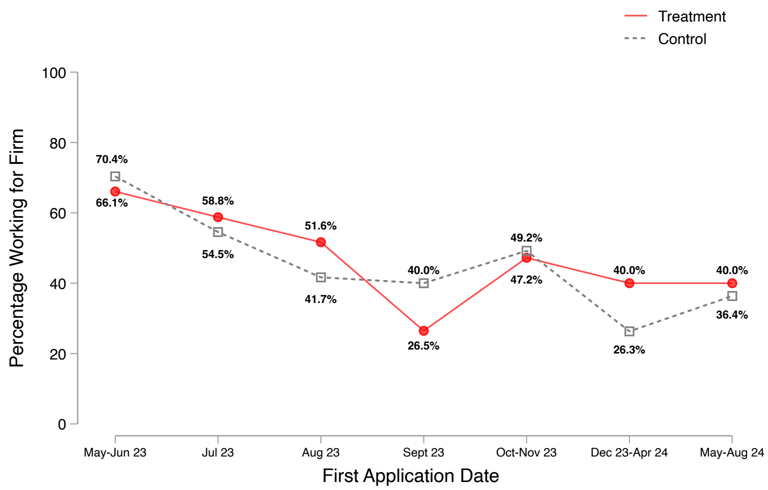
Notes: This figure presents coefficients from a regression of one's employment six months post-graduation on a set of independent variables, separately by gender. Each coefficient is from a multivariate regression conducted separately for men and women, in which all other listed variables are included as controls. The sample consists of students in the diagnostic sample (see Section 2.1 for details). The variables include cumulative GPA (measured in standard deviations), baseline reservation wage, baseline wage expectations, baseline belief about one's own probability of employment (transformed on a 0-1 scale and measured in standard deviations), preferred number of daily work hours, preference for onsite vs. remote work, an indicator for applying early (i.e. by the two-month follow-up), the total number of job applications submitted by the six-month follow-up (measured in standard deviations), and an indicator for having internship experience. All unbounded continuous variables are winsorized at the 2% level. Vertical brackets indicate statistical significance tests of equality of coefficients across gender, with the corresponding p-values reported. Figure A.12 presents the results of bivariate regressions. The horizontal bars show 95% confidence intervals. Corresponding means and standard deviations for the independent variables shown in the figure are provided in Table B.10.

Figure 4: Employment Beliefs about Self vs. Employment Beliefs about Peers

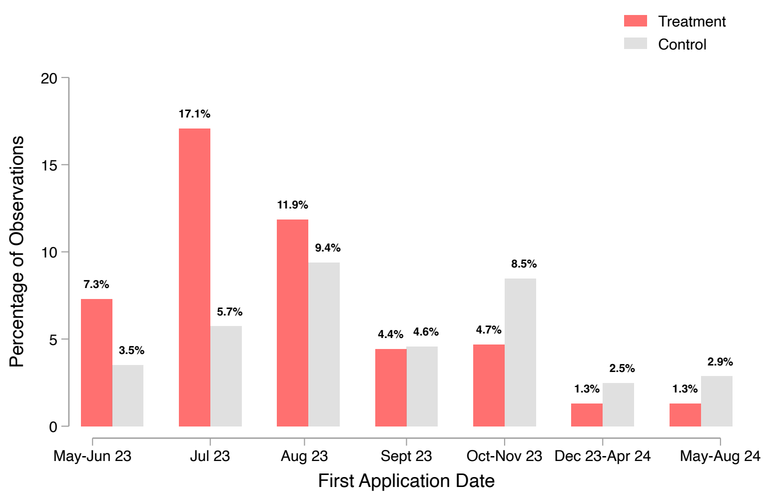


Notes: This figure presents respondents' beliefs about their own employment likelihood relative to beliefs about their peers' employment likelihood, separately by gender. The sample consists of students in the diagnostic sample (see Section 2.1 for details). Male beliefs are represented by the blue cluster of bars on the left. Female beliefs are represented by the red cluster of bars on the right. The first bar in each cluster shows average baseline beliefs about one's own employment likelihood. The second bar in each cluster shows average baseline beliefs about one's male peers' employment likelihood. The third bar in each cluster shows average baseline beliefs about one's female peers' employment likelihood. The average difference between beliefs about self and beliefs about male peers is shown directly above the second bar in each cluster. The average difference between beliefs about self and beliefs about female peers is shown directly above the third bar in each cluster. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 5: Explaining the Magnitude of our Treatment Effects



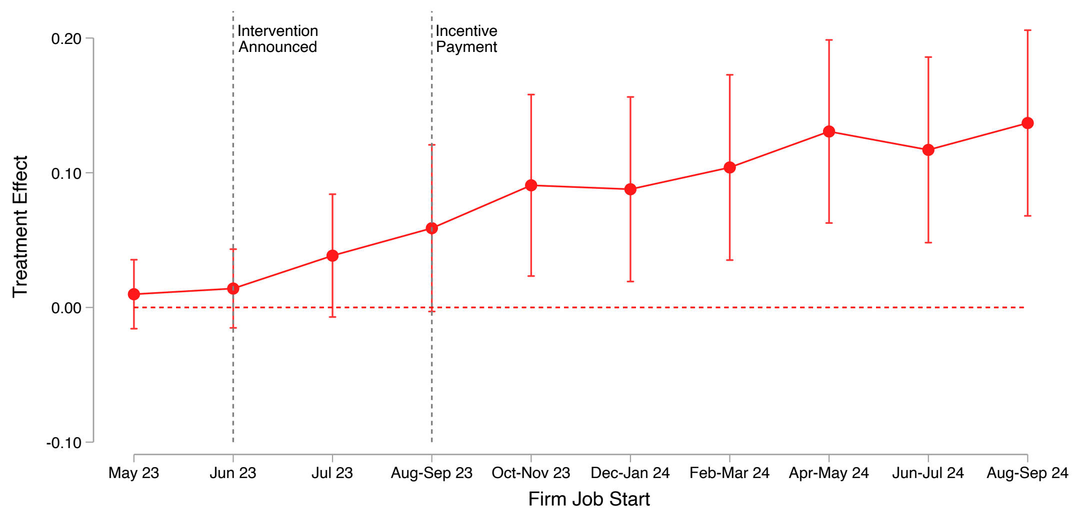
(a) Percent Working for Firm by First Application Date



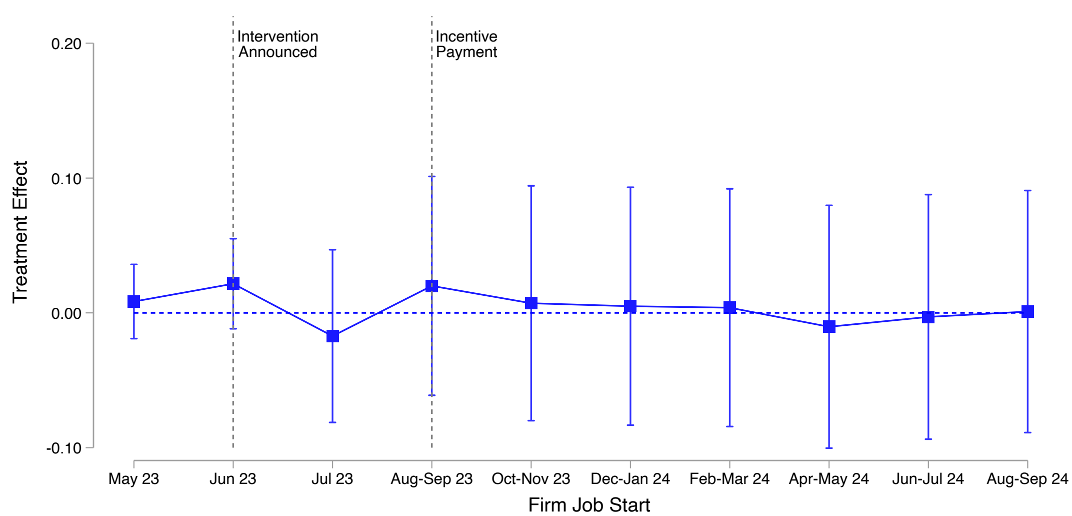
(b) Distribution of First Application Date

Notes: This figure presents the distribution of first application dates and the share of women working for a firm at the fourteen-month follow-up, separately by treatment status. The sample in Panel (a) consists of students in the experimental sample who were successfully surveyed at the six month follow-up (see Section 4.1 for details) and who provided the date of at least one job application. Panel (b) contains all female respondents at the 14 months follow-up conditional on being surveyed at the six month follow-up. Panel (a) shows the share of women in treatment and control who are working for a firm by the month of their first job application, by treatment status. For instance, 70.4% of the women in the control group who sent their first application in May or June 2023 are working for a firm in September 2024. Panel (b) presents a histogram of first application dates, by treatment status. For instance, 3.5% of women in the control group send their first application in May or June 2023. In Panel (a), treated women are represented by the solid, dark red line and control women are represented by the dashed, grey line. In Panel (b), treated women are represented by the darker red bars (left bar of each cluster), and control women are represented by the grey bars (right bar of each cluster).

Figure 6: Dynamic Treatment Effects on Firm Employment



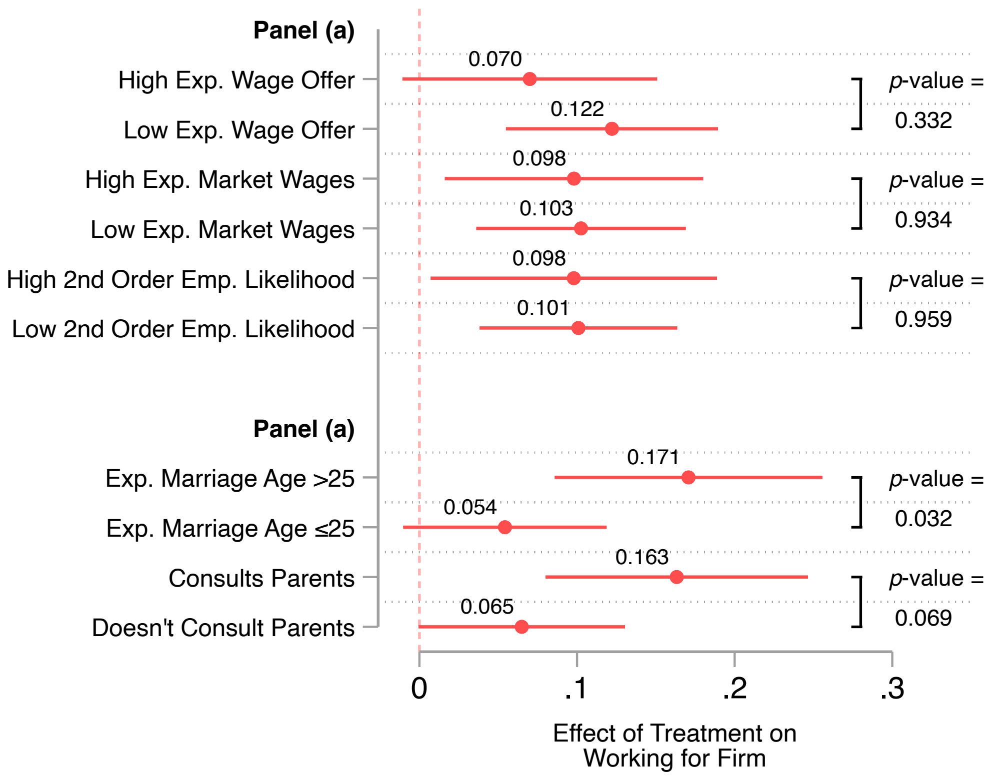
(a) Female: Treatment Effects on Firm Employment



(b) Male: Treatment Effects on Firm Employment

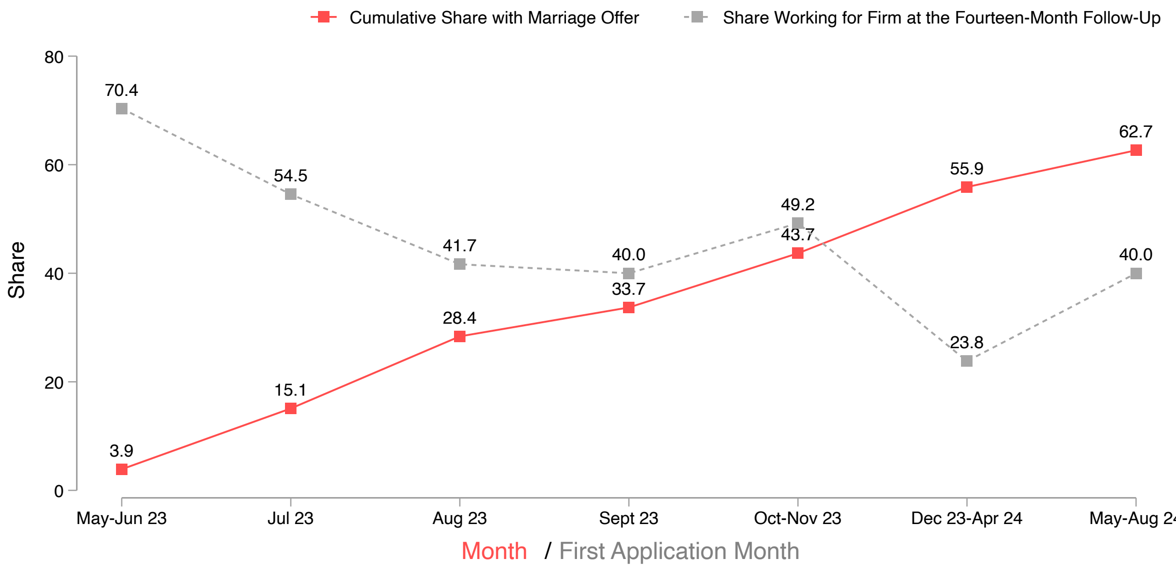
Notes: This figure presents the treatment effect on having ever worked (post-graduation) at a firm by a given date. Panel (a) shows treatment effects for women, and Panel (b) does the same for men. For instance, the last coefficient on the right in Panel (a) is the treatment effect on whether a female respondent has ever held a post-graduation firm job by the month of September 2024. The sample consists of students in the fourteen-months experimental sample, conditional on being successfully surveyed at the six month follow-up (see Section 4.1 for details). In both panels, the first dashed vertical line indicates when the intervention was announced, and the second dashed vertical line indicates the deadline by which treated students were required to submit proof of early applications. Both panels include a dashed horizontal line indicating where 0 lies on the y-axis. Vertical bars represent 95% confidence intervals.

