Wage Subsidies to Promote Female Hiring: Evidence from Pakistan*

Maurizio Bussolo, Jean Lee, Mahreen Mahmud, Nayantara Sarma, Anaise Williams

November 3, 2025

Abstract

Can employer-side wage subsidies increase women's hiring in low female labor force participation settings? We test this through a randomized experiment with 1,227 Pakistani firms on a national jobs platform. Treatment firms were offered a six-month wage subsidy determined via the Becker–DeGroot–Marschak mechanism. Firms in the treatment group are 11 percentage points more likely to hire a woman, with larger effects for male-only firms. After 18 months, treated firms remained more likely to employ a woman in the subsidized role, though firm-wide share of female employees did not change. Administrative job-posting data show reduced male preference among treated firms, suggesting emerging demand-side shifts.

*Bussolo: World Bank, mbussolo@worldbank.org; Lee: World Bank, jlee20@worldbank.org; Mahmud: University of Exeter, m.mahmud@exeter.ac.uk; Sarma: World Bank, nsarma@worldbank.org; Williams: Johns Hopkins University, awill137@jhmi.edu. We acknowledge the support of the Rozee.pk team based in Lahore in collecting the data and managing communication with firms. We would like to thank Ali Bakhtawar, Arushay Ali, Fouzia Ali, Sajwaar Khalid, Sameen Tariq, Shiraz Hassan and Zahra Riaz. This project would not have been possible without the support of Shahbaz Khan, Country Head - Enterprise Solutions at Rozee.pk. We also thank Tillmann Spindeldreier for excellent research assistance. We thank colleagues at the World Bank South Asia Chief Economist Office for guidance and inputs on conceptualization and process. We acknowledge financial support for this study from the World Bank South Asia Region Gender Innovation Lab. The study was approved by the University of Exeter Business School Research Ethics Committee (#eUEBS004187). The study is registered as AEARCTR-0009226 (Bussolo et al., 2024) and a pre-analysis plan was lodged there before the hiring followup survey data collection was complete.

1 Introduction

Globally, female labor force participation (FLFP) averages around 50 percent – over 20 percentage points lower than that of men – and is substantially lower in regions such as the Middle East and North Africa and South Asia (International Labour Organization, 2025; World Economic Forum, 2020). The supply-side correlates of women's participation – education and skills, norms and attitudes, transport and childcare access, and safety – are well documented (Heath et al., 2024; Borker, 2025; Jayachandran, 2021; Groh et al., 2016; Compton and Pollak, 2014; Garlick et al., 2025; McKelway, 2025; Dean and Jayachandran, 2019; Afridi et al., 2022). By contrast, less is known about demand-side frictions and policies that relax them, even though audit and field evidence documents gender gaps in hiring (Gentile et al., 2023; Jensen, 2012).

Motivated by the limited experimental evidence on demand-side interventions and the persistence of gender gaps in hiring, we evaluate an employer-targeted wage subsidy that lowers the cost of hiring women. We run the experiment in Pakistan, where FLFP is approximately 24 percent, one of the lowest rates globally. Our sample consists of 1,227 small, medium, and large firms, recruited via the country's largest online job portal. In a randomized controlled trial, firms with an active job opening were randomly offered a six-month subsidy for hiring a woman.[1] This design allows us to estimate the effect of a temporary women's wage subsidy on initial hiring and on women's longer-run employment after the subsidy expires, and examine whether it shifts revealed gender preferences in future job advertisements and hiring.

We partnered with Pakistan's largest online job platform, Rozee.pk, to recruit firms. The sampling frame comprised all vacancies posted on Rozee.pk between April and November 2022 that stated no gender preference, and for budget reasons, we restricted our sample to ads offering monthly salaries below PKR 80,000 (about USD 1,700 PPP). Vacancies largely involved technical and professional roles and, given the salary cap, were predominantly entry-level.

A novel feature of our design is the use of the Becker–DeGroot–Marschak (BDM) mechanism (Becker et al., 1964) to elicit firms' willingness to hire a woman.[2] In the treatment group, each firm states the minimum six-month subsidy it would accept to hire a woman; a subsidy is offered only if the stated amount does not exceed a computer-generated random draw. Our implementation yields a revealed-preference measure of the subsidy required to equalize hiring choices and lets us quantify how far a targeted subsidy can offset male-leaning hiring preferences. On average, treatment-group firms report needing about a 15 percent subsidy to hire a woman, well below the subsidy magnitudes

[1] Our experiment took place in a context of a significant gender pay gap, where the mean and median wages difference between men and women in paid employment is estimated to be from 25-30 percent depending on methods (ILO, 2025). This gender pay gap is among the highest in South Asia, as reported by the International Labour Organization (2025).

[2] BDM has been successfully used in other developing country contexts to study gender differences in willingness to pay for health and education goods (Dizon-Ross and Jayachandran, 2023), the effect of quality signaling on demand for solar lanterns (Coville et al., 2025), and demand for water filters in Ghana (Berry et al., 2020).

tested in prior wage-subsidy experiments on gender (McKenzie, 2017).

Our pre-specified intent-to-treat estimates show a 10.7 percentage points increase in the likelihood of a woman being hired. This is robust to the use of Lee (2009) bounds to account for differential attrition at the stage which hiring information was collected. Of the 657 firms in the treatment group, 70% were offered a wage subsidy – that is, their elicited minimum subsidy level in the BDM was below the randomly drawn number – and they were informed immediately. We instrument for the probability of being offered a subsidy with the randomly drawn subsidy level and the treatment effects, conditional on the elicited subsidy level required, are larger at 20.4 percentage points. Further, given the limited sample size, which prevents us from estimating treatment effects at each subsidy level, we estimate kernel-weighted IV to understand how the impact on hiring varies across the elicited subsidy level required. The effects are positive but not statistically significant at most subsidy levels. As expected, we see a downward trend in estimated impacts as the required subsidy level increases.

Secondary outcomes of interest, such as the gender composition of applicants and interviewees, do not show any significant differences although there may have been unmeasured changes in the quality of candidates and by gender.[3] We find that the subsidy may have enabled treated firms to pay higher wages and issue formal contracts: the reported salary of the hired candidate is around 7% higher in the treatment group and marginally significant, as is the likelihood of being offered a formal contract. This indicates that firms passed on part of the subsidy to the candidates.

We explore heterogeneity along two observable dimensions, job and firm characteristics, and, while we interpret the patterns cautiously, a plausible reading uses an information-frictions lens. Job features such as skill requirements, pay level, and customer contact proxy the uncertainty managers face about women's productivity and retention in a given role. When that uncertainty (and the downside risk of a 'mistake') is lower, for example, lower-skill, lower-pay roles, even a modest subsidy can compensate, yielding larger effects. We find that the intent-to-treat effects are higher by 10 percentage points in below median salary vacancies and by 18 percentage points in jobs that do not require an undergraduate degree. On the firm side, having no female employees at baseline signals little in-house information and fewer established processes for recruiting and onboarding women. In such male-only firms, the subsidy both insures against a potentially costly mismatch and facilitates learning by providing a credible first experience, and we find that the treatment effect is 17 percentage points higher.

Consistent with this interpretation, BDM bids, which can be read as an ex-ante risk premium for hiring a woman, are lowest where managers already have informative signals, proxied by gender-separate facilities or progressive attitudes. In addition, 'honest' causal forests (Wager and Athey, 2018) also highlight manager beliefs and baseline share of female employees as leading

[3]We would not expect the gender composition of the applicants to be different between treatment and control groups as a result of the intervention unless treatment firms encouraged female candidates to apply.

predictors of treatment effects. Taken together, the heterogeneity patterns and elicited bids suggest that the subsidy relaxes information frictions by reducing perceived risk and enabling firms to update their beliefs about female worker productivity. That said, alternative channels, such as setup frictions or taste-based reluctance, could also contribute, and our design does not permit a clean separation among them.

To assess longer-run impacts of the wage subsidy, we triangulate three data sources. First, in a survey of hired women, we find that the vast majority of those hired under the subsidy were retained or moved to better positions: Of the 131 women hired by subsidized firms, we successfully surveyed 112 one month after the six-month subsidy ended, and approximately 28% were no longer with the same firm; no respondent reported leaving because the subsidy had ended.[4] Second, the endline survey – conducted about 1.5 years after the baseline – reached only 27% of firms because many had shut down, relocated, or refused to participate amid Pakistan's post-COVID downturn. With high but balanced attrition, we find that treatment group firms were 13 percentage points more likely to still have a woman in the subsidized post; however, firm-wide female staffing was unchanged. Third, administrative data on job advertisements from our partner platform (covering study firms from 2020 to mid-2024) show evidence of longer-term shifts in gendered hiring preferences: treated firms were 5.5 percentage points less likely to specify a male preference in their ads after the intervention, with the sharpest decrease in firms that were originally all-male or had few female employees.

We contribute to the literature by providing, to our knowledge, the first experimental test of a *demand-side* wage subsidy designed to raise female hiring in a low- and middle-income country context.[5] While wage subsidies have mostly been deployed on the supply-side in high-income economies (Katz, 1996), firm-level field experiments have typically been limited in scale, especially for medium and large firms, due to implementation and recruitment frictions (Bandiera et al., 2011). Our design leverages a large online platform to recruit a diverse, though not fully representative, set of firms spanning sizes, industries, locations, managerial norms and preferences, and baseline gender composition. Closest antecedents include Groh et al. (2016) in Jordan, Abel et al. (2022) in Mexico, and Armand et al. (2026) in North Macedonia which subsidize workers; earlier firm-level trials in Argentina (Galasso et al., 2004) and South Africa (Levinsohn et al., 2014) saw minimal take-up due to regulatory concerns and de Mel et al. (2019) focused on micro-enterprises with no employees. Our approach differs: an employer-targeted *partial* subsidy, with levels elicited via a BDM mechanism, allowed firms to compete for qualified female candidates and achieved substantially higher take-up, alongside suggestive evidence of longer-run shifts in gendered hiring preferences.

[4]Of the 439 firms in the treatment group offered the subsidy, 247 hired a woman. 53% of these eligible firms provided the necessary paperwork to set it up. See Section 5.2 for details.

[5]Jaenichen and Stephan (2011) and Bruhn (2020) provide evidence using non-experimental methods on the effectiveness of wage subsidies paid to employers in the context of hard-to-place workers in Germany and post-pandemic Mexico.

2 Conceptual Framework

Firms may be reluctant to hire women because doing so increases the effective cost of filling a vacancy, either through beliefs about performance and team fit or through concrete organizational requirements. On the belief (expectations) side, managers may expect lower individual productivity, higher absence risk, or lower team productivity due to male–female interactions. On the organizational side, firms may face one-off or recurring outlays (e.g., setting up gender-separate amenities, revising HR processes, or additional supervision).[6] A separate, conceptually distinct channel is preference (taste-based) discrimination in the sense of Becker (1957), whereby decision-makers (or customers/co-workers) experience a disutility from employing women, raising the implicit cost even when expected productivity is the same.

It is helpful to distinguish *belief-based* discrimination, often called "statistical discrimination", from *taste-based* discrimination. In the former, employers act on beliefs about average productivity, retention, or team externalities for women relative to men. These beliefs may be correct (given the employer's information set and constraints) or biased/misperceived (e.g., stereotypes, outdated priors, or norms-driven misbeliefs). In either case, employers use beliefs, rather than direct taste, to screen. In the latter (taste-based), employers derive disutility from hiring women per se; even with identical expected productivity, hiring a woman reduces utility unless compensated.

