Paternalistic Discrimination*

Nina Buchmann
Stanford University

Carl Meyer
Stanford University

Colin D. Sullivan
Purdue University

June 27, 2024

We combine two field experiments in Bangladesh with a structural labor model to identify *paternalistic discrimination*, the differential treatment of two groups to protect one group, even against its will, from harmful or unpleasant situations. We observe hiring and application decisions for a night-shift job that provides worker transport at the end of the shift. In the first experiment, we use information about the transport to vary employers' perceptions of job costs to female workers while holding taste-based and statistical discrimination constant: Not informing employers about the transport decreases demand for female labor by 21%. Employers respond more to transport information than cash payments to female workers that enable workers to purchase transport themselves. In the second experiment, not informing applicants about the transport reduces female labor supply by 15%. In structural simulations, paternalistic discrimination has a larger effect on gender employment and wage gaps than taste-based and statistical discrimination.

*Buchmann: nina.buchmann@stanford.edu, Meyer: carl.meyer@stanford.edu, Sullivan: sulli360@purdue.edu.
We are indebted to B. Douglas Bernheim, Pascaline Dupas, Marcel Fafchamps, Muriel Niederle, and Alessandra Voena for invaluable guidance and support on this project. We thank Gonzalo Arrieta, Maxim Bakhtin, Kayleigh Barnes, Katy Bergstrom, Adrian Blattner, Luisa Cefala, Arun Chandrasekhar, José Ignacio Cuesta, Ricardo De la O, Cauê Dobbin, Zach Freitas-Groff, Liran Einav, Matthew Gentzkow, Suhani Jalota, Diego Jiménez, Leticia Juarez, Judd Kessler, Michelle Layvant, Melanie Morten, Kirby Nielsen, Sebastián Otero, Milena Wittwer, Tom Zohar, and seminar participants from the BRAC Institute of Governance & Development, the European Central Bank, Fudan University, Harvard University, the Ohio State University, Peking University, Princeton University, Purdue University, Stanford University, UC Berkeley, UCLA, the University of Chicago, the University of Pittsburgh, the University of Southern California, Warwick University and Yale University for constructive comments. We are immensely grateful to Shakil Ayan, Giovanni D'Ambrosio, Utkarsh Dubey, Md. Al Hasib, Tanvir Mahatab, Alexander Marchal, and Zhongheng Qiao for outstanding research assistance and to all our study participants for their time and attention. The data collection was done with the assistance of the Development Research Initiative (dRi). The Stanford IRB and the Dhaka University Institute of Health Economics IRB approved the research protocols. The experimental designs were pre-registered in the AEA RCT registry (AEARCTR-0010971). We gratefully acknowledge funding from the Center for Effective Global Action (CEGA) through its Psychology and Economics of Poverty (PEP) Initiative, the Gender, Growth and Labor Markets in Low Income Countries programme of IZA and FCDO, the George P. Shultz Dissertation Support Fund at SIEPR, the International Growth Centre, the National Science Foundation, Pascaline Dupas, the Stanford King Center on Global Development's Graduate Student Research Funding, the 'Structural Transformation and Economic Growth' (STEG) initiative, and the Weiss Fund for Research in Development Economics at the University of Chicago. The views expressed are those of the authors and not necessarily those of the Foreign, Commonwealth & Development Office (FCDO). All errors are our own.

1 Introduction

Economists traditionally distinguish between two forms of labor market discrimination: *taste-based discrimination*, a simple preference for hiring one group over the other (Becker, 1957), and *statistical discrimination*, a belief that one group is more productive than the other (Phelps, 1972; Arrow, 1973). However, the standard labor market model does not incorporate other-regarding employers, i.e., employers who care about their workers' wellbeing. Other-regarding employers may attempt to protect workers from physical injury, reputational damage, or long hours away from their families. This protective behavior could be the source of a third form of discrimination: Employers protect one group differently than another, potentially depriving the protected group of opportunities to build important skills and experiences.

In this paper, we define and test for *paternalistic discrimination*: the differential treatment of two groups to protect one group, even against its will, from harmful or unpleasant situations. In the labor market, paternalistic discrimination may lead employers to hire women over men for female-stereotyped jobs, to avoid promoting recent mothers to reduce workloads, or to fire single workers over workers with families.[1] Given global attitudes towards protecting women (Glick et al., 2000), this discrimination could be particularly prevalent against women. Consistently, governments around the world restrict women's employment paternalistically: 49 countries restrict women's work in hazardous jobs and 21 in night jobs (World Bank, 2023a).[2]

We combine a simple model, two field experiments, and structural estimates to measure paternalistic discrimination against women. First, we augment a standard labor market model with other-regarding employers, i.e., employers who value their workers' welfare. Second, we test the model's predictions using two labor market experiments in which we observe real hiring and application decisions for a night-shift job in Bangladesh. Finally, we estimate the model parameters and combine the results of both experiments to benchmark the importance of paternalistic discrimination and evaluate the effectiveness of potential labor market interventions.

The key innovation of our model is that employers internalize the welfare of their workers and thus hire fewer workers with a low perceived welfare. Building on traditional models of discrimination, our model incorporates a simple distaste for one particular gender (taste-based discrimination), beliefs about the profitability of hiring from a particular gender (statistical discrimination), and beliefs about the welfare of a particular gender (other-regarding discrimination). We distinguish between two possible types of other-regarding discrimination, committed

[1] Outside of the labor market, parents may be more protective of their daughters than their sons (Bezirganian and Cohen, 1992; Kuhle et al., 2015). In addition, women may receive different advice about educational pathways, careers, or investments (Bajtelsmit and Bernasek, 1996; Carlana, 2019; Gallen and Wasserman, 2021).

[2] For example, women are barred from working during the night in some jobs in Nigeria and from working in mining and underground construction in Thailand (The Labour Protection Act B.E. 2541, 2014; World Bank, 2023b). Similar laws limit women's employment in Argentina, Cameroon, China, the Republic of Korea, Saudi Arabia, and other countries (US Department of State, 2022a,b; World Bank, 2023b). In South Asia, such laws are prevalent: Women in India face different restrictions than men when working at night, performing hazardous or difficult tasks (such as lifting heavy objects), and selling alcohol (Anand and Kaur, 2022). In addition, all but 17 countries ban women from fighting in combat (Fitriani et al., 2016).

by either deferential or paternalistic employers. Deferential employers use *applicants' beliefs and preferences*, such as risk preferences, to evaluate worker welfare, while paternalistic employers use *their own beliefs and preferences for workers* to evaluate worker welfare. Our model yields five predictions, which we evaluate using two field experiments and structural estimation.