Figure 7: Heterogeneous Treatment Effects on Firm Employment



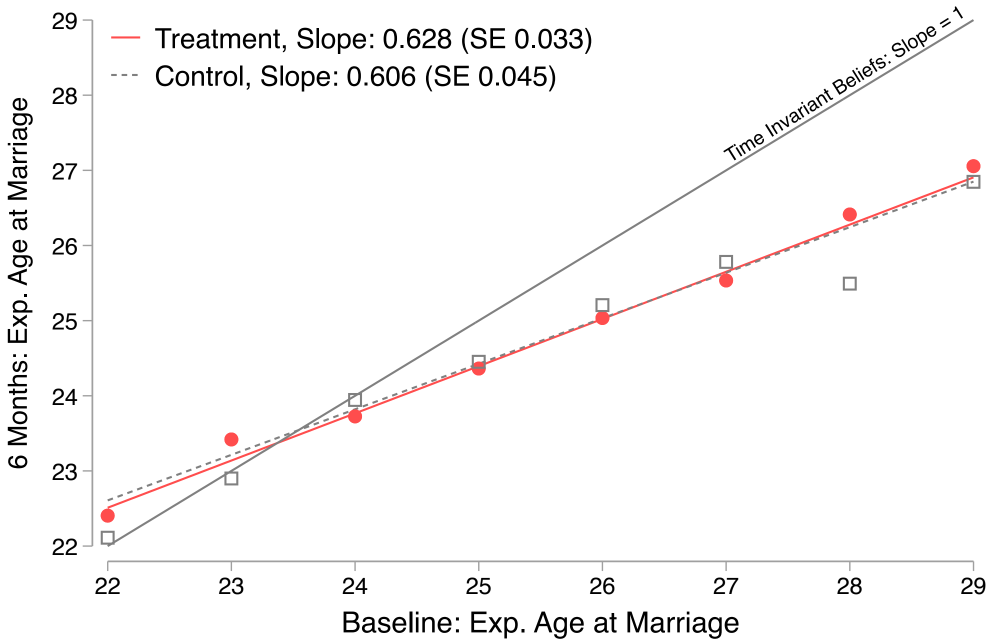
Notes: This figure presents heterogeneous treatment effects on women's likelihood of working for a firm, where firm employment excludes self-employment and unpaid work. Responses are pooled from the six- and fourteen-month experimental survey waves, with wave fixed effects, and standard errors are clustered at the individual level. Panel (a) examines heterogeneity based on women's baseline labor market beliefs, including their expected wage offer and second-order beliefs about peers' expected wages and employment likelihood. Each variable is classified as "High" (above the median) or "Low" (below the median) after winsorizing at the 2% level for unbounded continuous variables. Panel (b) explores heterogeneity based on familial involvement in job search, and marital expectations. Specifically, it includes whether women consult their parents during job search and whether they expect to marry after age 25 (the national median for college-educated women). Horizontal bars represent 95% confidence intervals, while the dashed vertical line marks zero on the x-axis. Vertical brackets indicate statistical significance tests between complementary groups, with the corresponding *p*-values reported.

Figure 8: The Relationship between the Job Market and the Marriage Market



Notes: This figure presents the relationship between marriage offers and employment rates of women over time. The sample consists of women in the control group of the fourteen-month experimental sample, conditional on being surveyed at the six month follow up (see Section 4.1 for details) excluding the small share of women who are already married or engaged at baseline for the line that plots the share with a marriage offer (darker red, solid line). The solid line presents the cumulative share of women with at least one post-graduation marriage offer by each (set of) month(s). For instance, 28.4% of women in our sample have received a marriage offer by August 2023, 33.7% by September 2023. The dashed line displays, as in Figure 5 Panel (a), shows the share of control women working for a firm at the fourteen-month follow-up based on whether their first application was sent in that (set of) month(s).

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



Notes: This figure presents the relationship between women's expected age of marriage at baseline and their updated expected age of marriage at six months, separately by treatment status. The sample consists of female students in the six-months experimental sample who were not already married or engaged by baseline (see Section 4.1 for details). Expected age at marriage is winsorized at the 2% level. Women in the treatment group are represented by the solid red line with square markers. Women in the control group are represented by the dashed gray line with circular markers. The solid gray 45-degree line represents time-invariant beliefs, indicating that the expected age of marriage is unchanged between baseline and the six-month follow-up. A slope less than one indicates that women who initially expected to marry later tend to revise their expectations downward over time.

Tables

Table 1: Descriptive Statistics for the Diagnostic Sample

	All	Male	Female	Diff.	P-value
	(1)	(2)	(3)	(4)	(5)
Nb. Obs.	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
Life Sciences / Pharmacy	12.1	5.1	21.6	-16.6	0.00
Sciences	13.2	5.8	23.2	-17.5	0.00
Humanities / Languages / Education	15.5	13.4	18.5	-5.1	0.03
Social Sciences	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 baseline descriptive statistics for respondents in the diagnostic sample (see Section 2.1 for details). Column 1 presents statistics for all respondents. Columns 2 and 3 provide statistics for male and female respondents, respectively. Column 4 shows the difference between genders (calculated as the male - female difference), and Column 5 provides the p-value from a test for equality of means between genders. The first row reports the number of observations in the sample. The rows for Age and GPA report mean values. The remaining rows report the percentage of students in each category. For instance, the share of all students that are married is 4.3%, and the share of male students in an engineering major is 39.2%. Age is winsorized at the 2% level.

Table 2: Descriptive Statistics for the Experimental Sample

	All	Male	Female	Diff.	P-value
	(1)	(2)	(3)	(4)	(5)
Nb. Obs.	1,442	516	926		
Age	22.7	23.2	22.5	0.7	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
Life Sciences / Pharmacy	12.6	10.1	14.0	-4.0	0.02
Sciences	27.0	29.7	25.6	4.1	0.10
Humanities / Languages / Education	26.3	27.9	25.4	2.5	0.30
Social Sciences	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 baseline descriptive statistics for respondents in the six-month experimental sample (see Section 4.1 for details). Column 1 presents statistics for all respondents. Columns 2 and 3 provide statistics for male and female respondents, respectively. Column 4 shows the difference between genders (calculated as the male - female difference), and Column 5 provides the p-value from a test for equality of means between genders. The first row reports the number of observations in the sample. The rows for Age and GPA report mean values. The remaining rows report the share of students in each category. For instance, the share of all students that are married is 4.4%, and the share of male students in an engineering major is 6.4%. Age is winsorized at the 2% level.

Table 3: Treatment Effects on Employment

	August 15	6 Months		14 Months	
	Has Applied	**Working**	**Working for Firm**	**Working**	**Working for Firm**
	(1)	(2)	(3)	(4)	(5)
Panel A: Female					
Treatment	0.255***	0.075**	0.102***	0.069**	0.095***
	(0.031)	(0.031)	(0.029)	(0.035)	(0.034)
Female Control Mean	0.316	0.336	0.253	0.518	0.416
Panel B: Male					
Treatment	0.177***	0.001	0.012	-0.019	-0.005
	(0.042)	(0.042)	(0.042)	(0.040)	(0.045)
Male Control Mean	0.370	0.551	0.378	0.719	0.561
Nb. Obs.	1,442	1,442	1,442	1,218	1,218

Notes: This table presents the treatment effects on labor market outcomes for students in the six- and fourteen-month experimental sample (see Section 4.1 for details). Students must be successfully surveyed at the six month wave for inclusion in the analysis. Panel A shows results for women, and Panel B shows results for men. Column 1 reports the treatment effect on a respondent's likelihood of applying to at least one job by August 15th. Column 2 reports the treatment effect on employment at six months. Column 3 reports the treatment effect on firm employment at six months. Columns 4 and 5 report the same outcomes as Column 2 and 3, respectively, but at fourteen months. Coefficients in the panels are estimated together in a single regression (pooling men and women) and all regressions control for the variables selected following the post-double-selection LASSO procedure. Table B.4 shows the results without the LASSO-selected controls. Control means are provided separately for each gender. The last row reports the number of observations. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: OLS and IV Estimates of Early Applications on Employment

	6 Months		14 Months	
	OLS Working for Firm	**2SLS Working for Firm**	**OLS Working for Firm**	**2SLS Working for Firm**
	(1)	(2)	(3)	(4)
Panel A: Female				
Applied Early (by Aug. 15th)	0.236***	0.383***	0.195***	0.308***
	(0.030)	(0.112)	(0.035)	(0.110)
Female, Not Applied Early, Mean	0.188	0.188	0.349	0.349
Panel B: Male				
Applied Early (by Aug. 15th)	0.225***	0.057	0.127***	-0.032
	(0.042)	(0.237)	(0.045)	(0.213)
Male, Not Applied Early, Mean	0.269	0.269	0.475	0.475
Nb. Observations	1,442	1,442	1,218	1,218

Notes: This table presents OLS and 2SLS treatment effects on early job application and firm employment. The sample consists of students in the six- and fourteen-month experimental sample (see Section 4.1 for details). Students must be successfully surveyed at the six month wave for inclusion in the analysis. Panel A shows results for women, and Panel B shows results for men. Column 1 reports the OLS estimates from regressing employment at six months on whether a respondent applied early. Column 2 uses the exogenous treatment as an instrumental variable and reports the 2SLS estimates of the effect of applying early on employment at six months. Columns 3 and 4 report the same outcomes as Columns 1 and 2, respectively, but at fourteen months. Coefficients in the panels are estimated together in a single regression (pooling men and women) and all regressions control for the variables selected following the post-double-selection LASSO procedure. Control means are provided separately for each gender. The last row reports the number of observations.

Table 5: Treatment Effects on Labor Market Intentions

	Intended Nb. Apps by Aug. 15	Intended Search Start (Days)	Emp. Likelihood in 6 Mo.
	(1)	(2)	(3)
Panel A: Female			
Treatment	0.593*	-20.440***	-0.885
	(0.324)	(2.658)	(1.031)
Female Control Mean	5.586	50.280	79.802
Panel B: Male			
Treatment	-0.580	-12.648***	0.743
	(0.559)	(3.632)	(1.271)
Male Control Mean	7.366	43.748	81.618
Nb. Obs.	1,417	1,417	1,442

Notes: This table presents treatment effects on self-reported job search intentions and employment expectations from the baseline survey. Responses were collected after the treatment group was informed about their eligibility for a financial incentive for submitting four early applications. The sample consists of students in the experimental sample, conditional on being successfully surveyed in the six month survey (see Section 4.1 for details). Panel A shows results for women, while Panel B shows results for men. Column 1 reports the treatment effect on the intended number of job applications the respondent intends to send by August 15th. Column 2 reports the effect on the intended job search start date, recorded as the number of days from the survey date until the student plans to begin applying. Column 3 reports the treatment effect on the respondent's self-reported likelihood of being employed in six months. Coefficients in the panels are estimated together in a single regression (pooling men and women) and all regressions control for the variables selected following the post-double-selection LASSO procedure. Control means are provided separately for each gender. The last row reports the number of observations. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Treatment Effects on Applications and Offers

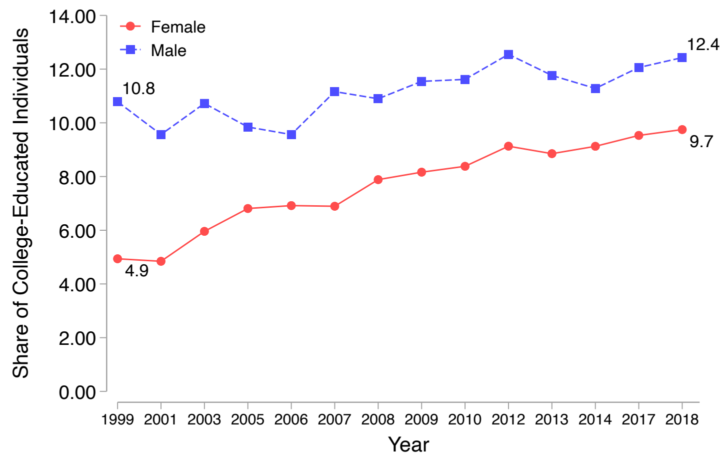
	6 Months		14 Months	
	Nb. Apps	**Nb. Offers**	**Nb. Apps**	**Nb. Offers**
	(1)	(2)	(3)	(4)
Panel A: Female				
Treatment	1.445**	0.155	1.710**	-0.000
	(0.736)	(0.163)	(0.713)	(0.097)
Female Control Mean	8.184	2.047	4.726	0.899
Panel B: Male				
Treatment	0.486	0.136	-0.387	-0.012
	(1.050)	(0.202)	(0.969)	(0.120)
Male Control Mean	9.565	1.941	6.516	0.959
Nb. Obs.	1,435	1,435	1,210	1,210

Notes: This table presents the treatment effects on intermediate labor market outcomes for students in the experimental sample, conditional on being surveyed in the six month wave (see Section 4.1 for details). Panel A shows results for women, and Panel B shows results for men. Column 1 reports the number of applications sent between graduation and the six-month follow up. Column 2 reports the number of job offers received between graduation and the six-month follow-up. Column 3 reports the number of applications sent in the six months prior to the fourteen-month follow-up. Column 4 reports the number of job offers received in the in the six months prior to the fourteen-month follow-up. Number of applications and number of offers are winsorized at the 2% level. Coefficients in the panels are estimated together in a single regression (pooling men and women) and all regressions control for the variables selected following the post-double-selection LASSO procedure. Control means are provided separately for each gender. The last row reports the number of observations. There are fewer observations in this table relative to Table 3 because the analysis is limited to cases with non-missing values for number of applications and offers. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

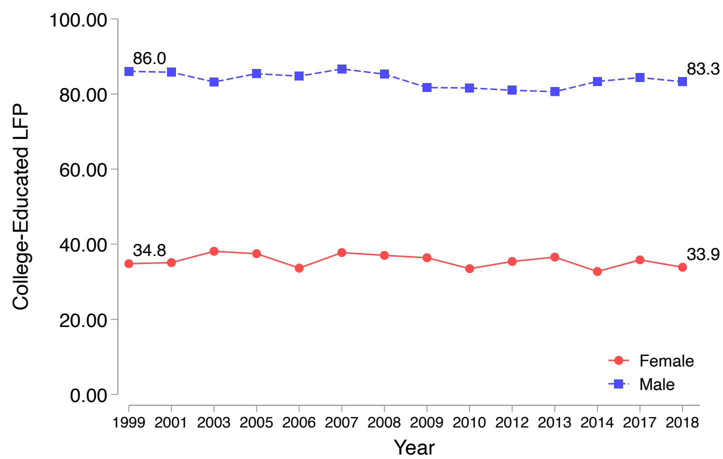
Appendix

A Appendix Figures

Figure A.1: College Education and Labor Force Participation in Pakistan



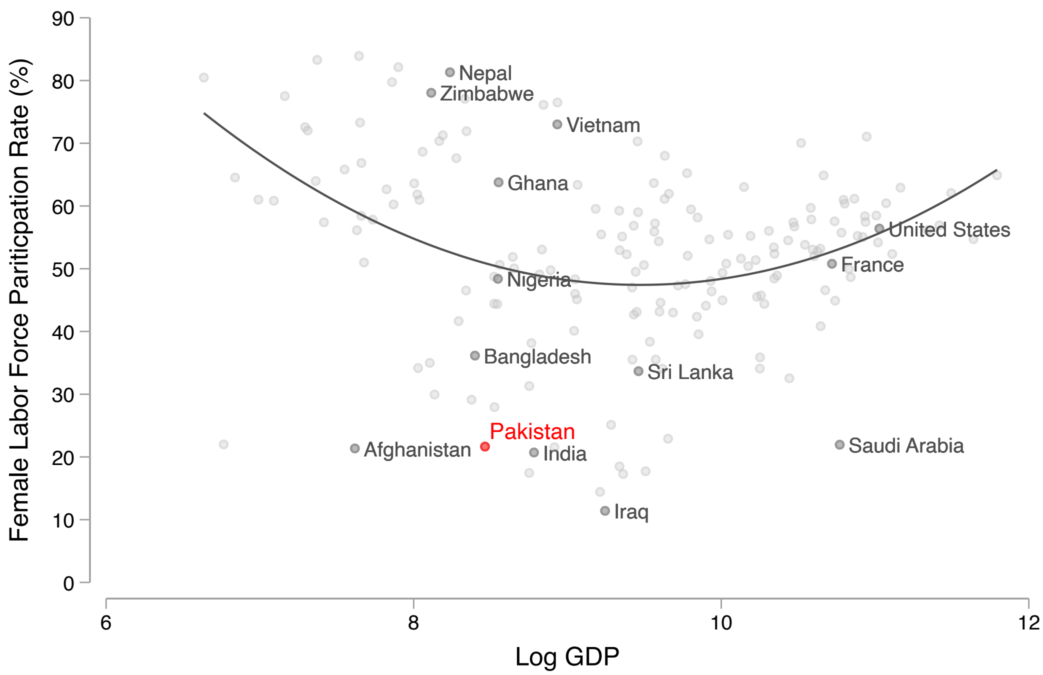
(a) Gender Gaps in College Education



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

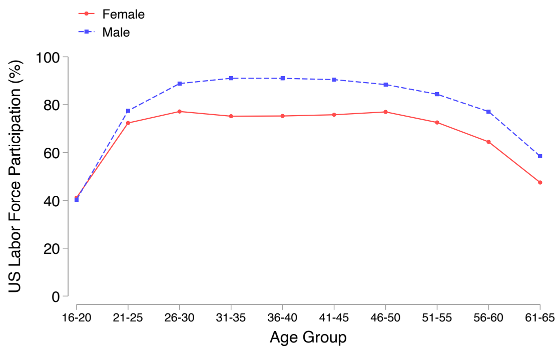
Notes: This figure presents trends in college education and labor force participation in Pakistan between 1999 and 2018. Panel (a) shows the share of individuals aged 22–35 who are college-educated, separately by gender. Panel (b) shows the labor force participation rates for college-educated individuals in the same age group, separately by gender. In both panels, men are represented by a dashed blue line with square markers, and women are represented by a solid red line with circular markers. Data for both panels is obtained from the Pakistan Labor Force Surveys.