A simple way to summarize the firm's hiring margin is:

$$\underbrace{\mathbb{E}\big[\text{productivity}_f\big]}_{\substack{\text{beliefs / expected}\\ \text{productivity}}} - \underbrace{w}_{\text{wage}} - \underbrace{c_{\text{org}}}_{\substack{\text{organizational}\\ \text{outlays}}} - \underbrace{d}_{\substack{\text{taste}\\ \text{penalty}}} + \underbrace{s}_{\text{subsidy}} \gtrless 0. \tag{1}$$

where c_{org} includes both one-off and recurring costs (e.g., gender-separate amenities, onboarding, supervision), d captures taste-based disutility, and s is a temporary wage subsidy. Belief-based channels operate through $\mathbb{E}[\text{productivity}_f]$ (including perceived team externalities), organizational frictions load into c_{org}, and taste loads into d.

Gender norms can influence the supply side (Jayachandran, 2021; Bursztyn et al., 2020). On the demand side, norms can (i) bias beliefs about women's productivity/retention, (ii) increase perceived team or customer backlash (raising effective organizational costs), or (iii) directly enter as a taste penalty. Potential employers who hold more regressive norms may therefore restrict female hiring even holding fundamentals fixed.

Within this framework, a temporary, employer-side wage subsidy can raise women's hiring via two proximate mechanisms. First (static cost-offset): by lowering the effective wage bill today, s compensates for organizational outlays c_{org} and/or for the perceived downside risk from

[6] Recent work by Gentile et al. (2023) using datasets from three sources in Pakistan finds that firms with gender-separate amenities are more willing to hire women.

uncertainty about productivity, inducing marginal hires in states where the inequality would otherwise be negative (McKenzie, 2017). Second (dynamic learning/adoption): if first-time hires generate information about match quality and team functioning, or catalyze set-up investments (e.g., amenities, recruitment processes), then posterior beliefs improve and some costs become sunk or amortized, potentially yielding persistence after s expires. By contrast, if reluctance is primarily taste-based and insensitive to experience, impacts may attenuate once the subsidy ends unless norms or preferences also shift.

This simple framework yields testable implications that connect to our design. *Job features* (skill requirements, pay level, customer contact) can proxy the perceived uncertainty and downside risk of hiring a woman: where the stakes of a mismatch are lower (e.g., lower-skill/lower-pay roles), a modest s should move the margin more. *Firm characteristics*, notably the baseline female share—index, prior experience, and information as well as the presence of organizational complements (amenities, processes): firms with no women at baseline may require a larger initial push but can exhibit big responses when s both insures the first hire and kick-starts learning. Finally, to the extent that norms and beliefs are mutable, information (through experience) and temporary subsidies can be complements: the subsidy enables the experiment; the experiment updates beliefs and lowers future costs, creating scope for longer-run gains, even as we remain agnostic ex ante about how much is information vs. taste.

3 Experimental Design and Intervention

3.1 Context and Sample

Similar to other countries in South Asia and the Middle East and North Africa, Pakistan has a large gender gap in labor market outcomes: in 2024, the male rate of employment was 80.3% while it was 24.3% for women (World Bank, 2024). This gap persists for urban employment rates across genders and for those with at least an undergraduate degree (Ahmed et al., 2024).

Rozee.pk, launched in 2007, is Pakistan's largest online job portal. Our sample frame comprises all firms posting eligible vacancies on Rozee.pk's between April and November 2022. Eligibility required: (i) no stated gender preference on the vacancy, (ii) a Pakistan-based position, and (iii) a posted monthly salary below PKR 80,000 (USD 1,700 PPP).[7]

We contacted 7,522 firms with active job advertisements that met these criteria, and 1,277 (17%) completed the baseline survey and agreed to participate. The study sample is diverse: firms employ on average about 115 individuals; 27% operate in ICT, and 93% are based in a major city (population >1 million) (Table 1). Hiring managers are predominantly male (84%). On average, 23% of the employees are women, with substantial heterogeneity: 20% of firms report no female

[7]The salary limit was initially PKR 100,000 and was revised downward after a few days of fieldwork. Hence, the upper limit of the salary band reported in Table 1 is PKR 100,000 and not PKR 80,000.

employees at baseline. Relatedly, roughly 80% report gender-separated toilets, whereas only 16% offer childcare services or creches.

Table 1: Baseline Summary Statistics

	Mean	SD	Min	Max	N
Panel A: Firm Characteristics					
Firm size	114.84	781.45	1.00	15,000	1,227
Share of female employees	0.23	0.20	0.00	1	1,204
Male only firm	0.19	0.39	0.00	1	1,227
Main city	0.93	0.25	0.00	1	1,227
Industry - Communications	0.11	0.31	0.00	1	1,224
Industry - Finance	0.10	0.30	0.00	1	1,224
Industry - Health, Education, Admin	0.08	0.27	0.00	1	1,227
Industry - Manufacturing & Retail	0.10	0.30	0.00	1	1,227
Industry - IT	0.27	0.44	0.00	1	1,224
Industry - Other	0.22	0.41	0.00	1	1,224
Industry - Not Reported	0.13	0.34	0.00	1	1,224
Separated toilets by gender	0.78	0.41	0.00	1	1,227
Childcare services or creches	0.16	0.36	0.00	1	1,227
Panel B: Job Characteristics					
Frequent interaction with public	0.46	0.50	0.00	1	1,226
Salary (upper limit PKR)	58145.88	24243.98	20000.00	100000	1,227
Requires undergraduate or above	0.67	0.47	0.00	1	1,227
Transportation (or allowance)	0.55	0.50	0.00	1	1,227
Health insurance	0.31	0.46	0.00	1	1,227
Flexible working hours	0.68	0.47	0.00	1	1,227
Possibility to work from home	0.60	0.49	0.00	1	1,227
Panel C: Manager Characteristics					
Age (years)	32.05	7.84	18.00	84	1,227
Male	0.84	0.37	0.00	1	1,227
Married	0.61	0.49	0.00	1	1,227
Masters or above	0.51	0.50	0.00	1	1,227

Notes: The table reports statistics from the baseline survey. Panel A contains indicator variables for firm characteristics, except for Firm size, which is the number of employees and Share of female employees, which is the ratio of total females to the total number of employees. Panel B contains indicator variables for job characteristics except the salary upper limit. Panel C contains indicator variables for manager characteristics, except for age, which is in years.

3.2 Sample representativeness

Our study draws on Rozee.pk job posts that state (i) no gender preference and (ii) a monthly salary below PKR 80,000. With the exception of this wage ceiling – imposed for budgetary reasons – these vacancies look very similar to the full set of Rozee ads without gender preferences: other job and firm attributes differ little, and any statistically significant gaps are economically negligible (Table

SA2).[8,9] Only 30 percent of all Rozee.pk advertisements specify a preferred gender, and those posts differ markedly in pay levels and sectors (Table SA3). We therefore focus on the remaining 70 percent of ostensibly gender-neutral vacancies.

It is important to recognize that Rozee.pk itself is not nationally representative. The platform oversamples younger and ICT-oriented firms located in Pakistan's three largest cities, and its users are disproportionately urban, educated, and young (Matsuda et al., 2019). As a result, our sample sits at the more progressive end of the employer spectrum. Finding a sizable treatment effect here suggests that a similar wage-subsidy program could have equal or even greater impacts in more conservative labor market segments, where baseline female employment is lower and existing amenities are scarcer.

3.3 Intervention

This randomized control trial tests the effects of a 6-month wage subsidy for hiring a woman. The subsidy, paid monthly to the employee, can be fully passed on, shared, or offset by the employer. Using the Becker-DeGroot-Marschak (BDM) method, firms report the minimum subsidy they would accept, and only those whose required level is at or below a randomly drawn number are eligible. As discussed in Dizon-Ross and Jayachandran (2023), it is then weakly optimal for firms to report truthfully the ex ante subsidy amount that would make them indifferent between hiring a woman and the status quo.

We tell the employer "*Imagine you were offered a wage subsidy, that is, a share of the employee's salary would be paid by us for 6 months. What is the lowest subsidy percentage you would be willing to accept to hire a qualified female candidate for this position?*". Employers are informed that if their stated subsidy exceeds the randomly drawn amount, they will not receive any wage support for hiring a female employee. To ensure comprehension of the BDM mechanism, they first complete a practice round using an unrelated example.

The manager was immediately informed of the outcome after providing their chosen subsidy level. If they won, they were informed of the randomly chosen subsidy level that they would be provided if they hire a woman for the position. The firms then proceeded to shortlisting, interviewing, and hiring. After confirming the gender of the hired candidate, the subsidy is paid monthly for up to 6 months as long as the woman continues to work in the advertised position. Note that hiring of a woman was not enforced, conditional on winning the wager, but that firms could choose after learning the subsidy and seeing the candidate pool to hire a woman or a man, given that ex ante expectations of productivity differentials, for example or other costs may not have matched those ex post.

[8]The reported salary in Table SA2 is different from what is reported in Table 1 as the former uses the portal data and the latter relies on hiring managers' reported salary band.

[9]If labor demand is less elastic at higher salary levels, then the effect on female hiring would be attenuated in high-paying jobs relative to those represented in our sample.

3.4 Randomization

We used a minmax t-stat re-randomization procedure to split the sample of hiring managers into two groups, i.e. treatment and control, following Bruhn and McKenzie (2009). We used the following baseline variables as covariates: firm characteristics (location, industry, sex-ratio), hiring manager characteristics (unincentivized subsidy required at baseline, age, gender, marital status, education, gender attitudes and norms) and educational qualifications required for the job. As fresh vacancies were posted continuously, firms were randomized every week so that the overall sample is balanced. Only one subsidy was offered per firm, even if they hired more than 1 person for the advertised post, and only one job-posting was considered per firm in our sample.

4 Data and Estimation Strategy

4.1 Data

All surveys were administered over the phone. The study design and the sample size for each survey are shown in Figure SA1. The baseline survey, administered from April - November 2022, was an approximately 20-minute phone survey with the position's hiring manager. Baseline questions cover details about the firm, the advertised job, the hiring manager, and questions on both personal attitudes and societal norms on gender in the workplace.

After the baseline survey, firms were randomized into the treatment and control groups. There are a total of 1,227 firms that completed the baseline survey and were randomized, of which 657 firms are in the treatment group and 570 firms are in the control group.[10] The sample is balanced across the two groups on observable characteristics at baseline (Table SA1).

The intervention was administered to treatment firms within one week after the baseline. We were able to reach 547 (83%) out of the 657 firms in the treatment group for the intervention.[11] Nearly 80% of these 547 firms (and 67% of the 657 firms in the treatment group) were offered the subsidy.

A hiring follow-up survey (HFS) was administered between May 2022 and August 2023 with 1,054 of the 1,227 firms. We made the first attempt to contact firms approximately one month after their baseline. If they had not yet made a hire, we continued follow-ups until they either filled the position or decided not to do so. The hiring follow-up survey asked whether the position was filled, the gender of the hired person, reasons for hiring a man versus a woman, and how many people they interviewed for the post by gender. This data was complemented by Rozee.pk portal data on

[10] We initially used the mean of a uniformly distributed random number to assign firms to treatment and control groups, which resulted in an unequal distribution favoring the treatment group. To ensure balance, we subsequently adopted a median split, which produced an equal allocation across groups.