We experimentally vary the perceived safety of a night-shift job to test the first theoretical prediction: Holding worker selection and productivity fixed, labor demand decreases in perceived job costs for workers. We recruit 495 *employers*, individuals with recent hiring experience, in Dhaka, Bangladesh. These employers make 4,950 hiring decisions (10 per employer) between one male and one female applicant for a job created by the research team: a one-time workshop and office job on the night shift. We randomly implement one hiring choice per employer and pay the employer based on the performance of their worker. We randomize whether we inform employers that workers receive free, safe transport home at the end of the shift. We find that employers who are not informed about the transport hire 21% fewer women.

The key feature of our design is that we hold taste-based and statistical discrimination constant across transport treatments. To hold constant the perceived selection of applicants willing to work, and thus taste-based selection, we inform employers that all applicants have applied for the job without knowing about the transport. In addition, we show every applicant-pair to several employers, allowing us to test whether information about the transport affects the hiring choices for the same woman compared to the same man. To hold constant the perceived productivity of applicants willing to work, and thus statistical discrimination, we inform employers that workers will only learn about the transport after completing the shift, i.e., that the transport cannot affect their attendance or on-the-job performance.[3] In addition, to ensure that differences in hiring are not explained by concerns about the employers' reputation, all hiring choices are private and anonymous. We limit our sample to employers who correctly respond to all comprehension checks and understand that the transport neither influences worker selection nor productivity. We also verify using survey questions that employers do not differentially base their hiring choices on taste, statistical, or reputation concerns between treatment arms.

We experimentally vary payments to workers to test the second theoretical prediction: Deferential employers respond differently from paternalistic employers to providing workers with cash payments and transport. Deferential employers respect workers' beliefs and preferences. They thus demand workers weakly more when workers receive sufficient cash to decide themselves whether to purchase safe transport, rather than when workers receive the transport itself. Paternalistic employers, on the other hand, may demand workers less with the cash than with transport if they believe that workers *should* purchase the transport but, when given the choice, would not. We cross-randomize employers into one of four cash treatments ("subsidies"): (i) female workers receive a surprise subsidy of 1,000 Bangladesh Taka (BDT, or USD 10)—an amount much larger than standard transport costs in our setting (Uber in Dhaka typically costs less than BDT 500 from our shift site and is easily available and considered safe), (ii) male workers receive a

[3] Note that there are no concerns about worker retention because there is only one shift.

surprise subsidy of BDT 1,000, (iii) employers receive a subsidy for hiring female workers of BDT 1,000, or (iv) neither employers nor workers receive a subsidy. We find that employers hire women significantly less with the female worker subsidies than the transport—even though workers prefer cash over the transport. This suggests that employers paternalistically prevent workers from making their own choices. Finally, consistent with the third theoretical prediction, employers react significantly more to the employer than either of the worker subsidies.

We test whether employers who score highly in a survey module on other-regarding preferences towards women react more to information about the transport to test the fourth theoretical prediction: The demand response to changes in perceived worker welfare increases in employers' other-regarding preferences. We find that more paternalistic employers react almost four times as much to the transport than less paternalistic employers. This difference suggests that our findings indeed reflect paternalistic preferences rather than, for example, experimenter demand effects, i.e, employers trying to satisfy the experimenters' preferences.[4]

Employers' response to information about the transport and female worker subsidies appears to be driven by employers disagreeing with both applicants' *beliefs* about job costs and applicants' *preferences*. Employers report that women underestimate the costs of working at night more than men—even though, in reality, employers overestimate night shift risks to women. Consistent with employers' beliefs, we find the strongest treatment effects against inexperienced women. We also find that employers who believe *women should be* rather risk-averse but not employers who believe *women are* rather risk-averse reduce hiring significantly more without transport. Finally, female employers, unlike male employers, do not report that women underestimate the costs of working at night more than men. Female employers also do not react to information about the transport.[5] Overall, these heterogeneous treatment effects offer additional evidence that (male) employers are paternalistic rather than deferential.

To evaluate the market's equilibrium behavior, we complement the demand-side experiment with a supply-side experiment, in which we randomly vary information about the transport to applicants. We find that randomly withholding information on the transport from potential applicants reduces the supply of female labor by 15%, significantly less than the demand reduction from employers. In particular, the reservation wages of the 770 *applicants*, both male and female, recruited from the same population but distinct from those in the demand-side experiment, increase by about BDT 200 (USD 2), which is much less than employers' valuation of female worker transport of BDT 1,400 (USD 14).

We estimate the model parameters and combine the results of the demand- and supply-side experiments to study the fifth theoretical prediction: Equilibrium wages decrease in perceived job costs if labor demand decreases more than supply. We construct the labor demand function by estimating preference parameters in a binary choice model using the hiring choices in the demand-side experiment. We construct the labor supply function by aggregating the reservation

[4] In addition, while experimenter demand effects should *decrease* when making the study purpose less salient, we find that treatment effects *increase* in a small "low-salience" control test.

[5] However, we consider these results as suggestive given our small sample of female employers.

wages in the supply-side experiment. We combine the two functions to construct equilibria for both genders with and without transport in counterfactual markets for our night-shift job.

We find that lack of transport reduces both female employment and wages. Moreover, we find that paternalistic discrimination has a greater effect on gender employment and wage gaps than taste-based or statistical discrimination in our setting. However, the welfare effects of eliminating paternalistic discrimination are ambiguous: If employers are well-informed about the dangers on the job, paternalism may help women avoid pitfalls; on the other hand, if women understand the risks, paternalism may restrict them from accessing beneficial opportunities. We simulate counterfactual policy interventions and find that the welfare effects similarly depend on the accuracy of beliefs about job costs for workers. Transport increases worker welfare more if employers have correct beliefs, while female worker subsidies increase total welfare more if applicants have correct beliefs.

The degree of paternalism revealed in our experiment suggests opportunities to increase female employment and wages in settings with strong gender norms. Although we model other-regarding employers, our model is observationally equivalent to one in which employers *act as if* they internalize workers' welfare to, for example, follow a protective social norm (Boudet, 2013; Jayachandran, 2021) or signal their identity as a protective person (Akerlof and Kranton, 2000).[6] Yet, policies that reduce paternalistic discrimination may also eliminate some self-regarding reasons to discriminate, by changing norms or eliminating socially acceptable covers for taste-based discrimination. In addition, while previous research has shown that the prevention of work-related dangers and unsafe transportation can increase *supply* of female labor (Field and Vyborny, 2022; Cheema et al., 2022; Grosset, 2024), our findings suggest that these policies can also increase *demand* for female labor. This implies that there are compounding benefits from policies that reduce women's perceived job costs (e.g., crime reduction programs and workplace safety regulation) or increase their perceived benefits (e.g., wage laws and subsidies).