Figure A.2: Female Labor Force Participation vs. (Log) GDP per Capita by Country (2018)

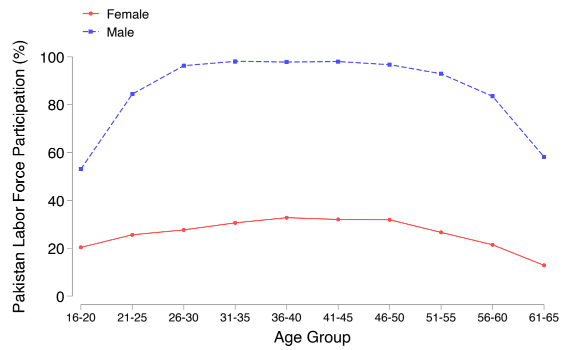


Notes: This figure compares log GDP per capita and female labor force participation across several countries, with Pakistan highlighted in red. The U-shaped relationship between GDP and female labor force participation is represented by a solid black line. Countries that lie directly on the U-shaped line are shown in gray (for example, Nigeria and France). Countries with similar levels of GDP per capita to Pakistan (for example, Nepal and Ghana) or similar levels of female labor force participation (for example, Saudi Arabia and India) are also shown in gray. Data are obtained from the World Bank (2018).

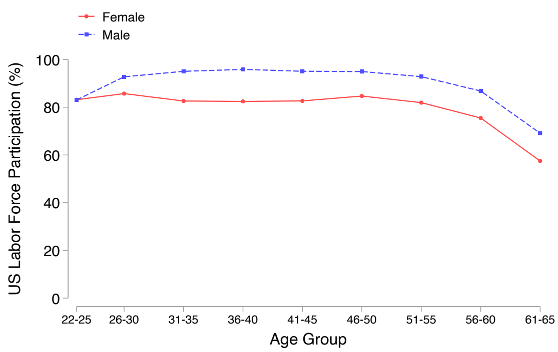
Figure A.3: Labor Force Participation by Age and Gender, US vs. Pakistan (2018)



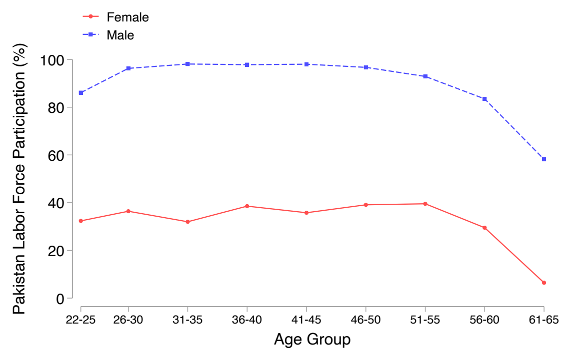
(a) U.S. Overall



(b) Pakistan Overall



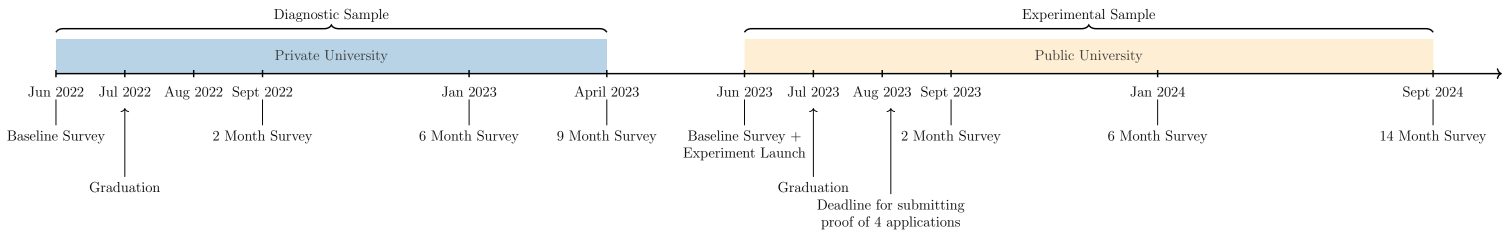
(c) U.S. College Graduates



(d) Pakistan College Graduates

Notes: This figure compares age, education, and labor force participation in the United States and Pakistan, separately by gender. Panels (a) and (b) show labor force participation for men and women aged 16 to 65 in the United States and Pakistan, respectively. Panels (c) and (d) show labor force participation for college-educated men and women aged 22 to 65 in the United States and Pakistan, respectively. In all panels, men are represented by a dashed blue line with circular markers and women are represented by a solid red line with circular markers. Data for the United States is obtained from the Current Population Survey (2018) and data for Pakistan is obtained from the Pakistan Labor Force Surveys (2018).

Figure A.4: Research Timeline

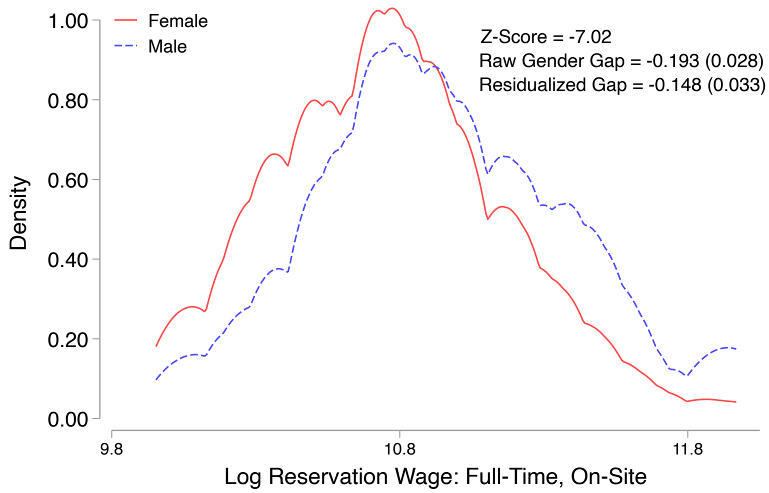


Notes: This figure presents the timing of research surveys relative to students' graduation timeline. The first surveys (represented in blue on the timeline) were conducted with the diagnostic sample described in Section 2.1 at the private university. The baseline survey was conducted in June 2022, one month prior to the end of the academic term. This is noted on the timeline as "graduation", although convocation ceremonies were scheduled at different times by different departments. This cohort of students was followed-up with two, six, and nine months later, in September 2022, January 2023, and April 2023, respectively. Insights from these follow-ups informed the intervention implemented one year later in the experimental sample described in Section 4.1 at the public university. These surveys are represented in yellow on the timeline. 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. This is noted again as "graduation" on the timeline. The deadline given to the treatment group to show proof of 4 applications was August 15, 2023. This cohort of students was followed-up with two, six, and fourteen months later, in September 2023, January 2024, and September 2024, respectively.

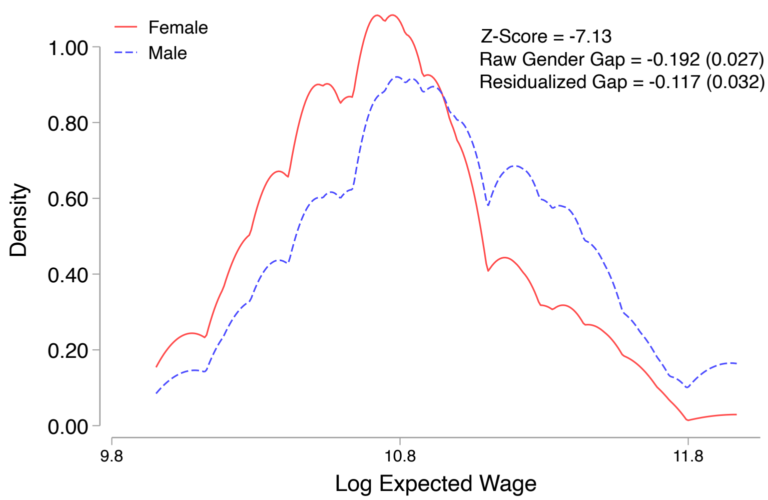
Figure A.5: Survey Incentives

Notes: This figure presents a photograph taken on June 9, 2022 at the private university in Lahore, Pakistan. It shows the setup of one of our food stands during baseline data collection. All students who completed the survey were given vouchers to redeem KFC meals and a bakery item from the food stand.

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



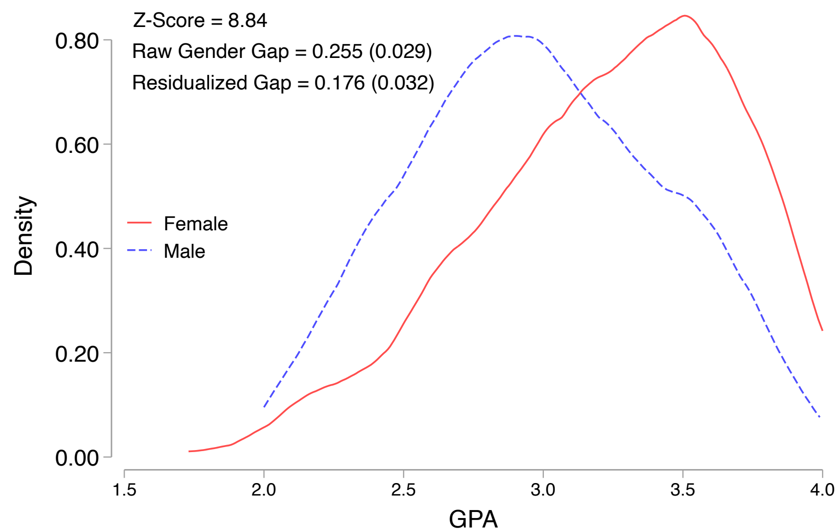
(a) Kernel Density: Log Reservation Wage



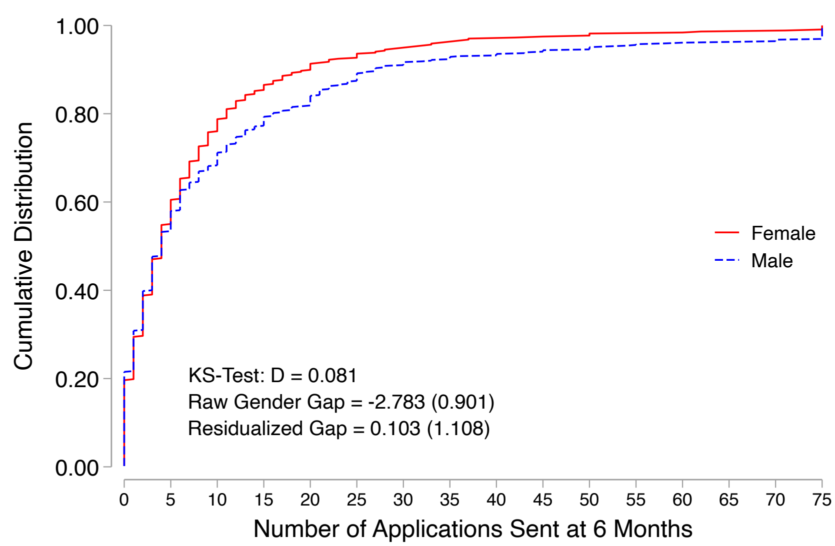
(b) Kernel Density: Log Expected Wage

Notes: This figure presents the distribution of baseline reservation and expected wages, separately by gender. The sample consists of students in the diagnostic sample (see Section 2.1 for details). Panel (a) shows the kernel density of log reservation wages for a full-time, on-site job. Panel (b) shows the kernel density of log expected wages for a full-time, on-site job for the respondent's preferred job title. Reservation and expected wages are winsorized at the 2% level. In both panels, women are represented by a solid red line, and men are represented by a dashed blue line. Both panels show raw and residualized gender gaps, calculated as the female - male difference, and the z-score for the difference between women and men. The residualized estimate controls for cumulative GPA, major, and industry of search.