[11] 566 agreed to be surveyed, but 19 had already hired in between the baseline and the intervention call; 91 were not reached or refused to participate in a second survey.

applicants who applied via the portal.

We also conducted a survey with the women hired through the subsidy and an endline survey with all sample firms. Details for these surveys are in sections 5.2 and 6.1. Finally, we obtained a dataset of 25,963 job advertisements posted from 2020 to 2024 by the firms in our experimental sample on the Rozee.pk job portal.

4.2 Findings from Becker-DeGroot-Marschak (BDM)

At baseline, all firms were asked, without incentivizing, the question outlined in Section 3.3 to gauge the minimum subsidy required to hire a woman. The average stated subsidy was 25.87%, roughly 10 percentage points higher than the incentivized responses under the BDM mechanism, which averaged 15% (Figure SA2). This gap underscores the value of BDM in eliciting more truthful (and lower), subsidy requirements from managers. It provides an empirically driven estimate of the level of subsidy needed to encourage hiring of women while previous studies have generally evaluated the impact of fixed subsidy levels (Groh et al., 2016; Abel et al., 2022).

One concern with the incentivized BDM is that managers might not have understood well the lottery mechanism. We do not find evidence of this from the survey questions asked after the intervention. The weakly dominant strategy in the BDM is to bid a percentage no more than the subsidy needed. To assess the credibility of responses, a LASSO regression was employed to identify firm and manager characteristics that correlate with the incentivized subsidy amount. Firms with gender separated toilet facilities and managers with less conservative perceptions of societal gender norms tended to report lower required subsidies. These patterns support the validity of the BDM responses and suggest they reflect meaningful constraints rather than random inputs.

4.3 Estimation strategy

Our main estimation strategy is the following:

$$Y_i = \alpha + \beta_1 * Treated_i + \epsilon_i \tag{2}$$

where Y_i is the outcome of interest for firm i and $Treated_i$ is an indicator for the firm being offered the treatment. The parameter of interest is β, i.e. the intent-to-treat estimate of a firm being in the treatment group. ϵ_{it} is a random error term.

There is 15% attrition in the hiring follow-up survey, and this is differential across the treatment and control groups by 3.8 percentage points (Table SA4). Therefore, as pre-specified, we present the lower and upper limits of treatment effects using the bounding procedure proposed in Lee (2009).

5 Short Term Effects

5.1 Hiring of women

We find a large and statistically significant effect of the treatment on the primary pre-specified outcome of interest: the intent-to-treat likelihood of hiring a woman. Firms in the treatment group are 10.7 percentage points more likely to hire a woman compared to those in the control group (Table 2, Panel A Column 1). This effect represents a meaningful shift in hiring behaviour. In column 2, we report the upper and lower bounds of treatment effects constructed following Lee (2009) to account for non-random attrition. The bounds, ranging from 8.6 to 13.3 percentage points, do not cross zero, indicating that the positive effect of the treatment on the likelihood of hiring women is robust. The result is also robust to the inclusion of controls (Table 2, Column 3). In Column 4, we further validate the robustness of the findings by employing Post-Double Selection LASSO to select the optimal set of control variables based on their predictive power for both treatment assignment and the outcome of interest.

Next, we go beyond intent-to-treat effects to understand how the impact varies by the minimum subsidy level required. The BDM elicits this level and then randomises whether the subsidy is offered or not, conditional on the subsidy level required. We estimate the linear treatment-on-the-treated effects by instrumenting the indicator for winning the subsidy by the random subsidy level drawn for those participating in the BDM, and condition on the elicited subsidy level.[12] We report these for the full sample, as well as separately for male-only and mixed-gender firms in Table SA5. We have a strong first stage (Table SA5 Panel A) and find larger effects of 20.4 percentage points on the likelihood of female hiring (Table SA5 Panel B) as compared to the ITT estimate of 11 percentage points. There are even larger (28 pp.) for male-only firms (Panel B, Column 2).

Similar to Berry et al. (2020), our sample size is not large enough to estimate the impact of being offered a subsidy (i.e., "winning") at each individual subsidy bid level. Therefore, we employ a grid based on the subsidy required as elicited through the BDM mechanism and estimate kernel-weighted linear two-stage least squares models. As expected, the first-stage results indicate that the randomly assigned subsidy level is positively correlated with the likelihood of being offered a subsidy (results not shown). The second-stage kernel IV estimates reveal that the point estimates are positive across most of the distribution of BDM-elicited subsidy requirements except beyond subsidy levels of 38% (Figure SA3).[13] The estimates are statistically insignificant at all required subsidy levels except showing marginal significance around 25%, after which the perceived cost may be sufficiently high that the subsidy no longer incentivizes the hiring of women.[14] We observe a clear downward trend in the estimated impact on the likelihood of hiring a woman as the required subsidy level increases.

[12] These are set to 0 for everyone who did not participate in the BDM.
[13] We estimate impacts between 10% and 40% subsidy bid levels, using 1% intervals, to minimize boundary bias.
[14] The effects are statistically significant between 15% and 25% subsidy levels when controls are included.

Table 2: Impacts on the Hiring of Women and the Hiring Process

Panel A: Hiring of Women

	(1)	(2)	(3)	(4)
Treated	0.107***		0.121***	0.118***
	(0.031)		(0.030)	(0.030)
Lower Bound		0.086***		
		(0.033)		
Upper Bound		0.131***		
		(0.032)		
Controls	No	No	Yes	Yes
Control mean	0.433	0.433	0.429	0.429
# obs	1,054	1,227	1,033	1033

Panel B: Hiring Process

	(1) Share of Women in Total Applicants	(2) Share of Women in Total Interviewed	(3) Log Salary	(4) Formal Contract
Treated	0.007	0.026	0.074*	0.053*
	(0.009)	(0.023)	(0.038)	(0.032)
Control mean	0.177	0.392	10.6	0.608
# obs	1,213	574	789	912

Notes: In Panel A the dependent variable in all columns is a binary indicator equal to 1 if the manager reported hiring a woman in the hiring follow-up survey, and 0 if they hired a man or ceased recruitment for the position. In both Panel A and B, Treated is an indicator variable for firms randomized to be offered the subsidy. The Lower Bound and Upper Bound in column (2) refer to the treatment effect bounds constructed using the Lee (2009) procedure. Results in column (3) include the following controls: frequent interaction with public, firm size, share of female employees, large city, male firm only, manager age, gender, marital status, education, norms score, job education requirement, industry dummies. In column (4) controls are selected using PDS lasso. All specifications use robust standard errors. In Panel B, he dependent variable in column 1 is the ratio of female applicants to total applicants, in column 2 the ratio of female applicants who were interviewed to the total number of candidates interviewed, in column 3 is the log of the final monthly salary for the hired candidate and in column 4 is an indicator variable of whether the hired candidate was issued a formal contract. All specifications use robust standard errors. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

5.2 Women hired on a subsidy

Of the 247 firms eligible for the subsidy, 131 (53%) completed the required paperwork. The primary reason firms failed to receive the subsidy was their inability or unwillingness to submit the required documentation.[15,16] While the rate of take-up of the subsidy is not very high, it is still sizeable compared to earlier work by Galasso et al. (2004) and Levinsohn et al. (2014), which found that almost no firms took up the subsidy offer due to not wanting to face labor regulations. We surveyed the hired women from these subsidy-receiving firms one month after the disbursements were completed. The primary motive of these surveys was to understand whether the women were retained after the subsidy period ended and if there were any unintended adverse consequences of providing the subsidy. We reached 112 women (86%) out of the 131 women hired in a subsidy-receiving firm. Of these, 31 (about 28%) were no longer at the same company, although 10 women (9%) reported leaving for better employment opportunities and an additional 10 women reported leaving for family obligations. Nearly all women said they were treated well by colleagues during the subsidy period, and only 2 reported not being treated well after the subsidy ended.

5.3 Mechanisms

Building on the conceptual framework in Section 2, we test whether the intervention changed the firms' hiring processes (search, screening, quality of jobs offered) and whether the impacts vary with job and firm characteristics.

5.3.1 Hiring process

The effect of the wage subsidy offer does not appear to occur through changes in firms' hiring process. During the hiring follow-up survey, we collected information on the gender composition of interviewees, the salary range of the hired candidates, and whether they were given a formal contract, supplemented by the gender composition of applicants from Rozee.pk database.[17] We prespecified these as secondary outcomes of interest. We find no treatment effects on the ratio of female to total applicants (Table 2 Panel B Column 1). About 18% of the applicants are women in the control group. We would not expect this to change significantly as a result of the treatment unless winning a subsidy results in firms actively encouraging women to apply. We conducted the intervention *after* the firms had advertised but *before* the candidate was hired and there is no evidence to show that treatment firms solicited more female applications.

[15]The required documents included the candidate's offer letter, bank account details to make the transfer and the company's National Tax Number (NTN) or Computerized National Identity Card (CNIC, if NTN not available). However, these criteria, other than the bank account number, were relaxed to "if available" to not dissuade informal firms from participating.

[16]One possible explanation is that hiring managers, rather than business owners, were responsible for the process and may have lacked the authority or motivation to complete the additional steps. Alternatively, the financial incentive may not have been compelling enough to justify the effort, highlighting challenges to scaling up such an intervention.

[17]This only includes those who applied via Rozee.pk. Candidates could have also directly applied for the job.

Since there are no treatment effects on the gender composition of applicants, we check how this interacts with treatment. We find that treated firms that have above median number of female applicants (>4 female applicants) have significantly larger likelihood of hiring a woman (Figure 1). This may reflect lower search costs when firms receive more female applicants.

There is no significant impact of the treatment on the share of women in the total number of candidates who were interviewed for the job (Table 2 Column 2). Hence, it does not appear that treated firms are adjusting their behavior and interviewing more women than men in order to find a suitable female candidate. Consistent with the rates of hiring of women we see in our dataset, the average share of women being interviewed is around 40%. We only have data on the gender breakdown of candidates interviewed for 574 firms due to a survey skip error. Of these 574 firms, 250 are from the control group and report female hiring rates (48%) close to the overall rate in the control group of 43%, suggesting that our data on the gender composition of interviewees do not come from a highly selective subset. Data also show internal consistency: firms that ultimately do not hire a woman interview a lower share of female candidates (24%), while those that do hire a woman interview a significantly higher proportion (52%).

We examine two aspects of job quality, salary and formal contracts, and find modest improvements due to the subsidy (Table 2, Columns 3-4). Salary data was available for about 800 firms, showing a marginally significant 7.4% increase in hiring salary. This suggests firms may have passed on some of the subsidy, possibly due to high reservation wages or salary-sensitive labor supply. Additionally, the likelihood of offering a formal contract rose by 5.3 percentage points ($p = 0.100$), despite no requirement to do so.