We contribute to three separate strands of literature. First, we contribute to the literature on discrimination by defining a novel form of discrimination. A large body of literature measures taste-based and statistical discrimination on a variety of characteristics (e.g., Goldin and Rouse, 2000; Bertrand and Mullainathan, 2004; Levitt, 2004; Gneezy et al., 2012; Glover et al., 2017; Baert, 2018; Bohren et al., 2019; Kessler et al., 2019; Dianat et al., 2022; Kline et al., 2022; Chan, 2022; Macchi, 2023; Adamovic and Leibbrandt, 2023) and describes behavioral foundations of discriminatory beliefs (e.g., Bordalo et al., 2016; Esponda et al., 2023). Paternalistic discrimination differs from taste-based discrimination, as it varies with workers' perceived welfare.[7] It also differs from statistical discrimination in two key ways. First, it is driven by other-

[6] One experimental result suggests that employer behavior is not driven strictly by adherence to norms: employers increase female hiring when women receive a surprise cash subsidy. This behavior appears more consistent with our model of other-regarding preferences.

[7] Our experimental evidence cannot be explained by a distaste parameter that is constant across jobs, so we reject the simplest formulation of the traditional model. As a result, we describe paternalistic discrimination as a novel form of discrimination. The alternative interpretation—that paternalistic discrimination is a component of taste-based discrimination that varies with job characteristics—is equally valid.

regarding rather than self-regarding motives—indeed, employers may forego profits to indulge their paternalism. Second, unlike statistical discrimination, it does not require any uncertainty, as paternalistic employers may overrule applicants' preferences even if job costs are known with certainty.[8] We consider paternalistic discrimination most closely related to benevolent sexism (Glick and Fiske, 1997) in the psychology literature—idealized, seemingly positive but stereotypical views of women (e.g., that women should be cherished and protected, Glick et al., 2000; Fraser, 2015; Shnabel et al., 2016; Glick and Raberg, 2018).[9] Our paper draws on the benevolent sexism literature to formalize the first economic model of paternalistic discrimination.[10]

Second, we contribute to the literature on paternalism and other-regarding preferences by highlighting the role of other-regarding preferences in hiring. Paternalism—limiting the options available to others for their own benefit—drives support for many policies, including banning "repugnant" transactions (e.g., Leider and Roth, 2010; Elías et al., 2023), regulating addictive products (e.g., Allcott et al., 2019a,b; DeCicca et al., 2022), and protecting boundedly rational or time-inconsistent consumers (e.g., Allcott and Taubinsky, 2015; Allcott et al., 2021). Researchers have also explored the drivers of and responses to paternalism (Uhl, 2011; Ambuehl et al., 2021; Bartling et al., 2023), and the relationship between paternalism and altruism (Jacobsson et al., 2007). Most relevant to our setting, other-regarding preferences also drive behavior in the workplace, including wage setting (Akerlof, 1982), workers' effort (Bandiera et al., 2005; Asad et al., 2023), resource allocation (Bandiera et al., 2009; Hjort, 2014), hiring (Dhillon et al., 2020) and layoff decisions (Guenzel et al., 2023). To our knowledge, our paper is the first to consider how other-regarding preferences differentially affect men and women in the workplace.

Third, we contribute to the growing literature on the barriers to female employment, particularly in low-income countries. Among other factors, this literature considers social norms (e.g., Fernández, 2013; Bertrand et al., 2015; Bernhardt et al., 2018, 2019; Bursztyn et al., 2020; Field et al., 2021; McKelway, 2023; Bursztyn et al., 2023; see Jayachandran, 2021 for an overview), safety (Chaudhary et al., 2021; Field and Vyborny, 2022; Siddique, 2022; Becerra and Guerra, 2023), and work location (Ho et al., 2024; Jalota and Ho, 2024) as barriers to female labor supply. Part of this literature specifically examines how unsafe transport restricts women's physical mobility (e.g., Kondylis et al., 2020; Aguilar et al., 2021; Borker et al., 2022; Field and Vyborny, 2022; Cheema et al., 2022; Christensen and Osman, 2023; Becerra and Guerra, 2023). Efforts to study discrimination in South Asia have focused primarily on India, examining differential

[8] Paternalistic discrimination could also result from and induce systemic discrimination (Bohren et al., 2022). Paternalistic discrimination could result from systemic discrimination that drives differences in non-gender characteristics that correlate with gender, such as transport costs. Paternalistic discrimination could induce systemic discrimination by depriving women of opportunities to build important skills and experiences, limiting their human capital accumulation and reducing their promotability later on.

[9] In line with theories of benevolent sexism, economists have found that advisors withhold negative feedback from female advisees (Coutts et al., 2024) and that male evaluators attribute bad outcomes to bad luck more often for women than for men (Erkal et al., 2023).

[10] U.S. law treats benevolent discrimination as any other kind of discrimination (U.S. EEOC, 2007, 2022). A separate but related concept in the law literature is *benign discrimination*, which generally refers to discriminatory policies designed to benefit minority or marginalized groups (see, for example, Evans 1974 and Patty 1989).

treatment based on caste (Banerjee et al., 2009; Ito, Takahiro, 2009; Siddique, 2011), religion (Thorat and Attewell, 2007), and gender (Choudhury, 2015; Islam et al., 2021). This paper is the first to study how paternalism restricts women's employment, particularly in South Asia.

The rest of the paper proceeds as follows. Section 2 describes the labor model with other-regarding employers and section 3 describes the empirical setting. Sections 4 and 5 present the demand-side experiment with employers and the supply-side experiment with applicants. Section 6 combines the results from the two experiments in an equilibrium model and evaluates a series of counterfactuals. Section 7 concludes.

2 A Labor Market Model with Other-Regarding Employers

In this section, we augment a standard labor market model with other-regarding employers, i.e., employers internalizing their workers' perceived on-the-job welfare, and outline the resulting comparative static predictions. First, we describe two markets with other-regarding employers, one for male workers and one for female workers, and define paternalistic discrimination. Second, we evaluate how increases in workers' perceived job costs affect labor supply and demand and how the demand effects vary with employers' other-regarding preferences. Finally, we investigate how increases in workers' perceived job costs affect equilibrium wages.