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



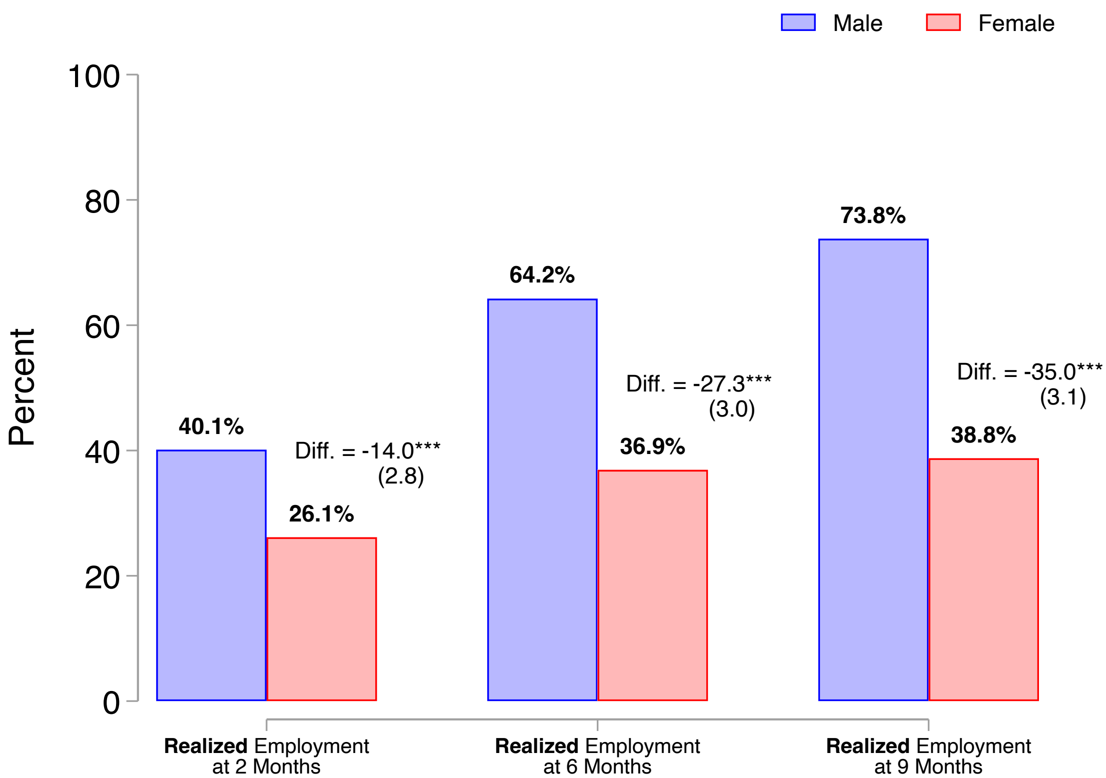
(a) Kernel Density: GPA



(b) CDF: Job Applications at Six Months

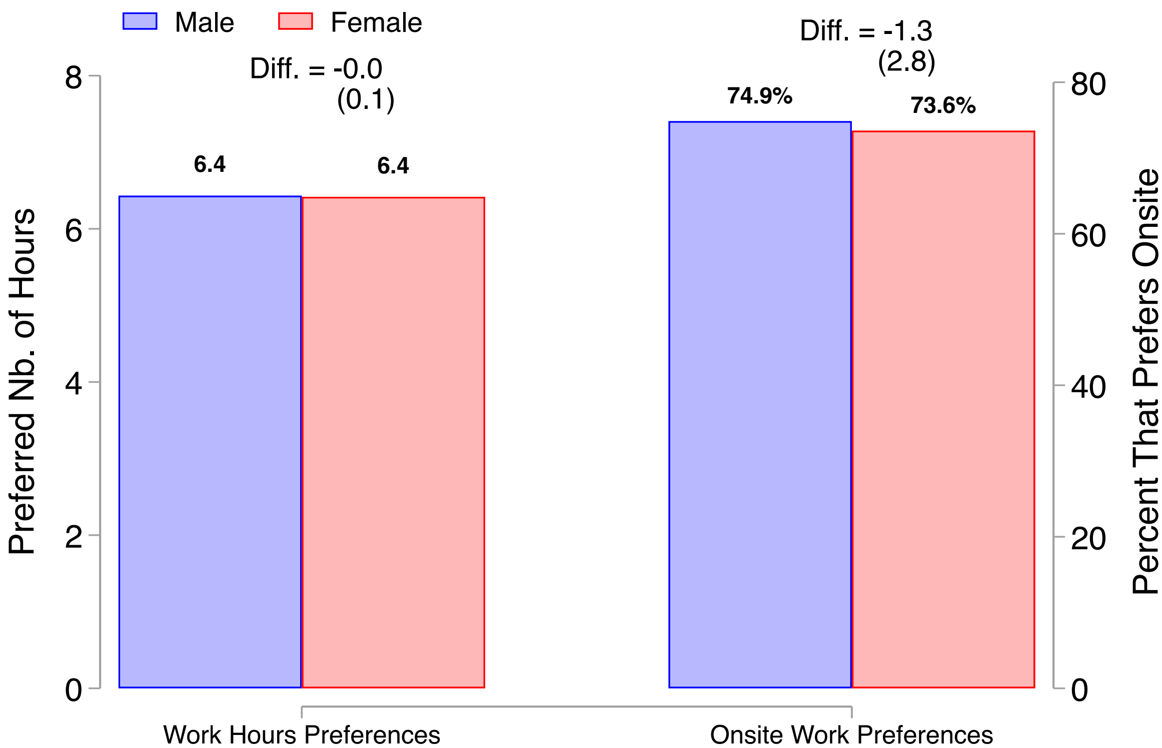
Notes: This figure presents the distribution of baseline GPA and number of job applications sent by the six-month follow-up, separately by gender. The sample consists of students in the diagnostic sample (see Section 2.1 for details). Panel (a) shows the kernel density of cumulative GPA. Panel (b) shows the cumulative distribution of the number of job applications sent by the six-month follow-up. Number of job applications is winsorized at the 2% level. In both panels, women are represented by a solid red line, and men are represented by a dashed blue line. Both panels show raw and residualized gender gaps, calculated as the female - male difference, and the z-score for the difference between women and men. The residualized estimate in Panel (a) controls for major. The residualized estimate in Panel (b) controls for cumulative GPA, major, industry of search, preference for onsite vs. remote work, number of preferred daily work hours, and internship experience.

Figure A.8: Employment Rates By Gender Across Survey Waves



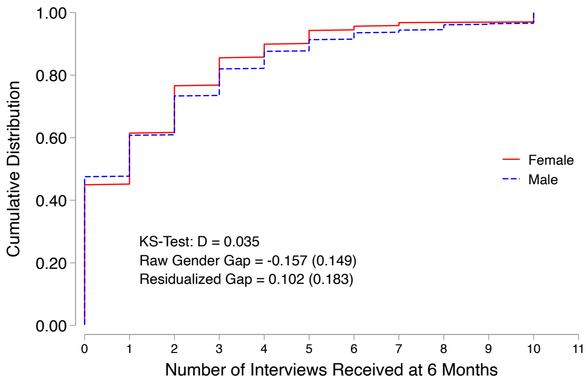
Notes: This figure presents employment rates, separately by gender and survey wave. The sample consists of students in the diagnostic sample (see Section 2.1 for details). Employment is shown for the two, six, and nine-month follow-up surveys by the left, middle, and right clusters of bars, respectively. Men are represented in blue (bars 1, 3, and 5) and women are represented in red (bars 2, 4, and 6). The average difference between male and female employment is shown directly above the female bar for each follow-up survey. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.9: Supply-Side Factors III: Preferred Work Arrangements

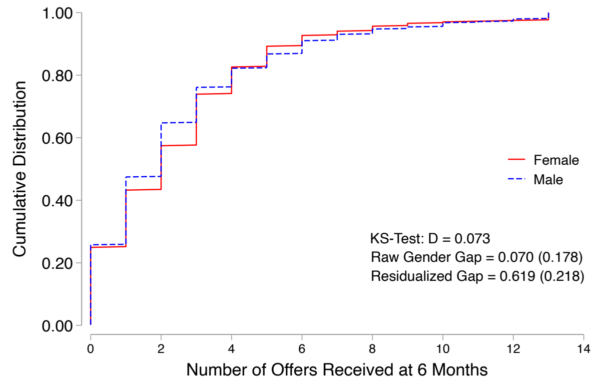


Notes: The figure presents information on baseline preferred work arrangements, separately by gender. The sample consists of students in the diagnostic sample (see Section 2.1 for details). The left cluster of two bars shows the average number of preferred work hours. The right cluster of two bars shows the percentage of respondents that prefer to work onsite rather than work remotely. Men are represented in blue (bars 1 and 3) and women are represented in red (bars 2 and 4). The average difference between male and female responses is shown directly above the female bar. The variable for preferred hours of work in a day is winsorized at the 2% level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

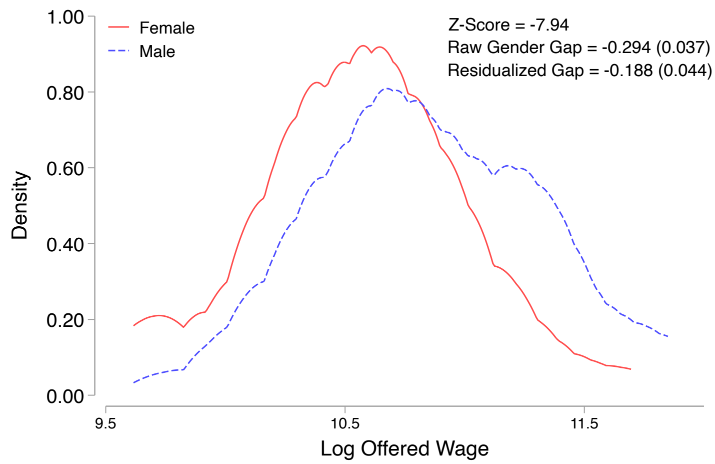
Figure A.10: Demand-Side Factors: Interviews, Offers and Offered Wages



(a) CDF: Job Interviews at Six Months



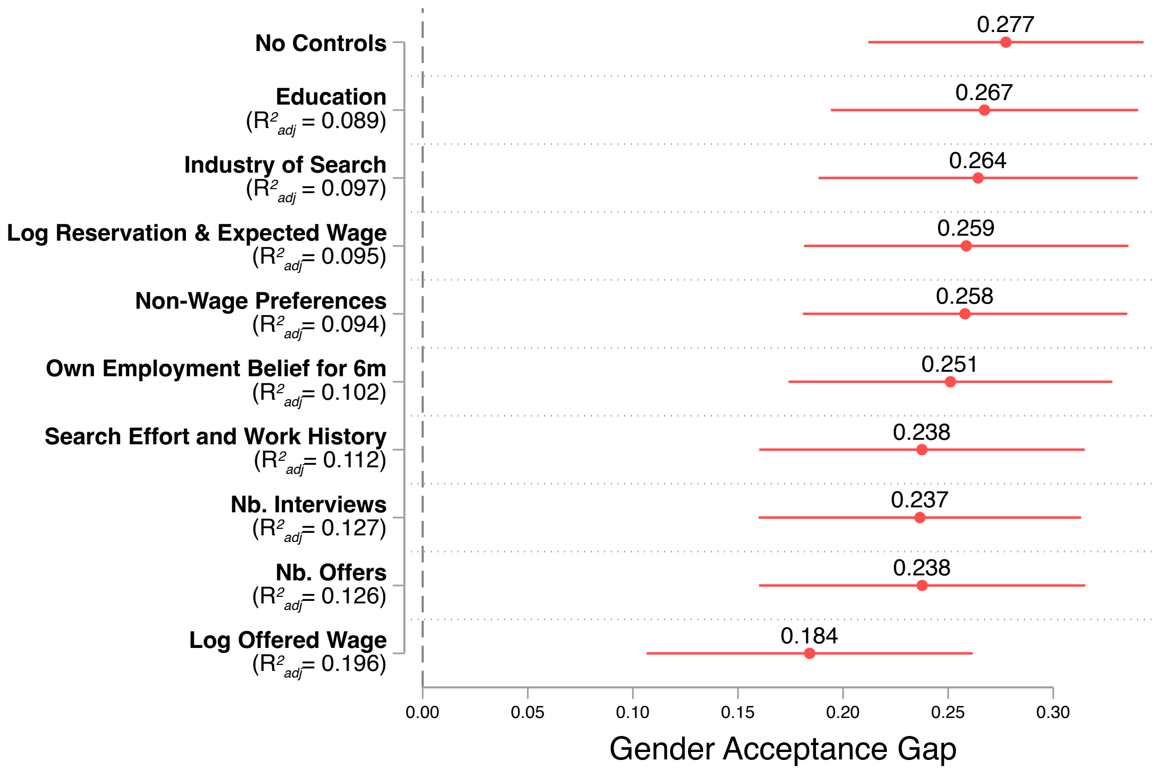
(b) CDF: Job Offers at Six Months



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

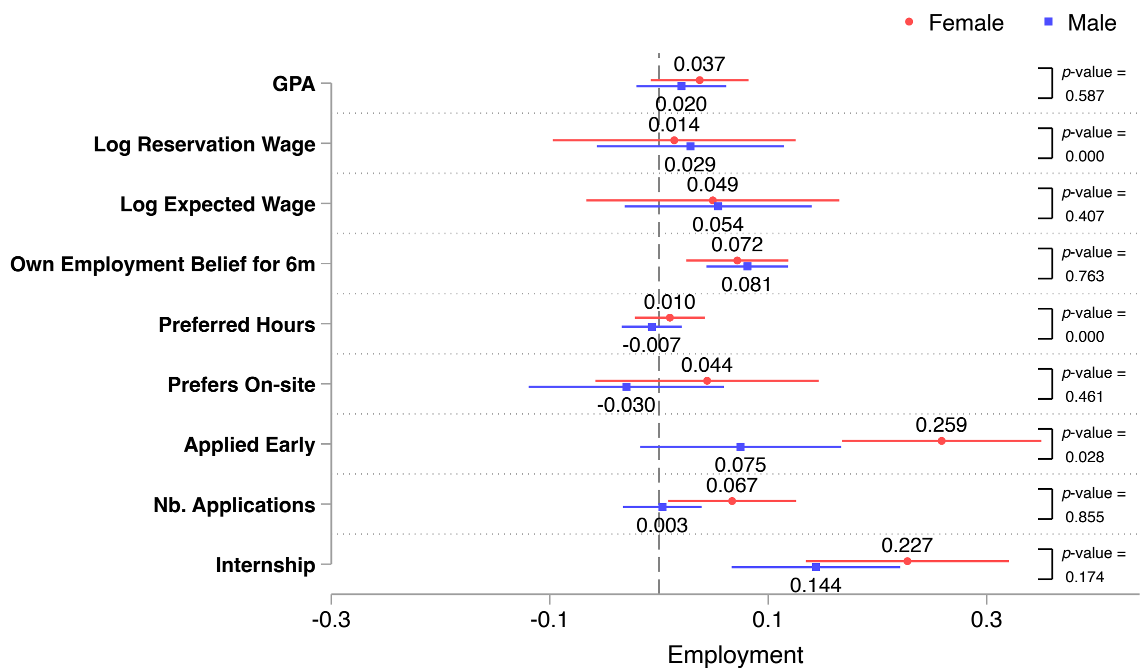
Notes: The figure presents the distribution of job interviews, job offers, and offered wages received by the six-month follow-up, separately by gender. The sample consists of students in the diagnostic sample (see Section 2.1 for details). Panel (a) shows the cumulative distribution of the number of job interviews received by the six-month follow-up. Panel (b) shows the cumulative distribution of the number of job offers received by the six-month follow-up. Panel (c) shows the kernel density of log offered wage at six months for respondents who reported at least one job offer. Offered wage represents the highest wage offer a student has received for a job (regardless of whether they have accepted it). Number of interviews, number of offers, and offered wage are winsorized at the 2% level. In all three panels, women are represented by a solid red line, and men are represented by a dashed blue line. All three panels show raw and residualized gender gaps, calculated as the female - male difference. Panels (a) and (b) show the KS-test for differences in the raw distributions of female and male interviews and offers, respectively. Panel (c) reports the z-score for the difference in mean log offered wages between women and men. The residualized estimate in Panels (a) and (b) controls for cumulative GPA, major, industry of search, preference for onsite vs. remote work, number of preferred daily work hours, internship experience, reservation wage, and expected wage. The residualized estimate in Panel (c) controls for cumulative GPA, major, and industry of search.