5.3.2 Job and firm characteristics

We proxy cross-role variation in job characteristics through measures of skill intensity, compensation level, and customer-facing involvement. Treated firms are more likely to hire a woman in lower-skill and lower-pay roles: the treatment effect is larger by 17 percentage points for jobs that do not require an undergraduate degree, and by 10 percentage points (marginally significant) for jobs with below-median salaries ($<$ PKR 50,000) (Figure 1).[18] At lower salary levels, the perceived downside risk of a mismatch is likely smaller, which can make a modest incentive more effective; by contrast, in higher-skill, higher-pay posts, we do not observe comparable increases in female hiring. We also examine the largest sector in our sample – ICT – and find no differential response relative to other sectors.

Turning to firm characteristics, we use baseline gender composition and the presence of organizational complements (e.g., amenities, HR processes) to proxy for prior exposure/experience.

[18]Salary information is collected in categorical ranges from below PKR 20,000 to above PKR 100,000 with PKR 5,000 increments at baseline; we use the upper bound of the reported range. This baseline salary may differ from the portal-stated salary used for study eligibility.

About 19% of firms had no female employees at baseline. We find that male-only firms exhibit a 17 percentage-point larger treatment effect relative to firms with any female employees at baseline (Figure 1). Mixed-gender firms also respond positively, though the magnitude is smaller (about 7 percentage points).[19] Descriptively, male-only firms are smaller on average (about 16 employees) than mixed-gender firms (about 140 employees), but we do not find systematic heterogeneity by firm size per se.

We do not find any evidence of heterogeneity in treatment effects by other characteristics such as manager characteristics, including their perceptions of societal beliefs about women working and un-incentivized subsidy levels needed to hire a woman (Figure SA4). To probe higher-dimensional patterns, we use 'honest' causal random forests, and appendix Figure SA5 plots the distribution of Conditional Average Treatment Effects (CATEs) across firms (given firm, job-posting, and manager characteristics), which confirms meaningful heterogeneity. Manager beliefs and the baseline share of female employees are identified as the two most important splitters (Figure SA6). Treatment effects display a U-shaped relationship with the manager's gender-attitude index: the most progressive and the most regressive managers respond strongly to the subsidy, while the middle tercile shows little change. One interpretation consistent with the framework is that highly progressive managers face modest perceived barriers, so even a small incentive moves the margin; highly regressive managers face larger barriers, but the subsidy is sufficient to induce a first hire (and may facilitate learning), while moderately regressive managers face barriers of both kinds such that a partial subsidy is insufficient. Consistent with this view, treatment effects are also larger in firms with the lowest baseline share of female employees, mirroring the result for male-only firms.

6 Longer Term Effects

6.1 Endline survey

An endline survey was conducted between October 2023 and March 2024, approximately 1.5 years after the baseline and after the full 6 months for all subsidy-receiving firms finished. Out of 1,227 sample firms, we reached only 329 (27%) while 34% of the firms refused the survey and 39% were unreachable. High attrition may stem from Pakistan's worsening economic conditions, widespread firm closures post-COVID, and the inherent difficulty of firm-level follow-ups. Using Rozee.pk portal data from January 2020 to May 2024, we confirm a contraction in job advertisements (Figure SA8) around the time of our endline, reaching levels observed during the peak of the pandemic. Due to the high attrition and challenges reaching firms, we did not register an updated pre-analysis plan for this round. Nevertheless, attrition is balanced, (Table SA6) and we interpret the results with caution.

[19] We assess whether the stronger response among male-only firms is driven by a higher propensity to advertise lower-paid, lower-education jobs; this does not appear to be the case.

Figure 1: Heterogeneous Treatment Effects on Female Hiring by Firm and Job Characteristics



Note: Each point corresponds to the coefficient of Treatment x Dimension at Baseline. Treatment is an indicator variable for firms that are randomized to be offered the subsidy. The dependent variable is an indicator variable for if a woman has been hired. Frequent Interaction with Public is an indicator for the job requiring frequent public interaction. Requires Undergraduate or below is an indicator for the job requiring less than an undergraduate degree. Below Median Salary is an indicator for the job posting having a below median ($<$ PKR 50,000) upper salary level. Information Technology Firm is an indicator for the firm being in IT sector. Above Median Firm size is an indicator for firms with more than 16 employees. Male Only Firm is an indicator for firms with only male employees. Above Median Share of Female Employees is an indicator for firms with above median ($>$ 21%) female to total employees in the firm. Above Median Number of Female Applicants is an indicator for firms with above median ($>$ 4) female applicants to the job posting.

We find a 13 percentage point greater likelihood of a woman continuing to work in the advertised position in the treatment group (Table SA7).[20] However, we do not find evidence of broader changes at the firm level – there is no significant impact on the share of women in overall firm employees, in the share of women in recent hires (last 6 months), or in the share of women in the respondent manager's team.

To examine the role of team dynamics to explain the lack of firm-wide effects,we look at impacts on turnover, cooperation, workflow efficiency, and manager job satisfaction (Table SA8). None of these outcomes differ between treatment and control firms, so we find no evidence of adverse spillovers. The more plausible explanation is structural: the program subsidized and verified only one vacancy, removing the barrier for that position but leaving the remaining fixed costs and norm-based frictions untouched. In short, the subsidy can tip a single woman over the threshold, yet shifting the overall gender composition would require either subsidizing additional posts or

[20] We also asked for how long the hired employee stayed in the position and find that it was for an average of around 7.5 months. There is no significant difference in this between the treatment and control groups.

pairing the wage support with an intervention that addresses residual firm-level constraints.

6.2 Job advertisements

To test for longer-run impacts of the treatment on gender preferences, we analyze administrative data from Rozee.pk on the universe of job postings listed by the companies in our sample from January 2020 to April 2024. There are 25,863 job advertisements during this period by the sample firms, and the characteristics of these jobs are described in Table SA9. As is common in Pakistan and in many developing country labor markets, firms are free to specify a preference for a man (14%), a woman (7%) or no gender preference (79%), as described in Gentile et al. (2023).

The primary outcomes of interest are indicators for whether firms specify a preference for hiring a man ("male preference") and whether firms specify a preference for hiring a woman ("female preference"). We argue that the frequency with which firms choose to actively advertise for a man or a woman are an indicator of gender preferences, and possibly bias, at the firm.

Because we observe job postings for firms in our sample before, during and after the study, we are able to estimate the impacts of being randomized into treatment on these outcomes using a difference-in-difference design. We augment equation 2 with $Post_{it}$, an indicator for whether the job advertisement was posted after the baseline date[21] and the interaction of $Treated_i$ and $Post_{it}$ which captures the impact on job advertisement being posted after the baseline. We include year and month fixed effects and control for job post characteristics.

We find that assignment to treatment reduces the likelihood that firms express a male preference in job postings and that this is robust to the inclusion of a rich set of controls (Table 3, columns 1-3). Point estimates suggest that assignment to treatment reduces the likelihood of male preference in a job advertisement by 5.1 - 5.6 percentage points, which reflects a 36% - 40% decline relative to the control mean. When looking at male-only firms (column 4), the effects are larger, with assignment to treatment leading to an 8.1 percentage point reduction. While there are no overall effects on firms expressing a female preference in a job advertisement, we see some evidence for male-only firms with a 7 percentage point significant increase in female preference. The treatment on the treated effects are even larger, suggesting that the effects are driven by firms that were offered a subsidy (discussed in Supplemental Appendix 7).

7 Discussion

We test whether temporary wage subsidies can increase women's employment in a field experiment in Pakistan. The intervention raised female hiring by 10.7 percentage points, with larger effects

[21] Although there was some lag between the baseline and the intervention, we use the baseline date for all firms to define the "post" indicator because it is observed for all firms, whereas the intervention date is only observed for treated firms who participated in the intervention.

Table 3: Treatment Effects on Gender Preference in Job Advertisements

Panel A: Male preference				
	(1)	(2)	(3)	(4)
Post	0.018	0.044**	0.045**	0.133**
	(0.016)	(0.020)	(0.020)	(0.060)
Post*Treated	-0.051**	-0.051**	-0.056**	-0.081*
	(0.025)	(0.025)	(0.025)	(0.046)
Control mean	0.14	0.14	0.141	0.108
Number of Obs	25,484	25,484	25,354	1932
Company Fixed Effects	Yes	Yes	Yes	Yes
Year and Month Fixed Effects	No	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Sample	All	All	All	Male Only Firms
Panel B: Female preference				
	(1)	(2)	(3)	(4)
Post	0.010	0.016	0.012	0.009
	(0.012)	(0.014)	(0.014)	(0.022)
Post*Treated	-0.003	-0.002	0.002	0.069**
	(0.015)	(0.015)	(0.015)	(0.032)
Control mean	0.072	0.072	0.072	.0367
Number of Obs	25,484	25,484	25,354	1932
Company Fixed Effects	Yes	Yes	Yes	Yes
Year and Month Fixed Effects	No	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Sample	All	All	All	Male Only Firms

Notes: The dependent variable in Panel A is an indicator for a firm expressing a preference for a man in a given job advertisement. The dependent variable in Panel B is an indicator for a firm expressing a preference for a woman in a given job advertisement. Column 1 shows specifications including only firm fixed effects. Column 2 shows specifications adding month and year fixed effects, and column adds a further set of controls for job characteristics, including controls for above median salary, job type, education requirement, job shift and industry. Column 4 shows results for male-only firms at baseline. Treated is an indicator variable for firms randomized to be offered the subsidy, and this overall treatment assignment indicator is dropped from the specifications because all specifications include firm fixed effects. Post is an indicator of whether the job advertisement in question was listed after the baseline survey, and the coefficient of interest is on the interaction between Treated and Post. Male only firms refer to firms that did not have any female employees at the study baseline. All specifications use cluster robust standard errors, clustered at the firm level. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

among firms that initially employed few or no women. Evidence suggests retention of female hires 18 months later, without broader firm-wide changes in composition or attitudes, which may require a more intense intervention than one more woman being hired. Administrative data on longer-run job advertisements reveal shifts in hiring preferences by gender, indicating the potential for short-term subsidies to reduce hiring biases and expand women's employment at relatively low cost. Using a Becker–DeGroot–Marschak (BDM) mechanism, we find that firms require, on average, a 15 percent subsidy to hire a woman, smaller than in previous studies. Future research should examine the effectiveness of such incentives among firms explicitly preferring men and explore general-equilibrium and behavioral spillovers from wage subsidy programs.

References

Abel, M., Carranza, E., Geronimo, K., and Ortega, M. (2022). Can temporary wage incentives increase formal employment? experimental evidence from mexico. *IZA Discussion Paper No. 15740*.

Afridi, F., Dhillon, A., Roy, S., and Sangwan, N. (2022). Social networks, gender norms and women's labor supply: Experimental evidence using a job search platform. *IZA Discussion Paper No. 15767*.

Ahmed, H., Mahmud, M., Said, F., and Tirmazee, Z. (2024). Encouraging female graduates to enter the labor force: Evidence from a role model intervention in pakistan. *Economic Development and Cultural Change*, 72(2):919–957.

Armand, A., Carneiro, P., Tagliati, F., and Xia, Y. (2026). Can subsidized employment tackle long-term unemployment? experimental evidence from North Macedonia. *Journal of Development Economics*, 178:103598.

Bandiera, O., Barankay, I., and Rasul, I. (2011). Field experiments with firms. *Journal of Economic Perspectives*, 25(3):63–82.