2.1 Setup

Market Structure We study two markets, one for each gender $g \in \{m, f\}$.[11] A unit mass of price-taking employers, indexed by k, demand labor in the two markets. A unit mass of male ($g = m$) and a unit mass of female ($g = f$) price-taking workers, indexed by i, supply labor. We use the superscripts A and E to denote preferences and beliefs of workers (i.e., applicants) and employers, respectively. The mass of gender g workers supplying labor is given by L_g^S, and the mass of gender g workers demanded by employers is given by L_g^D. Wages for gender g labor and the quantity of hired gender g labor are determined in equilibrium. w_m^* and w_f^* are the equilibrium wages that equate the labor supply and labor demand for both genders simultaneously. L_m^* and L_f^* are the equilibrium quantities of both genders at these wages.

Workers' Problem Workers supply their labor if their expected utility from working is weakly positive.[12] Worker i's expected money-metric utility from working for employer k depends on the wage w_g, the expected costs c_{igk}, and the disutility associated with the cost of working $u_i^A(\cdot)$, where u_i^A is continuously differentiable and monotonically increasing in c_{igk}, with $u_i^A(0) = 0$.[13] The cost of worker i of gender g working for employer k is given by $c_{igk} = c_g + c_i + c_{kg}$, the sum

[11] The setup generalizes to more groups, each having a separate market.

[12] We study extensive margin decisions to highlight worker selection into different jobs. However, the setup generalizes to intensive margin decisions, for example, by considering every worker supplied as a time unit.

[13] Following Rabin 2013, we treat utility as linear in the relatively small one-day salaries; we allow agents to have risk preferences over costs which may be at a larger scale (e.g., sexual assault).

of (i) a gender-specific constant cost c_g, known to the worker and the employer, (ii) the worker-specific cost c_i, known only to the worker, and (iii) the employer-gender-specific cost c_{kg}, known only to the employer. The employer- and worker-specific costs c_{kg} and c_i follow distributions h_g^K and h^I with CDFs H_g^K and H^I and means $\bar{c}_g^K$ and $\bar{c}_g^I$. Workers rely on their cost assessments and do not attempt to learn about the employer-gender-specific costs from employers' hiring decisions.[14] We assume that applications are costless and that the outside option has zero value such that applicant i of gender g supplies labor if and only if:

$$\mathcal{W}_i^A = \mathbb{E}_i[w_g - u_i^A(c_{igk})] \geq 0. \tag{1}$$

Employers' Problem Employers decide how much male and female labor to demand to maximize their expected utility. Employer k's expected utility is linear and additively separable in (i) d_{kg}, the non-pecuniary benefits of hiring gender g labor (i.e., taste), (ii) $Y^E(L_{kf}, L_{km}) - L_{kf}w_f - L_{km}w_m$, the expected profits of hiring L_{kf} female and L_{km} male workers at wages w_f and w_m, and (iii) fraction $\alpha_{kg} \in (0,1)$ of the expected on-the-job welfare of the worker, $\mathcal{W}_{kg}$ (henceforth "welfare").[15] The expected production function Y^E is non-negative, concave (see appendix section B.1) and, akin to our empirical setting, not a function of costs, wages, or the selected pool of applicants.[16] Employers understand selection, realizing that the pool of applicants consists only of workers who believe the job will yield positive utility, i.e., for whom $\mathcal{W}_{kg}^A \geq 0$.

We differentiate between two possible types of other-regarding employers:

Definition 1. Deferential employers internalize their perception of workers' perception of welfare, $\mathcal{W}_{kg}^{E:A} = \mathbb{E}_k[\mathbb{E}_i[w_g - u_i^{E:A}(c_{igk})] | \mathbb{E}_i[u_i^{E:A}(c_{igk})] \leq w_g]$. Paternalistic employers internalize their own perception of workers' welfare, $\mathcal{W}_{kg}^E = \mathbb{E}_k[w_g - u_{ki}^E(c_{ikg}) | \mathbb{E}_i[u_i^{E:A}(c_{igk})] \leq w_g]$.

We denote the employer's second-order belief about u_i^A by $u_i^{E:A}$ and the employer's risk preferences for workers by u_{ki}^E. $u_i^{E:A}$ and u_{ki}^E follow the same functional form assumptions as u_i^A.

Other-regarding employer k thus maximizes the following objective function v_{kg}^E:

$$\max_{L_{kf}, L_{km}} \underbrace{\sum_{g \in \{f,m\}} L_{kg} d_{kg}}_{\text{Taste utility}} + \underbrace{Y^E(L_{kf}, L_{km}) - \sum_{g \in \{f,m\}} L_{kg} w_g}_{\text{Profit}} + \underbrace{\sum_{g \in \{f,m\}} L_{kg} \alpha_{kg} \mathcal{W}_{kg}}_{\text{Other-regarding utility}}, \tag{2}$$

with $\mathcal{W}_{kg} \in \{\mathcal{W}_{kg}^{E:A}, \mathcal{W}_{kg}^E\}$.[17]

[14] We assume away that sophisticated workers apply for costly jobs, anticipating that paternalistic employers will protect them from mistakes. This assumption is consistent with empirical evidence: Anticipating discrimination to make inferences about job costs requires extensive contingent reasoning, which a large literature suggests is rare; see Niederle and Vespa (2023) for an overview. Since workers do not anticipate paternalistic discrimination, employers do not worry about workers' preferences for being "paternalized."

[15] Our model is also flexible enough to allow the other-regarding utility to vary with ability, for example, by considering high- and low-skilled workers as separate groups with different welfare weights α_{kg}. We do not distinguish between different channels for other-regarding preferences, such as warm glow or guilt aversion; these different sources of other-regarding behavior are observationally equivalent in our model.

[16] We relax this assumption in the structural model.

[17] Assuming the outside option has zero value, equation 2 is the same if the employer internalizes the welfare of *only* hired or of hired *and* non-hired workers.

2.2 Defining Discriminatory Preferences

We define discriminatory preferences leading to preferential treatment of men over women for a given set of wages (w_f, w_m) and hiring levels (L_{kf}, L_{km}) as follows:

1. *Taste-based discrimination*: $d_{km} > d_{kf}$. The employer receives more (or less negative) non-pecuniary returns from hiring male than female workers.

2. *Statistical discrimination*: $\frac{\partial Y^E}{\partial L_{km}} > \frac{\partial Y^E}{\partial L_{kf}}$. The employer expects to receive higher revenues from the marginal male than the marginal female worker.