Figure A.11: Explaining the Gender Acceptance Gap Six Months After Graduation



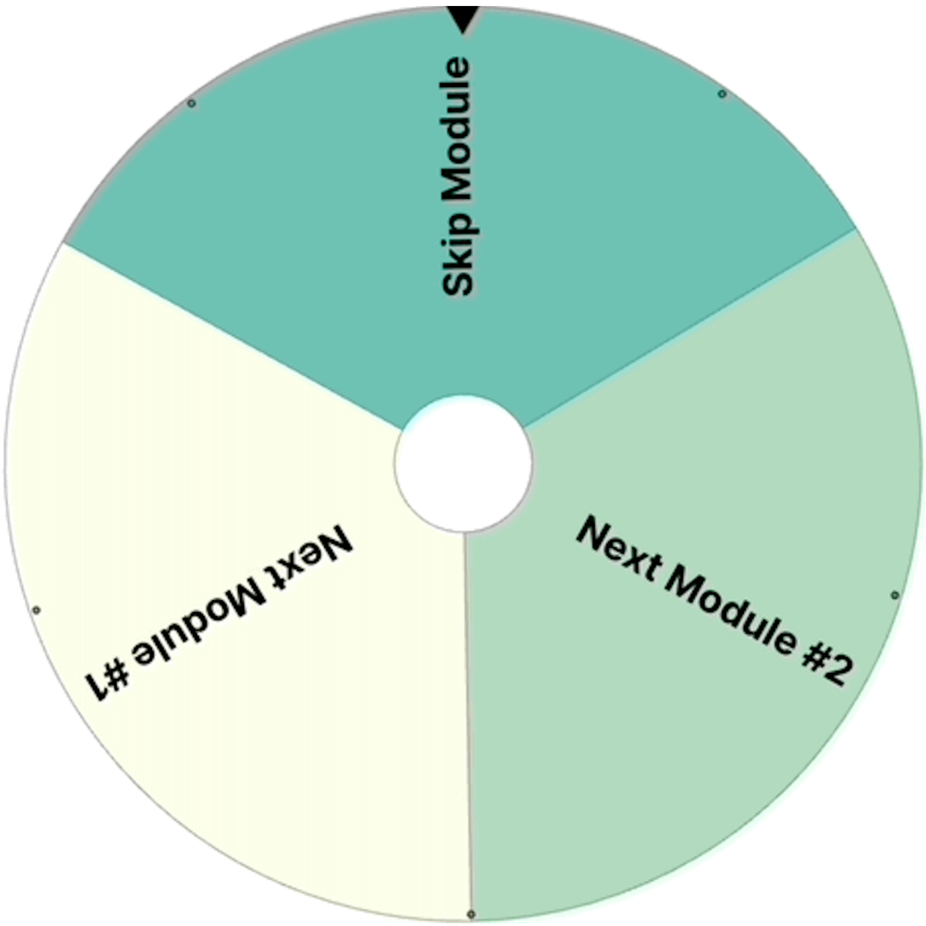
Notes: This figure presents the difference between male and female likelihood of having accepted a job offer six months post-graduation, with controls for observable characteristics added incrementally. The sample consists of students in the diagnostic sample (see Section 2.1 for details). Each row reports the gap between male and female job offer acceptance after controlling for the variable listed in that row and all rows above. The Education controls include cumulative GPA and major fixed effects. The Industry of Search controls incorporate fixed effects derived from the semantic classification of a respondent's preferred job title into Standard Occupational Classification (SOC) codes. The Reservation and Expected Wage controls include respondents' wage expectations at baseline. The Non-Wage Preferences controls include preferences regarding onsite vs. remote work and the number of preferred daily work hours. The Own Employment Belief for 6 Months control includes baseline belief about one's probability of employment. The Search Effort and Work History controls include the total number of job applications sent by the six-month follow-up, an indicator for applying early, and an indicator for internship experience. The Nb. Interviews control includes the number of interviews received by the six-month follow-up. The Nb. Offers control includes the number of job offers received by the six-month follow-up. The Offered Wage control includes the highest wage offer a student has received for a job (regardless of whether they have accepted it). All unbounded continuous variables are winsorized at the 2% level. Each coefficient (except for the first row) is from a multivariate regression that includes all preceding rows' variables as controls. The first row presents the coefficient from a bivariate regression of an indicator for accepting a job offer on the female indicator. Adjusted R-squared values are shown for each row. The horizontal bars show 95% confidence intervals. The mean acceptance rate for men is 73.6%.

Figure A.12: Timing Distinctively Predicts Women's Employment (Bivariate Estimates)



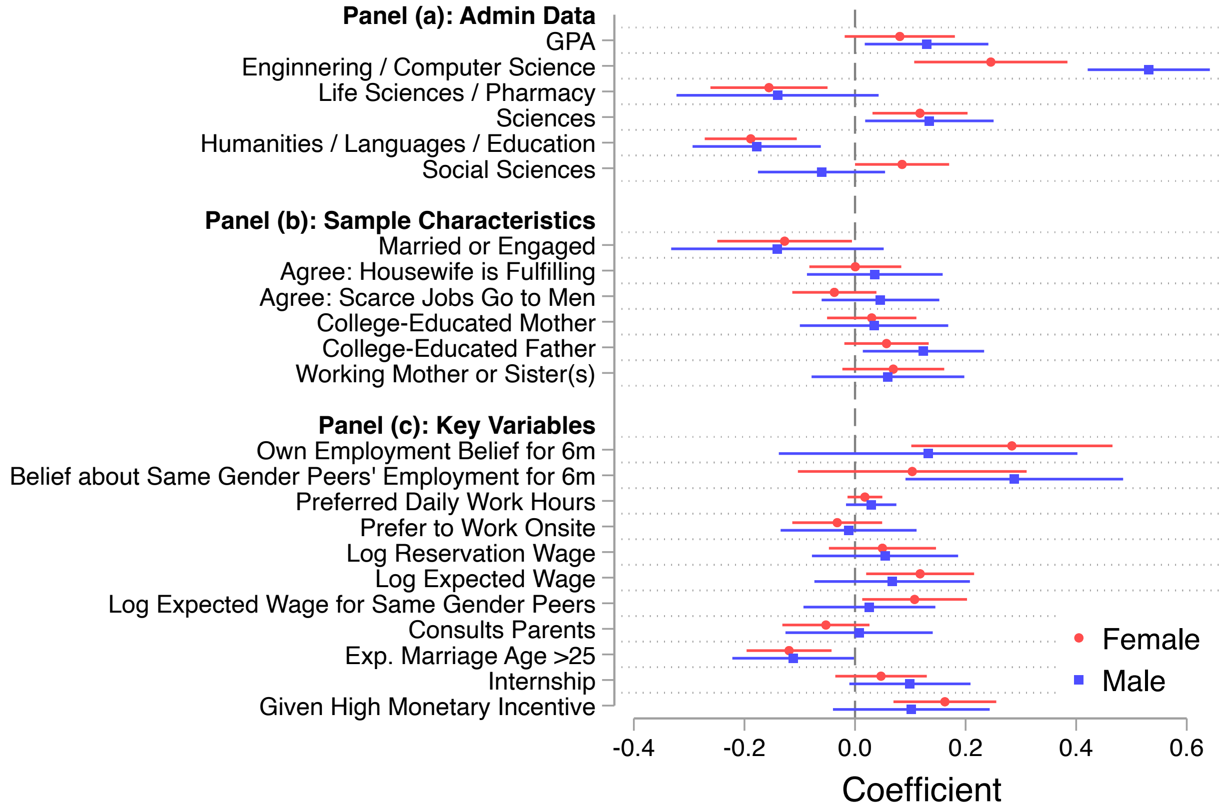
Notes: This figure presents the probability of employment six months post-graduation as explained by a set of independent variables, separately by gender. The sample consists of students in the diagnostic sample (see Section 2.1 for details). The predictors include cumulative GPA (measured in standard deviations), baseline wage expectations, baseline belief about the probability of employment (on a 0-1 scale, measured in standard deviations), preferred daily work hours, preference for onsite vs. remote work, an indicator for applying early, the total number of job applications submitted by the six-month follow-up (measured in standard deviations), and an indicator for having internship experience. All unbounded continuous variables are winsorized at the 2% level. Each coefficient is from a bivariate regression conducted separately for men and women. Figure 3 presents the results of a multivariate regression for each predictor of employment, rather than bivariate regressions. The horizontal bars show 95% confidence intervals. Corresponding means and standard deviations for the independent variables shown in the figure are provided in Table B.10.

Figure A.13: Wheel Shown to Students in the Experimental Sample



Notes: This figure presents the wheel shown to students during the baseline survey in the experimental sample described in Section 4.1 at the public university. The wheel was spun on a tablet, with both the student and the enumerator able to observe the outcome. The wheel randomly assigned students to one of two treatment groups or the control group. Students in the control group were informed that the survey would proceed to the final module, while those in the treatment groups were informed about the incentives for early application.

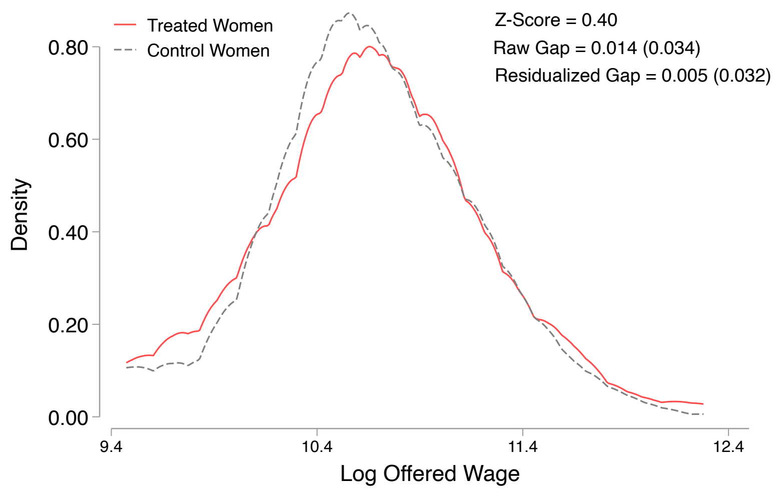
Figure A.14: Characteristics of the Treated



Notes: This figure compares observable characteristics of students who were offered and took up treatment and those who were offered but did not take up treatment, separately by gender. The sample consists of students in the six-months experimental sample (see Section 4.1 for details). Panel (a) shows administrative data, such as GPA and college major. Panel (b) shows survey responses about demographic background, such as marital status, agreement with two statements that reflect gender values, parental education, and female family members' work status. Panel (c) shows survey responses which may determine of labor force participation, such as employment beliefs about oneself and others, work arrangement preferences, wage expectations, and whether the student was offered the high-monetary incentive. All unbounded continuous variables are winsorized at the 2% level. Each coefficient is from a bivariate regression conducted separately for men and women. The horizontal bars show 95% confidence intervals.

Figure A.15: Wage Distribution for Treated and Control Women

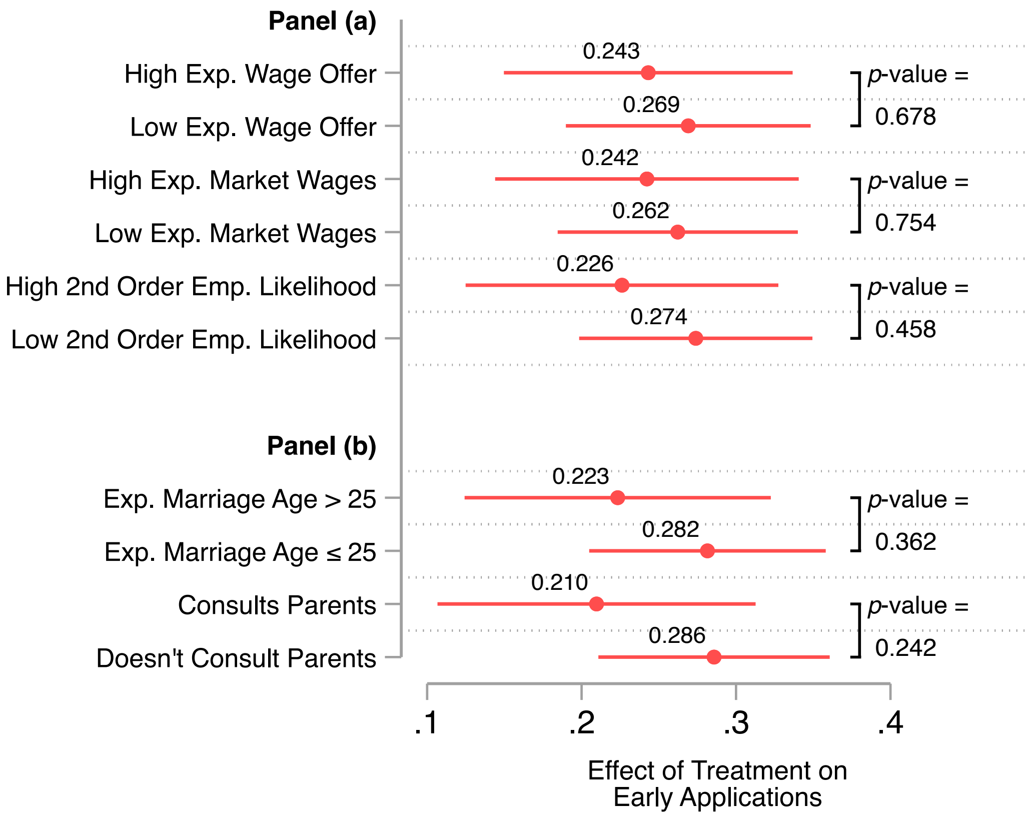
(a) Log Current Wage by Treatment



(b) Log Offered Wage by Treatment

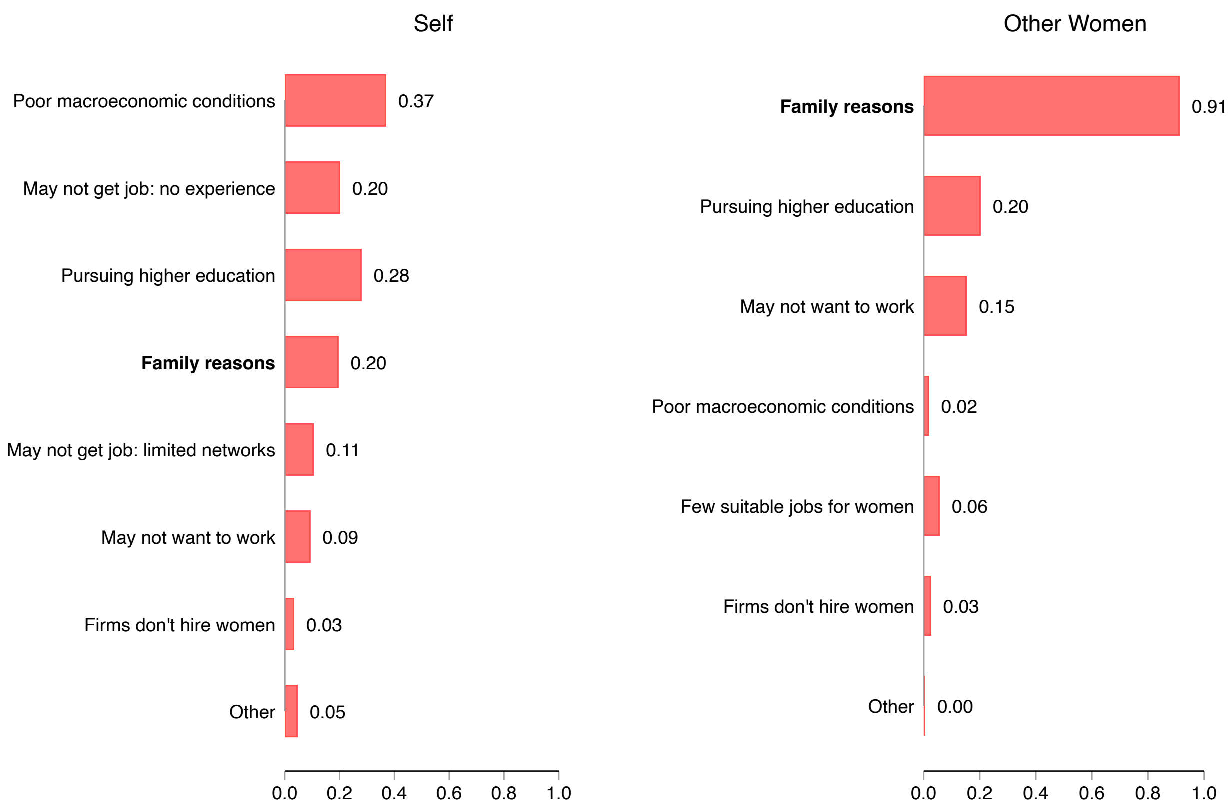
Notes: This figure presents the distribution of current and offered wages for women at the six and fourteen-month follow-ups, separately by treatment status. The sample consists of students in the experimental sample (see Section 4.1 for details) who provided a current or offered wage, respectively. Panel (a) shows kernel density of log current wage and Panel (b) shows kernel density of log offered wage. Current wage and offered wage are winsorized at the 2% level. In both panels, treated women are represented by a solid red line, and control women are represented by a dashed gray line. Both panels show raw and residualized gaps, calculated as the treatment - control difference, and the z-score for the difference between treated women and control women. The residualized estimate controls for cumulative GPA, major, and industry of search.