Becker, G., Degroot, M., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232.

Becker, G. S. (1957). *The economics of discrimination*. chicago: University of Chicago.

Berry, J., Fischer, G., and Guiteras, R. (2020). Eliciting and utilizing willingness to pay: Evidence from field trials in northern Ghana. *Econometrica*, 128(4).

Borker, G. (2025). Safety first: Perceived risk of street harassment and educational choices of women. Conditionally Accepted, American Economic Review.

Bruhn, M. (2020). Can wage subsidies boost employment in the wake of an economic crisis? Evidence from Mexico. *The Journal of Development Studies*, 56(8):1558–1577.

Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review*, 110(10):2997–3029.

Bussolo, M. et al. (2024). Evaluating demand side constraints to female employment among firms in Pakistan through experimental provision of a targeted wage subsidy. AEA RCT Registry. October 12. https://doi.org/10.1257/rct.9226-4.0.

Compton, J. and Pollak, R. A. (2014). Family proximity, childcare, and women's labor force attachment. *Journal of Urban Economics*, 79:72–90.

Coville, A., Graff Zivin, J., Reichert, A., and Reitmann, A.-K. (2025). Quality signaling and demand for renewable energy technology: Evidence from a randomized field experiment. *Journal of Development Economics*, 176:103514.

de Mel, S., McKenzie, D., and Woodruff, C. (2019). Labor drops: Experimental evidence on the return to additional labor in microenterprises. *American Economic Journal: Applied Economics*, 11(1):202–235.

Dean, J. T. and Jayachandran, S. (2019). Changing family attitudes to promote female employment. *AEA Papers and Proceedings*, 109:138–142.

Dizon-Ross, R. and Jayachandran, S. (2023). Detecting mother-father differences in spending on children: A new approach using willingness-to-pay elicitation. *American Economic Review: Insights*, 5(4):445–59.

Galasso, E., Ravallion, M., and Salvia, A. (2004). Assisting the transition from workfare to work: A randomized experiment. *ILR Review*, 58(1):128–142.

Garlick, R., Field, E., and Vyborny, K. (2025). Women's mobility and labor supply: Experimental evidence from Pakistan. *IZA Discussion Paper No. 17883*.

Gentile, E., Kohli, N., Subramanian, N., Tirmazee, Z. S., and Vyborny, K. (2023). Barriers to entry: Decomposing the gender gap in job search in urban pakistan. Technical report, ADB Economics Working Paper Series.

Groh, M., Krishnan, N., McKenzie, D., and Vishwanath, T. (2016). Do wage subsidies provide a stepping-stone to employment for recent college graduates? Evidence from a randomized experiment in Jordan. *Review of Economics and Statistics*, 98(3):488–502.

Heath, R., Borker, G., Fitzpatrick, A., Keats, A., McKelway, M., Menzel, A., Molina, T., and Sharma, G. (2024). Female labour force participation. *VoxDevLit*, 11.

International Labour Organization (2025). ILO Modelled Estimates and Projections (ILOEST) database. *ILOSTAT*. Accessed January 07, 2025.

Jaenichen, U. and Stephan, G. (2011). The effectiveness of targeted wage subsidies for hard-to-place workers. *Applied Economics*, 43(10):1209–1225.

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review*, 69(3):576–595.

Jensen, R. (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.

Katz, L. F. (1996). Wage subsidies for the disadvantaged. *NBER Working Paper 5679*.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3):1071–1102.

Levinsohn, J., Rankin, N., Roberts, G., Schöer, V., et al. (2014). Wage subsidies and youth employment in South Africa: Evidence from a randomised control trial. *Stellenbosch Economic Working Papers: 02/14*, 2:14.

Matsuda, N., Ahmed, T., and Nomura, S. (2019). Labor market analysis using big data: The case of a Pakistani online job portal. *World Bank Policy Research Working Paper*, (9063).

McKelway, M. (2025). Women's self-efficacy and economic outcomes: Experimental evidence from

India. *Unpublished Working Paper*.

McKenzie, D. (2017). How effective are active labor market policies in developing countries? A critical review of recent evidence. *The World Bank Research Observer*, 32(2):127–154.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

World Bank (2024). Gender data portal: Pakistan. Available at: https://genderdata.worldbank.org/en/economies/pakistan; Accessed: 2025-07-29.

World Economic Forum (2020). Global Gender Gap Report 2020. Technical report, World Economic Forum.

Supplemental Appendix

Title: Wage Subsidies to Promote Female Hiring: Evidence from Pakistan

Authors: Maurizio Bussolo, Jean Lee, Mahreen Mahmud, Nayantara Sarma, Anaise Williams

A Additional Tables and Figures

Table SA1: Baseline Summary Statistics by Treatment Status

Variable	Control	Treatment	Treatment - Control
Firm size	112.530	116.816	4.286
	(795.150)	(769.972)	(44.749)
Share of female employees	0.229	0.238	0.010
	(0.197)	(0.199)	(0.011)
Male only firm	0.204	0.184	-0.019
	(0.403)	(0.388)	(0.023)
Main city	0.935	0.932	-0.004
	(0.247)	(0.253)	(0.014)
Industry - Communications	0.106	0.107	0.001
	(0.308)	(0.309)	(0.018)
Industry - Finance	0.106	0.098	-0.008
	(0.308)	(0.297)	(0.017)
Industry - Health, Education, Admin	0.088	0.076	-0.012
	(0.283)	(0.265)	(0.016)
Industry - Manufacturing and Retail	0.096	0.099	0.002
	(0.296)	(0.299)	(0.017)
Industry - IT	0.271	0.264	-0.007
	(0.445)	(0.441)	(0.025)
Industry - Not Reported	0.118	0.140	0.022
	(0.323)	(0.348)	(0.019)
Industry - Other	0.215	0.216	0.002
	(0.411)	(0.412)	(0.024)
Separated toilets by gender	0.777	0.791	0.014
	(0.416)	(0.407)	(0.024)
Childcare services or creches	0.147	0.164	0.017
	(0.355)	(0.371)	(0.021)
Frequent interaction with public	0.477	0.447	-0.031
	(0.500)	(0.498)	(0.029)
Salary (upper limit PKR)	58,789.473	57,587.520	-1,201.955
	(24,899.260)	(23,665.646)	(1,387.875)
Candidate BA education and above	0.663	0.674	0.011
	(0.473)	(0.469)	(0.027)
Transportation (or allowance)	0.582	0.519	-0.063**
	(0.494)	(0.500)	(0.028)
Health insurance	0.351	0.282	-0.069***
	(0.478)	(0.450)	(0.027)
Flexible working hours	0.700	0.665	-0.035
	(0.459)	(0.472)	(0.027)
Possibility to work from home	0.607	0.591	-0.016
	(0.489)	(0.492)	(0.028)
Age	32.128	31.976	-0.152
	(8.218)	(7.502)	(0.449)
Male	0.847	0.833	-0.015
	(0.360)	(0.374)	(0.021)
Married	0.618	0.604	-0.013
	(0.486)	(0.489)	(0.028)
Masters or above	0.504	0.516	0.012
	(0.500)	(0.500)	(0.029)
# obs	570	657	1,227

Note: Sample includes firms that completed the baseline survey and were randomized into treatment or control groups, regardless of whether the firm was reached for the intervention.

Table SA2: Our Sample Comparison to Rozee.pk No-Gender Preference Jobs

	(1) Other non-pref jobs (n=27,576)	(2) Study sample (n=1,227)	(3) Difference (1) - (2)	(4) Standard Error
Max salary	89813.32	45264.77	44548.55***	(4259.30)
Industry - Communications	0.12	0.11	0.01	(0.01)
Industry - Finance	0.11	0.10	0.01	(0.01)
Industry - Health, Education, Admin	0.11	0.08	0.03***	(0.01)
Industry - Manufacturing & Retail	0.09	0.10	-0.01	(0.01)
Industry - IT	0.28	0.27	0.01	(0.01)
Industry - Other	0.20	0.22	-0.01	(0.01)
Industry - Not Reported	0.00	0.13	-0.13***	(0.00)
Main city	0.91	0.93	-0.03***	(0.01)
Candidate BA education and above	0.79	0.67	0.12***	(0.01)
Job type- contract or freelance	0.05	0.02	0.03***	(0.01)
Job type- full time/ perm	0.91	0.94	-0.03***	(0.01)
Job type- part time	0.01	0.01	0.00	(0.00)

Note: The statistics reported in this table are based on administrative data from the Rozee.pk portal for the April 2022 - Nov 2022 period. The sample is restricted to job advertisements that do not specify the gender of the candidate. Other non-pref jobs refer to all job advertisements that did not specify a preferred gender but are not part of the study sample. The study sample refers to the characteristics of the firms which are part of this study. Salary information is missing for 53.6% for those not included in our sample, 0% for sample; Industry is missing for 10.5% for non-sample and 0.2% for those that are included in our sample. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

Table SA3: Job Characteristics by Gender Preference on the Advertisement

	All (n=39,801) 100 %	No preference (n=28,801) 72.5 %	Male (n=7,350) 18.3 %	Female (n=3,650) 9.2 %
Max salary	80879.55 (83996.31)	88199.15 (89761.32)	67348.57 (71851.74)	51926.23 (39397.55)
Industry - Communications	0.10 (0.31)	0.12 (0.32)	0.07 (0.25)	0.09 (0.28)
Industry - Finance	0.10 (0.30)	0.11 (0.32)	0.05 (0.22)	0.09 (0.29)
Industry - Health, Education, Admin	0.12 (0.32)	0.11 (0.32)	0.12 (0.32)	0.18 (0.39)
Industry - Manufacturing and Retail	0.12 (0.32)	0.09 (0.28)	0.26 (0.44)	0.11 (0.31)
Industry - IT	0.25 (0.43)	0.28 (0.45)	0.20 (0.40)	0.12 (0.32)
Industry - Other	0.20 (0.40)	0.20 (0.40)	0.19 (0.39)	0.23 (0.42)
Industry - Not Reported	0.00 (0.06)	0.01 (0.07)	0.00 (0.00)	0.00 (0.00)
Main city	0.90 (0.30)	0.91 (0.29)	0.86 (0.34)	0.93 (0.26)
Candidate BA education and above	0.75 (0.43)	0.79 (0.41)	0.63 (0.48)	0.70 (0.46)
Job type- contract or freelance	0.04 (0.20)	0.05 (0.21)	0.03 (0.16)	0.03 (0.17)
Job type- full time/ perm	0.92 (0.27)	0.91 (0.28)	0.95 (0.21)	0.94 (0.24)
Job type- part time	0.01 (0.10)	0.01 (0.11)	0.00 (0.07)	0.01 (0.11)

Note: The statistics reported in this table are based on administrative data from the Rozee.pk portal for the April 2022 - Nov 2022 period. All differences are statistically significant (p=0.001) across no pref/ male/female jobs. We use OLS for max salary and chi-squared test for binary measures. The industry is missing for 8.5% of the jobs with no pref, 12.1% for male-only jobs and 18.3% for female-only jobs.