3. *Other-regarding discrimination*: $\alpha_{km}\mathcal{W}_{km} > \alpha_{kf}\mathcal{W}_{kf}$. The employer expects to receive higher other-regarding utility from the marginal male than the marginal female worker. This is deferential if employers use their perception of workers' perception of worker welfare and paternalistic if employers use their own perception of worker welfare.

Other-regarding discrimination arises because an employer places a different welfare weight on men than women or because an employer expects men's welfare to be different than women's. Three mechanisms could explain why employers may expect men's welfare to be different even in the same role and with the same wages: employers may (i) believe men have different costs than women (using either the employers' first- or second-order beliefs), (ii) have different risk preferences for men and women $\left(u^E_{km} \neq u^E_{kf}\right)$, and/or (iii) engage in selection neglect, i.e., they do not condition on $\left(\mathcal{W}^{E:A}_{kg}\right)$ and thus compare the average man and woman in the population.

We consider other-regarding discrimination distinct from taste-based discrimination because, unlike taste-based discrimination, it varies with perceptions of job costs c_{igk}. We consider other-regarding discrimination distinct from statistical discrimination because, unlike statistical discrimination, it can arise even without uncertainty, i.e., when $\mathcal{W}_g = \mathcal{W}^{E:A}_{kg} = \mathcal{W}^{A}_{kg}$.[18]

2.3 Comparative Statics in Gender-Specific Costs and Wages

In this section, we investigate how labor supply and labor demand by other-regarding employers react to changes in gender-specific costs and wages and how equilibrium wages react to changes in gender-specific costs.

2.3.1 Labor Supply

Workers' perceived welfare is increasing in wages and decreasing in costs. Therefore, workers are less willing to supply their labor if they pay higher gender-specific costs, and more willing to supply their labor if they earn higher gender-specific wages.

[18] Note that other-regarding discrimination can only lead to restricting the employment opportunities of willing workers as employers cannot force workers to apply who do not want to apply. Furthermore, other-regarding discrimination can persist in repeated markets if employers do not learn about costs (e.g., if they never observe women working the night shift) or disagree with workers' preferences.

2.3.2 Labor Demand

Other-Regarding Employers and Costs First, we assess the demand response to changes in gender-specific costs by other-regarding employers (appendix B.2). An increase in c_g has two effects on employers' other-regarding utility: (i) a *direct effect*: the job cost increases, reducing employers' perception of worker welfare, and (ii) a *selection effect*: workers with smaller worker-specific cost self-select into the job, increasing employers' perception of worker welfare. Holding selection and productivity constant, an increase in gender-specific costs unambiguously reduces employers' perception of worker welfare and reduces the labor demand of other-regarding but not non-other-regarding employers.

Prediction 1 (Other-Regarding Employers)**.** *Holding selection and productivity constant, employers are other-regarding if and only if labor demand is decreasing in gender-specific costs.*

Deferential and Paternalistic Employers and Costs Second, we assess how the demand response to changes in gender-specific costs differs between deferential and paternalistic employers (appendix B.3). If employers are deferential (i.e., if they internalize their perception of the workers' perception of worker welfare, $\mathcal{W}_{kg}^{E:A}$), then their other-regarding utility is weakly lower when workers receive an amenity rather than a cash payment that allows them to afford the amenity. Workers are weakly better off receiving the cash payment, as they can use their own valuation of the amenity to decide whether to purchase it. Therefore, if employers demand less labor with the cash payment than the amenity, they must be paternalistic (i.e., they must use their own beliefs or preferences to evaluate worker welfare, $\mathcal{W}_{kg}^{E}$).

Prediction 2 (Deferential and Paternalistic Employers)**.** *Holding selection and productivity constant, the labor demand of deferential employers is increasing weakly more in cash payments to workers than in equally (or lower) priced worker amenities. If labor demand is increasing less, employers are paternalistic.*

Other-Regarding Employers and Wages Third, we assess the demand response to changes in gender-specific wages by other-regarding employers (appendix B.4). An increase in w_g reduces employers' profit and has two effects on employers' other-regarding utility: (i) *a direct effect*: the wage increases, increasing the employer's perception of worker welfare, (ii) a *selection effect*: workers with higher worker-specific cost self-select into the job, decreasing the employer's perception of worker welfare. Holding selection and productivity constant, an increase in gender-specific wages reduces the labor demand of other-regarding employers (as $\alpha_{kg} < 1$).

Prediction 3 (Wages)**.** *Holding selection and productivity constant, the labor demand of other-regarding employers is decreasing in gender-specific wages.*

Heterogeneity in α_{kg} Finally, the demand response to gender-specific costs and wages by other-regarding employers changes in α_{kg} (appendix B.5). Employers who weigh workers' welfare highly experience a high loss in other-regarding utility from higher gender-specific costs. By contrast, their profit loss from higher wages is offset by a larger other-regarding utility gain. Therefore, an increase in the welfare weight increases the demand response to gender-specific costs and decreases the demand response (in absolute terms) to gender-specific wages.

Prediction 4 (Heterogeneity)**.** *Holding selection and productivity constant, larger other-regarding preferences α_{kg} result in a larger demand response to changes in gender-specific costs and a smaller demand response to changes in gender-specific wages.*

2.3.3 Equilibrium Wages

The equilibrium wage response to gender-specific costs depends on the magnitude of the cost elasticity of demand relative to that of supply (appendix B.6). An increase in c_g decreases labor demand and supply, thus reducing the equilibrium labor quantity. However, an increase in c_g can increase or decrease equilibrium wages depending on the relative magnitudes of the demand and supply elasticities. If the elasticity of demand is sufficiently large relative to that of supply, equilibrium wages decrease because the downward pressure on wages from the decrease in labor demand dominates the upward pressure on wages from the decrease in labor supply.

Prediction 5 (Equilibrium Wages)**.** *Equilibrium wages are decreasing in gender-specific costs if the ratio of the demand and supply elasticities with respect to gender-specific costs is larger than a cutoff that depends on the substitutability of male and female labor and the demand and supply elasticities with respect to wages of the other gender.*

The equilibrium labor quantity and wages of the other gender do not respond to increases in gender-specific costs if male and female workers are additively separable in the production function, increase if they are substitutes, and decrease if they are complements (appendix B.7).

The remainder of the paper explores theoretical predictions 1 to 5 in three steps. First, we experimentally vary employers' beliefs about job costs for female workers to study the effect of paternalism on the demand for female labor. Second, we experimentally vary applicants' beliefs about job costs to study the effect of these costs on female labor supply. Finally, we estimate the theoretical model by combining the results of both experiments to predict the equilibrium wage effects of changing job costs for female workers.