Figure A.16: Heterogeneous Treatment Effects on Early Application



Notes: This figure presents heterogeneous treatment effects on women's likelihood of applying to jobs early, defined as submitting at least one application by August 15—the deadline for treated students to qualify for the financial reward. Responses are pooled from the six- and fourteen-month experimental survey waves, with wave fixed effects, and standard errors are clustered at the individual level. Panel (a) shows heterogeneity based on women's baseline labor market beliefs, that is their expected wage offer and second-order beliefs about peers' expected wages and employment likelihood. Each variable is classified as "High" (above the median) or "Low" (below the median) after winsorizing at the 2% level for unbounded continuous variables. Panel (b) shows heterogeneity based on familial involvement in job search, and marital expectations. Specifically, it includes whether women consult their parents during job search and whether they expect to marry after age 25 (the national median for college-educated women). Horizontal bars represent 95% confidence intervals, while the dashed vertical line marks zero on the x-axis. Vertical brackets indicate statistical significance tests between complementary groups, with the corresponding *p*-values reported.

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



Notes: This figure presents women's beliefs at baseline about why they and other women in their class may not work in the future. The sample consists of all female students in the control group at the public university who were successfully interviewed in both the baseline and six month follow-up (see Section 4.1 for details). The left panel ("Self") reports responses to the question: "Why do you think there is a chance you may not work?" This question was asked to all women who reported less than a 100% likelihood of being employed six months after graduation (75.6% of respondents). The right panel ("Other women") presents responses to the question: "Out of the remaining XX women, why do you think they are not working?" This question was asked to all women who previously reported expecting a non-zero share of their female peers to be unemployed (97.0% of respondents). Respondents provided open-ended answers, which enumerators then categorized into pre-defined bins or coded under "Other" if none applied. Enumerators could select multiple responses on behalf of respondents. An answer was coded as "Family reasons" if it corresponded to the responses: "I may not get permission from family to work" or "I am getting married, have a baby, or focus on my family."

B Appendix Tables

Table B.1: Attrition in the Diagnostic Sample: Baseline, Six-Month, and Nine-Month Follow-Ups

		6m Follow-Up				**9m 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. Obs.	1,493	1,029	464			910	583		
Panel A: Administrative Data									
Female	43.87	42.66	46.55	-3.89	0.16	42.53	45.97	-3.44	0.19
GPA	3.07	3.09	3.04	0.04	0.05	3.07	3.07	-0.00	0.81
Age	22.42	22.47	22.31	0.16	0.01	22.45	22.38	0.07	0.32
Majors:									
Engineering / Computer Science	25.18	26.24	22.84	3.39	0.32	24.40	26.42	-2.02	0.16
Life Sciences / Pharmacy	14.00	12.15	18.10	-5.96	0.01	12.64	16.12	-3.49	0.11
Sciences	13.33	13.22	13.58	-0.36	0.88	15.71	9.61	6.11	0.00
Humanities / Languages / Education	13.40	15.55	8.62	6.93	0.00	11.65	16.12	-4.48	0.02
Social Sciences	34.09	32.85	36.85	-4.01	0.10	35.60	31.73	3.87	0.15
Panel B: Survey Responses (Sample Characteristics)									
Married	4.29	4.28	4.31	-0.03	0.92	3.96	4.80	-0.85	0.51
Engaged	6.90	6.61	7.54	-0.93	0.58	6.37	7.72	-1.35	0.37
College-Educated Mother	41.06	41.01	41.16	-0.15	0.94	39.23	43.91	-4.68	0.07
College-Educated Father	54.19	53.16	56.47	-3.31	0.25	52.75	56.43	-3.68	0.17
Panel C: Survey Responses (Key Variables)									
Own Employment Belief for 6m	74.42	74.77	73.66	1.11	0.46	74.87	73.73	1.14	0.42
Belief about Female Peers' Employment for 6m	50.96	50.81	51.31	-0.50	0.76	50.91	51.04	-0.13	0.97
Preferred Daily Work Hours	6.39	6.39	6.39	-0.01	0.98	6.40	6.37	0.02	0.73
Prefer to Work Onsite	72.49	74.33	68.41	5.92	0.02	74.58	69.22	5.36	0.03
Reservation Wage	54.72	55.40	53.20	2.20	0.21	54.42	55.19	-0.77	0.41
Expected Wage	54.42	55.11	52.88	2.23	0.21	55.05	53.42	1.63	0.37

Notes: This table compares attritors and non-attritors in the diagnostic sample (see Section 2.1 for details). The first row reports the number of observations at baseline, the six-month follow-up, and the nine-month follow-up, separately by attrition status. Panel A shows administrative data such as GPA, age, and college major. Panel B shows survey responses about demographic background such as marital status and parental education. Panel C shows survey responses which may determine labor force participation, such as employment beliefs about oneself and others, work arrangement preferences, and wage expectations. All unbounded continuous variables are winsorized at the 2% level. Column 1 reports the distribution of characteristics at baseline for the entire diagnostic sample. Columns 2 and 3 report the distribution for respondents who answered the six-month follow-up survey ("non-attritors") and those who did not ("attritors"). Column 4 reports the differences between the two groups, and Column 5 reports the p-values from testing the differences after controlling for gender. Columns 6 through 9 show the same analyses as Columns 2 through 5, respectively, for the nine-month follow-up survey.

Table B.2: Experimental Sample Treatment Balance: Baseline, Six-Month, and Fourteen-Month Follow-Ups

	Baseline				**6m Follow-Up**				**14m 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. Obs.	939	1,008			688	754			582	636		
Panel A: Administrative Data												
Female	65.18	66.27	-1.09	0.61	63.08	65.25	-2.17	0.39	62.03	63.84	-1.81	0.51
GPA	3.31	3.31	0.00	0.92	3.32	3.33	-0.01	0.88	3.33	3.33	0.00	0.76
Majors:												
Engineering / Computer Science	8.20	6.45	1.75	0.14	7.56	6.63	0.93	0.48	8.08	6.60	1.47	0.32
Life Sciences / Pharmacy	12.67	11.90	0.77	0.58	13.08	12.20	0.88	0.58	13.92	10.85	3.07	0.11
Sciences	25.13	27.28	-2.15	0.27	25.15	28.78	-3.63	0.11	25.77	29.87	-4.10	0.11
Humanities / Languages / Education	28.01	26.09	1.92	0.34	28.49	24.27	4.22	0.07	27.49	23.90	3.59	0.16
Social Sciences	25.99	28.27	-2.29	0.26	25.73	28.12	-2.39	0.31	24.74	28.77	-4.03	0.11
Panel B: Survey Responses (Sample Characteristics)												
Married	4.37	5.36	-0.99	0.31	3.92	4.77	-0.85	0.43	3.95	4.40	-0.45	0.69
Engaged	5.01	3.87	1.14	0.21	4.65	3.18	1.47	0.14	4.30	3.30	0.99	0.36
College-Educated Mother	28.43	28.57	-0.14	1.00	27.62	28.12	-0.50	0.93	26.63	27.99	-1.36	0.66
College-Educated Father	43.98	42.56	1.42	0.50	43.17	42.18	0.99	0.64	41.75	40.57	1.19	0.63
Panel C: Survey Responses (Key Variables)												
Own Employment Belief for 6m	79.99	79.54	0.46	0.65	80.47	80.14	0.33	0.82	80.43	80.26	0.16	0.94
Belief about Female Peers' Employment for 6m	51.23	50.70	0.52	0.54	51.15	50.08	1.07	0.28	51.37	49.91	1.47	0.19
Preferred Daily Work Hours	7.30	7.34	-0.03	0.46	7.29	7.36	-0.07	0.16	7.33	7.38	-0.05	0.37
Prefer to Work Onsite	69.29	67.49	1.80	0.43	69.95	67.81	2.14	0.45	70.14	66.83	3.31	0.26
Reservation Wage	53.48	52.52	0.96	0.42	53.02	52.72	0.30	0.92	52.56	52.57	-0.01	0.91
Expected Wage	62.03	60.54	1.49	0.26	62.47	60.85	1.63	0.33	62.34	61.10	1.24	0.49

Notes: This table compares treatment and control groups in the experimental sample (see Section 4.1 for details). The first row reports the number of observations at baseline, the six-month follow-up, and the fourteen-month follow-up, separately by treatment status. Panel A shows administrative data such as GPA, age, and college major. Panel B shows survey responses about demographic background such as marital status and parental education. Panel C shows survey responses which may determine labor force participation, such as employment beliefs about oneself and others, work arrangement preferences, and wage expectations. All unbounded continuous variables are winsorized at the 2% level. Column 1 reports the distribution of characteristics at baseline for the control group. Column 2 reports the distribution of characteristics at baseline for the treatment group. Column 3 reports the difference between the two groups, and Column 4 reports the p-values from testing the differences after controlling for gender. Columns 5 through 8 and Columns 9 through 12 show the same analyses as Columns 1 through 4 for the six-month and nine-month follow-up surveys, respectively.

Table B.3: Attrition in the Experimental Sample: Baseline, Six-Month and Fourteen-Month Follow-Ups

		6m Follow-Up				14m 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. Obs.	1,947	1,442	505			1,218	729		
Percentage Treated	51.77	52.29	50.30	1.99	0.42	52.22	51.03	1.19	0.58
Panel A: Administrative data									
Female	65.74	64.22	70.10	-5.88	0.01	62.97	70.37	-7.40	0.00
GPA	3.31	3.33	3.27	0.05	0.00	3.33	3.28	0.05	0.00
Age	22.73	22.74	22.71	0.03	0.92	22.71	22.76	-0.05	0.15
Majors:									
Engineering / Computer Science	7.29	7.07	7.92	-0.85	0.54	7.31	7.27	0.04	0.98
Life Sciences / Pharmacy	12.28	12.62	11.29	1.33	0.34	12.32	12.21	0.11	0.78
Sciences	26.25	27.05	23.96	3.09	0.19	27.91	23.46	4.46	0.04
Humanities / Languages / Education	27.02	26.28	29.11	-2.83	0.22	25.62	29.36	-3.74	0.07
Social Sciences	27.17	26.98	27.72	-0.75	0.73	26.85	27.71	-0.86	0.66
Panel B: Survey Responses (Sample Characteristics)									
Married	4.88	4.37	6.34	-1.97	0.11	4.19	6.04	-1.85	0.08
Engaged	4.42	3.88	5.94	-2.06	0.10	3.78	5.49	-1.71	0.12
College-Educated Mother	28.51	27.88	30.30	-2.42	0.48	27.34	30.45	-3.11	0.31
College-Educated Father	43.25	42.65	44.95	-2.30	0.47	41.13	46.78	-5.64	0.03
Panel C: Survey Responses (Key Variables)									
Own Employment Belief for 6m	79.76	80.30	78.20	2.10	0.08	80.34	78.78	1.56	0.18
Belief about Female Peers' Employment for 6m	50.96	50.59	52.00	-1.41	0.23	50.61	51.54	-0.93	0.43
Preferred Daily Work Hours	7.32	7.33	7.29	0.04	0.91	7.36	7.25	0.10	0.24
Prefer to Work Onsite	68.36	68.83	67.01	1.82	0.61	68.40	68.29	0.11	0.78
Reservation Wage	52.98	52.86	53.33	-0.47	0.42	52.56	53.68	-1.12	0.12
Expected Wage	61.26	61.62	60.23	1.40	0.62	61.69	60.54	1.15	0.83

Notes: This table compares attritors and non-attritors in the experimental sample (see Section 4.1 for details). The first row reports the number of observations at baseline, the six-month follow-up, and the nine-month follow-up, separately by attrition status. Panel A shows administrative data such as GPA, age, and college major. Panel B shows survey responses about demographic background such as marital status and parental education. Panel C shows survey responses which may determine labor force participation, such as employment beliefs about oneself and others, work arrangement preferences, and wage expectations. All unbounded continuous variables are winsorized at the 2% level. Column 1 reports the distribution of characteristics at baseline for the entire experimental sample. Columns 2 and 3 report the distribution for respondents who answered the six-month follow-up survey ("non-attritors") and those who did not ("attritors"). Column 4 reports the differences between the two groups, and Column 5 reports the p-values from testing the differences after controlling for gender. Columns 6 through 9 show the same analyses as Columns 2 through 5, respectively, for the fourteen-month follow-up survey.