Table SA4: Attrition in the Hiring Followup Survey

	(1)	(2)	(3)
Treated	-0.038*		-0.110
	(0.020)		(0.156)
Firm size		0.000	0.000
		(0.000)	(0.000)
Share of female employees		0.012	0.154
		(0.062)	(0.102)
Male only firm		0.050	0.080
		(0.034)	(0.051)
Main city		0.052	-0.049
		(0.036)	(0.069)
Industry - Communications		0.089**	0.025
		(0.040)	(0.066)
Industry - Finance		0.029	-0.109**
		(0.037)	(0.052)
Industry - IT		0.083**	-0.017
		(0.033)	(0.051)
Separated toilets by gender		-0.037	-0.102**
		(0.027)	(0.042)
Childcare services or creches		-0.003	0.001
		(0.028)	(0.043)
Age		0.004**	0.007**
		(0.002)	(0.003)
Manager Masters or above		-0.019	-0.053
		(0.021)	(0.034)
Treated=1 × Firm size			0.000
			(0.000)
Treated=1 × Share of female employees			-0.225*
			(0.129)
Treated=1 × Male only firm			-0.046
			(0.068)
Treated=1 × Main city			0.168**
			(0.077)
Treated=1 × Industry - Communications			-0.041
			(0.081)
Treated=1 × Industry - Finance			0.089
			(0.071)
Treated=1 × Industry - IT			0.017
			(0.065)
Treated=1 × Separated toilets by gender			0.115**
			(0.054)
Treated=1 × Childcare services or creches			0.002
			(0.057)
Treated=1 × Frequent interaction with public			0.106**
			(0.045)
Treated=1 × Salary (upper limit PKR)			-0.000
			(0.000)
Treated=1 × Candidate BA education and above			-0.066
			(0.047)
Treated=1 × Transportation (or allowance)			-0.004
			(0.044)
Treated=1 × Health insurance			0.000
			(0.048)
Treated=1 × Flexible working hours			-0.036
			(0.048)
Treated=1 × Possibility to work from home			-0.020
			(0.046)
Treated=1 × Age			-0.005
			(0.004)
Treated=1 × Manager Masters or above			0.068
			(0.043)
Control Mean	0.16	0.16	0.16
# obs	1227	1200	1200

Note: All columns show regressions of a firm-level indicator for not being surveyed at the hiring follow-up survey on treatment indicator (column 1); baseline characteristics (column 2); and a treatment indicator, baseline characteristics, and their interactions (column 3). Treated is an indicator variable for firms randomised to be offered the subsidy. Baseline characteristics are described in Table 1 notes. All specifications use robust standard errors. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Table SA5: Treatment on the Treated Effects on the Hiring of Women

	All firms	Male only firms	Mixed-gender firms
	(1)	(2)	(3)
Panel A	First-stage dependent variable: Won intervention		
BDM subsidy random draw	0.028***	0.030***	0.028***
	(0.001)	(0.001)	(0.001)
Kleibergen-Paap F-stat	3164.838	685.5853	2531.596
Panel B	Second-stage dependent variable: Hired a woman		
Won intervention	0.204***	0.283***	0.183***
	(0.036)	(0.082)	(0.040)
Control mean	.412	.264	.473
# obs	1054	193	861

Notes: The dependent variable in the first-stage (Panel A) is an indicator of whether the firm won i.e. was offered the subsidy. The dependent variable in the second-stage (Panel B) is an indicator variable for the manager reporting hiring a woman in the hiring follow-up survey, as compared to hiring a man or no longer looking to fill the position. BDM subsidy random draw is the value of the randomly drawn subsidy. This is set to 0 for firms that did not participate in the BDM. Column 2 restricts the sample to male-only firms at baseline and Column 3 restricts the sample to the female-only or mixed-gender firms at baseline. All specifications use robust standard errors. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Table SA6: Attrition in the Endline Survey

	(1)	(2)	(3)
Treated	-0.003	-0.003	0.253
	(0.025)	(0.026)	(0.177)
Firm size		0.000***	0.000***
		(0.000)	(0.000)
Share of female employees		0.019	0.068
		(0.084)	(0.117)
Male only firm		0.011	0.025
		(0.042)	(0.061)
Main city		0.063	0.120
		(0.056)	(0.087)
Industry - Communications		-0.001	0.094
		(0.047)	(0.070)
Industry - Finance		-0.059	0.048
		(0.049)	(0.072)
Industry - IT		-0.049	0.033
		(0.037)	(0.057)
Separated toilets by gender		-0.030	-0.006
		(0.032)	(0.048)
Childcare services or creches		-0.037	-0.077
		(0.037)	(0.057)
Age		-0.000	0.000
		(0.002)	(0.003)
Masters or above		-0.039	-0.042
		(0.028)	(0.043)
Treated=1 × Firm size			0.000
			(0.000)
Treated=1 × Share of female employees			-0.092
			(0.168)
Treated=1 × Male only firm			-0.023
			(0.085)
Treated=1 × Main city			-0.093
			(0.113)
Treated=1 × Industry - Communications			-0.172*
			(0.094)
Treated=1 × Industry - Finance			-0.195*
			(0.099)
Treated=1 × Industry - IT			-0.153**
			(0.075)
Treated=1 × Separated toilets by gender			-0.053
			(0.065)
Treated=1 × Childcare services or creches			0.079
			(0.076)
Treated=1 × Age			-0.001
			(0.004)
Treated=1 × Masters or above			0.006
			(0.057)
Control Mean	0.73	0.73	0.73
# obs	1227	1201	1201

Note: All columns show regressions of a firm-level indicator for not being surveyed at the endline survey on treatment indicator (column 1); baseline characteristics (column 2); and a treatment indicator, baseline characteristics, and their interactions (column 3). Treated is an indicator variable for firms randomised to be offered the subsidy. Baseline characteristics are described in Table 1 notes. All specifications use robust standard errors. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Table SA7: Treatment Effects on Firms, Endline

	(1) Position Filled by Woman	(2) Share of Female Employees	(3) Share of Women Hired to Total Hires	(4) Share of Women in Team
Treated	0.129**	-0.012	0.031	0.009
	(0.058)	(0.022)	(0.034)	(0.025)
Control mean	.417	.3	.316	.331
# obs	300	325	283	315

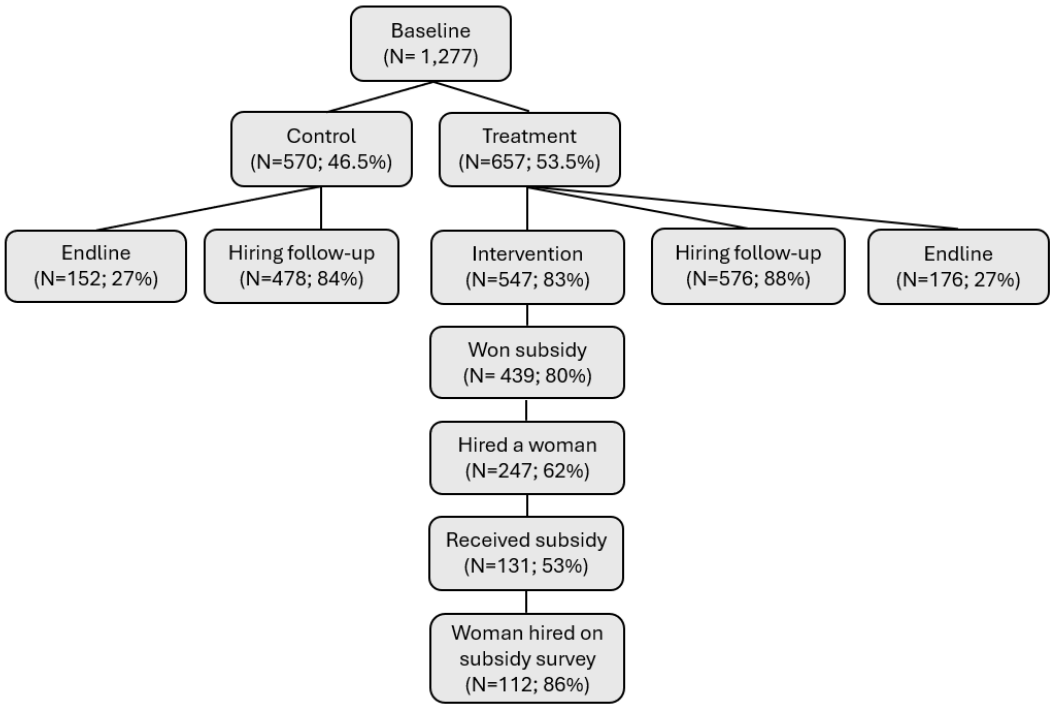
Note: Treated is an indicator variable for firms randomised to be offered the subsidy. Position filled by woman is an indicator for if the job advertised at baseline is filled with a woman. Share of female employees is the ratio of the total number of female employees to total number of employees. Share of women hired to total hires is the ratio of women hired as a proportion of all hiring in the last 6 months. Missing observations are due to firms not hiring in the last 6 months. Share of women in team is the ratio of women in the respondent managers' team. All specifications use robust standard errors. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Table SA8: Treatment Effects on Manager Responses, Endline

	(1) Team Performance Index	(2) Manager Well-being	(3) Perceived Norms	(4) Own Norms
Treated	0.034	-0.012	0.302	0.080
	(0.069)	(0.130)	(0.263)	(0.153)
Control mean	-.0183	1.96	29.4	21.6
# obs	327	328	328	328

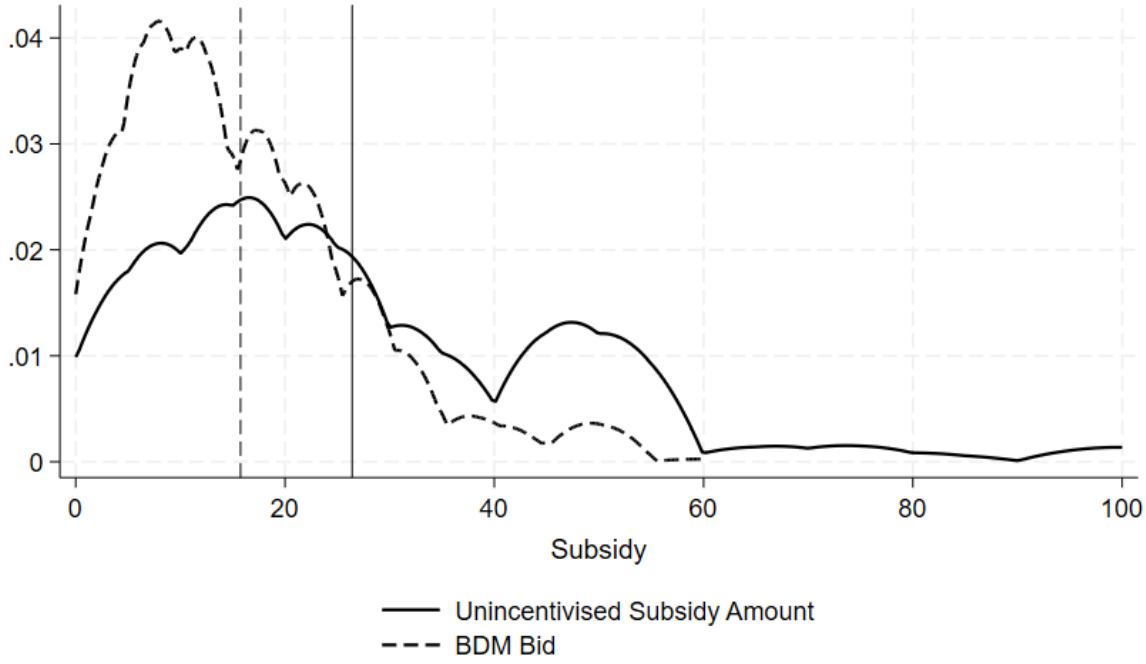
Note: Treated is an indicator variable for firms randomised to be offered the subsidy. Team performance index is a standardized index of manager's perception of how good the team is on the following dimensions: behavior with each other, honesty levels, sharing responsibility, and communication. Manager Wellbeing is a sum of indicator variables for the manager experiencing the following: conflict at work, stress, over-working, impact on sleep. Sum belief (higher prog.). Perceived norms are managers perceptions about what most people in society believe about women working outside the home (higher values indicate progressive perceptions). Own norms are manager's own beliefs about women working outside the home (higher values indicate progressive beliefs). Own norms and All specifications use robust standard errors. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Figure SA1: Study Design and Sample