3 Setting

We empirically test theoretical predictions 1 to 5 in two experiments in Dhaka, Bangladesh, in which we sequentially measure the labor demand and supply responses to exogenously varying the perceived job costs to workers. Around 40% of Bangladesh's population lives in urban areas, and about one sixth lives in Dhaka. Dhaka also accounts for one fifth of Bangladesh's GDP and nearly one half of its formal employment (World Bank DataBank, 2023).

Women in Bangladesh struggle to access the labor market, particularly male-dominated occupations (BDHS, 2016; BBS, 2021).[19] About 40% of working-age women in Bangladesh are employed, compared to about 80% of working-age men (World Bank DataBank, 2023). Women work predominantly in agriculture and industrial production, particularly in the garment sector, where they comprise 80% of the workforce, while men work predominantly in services (Farole et al., 2017; Quayyum, 2019). Women also earn less than men, especially in urban areas (USD 171 versus USD 133 per month, BBS (2018)).[20]

Gender norms contribute to Bangladesh's large gender employment and wage gaps. Women live under the protective guardianship of their fathers in childhood, their husbands and fathers-in-law in marriage, and their sons in widowhood (White, 2017). In addition, *purdah*—meaning literally veil or curtain, but used figuratively to describe women's seclusion from the public sphere—limits women's access to the labor market (Lata et al., 2021).

These protective attitudes toward gender carry over into economic behavior and policies. In the 2018 Bangladesh World Value Survey, 76% of respondents agreed that "men should have more right to a job than women," and 67% that "men make better business executives than women." Results were broadly similar among men and women and in urban and rural areas (World Values Survey, 2018). Bangladesh law does not prohibit discrimination on the basis of gender, nor does it mandate equal pay for women and men (World Bank DataBank, 2023). Women in Bangladesh are also legally restricted from operating or cleaning certain types of machinery, carrying heavy items, or working underwater or underground (Bangladesh Labour Act, 2006). While the Bangladesh Labor Act of 2006 lifted a prohibition on women working at night, employers are still required to obtain women's consent to work shifts between 8:00 P.M. and 6:00 A.M.; written consent is not required of men.

Safe transport represents a special concern for women in Dhaka. Women report high rates of physical harassment, such as groping, driver misconduct, and discomfort from overcrowding and crush loading (Rahman, 2010; Aachol Foundation, 2022; Kabir and Islam, 2023). These problems have led providers to establish women-only bus service routes in recent years, though these services offer limited routes and hours (Naher, 2022).

Given this environment, we took several precautions to ensure the ethical conduct of the experiment. We obtained written consent from all night-shift workers and calibrated the payments with input from our local partners to compensate workers for the inconvenience of a night-shift job without being coercive (Ambuehl et al., 2022). In addition, we provided transport to all workers at the end of the shift through a private transportation firm with vetted drivers.

[19] We restrict our discussion to differences between male and female workers, though Bangladesh's transgender or "hijra" population faces additional barriers to economic participation, as well as poor access to healthcare and exclusion from civic life (Al-Mamun et al., 2022).

[20] A recent report by the International Labor Organization finds that the factor-weighted mean hourly wage for women is higher than that of men in Bangladesh—a sole outlier among countries studied in the report (International Labour Organization, 2018). However, this finding does not appear to be robust to alternative model specifications (Rahman and Al-Hasan, 2022).

4 The Hiring Experiment: Job Costs and Labor Demand

To measure the labor demand response to variations in gender-specific costs and wages according to predictions 1 to 4 of the model, we conduct a "hiring experiment" with 495 *employers*, individuals with hiring experience in the previous three years, in Dhaka, Bangladesh. Enumerators recruit employers equally from three industries, selected based on recruitment feasibility, different perceived costs to female workers, and high levels of urban employment: manufacturing, retail/wholesale and services, and education (additional information on these industries is provided in appendix C.1).[21] Enumerators recruit employers in person between April 2023 and August 2023 by asking businesses whether any individual with hiring responsibility is interested in participating in the experiment on the spot or later. The vast majority of employers in our experiment (94%) are men, in line with estimates in the broader Bangladesh workforce (89% of managers are men, according to official statistics (BBS, 2018); see table 1 for overall summary statistics and appendix table A.1 for summary statistics by industry). Employers are, on average, 32 years old, 59% are married, and 45% have at least one child. Furthermore, 42% have at least a Bachelor's degree. On average, their businesses have nine male and six female employees, and they have made 27 hiring decisions in the previous three years.

Table 1: Employer Characteristics (N=495)

	Mean	SD
Female (%)	6.3	24.3
Age (Years)	31.5	7.8
Married (%)	58.5	49.3
Children (%)	45.2	49.8
Bachelor's (%)	42.2	49.4
# Male Employees	8.9	24.3
# Female Employees	6.1	41.5
# Hiring Decisions Last 3 Years	27.2	235.4

Notes: The table shows the means and standard deviations of characteristics of employers recruited for the hiring experiment. *Children* is an indicator equal to 1 if the applicant has at least one child. *Bachelor's* is an indicator equal to 1 if the applicant has at least a Bachelor's degree.

"Employers" make hiring decisions for a job created by the research team: a one-shift, three-

[21] We excluded agriculture, a primary employment sector, due to low recruitment feasibility in Dhaka. We asked 80 employers the following question for eight randomly selected applicants about a hypothetical job in their industry: "How dangerous or unpleasant or socially unacceptable do you think this job is for [applicant name], including their commute from and to their home, with 0 indicating a very safe job, equivalent to working from home, and 10 indicating that the job is very dangerous or very unpleasant or very socially unacceptable." The average response for female applicants was 2.5 in manufacturing, 1.0 in retail and services, and 0.3 in education.

hour Excel workshop followed by a stock market analysis task from 7:00 P.M. to midnight (see appendix figure C.1 for a photograph of the shift), with free and safe worker transport home in private six-seater cars after the shift. One supervisor accompanied each car.[22] The applicant pool consists of 764 male and 575 female applicants aged 18 to 60, recruited at booths on 11 university campuses between February and April 2023.[23] Applicants take two 12-minute back-to-back Excel screening tests incentivized with BDT 2 (USD 0.02) per correct answer for a total of up to BDT 40 (USD 0.4).[24] On average, male applicants in the hiring experiment are 25 years old and female applicants are 24 years old (table 2). Around one fifth are married (19% of male and 23% of female applicants), and 12% have children. Female applicants are slightly less experienced than male applicants (89% have up to three years of experience versus 80% of male applicants) but have similar education (36% have a Bachelor's degree versus 39% of male applicants) and Excel screening scores (25%, on average, versus 23% among male applicants).[25]