Table B.4: Treatment Effects on Employment (without Lasso Controls)

	6 Months		14 Months	
	Working	**Working for Firm**	**Working**	**Working for Firm**
	(1)	(2)	(3)	(4)
Panel A: Female				
Treatment	0.066**	0.098***	0.061*	0.087**
	(0.032)	(0.030)	(0.036)	(0.036)
Female Control Mean	0.336	0.253	0.518	0.416
Panel B: Male				
Treatment	0.006	0.015	-0.011	0.004
	(0.044)	(0.043)	(0.043)	(0.047)
Male Control Mean	0.551	0.378	0.719	0.561
Nb. Obs.	1,442	1,442	1,218	1,218

Notes: This table presents the treatment effects on labor market outcomes for students in the six and fourteen-month experimental sample (see Section 4.1 for details). Panel A shows results for women, and Panel B shows results for men. Column 1 reports the treatment effect on employment at 6 months. Column 2 reports the treatment effect on firm employment at six months. Columns 3 and 4 report the same outcomes as Column 1 and 2, respectively, but at fourteen months. Coefficients in the panels are estimated together in a single regression (pooling men and women) and are estimated without controlling for the LASSO-selected variables (such that the only independent variables are gender, treatment and their interaction). Table 3 shows the results with the LASSO-selected controls. Control means are provided separately for each gender. Control means are provided separately for each gender. The last row reports the number of observations. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.5: Treatment Effects of the Experiment on Treated vs. Non-Treated

	6 Months		14 Months	
	Working	**Working for Firm**	**Working**	**Working for Firm**
	(1)	(2)	(3)	(4)
Panel A: Female				
Treatment on Treated	0.136***	0.173***	0.116***	0.152***
	(0.037)	(0.036)	(0.039)	(0.040)
Treatment on Non-Treated	0.002	0.019	0.005	-0.001
	(0.038)	(0.036)	(0.046)	(0.045)
Female Control Mean	0.336	0.253	0.518	0.416
Panel B: Male				
Treatment on Treated	0.079	0.103**	0.041	0.081
	(0.051)	(0.052)	(0.045)	(0.052)
Treatment on Non-Treated	-0.073	-0.073	-0.085	-0.104*
	(0.053)	(0.050)	(0.053)	(0.057)
Male Control Mean	0.551	0.378	0.719	0.561
Nb. Obs.	1,442	1,442	1,218	1,218

Notes: This table presents the treatment effects on labor market outcomes, separately by whether the student was offered treatment and took it up (treated) or was offered treatment but did not take it up (non-treated). The sample consists of students in the experimental sample (see Section 4.1 for details). Panel A shows results for women, separately for treated and non-treated, and Panel B shows results for men, separately for treated and non-treated. Column 1 reports the treatment effect on employment at 6 months. Column 2 reports the treatment effect on (paid) firm employment at six months. Columns 3 and 4, respectively, report the same outcomes as Column 1 and 2, but at fourteen months. Coefficients in the panels are estimated together in a single regression (pooling men and women) and all regressions control for the variables selected following the post-double-selection LASSO procedure. Control means are provided separately for each gender. The last row reports the number of observations. Table 3 shows the same table with overall treatment effects rather than treatment effects by take-up status. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: Treatment Effects on Applications at the Extensive Margin

	6 Months		14 Months	
	Has Applied	**Has Offer**	**Has Applied**	**Has Offer**
	(1)	(2)	(3)	(4)
Panel A: Female				
Treatment	0.083***	0.055*	0.084***	0.040
	(0.024)	(0.030)	(0.020)	(0.029)
Female Control Mean	0.786	0.656	0.877	0.777
Panel B: Male				
Treatment	0.027	0.012	0.021	-0.006
	(0.031)	(0.040)	(0.022)	(0.035)
Male Control Mean	0.842	0.672	0.936	0.840
Nb. Obs.	1,435	1,435	1,210	1,210

Notes: This table presents the treatment effects on applications at the extensive margin for students in the experimental sample (see Section 4.1 for details). Panel A shows results for women, and Panel B shows results for men. Columns 1 and 2 report the treatment effect on a respondent's likelihood of applying for at least one job and receiving at least one job offer by the six-month follow-up, respectively. Columns 3 and 4 show the same outcomes as Columns 1 and 2, respectively, but at the fourteen-month follow-up. Coefficients in the panels are estimated together in a single regression (pooling men and women) and all regressions control for the variables selected following the post-double-selection LASSO procedure. Control means are provided separately for each gender. The last row reports the number of observations. There are fewer observations in this table relative to Table 3 because the analysis is limited to cases with non-missing values for number of applications and offers. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Distribution Across Top Three Occupations by Gender and Treatment

	Treatment (1)	**Control** (2)	**Diff** (3)	**P-value** (4)
Panel A: Women				
Teachers	23.80	24.11	-0.31	0.92
Software and I.T.	12.23	12.89	-0.66	0.77
Marketing and Sales	9.17	9.31	-0.14	0.94
Panel B: Men				
Software and I.T.	14.67	15.76	-1.09	0.77
Teachers	14.67	15.22	-0.54	0.88
Marketing and Sales	9.24	10.33	-1.09	0.73

Notes: This table presents the three most common occupations at the fourteen-month follow-up, separately by gender and treatment status. The sample consists of students in the experimental sample (see Section 4.1 for details) who are employed by the six-month follow-up survey. Panel A shows the top three occupations for women. Panel B shows the top three occupations for men. Column 1 reports the share of respondents in the treatment group in each occupation. Column 2 reports the share of respondents in the control group in each occupation. Column 3 reports the differences between the two groups, and Column 4 reports the p-values from testing the differences.

Table B.8: **Log Reservation Wages Over Time**

	Log Reservation Wage
	(1)
Six Month Response	-0.145***
	(0.027)
Fourteen Month Response	-0.071**
	(0.033)
Treatment	-0.024
	(0.020)
Treatment × Six Month Response	-0.002
	(0.038)
Treatment × Fourteen Month Response	0.007
	(0.044)
Female Control Mean	3.867
Nb. Obs.	2,142

Notes: This table reports time trends and treatment effects from the experimental intervention on log reservation wages among female students. The sample is restricted to women who were observed in the baseline, six months and fourteen months follow-up waves. The dependent variable is the log of the self-reported reservation wage, constructed using two survey questions tailored to current employment status. For employed respondents, the survey asks how much less than their current salary they would be willing to accept if their current job were to shut down. Their reservation wage is computed as their current salary minus this amount. For unemployed respondents, the survey first asks them to imagine a full-time, on-site job in their field and estimate the wage they would be offered. It then asks how much less than this offer they would be willing to accept. Their reservation wage is calculated as the expected offer minus this stated amount. The specification includes indicators for the six- and fourteen-month follow-up waves, as well as interactions between treatment and wave. The baseline wave serves as the omitted category. Coefficients are estimated using the post-double-selection LASSO procedure, which selects control variables from a high-dimensional set of baseline covariates. Control means are reported for the female control group in the baseline wave. The final row reports the number of observations used in the estimation. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.9: Treatment Effects of the Experiment on Marriage Outcomes

	6 Months			14 Months	
	Married or Engaged	**Nb. Marriage Offers**	**Highly Educated Groom**	**Married or Engaged**	**Nb. Marriage Offers**
	(1)	(2)	(3)	(4)	(5)
Treatment	0.011	0.059	0.007	0.024	-0.083
	(0.021)	(0.162)	(0.034)	(0.030)	(0.311)
Female Control Mean	0.098	1.545	0.293	0.175	3.266
Nb. Obs.	806	752	741	697	594

Notes: This table presents the estimated effects of the treatment on marriage outcomes. The sample consists of female students in the experimental sample (see Section 4.1 for details), excluding those who are already married or engaged at baseline. Column 1 reports the treatment effects on being married or engaged by the six-month follow-up. Column 2 reports the treatment effects on number of marriage offers received by the six-month follow-up. Column 3 reports the treatment effects on the whether the highest education level among all received offers is a Master's degree or higher, which proxies quality of marriage offers. Columns 4 and 5 show the same outcomes as Columns 1 and 2, respectively, but at the fourteen-month follow-up. There is no equivalent version of Column 3 because the question was not asked at the fourteen-month follow-up. Number of marriage offers is winsorized at the 2% level. All regressions control for the variables selected following the post-double-selection LASSO procedure. The female control mean is provided in the penultimate row, and the last row reports the number of observations. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.10: Means and Standard Deviations for Diagnostic Predictors of Employment

	Female		Male	
	Mean	**SD**	**Mean**	**SD**
	(1)	(2)	(3)	(4)
GPA	3.23	0.46	2.98	0.45
Log Reservation Wage	10.71	0.42	10.90	0.46
Log Expected Wage	10.70	0.40	10.90	0.46
Preferred Daily Work Hours	6.40	1.41	6.38	1.42
Prefers On-site	0.74	0.44	0.75	0.43
Own Employment Belief for 6m	0.72	0.22	0.77	0.24
Applied Early	0.73	0.45	0.72	0.45
Nb. Applications	7.86	12.18	10.64	16.70
Internship	0.38	0.49	0.52	0.50

Notes: This table presents the mean and standard deviations for the diagnostic predictors of employment shown in Figure 3 and Figure A.12. The sample consists of students in the diagnostic sample (see Section 2.1 for details). All unbounded continuous variables are winsorized at the 2% level.

C External Validity of Diagnostic Results

In this section we establish the external validity of our diagnostic findings from the private university by replicating them at Pakistan's oldest and largest public university. This new setting attracts students from more diverse socioeconomic and geographic background than the private university, making it an attractive setting for testing external validity. We establish external validity by showing that our diagnostic insights from Figures 1-2 hold when produced using the control group in our experimental sample.[33] We also present additional data suggesting that women not only expect to work post-graduation, but want to.

Figure C.1 Panel (a) compares baseline employment expectations with realized outcomes by gender in the control group of the experimental sample, replicating the analysis from Figure 1 Panel (a). Consistent with the patterns observed in the diagnostic sample, men and women in the control group of the experimental sample exhibit high and similar expectations for future employment at baseline: 79.8% of women and 81.6% of men expect to be employed within six months of graduation. The modest gender gap in expectations (1.8 pp) is comparable to the gap in the diagnostic sample (5.2 pp), 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—46.2 pp below their expectations, and 21.5 pp lower than men's realized employment rate. This overestimation of future employment among women mirrors the diagnostic sample, where 36.9% of women are employed six months post-graduation, 34.9 pp below their average baseline expectations and 27.3 pp lower than men's realized employment rate.

Figure C.1 Panel (b) shows that the relationship between baseline employment beliefs and realized outcomes in the experimental sample mirrors the analysis in Figure 1 Panel (b): both genders have inaccurate beliefs across the full distribution, and markedly different intercepts. Slopes are higher in the experimental sample than the diagnostic sample, but remain far from 1 (0.52 for women and 0.62 for men in the experimental sample, versus 0.31 for women and 0.35 for men in the diagnostic sample). Slopes are also less precisely

33. The treatment group is excluded as our treatment directly impacts some of the outcomes of interest, such as realized employment.

estimated, due to a smaller sample size, which results from restricting our analysis to the control group.

Figure C.2 compares the baseline first and second order employment beliefs of men and women in the experimental sample, replicating insights from Figure 4. In the diagnostic sample, we saw that even though men and women have inaccurate beliefs about their 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 diagnostic sample. 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%. 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 remarkably close the second order beliefs reported in the diagnostic sample (50.2% by men and 51.6% by women).

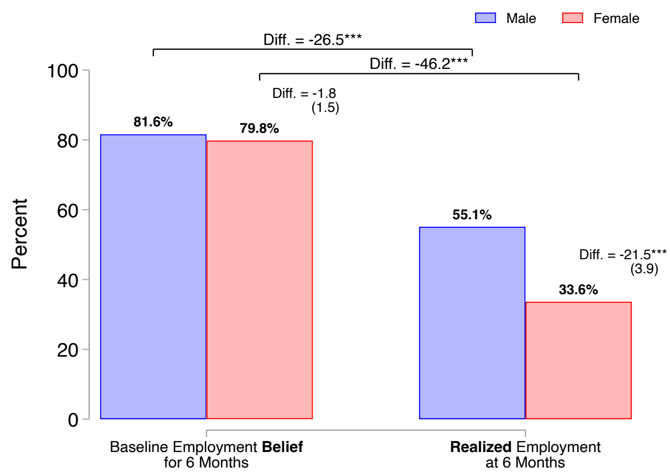
Figure C.3 follows Figure 2 in analyzing whether common demand- and supply-side barriers to female employment explain the gender employment gap observed six months after graduation. In the diagnostic sample, the raw gender gap in employment was 28.9 pp. In the control group of the experimental sample, the employment gap is smaller but still substantially large at 21.5 pp. As in Figure 2, 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 employment gap in the control group of the experimental sample only modestly, by 2.4 pp. Adding controls for job search effort and work history further narrows the gap by just 0.8 pp. Finally, accounting for demand-side factors (number of interviews, job offers, and offered wages) reduces the gap by another 0.3 pp. Taken together, these controls only reduce the raw gap by 3.5 pp. This is similar to the diagnostic sample, where the gap shrinks from 27.3 pp to 24.2 pp (3.6 pp), such that the key takeaway remains: even after controlling for student characteristics, job preferences, and demand-side factors (especially the number of job offers), the gender employment gap persists.

Beyond replication of the patterns already documented at the private university, we collect additional data at the public university in support of the fact that women at the time they graduate from college *want* to work. First, we explicitly ask women if they want to work on a paid basis within six months of graduation and, separately, two years after graduation.[34] Of the women surveyed, 97.6% and 98.4% reported that they want to be working both six months and two years after graduation, respectively. This suggests that women want to start working soon after graduation and they also want to continue working for a while after graduation. We also ask women who reported less than a 100% likelihood of working within six months, why they think they may not work. The first panel of Figure C.4 titled "Self" reports the distribution of their responses. The most prevalent barrier, cited by 37% of women, was poor macroeconomic conditions that may limit job availability. Importantly, only 9% told us that they may not work because they may not *want* to work six months from now. Even when considering respondents' beliefs about other women (second panel in Figure C.4 titled "Other Women"), only 15% of women reported that other women may not work post-graduation because they may not want to work. Finally, we find that there is no meaningful difference in women's stated likelihood of working in six months by their levels of familial wealth, suggesting that financial necessity may not be the primary driver of women's intention to work. At baseline, we measure ownership of five assets and amenities: car, motorbike, internet connection, laptop and smartphone. As shown in Table 2, most families own all of these assets, but there is more variation in car ownership across households. Accordingly, Figure C.5 shows baseline beliefs about work separately by whether a woman's family owns a car. There are no systematic differences in women's responses by family car ownership (78.5% and 78.3%). There are also no significant differences in realized employment outcomes six months later by car ownership. This suggests that financial constraints are not the primary reason driving women's willingness to work and that other motives (e.g., independence, desire for a career; bargaining power etc...) prevail.

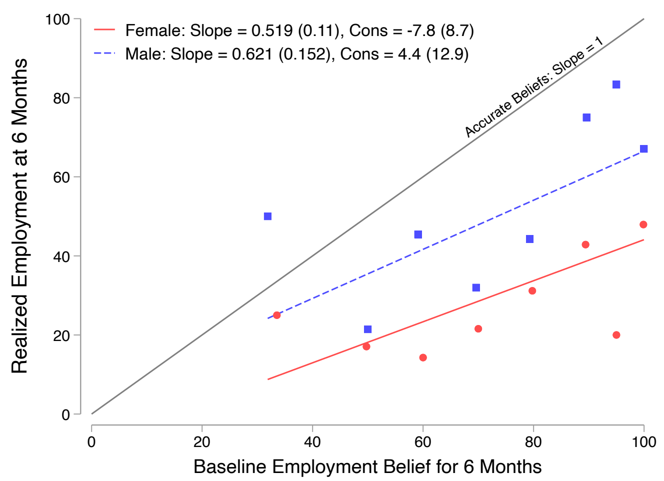
34. The exact questions were: "Do you want to be working (e.g., for a firm or in your own/family business) in the next **six months** after you graduate?" and "Do you want to be working (e.g., for a firm or in your own/family business) on a paid basis **2 years** from now?"