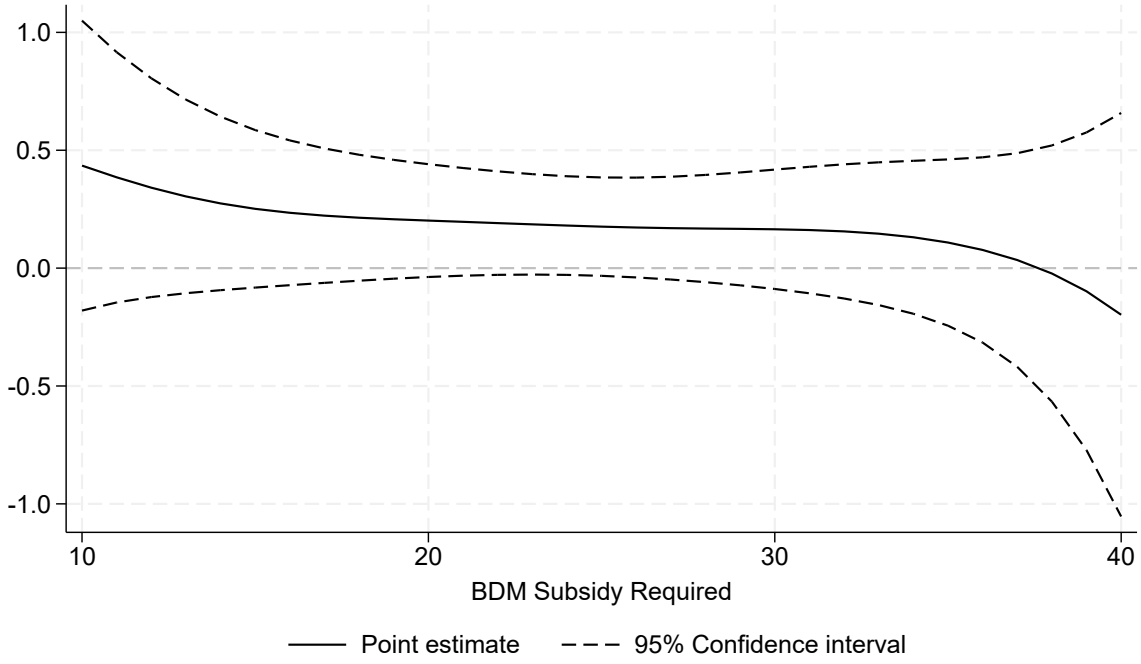
Note: The figure illustrates the study design and the sample size at each stage of data collection. The reported percentages represent the proportion of the sample size from the preceding level.

Figure SA2: Distribution of Unincentivized and Incentivized Subsidy (%) Required



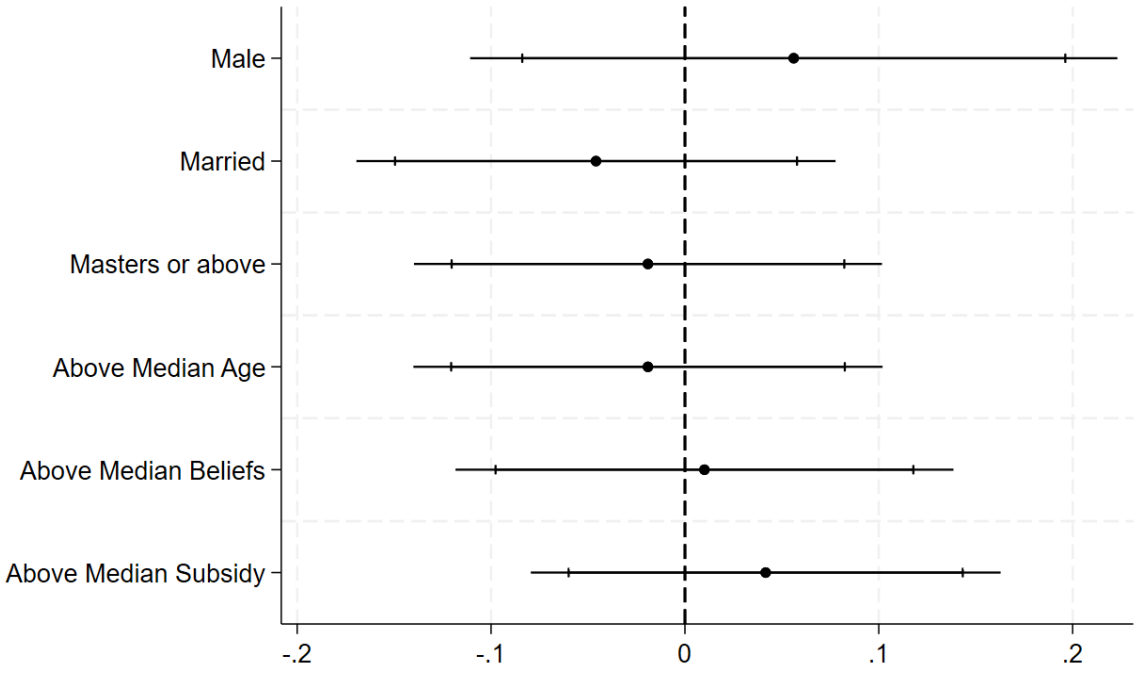
Note: The figure plots the unincentivised level of subsidy as a percentage of wage paid required to hire a female from the baseline survey and the incentivised level required (BDM Bid) from the intervention survey. The vertical lines show the mean value for each variable.

Figure SA3: Kernel IV Estimates of Hiring of Women



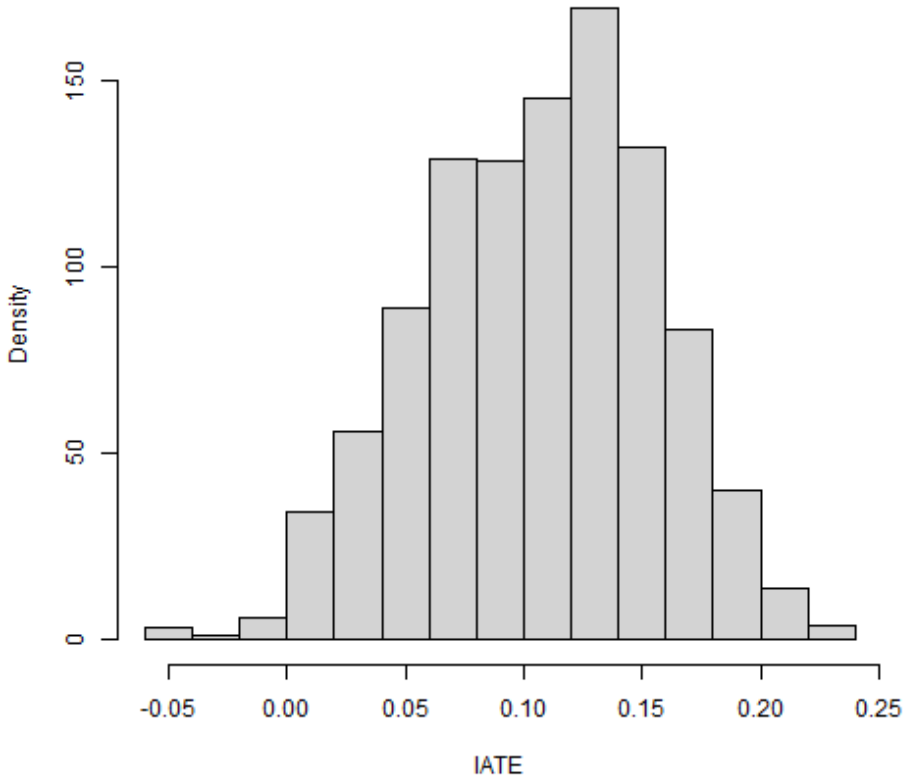
Note: The figure shows Kernel IV estimates of treatment effects on the probability of a woman being hired as a function of the subsidy level required elicited in the BDM. Estimates use linear two-stage least squares at the subsidy level required of 10 to 40 in intervals of 1, with observations weighted by their distance from this point using Epanechnikov kernel. The instrument is the randomly drawn subsidy in the BDM.

Figure SA4: Heterogeneous Treatment Effects on Hiring by Manager Characteristics



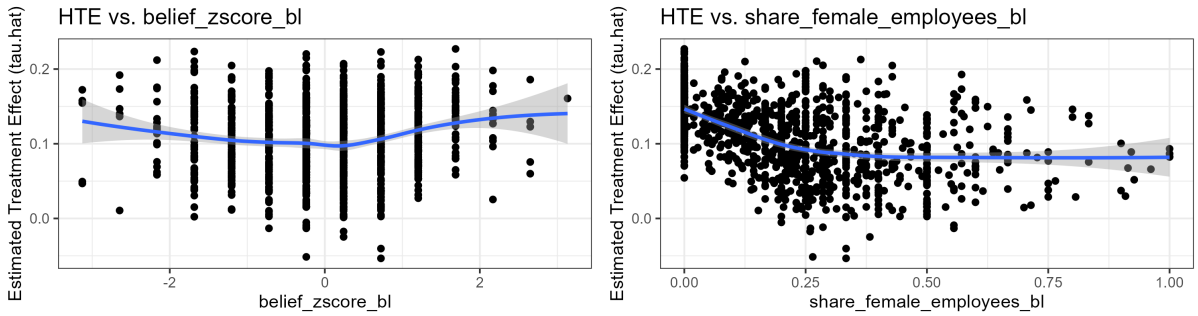
Note: Each point corresponds to the coefficient of Treatment x Dimension at Baseline. Treatment is an indicator variable for firms that are randomised to be offered the subsidy. The dependent variable is an indicator variable for if a woman has been hired. Male, Married, and Masters or above are indicator variables for the manager being male, married and having masters or above respectively. Above Median Beliefs are if they have above median beliefs about societal norms regarding women working. Above Median Subsidy is if the reported unincentivised subsidy value is higher than the median.