Table 2: Applicant Characteristics by Gender in Hiring and Application Experiments

	Hiring Experiment				Application Experiment			
	Male (N=764)		Female (N=575)		Male (N=354)		Female (N=344)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age (Years)	24.8	6.3	23.6	6.0	25.9	7.8	22.9	6.4
Married (%)	19.3	39.5	22.8	42.0	26.3	44.1	23.5	42.4
Children (%)	12.1	32.6	11.8	32.2	18.4	38.8	13.5	34.2
Bachelor's (%)	39.1	48.8	35.7	48.0	14.3	35.1	8.7	28.2
$\leq$ 3 Years Work Exp (%)	80.0	40.0	89.0	31.3	72.1	44.9	88.9	31.4
Excel Screening Score (%)	22.6	12.5	24.5	13.7	24.8	11.5	26.3	12.1

Notes: The table shows the means and standard deviations of characteristics of applicants recruited for the hiring experiment and the application experiment. *Children* is an indicator equal to 1 if the applicant has at least one child. *Bachelor's* is an indicator equal to 1 if the applicant has at least a Bachelor's degree.

4.1 Hiring Experiment Design

Employers make real hiring decisions and are randomized into different treatment conditions that experimentally vary the perceived job costs for workers, the payoffs received by workers, and the payoffs received by employers. This allows us to test whether employers hire fewer women

[22] The cars were mixed-gender. However, this information was not communicated to employers. Training opportunities at night are common in Bangladesh. In addition, night-shift work is becoming increasingly common as many outsourcing firms work European or US hours (Mamun et al., 2019).

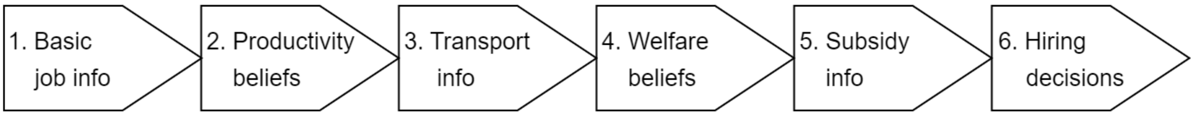
[23] 579 male and 399 female applicants come from the hiring experiment and 185 male and 176 female applicants from the application experiment (see section 5). In the hiring experiment, we initially targeted 585 men and 405 women to construct 45 hiring pools of nine female and 13 male applicants each (we oversampled men to make hiring choices more realistic); however, five hiring pairs were excluded from the sample as the male applicant was miscoded as female and one male and one female applicant revoked their consent. In addition, one male worker revoked their consent in the application experiment. We conducted recruitment on university campuses anticipating a high concentration of job seekers and that paternalistic discrimination may be particularly consequential for job seekers early in their careers. We do not restrict participation to university students.

[24] The tests were designed based on a scoping survey with 20 office employers about desired Excel skills.

[25] The experiment design does not require balance across genders.

when the job costs are perceived to be high (prediction 1)—even when workers can afford safe transport themselves (prediction 2). The variation in payoffs to workers and employers allows us to test whether employers react more to profit than to other-regarding concerns (prediction 3). The experiment is conducted in six stages and takes an average of 64 minutes (figure 1).

Figure 1: Stages of the Hiring Experiment



Notes: The figure shows the six stages of the hiring experiment, described in detail below. 1) We provide employers with detailed information about the job. 2) We elicit employers' beliefs about the on-the-job productivity of a subset of applicants. 3) We provide employers with the transport information. 4) We elicit employers' beliefs about the on-the-job welfare of a subset of applicants. 5) We provide employers with the subsidy (random employer and worker cash payments) information. 6) We elicit employers' hiring decisions.

1. **Basic job information**: In the first stage, we provide employers with basic information about the job. Employers receive the following information about the hiring process: (i) Applicants have applied for a one-day Excel workshop and job from 7:00 P.M. to midnight and completed an Excel screening test. (ii) Recruited workers will be compensated with BDT 1,500 (USD 15) and receive an Excel workshop completion certificate. (iii) We hire one worker based on each employer's decisions. (iv) Employers receive a base compensation of BDT 500 (USD 5) for their time as well as BDT 5 (USD 0.05) per task completed on the job (out of 100 possible tasks) by their recruited worker.

2. **Productivity beliefs elicitation**: In the second stage, we elicit employers' incentivized beliefs about the on-the-job productivity of four randomly selected applicants (two male-female pairs). Employers predict the number of tasks that each applicant will complete if hired based on first name, gender, marital status, education, years of experience, and Excel screening test score.[26] Employers are informed that two of these applicants are randomly selected for hire (in the application experiment described in section 5) and that the employers receive a bonus payment for correctly predicting the productivity of these applicants. Employers guess (i) the probability that each applicant shows up to the shift (incentivized using the binarized scoring rule, see Hossain and Okui 2013), and (ii) the number of completed tasks conditional on attendance (incentivized with BDT 10, USD 0.1, for guesses within five percentage points from the truth). To reduce the risk of strategic misreporting, we elicit employers' productivity beliefs before randomizing them to a treatment group. We also verify that the predictions of the 495 *Hiring* employers do not differ from those of 80 separately recruited *Prediction-Only* employers who make no hiring choices and therefore have no incentive to adjust their predictions to their hiring choices.

[26] Because of a translation mistake into Bangla, employers were shown "3 years of work experience" instead of ">3 years of work experience" when an applicant had >3 years of work experience.

3. **Transport information randomization**: In the third stage, we randomize employers into one of two transport treatments that experimentally vary employers' perception of workers' job costs while holding constant the perceived worker selection and productivity:

   (a) *Transport* (50%): Employers are informed about the transport with supervisors.

   (b) *No Transport* (50%): Employers are not informed about any transport.