External Validity Figures

Figure C.1: Baseline Employment Belief vs. Realized Employment: Experiment Control Sample



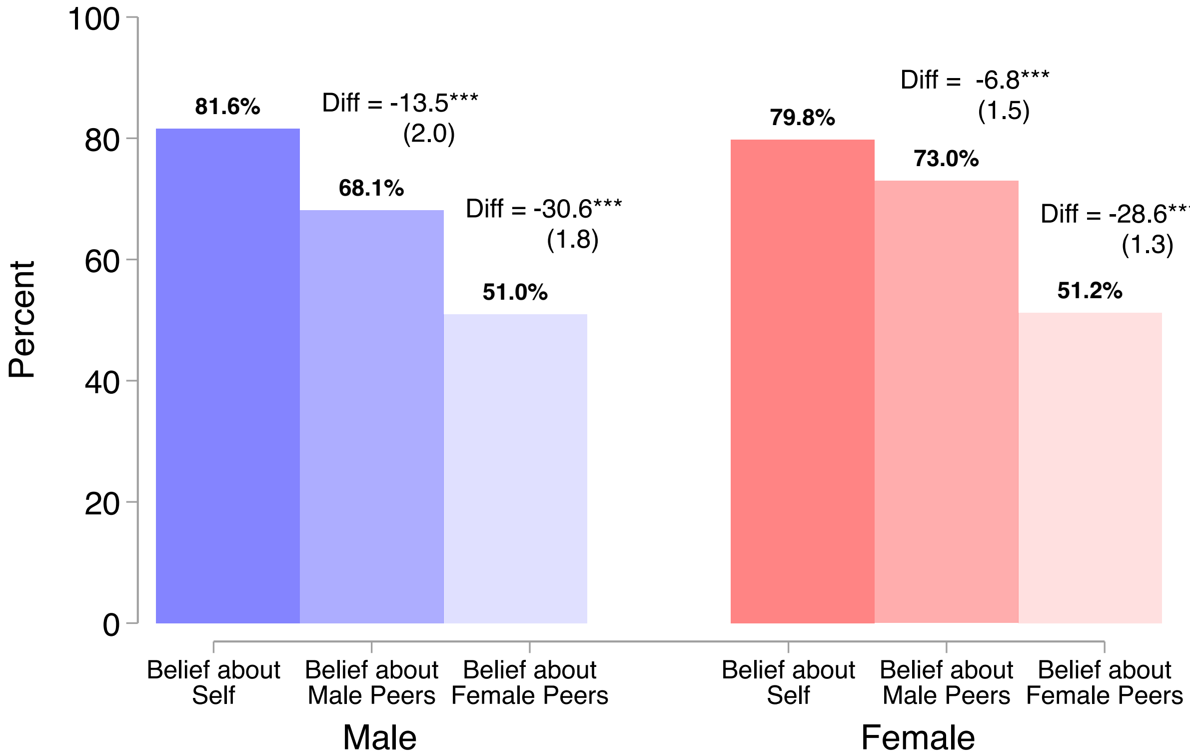
(a) Mean Levels: Expected vs. Realized Employment



(b) Binned Scatter: Expected vs. Realized Employment

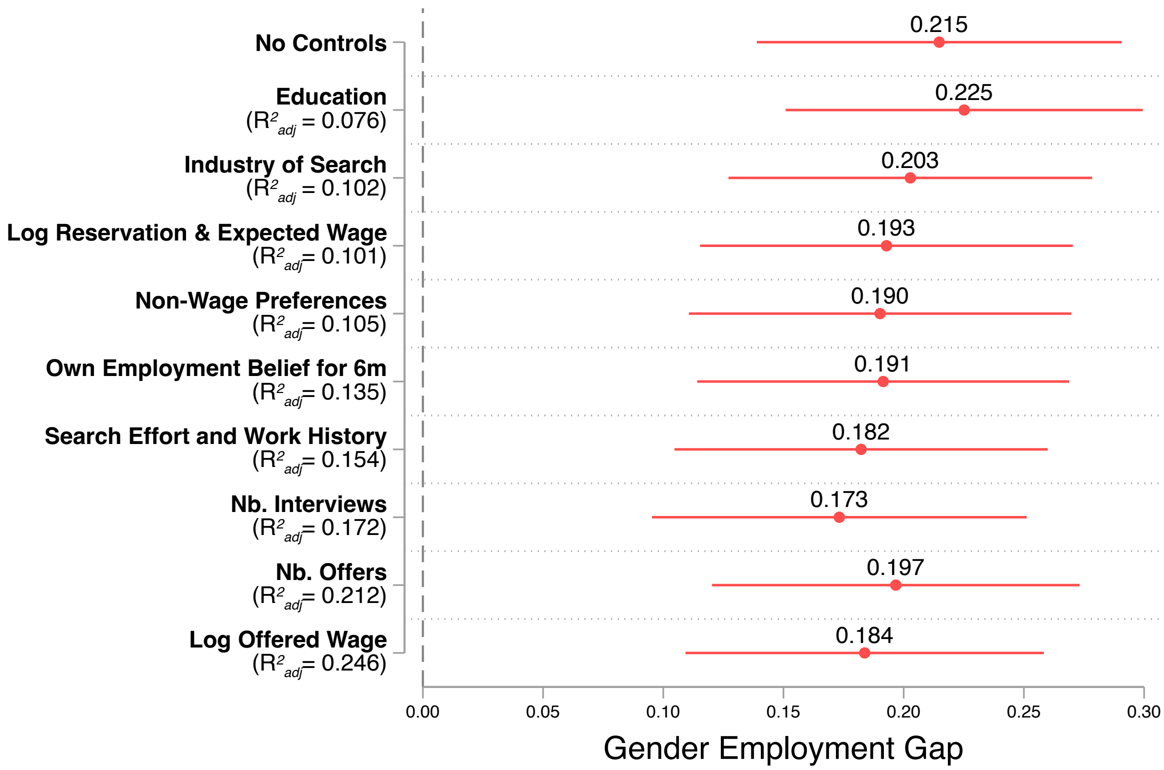
Notes: This figure replicates Figure 1 with the control group of the experimental sample (see Section 4.1 for details). Panel (a) shows the average self-assessed likelihood of employment six months after graduation (left cluster of two bars) and realized employment at six months (right cluster of two bars), separately by gender. Men are represented in blue (bars 1 and 3) and women are represented in red (bars 2 and 4). Panel (b) shows a binned scatter plot of baseline employment beliefs against realized employment, separately by gender. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

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



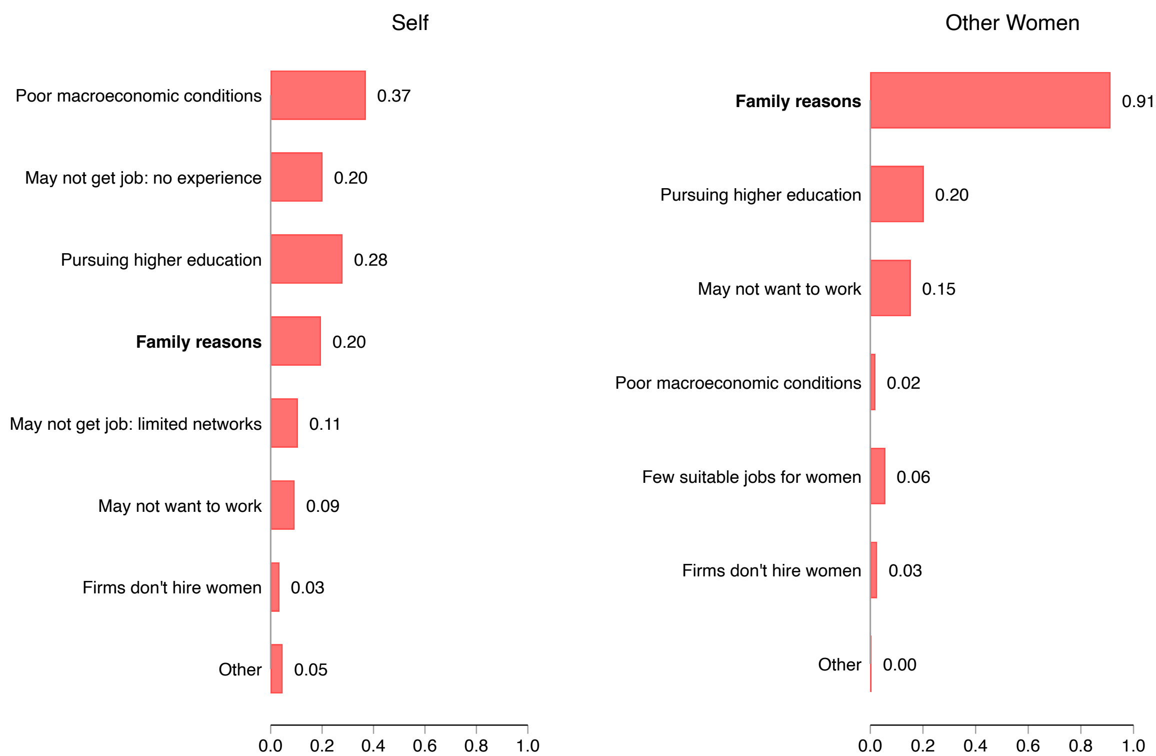
Notes: This figure replicates Figure 4 with data from the control group of the experimental sample (see Section 4.1 for details). Male beliefs are represented by the blue cluster of bars on the left. Female beliefs are represented by the red cluster of bars on the right. Bars 1 and 4 show average baseline beliefs about oneself. Bars 2 and 5 show average baseline beliefs about male peers. Bars 3 and 6 show average baseline beliefs about female peers. The average difference between beliefs about self and beliefs about male peers is shown directly above the second bar in each cluster. The average difference between beliefs about self and beliefs about female peers is shown directly above the third bar in each cluster. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure C.3: Explaining the Employment Gap at Six Months: Experiment Control Sample



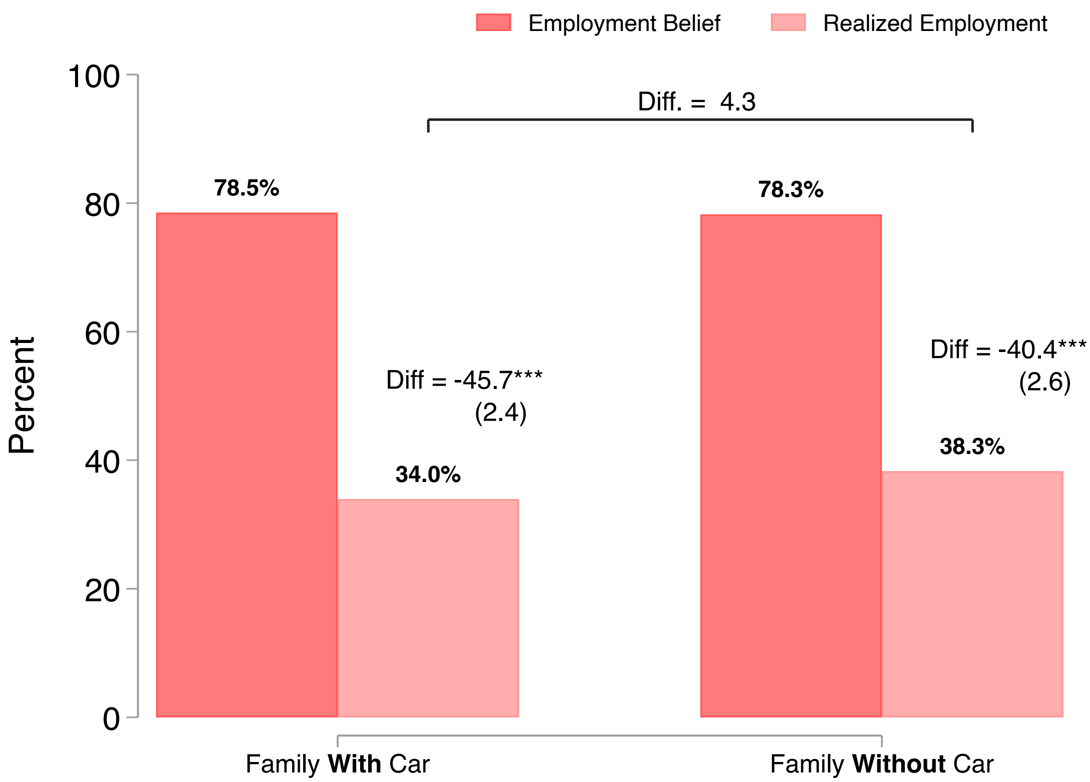
Notes: This figure replicates Figure 2 using data from the control group of the experimental sample (see Section 4.1 for details). Each row reports the gap between male and female employment after controlling for the variable listed in that row and all rows above. The Education controls include cumulative GPA and major fixed effects. The Industry of Search controls incorporate fixed effects derived from the semantic classification of a respondent's preferred job title into Standard Occupational Classification (SOC) codes. The Reservation and Expected Wage controls include respondents' wage expectations at baseline. The Non-Wage Preferences controls include preferences regarding onsite vs. remote work and the number of preferred daily work hours. The Own Employment Belief for 6 Months control includes baseline belief about one's probability of employment. The Search Effort and Work History controls include the total number of job applications sent by the six-month follow-up, an indicator for applying early, and an indicator for internship experience. The Nb. Interviews control includes the number of interviews received by the six-month follow-up. The Nb. Offers control includes the number of job offers received by the six-month follow-up. The Offered Wage control includes the highest wage offer a student has received for a job (regardless of whether they have accepted it). All unbounded continuous variables are winsorized at the 2% level. Adjusted R-squared values are shown for each row. The horizontal bars show 95% confidence intervals.

Figure C.4: Potential Reasons for Remaining Unemployed: Self vs. Peers



Notes: This figure presents women's beliefs at baseline about why they and other women in their class may not work in the future. The sample consists of all female students in the control group at the public university who were successfully surveyed at baseline and the six months follow-up (see Section 4.1 for details). The left panel ("Self") reports responses to the question: "Why do you think there is a chance you may not work?" This question was asked to all women who reported less than a 100% likelihood of being employed six months after graduation (75.6% of respondents). The right panel ("Other women") presents responses to the question: "Out of the remaining XX women, why do you think they are not working?" This question was asked to all women who previously reported expecting a non-zero share of their female peers to be unemployed (97.0% of respondents). Respondents provided open-ended answers, which enumerators then categorized into pre-defined bins or coded under "Other" if none applied. Enumerators could select multiple responses on behalf of respondents. An answer was coded as "Family reasons" if it corresponded to the responses: "I may not get permission from family to work" or "I am getting married, have a baby, or focus on my family."

Figure C.5: Employment Beliefs and Realized Outcomes by Wealth



Notes: This figure presents heterogeneity by familial wealth in women's beliefs about employment in six months, and their realized employment in six months. The sample consists of all female students in the control group (see Section 4.1 for details). Women whose families own a car are represented by the left cluster of two bars, and women whose families do not are represented by the right cluster of two bars. The darker red bars (bars 1 and 3) represent employment beliefs, and the lighter red bars (bars 2 and 4) represent realized employment. The average difference between employment beliefs and realized employment for each group of women is shown directly above the realized employment bar in each cluster. The average difference in realized employment for women whose families do and do not own a car is shown above the brackets at the top of the figure. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.