Figure SA5: Histogram of Firm Average Treatment Effects



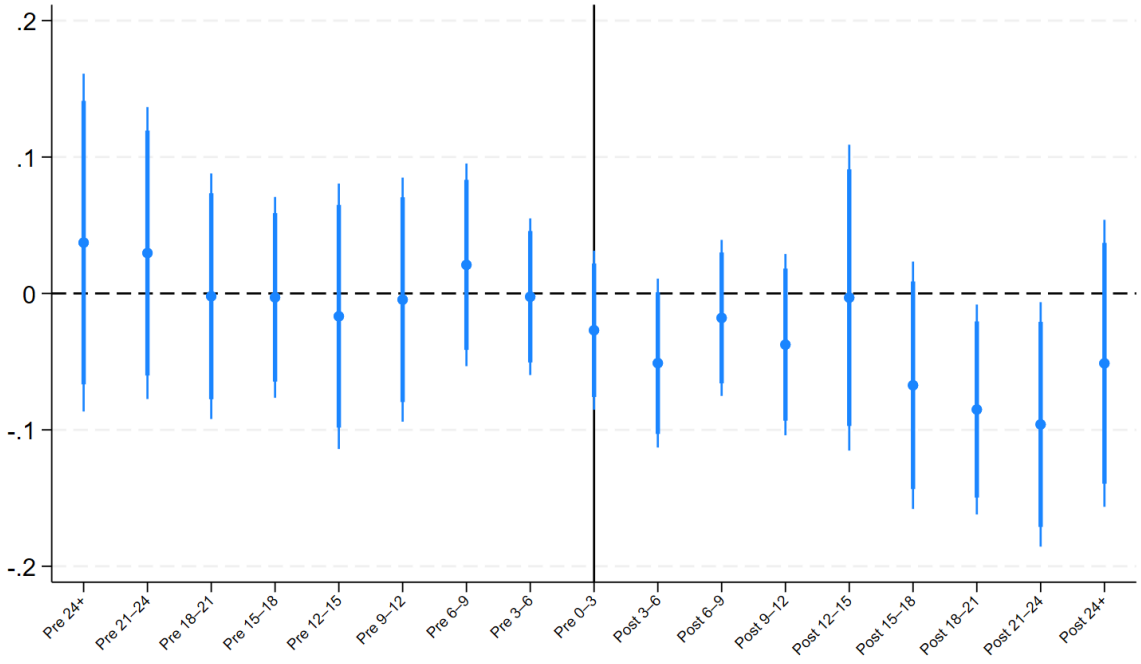
Note: The figure presents a histogram of individual firm-level treatment effect estimates (IATEs) obtained using the Generalized Random Forest (GRF) algorithm implemented by the causal_forest function in the grf R package. IATEs represent the conditional average treatment effect for each unit, given its covariate profile consisting of firm, manager, and job characteristics. Estimates are computed using out-of-bag predictions to mitigate overfitting with k-fold sample splitting (10 folds).

Figure SA6: Heterogeneous treatment effects over top 2 important covariates



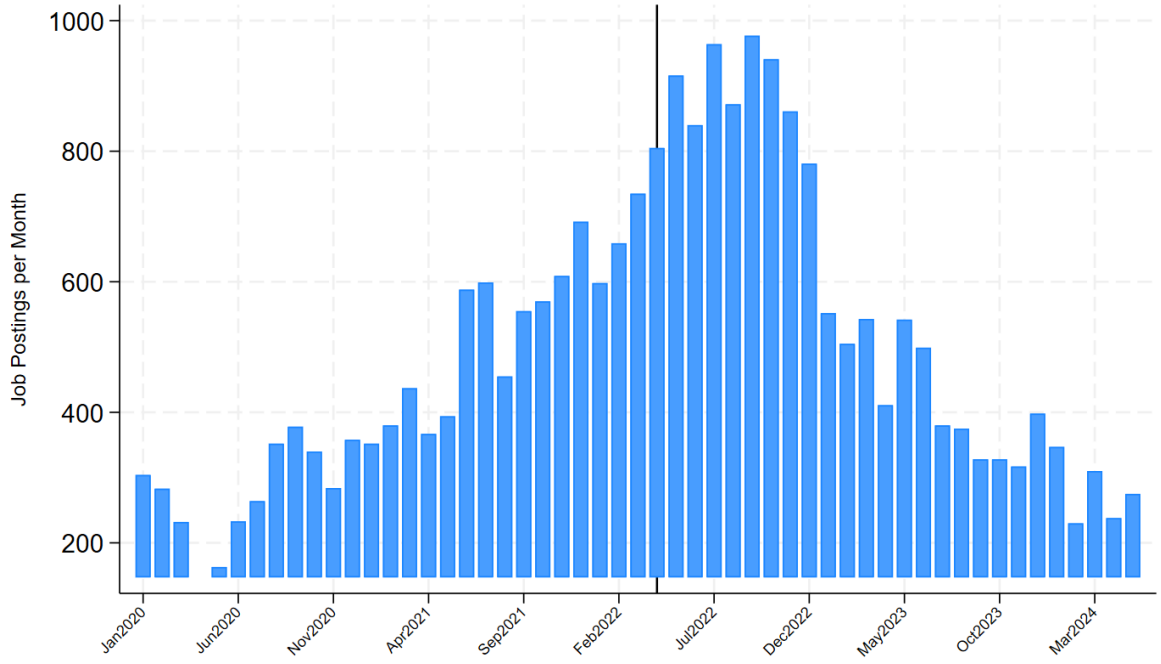
Note: The figure plots the estimated treatment effects over the distribution of the top 2 most important variables, identified by the weighted-sum of the number of times the variable was used to split nodes across all trees, reflecting how much it contributes to identifying treatment effect heterogeneity.

Figure SA7: Treatment Effects on Gender Preference in Job Advertisements



Note: The dependent variable is an indicator for a firm expressing a preference for a man in a given job advertisement. Each point corresponds to the difference in that 3-month period in the dependent variable value for the treated firms as compared to the control firms with the 3 month period after baseline during which the treatment firms received the intervention as the reference category, estimated using equation **??**. We show both the 90th and 95th confidence intervals, and the black line indicates the 3-month period before baseline.

Figure SA8: Job Postings on Rozee.pk



Note: The figure shows the number of job postings on the Rozee.pk portal each month from January 2020 to May 2024. The black line indicates the month when the intervention for the study started. Note that due to the Covid-19 pandemic, there were no job postings in April 2020.

B Rozee.pk Portal Job Advertisements Data

.1 Summary Statistics

Table SA9 presents summary statistics for all jobs on the Rozee.pk portal posted by the study sample of firms between January 2020 and May 2024.

Table SA9: Characteristics of Jobs Posted by Study Sample Firms (January 2020 - May 2024)

	Mean	SD	Min	Max	N
Candidate BA education and above	0.82	0.38	0.00	1	25,863
Full time	0.94	0.24	0.00	1	25,863
Entry level or intern position	0.31	0.46	0.00	1	25,837
Minimum required years of experience	2.24	2.33	0.00	30	25,813
Salary (PKR)	97788.47	1.2e+05	0.00	1000000	25,805
If a dedicated vehicle is available	0.08	0.27	0.00	1	3,465
If transport support is available	0.21	0.41	0.00	1	3,465
If flexible work arrangements are possible	0.30	0.46	0.00	1	3,465
If no overnight travel is required	0.75	0.43	0.00	1	3,465
Day shift	0.76	0.43	0.00	1	25,863
Communications	0.06	0.24	0.00	1	25,863
Health, Education, Admin	0.19	0.39	0.00	1	25,863
Finance	0.10	0.30	0.00	1	25,863
Manufacturing & Retail	0.14	0.35	0.00	1	25,863
IT	0.26	0.44	0.00	1	25,863
Other	0.23	0.42	0.00	1	25,863
Female friendly facilities	0.11	0.31	0.00	1	3,465
Female share of employees	0.27	0.18	0.00	1	25,808

Note: This table reports the characteristics of all jobs advertised between January 2020 and May 2024 by the firms that completed the study baseline survey. A firm is considered female-friendly if it offers 2 or more of the following: flexible work, transport, vehicle, or no overnight travel.

.2 Robustness of ITT estimates

Treatment-on-the-Treated Effects

To gauge the magnitudes of the impact of receiving a subsidy offer on firms' future gender preferences in hiring, we estimate treatment-on-the-treated (TOT) effects using a panel data instrumental variables specification. In particular, we instrument for whether or not a firm received an offer of a subsidy for hiring a woman and the interaction with the indicator for being after the baseline with the indicator for the assignment to treatment and the interaction with the indicator for being after the baseline.

We estimate the following TOT specification:

$$Y_{it} = \alpha_i + \beta SubsidyOffer_i + \gamma Post_{it} + \delta SubsidyOffer_i * Post_{it} + \epsilon_{it}$$

where $SubsidyOffer_i$ is an indicator for if the firm participated in the intervention and won a subsidy offer in the Becker-Degroot-Marschak lottery. We instrument for $SubsidyOffer_i$ and $SubsidyOffer_i * Post_{it}$ with $SubsidyAssignment_i$ and $SubsidyAssignment_i * Post_{it}$. The coefficient of interest is δ, the instrumented coefficient on the interaction between $SubsidyOffer_i$ and $Post_{it}$.

We find that as anticipated the instrumented TOT effects are larger than the estimated ITT effects (Table SA10), because the participation in the intervention was less than 100 percent and also because use of the BDM effectively lotteried some firms out of receiving the subsidy offer. Taking this into account, the effect of receiving a subsidy offer is to reduce male preference by 16.9 percentage points, or to effectively completely eliminate male preference relative to the control group mean. These effects are marginally statistically significant but very economically significant. There is no measured effect of the intervention on female preference on average. This is a striking result, and below in subsections 6.4 and 6.5, in the discussions of effects estimated from the variation within the BDM lottery and dynamic effects we explore whether or not this reflects effects that could be considered experimenter demand effects (e.g. if firms thought a future subsidy offer for women would be forthcoming) or if it reflects longer-term changes in attitudes and behaviors.

Table SA10: Treatment-on-the-treated Effects on Gender Preference in Job Advertisements

	Male preference			Female preference		
	(1)	(2)	(3)	(4)	(5)	(6)
SubsidyOffer*Post	-0.169*	-0.188*	-0.089*	-0.006	0.007	0.076**
	(0.098)	(0.101)	(0.050)	(0.050)	(0.049)	(0.034)
Post	0.049**	0.052**	0.132**	0.016	0.011	0.009
	(0.024)	(0.024)	(0.060)	(0.015)	(0.015)	(0.022)
Control mean	.14	.141	.108	.0722	.0719	.0367
Number of Obs	25484	25354	1932	25484	25354	1932
Company FE	Yes	Yes	Yes	Yes	Yes	Yes
Year and Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Sample	All	All	Male only	All	All	Male only

Note: The dependent variable in columns (1)-(3) is an indicator for whether or not the firm expressed a preference for a man in a given job advertisement. The dependent variable in columns (4)-(6) is an indicator for whether or not the firm expressed a preference for a woman in a given job advertisement. SubsidyOffer is an indicator variable for firms who were offered the subsidy, and this overall treatment assignment indicator is dropped from the specifications because all specifications include firm fixed effects. Post is an indicator for whether the job advertisement in question was listed after the baseline survey and the coefficient of interest is on the interaction between SubsidyOffer and Post. SubsidyOffer*Post is instrumented with SubsidyAssignment*Post, where SubsidyAssignment is an indicator for randomized assignment to the treatment arm. All specifications use cluster robust standard errors, clustered at the firm level. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.