   The randomization allows us to test theoretical prediction 1: Demand for female labor is lower without safe transport. To hold constant the perceived selection of applicants across treatments, we inform employers that all applicants have applied for the job without knowing about the transport. To hold constant the perceived productivity of applicants across treatments, we inform employers in the *Transport* treatment that workers will only learn about the transport after completing the shift, i.e., that the transport cannot affect their attendance or on-the-job performance. To hold constant liability or reputation concerns across treatments, we ensure employers that all hiring choices are private and anonymous.[27] To hold constant beliefs about applicants' beliefs across treatments and ensure that employers in the *No Transport* treatment know that applicants do not expect transport, we inform employers in both treatments "Aside from the job description before, no other benefits (such as flexible hours, work-from-home, [transport], or future employment) are offered to any applicant." ("transport" is only included in the *No Transport* treatment). We verify comprehension of the experimental set-up with five comprehension questions administered after the treatment assignments. We also find that (i) employers believe workers to be very likely to take the transport (i.e., high compliance), and (ii) no evidence for information spillovers (i.e., no contamination).[28]

4. **Cost beliefs elicitation:** In the fourth stage, we elicit employers' beliefs about the job costs (including commute) for the applicants for whom they also made productivity predictions. Employers report the job costs in terms of danger, unpleasantness, and social acceptability on a scale from 0 to 10. To reduce anchoring, we do not inform employers of applicants' reported costs, nor do we attach any experimental incentives to the elicitation in order to reduce strategic reporting.[29] We find no significant differences between *Hiring* and *Prediction-Only* employers, suggesting that our results are not driven by strategic misreporting to justify hiring decisions.

[27] We also find that only one employer chooses reputation concerns as one of their drivers for their hiring choices.

[28] Only one *Transport* employer believes that applicants will not take the transport. To prevent information spillovers (i.e., employers in the *No Transport* condition learning about the transport from previous workers), we started the shifts only after roughly half (57%) of the hiring experiment was completed (results are robust to restricting the sample to all employers surveyed before this shift). Only six employers in the *No Transport* treatment believe that applicants will get home by provided transport (three of these are excluded from the analysis due to incorrectly answering understanding questions used for screening comprehension). The vast majority of *No Transport* employers (98%) believe applicants will use public transport or a ride share (Uber, CNG, Rikshaw).

[29] For example, if we promised to convey the response as advice to the applicant, employers with a strong distaste for hiring women might misleadingly report a high cost.

5. **Subsidy randomization**: In the fifth stage, we cross-randomize employers into one of five subsidy treatments that experimentally vary the cash payments received by workers and employers while holding constant worker selection and productivity:

   (a) *No Subsidy* (40%): Male and female workers receive BDT 1,500 (USD 15) for completing the shift. Employers receive BDT 500 (USD 5) for hiring any worker.

   (b) *Male Worker Subsidy* (20%): Male workers receive BDT 2,500 (USD 25), and female workers BDT 1,500 (USD 15) for completing the shift. Employers receive BDT 500 (USD 5) for hiring any worker.

   (c) *Female Worker Subsidy* (20%): Male workers receive BDT 1,500 (USD 15), and female workers BDT 2,500 (USD 25) for completing the shift. Employers receive BDT 500 (USD 5) for hiring any worker.

   (d) *Employer Subsidy for Hiring Women* (19%): Male and female workers receive BDT 1,500 (USD 15) for completing the shift. Employers receive BDT 500 (USD 5) if their hired worker is a man, and BDT 1,500 (USD 15) if it is a woman.

   (e) *Employer Subsidy for Hiring Men* (1%): Male and female workers receive BDT 1,500 (USD 15) for completing the shift. Employers receive BDT 1,500 (USD 15) if their hired worker is a man, and BDT 500 (USD 5) if it is a woman.

   The randomization allows us to test theoretical predictions 2 and 3: Demand for female labor is higher with safe transport than with subsidies paid to female workers, and labor demand increases more in subsidies paid to employers than to workers. Qualitative interviews suggest that it is common knowledge that workers can afford an Uber (typically costing $\leq$ BDT 500 or USD 5) or professional car service (costing $\leq$ BDT 800 or USD 8) using the subsidy of BDT 1,000 (USD 10). We did not inform employers that workers can use the subsidy to purchase transport from us to avoid deception (since we provide transport to all workers for ethical reasons).

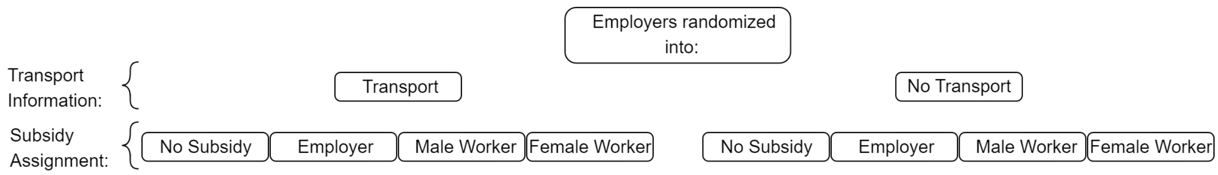
   We take the following steps to hold constant both perceived worker selection and productivity across subsidy treatments. First, to ensure that employers understand that subsidies are due to chance, employers draw a piece of paper to determine their treatment assignment. Second, to prevent asymmetrical subsidies that could signal gender differences in qualification, the *Employer Subsidy for Hiring Men* is included (enumerators describe all treatments but not the relative frequencies to the employers). We exclude the six employers assigned to this treatment. Third, to hold constant perceived selection and productivity, employers are informed that workers will be surprised by the subsidies after the shift.

6. **Hiring**: In the sixth stage, employers make twelve hiring decisions, each between two randomly selected applicants. For each applicant, employers are shown the same characteristics as employers in the prediction questions (see appendix figure C.2 for the experimental

interface). Each employer makes decisions for two male–male pairs and ten mixed-gender pairs in random order. We did not include two female–female pairs because we wanted to keep the female-to-male ratio similar to that observed in labor markets in Bangladesh. Two mixed-gender pairs were already included in the productivity and cost beliefs elicitation stages. The remaining 11 pairs are shown to 11 employers each, implementing a different pair per employer so that we implement one hiring choice per pair (see appendix C.3 for a description of the matching process). Employers are informed that one of their decisions will be implemented and that their identity will not be revealed to any workers.

We make several design choices to reduce experimenter demand effects across treatments (figure 2). First, to make it harder for participants to infer the study purpose, all treatments are assigned across employers. Second, all interviews are conducted privately and anonymously. Third, to ensure that the subsidies do not signal experimenters' preference to hire workers of a specific gender, participants are informed about all subsidies regardless of their assignments.

Figure 2: Treatments: Varying Information about Transport and Worker and Employer Pay



Notes: Employers are cross-randomized into two transport and four subsidy treatments (we exclude the employers assigned to the *Employer Subsidy for Hiring Men*, a treatment we only include for perceptions of fairness and symmetry). First, we randomize employers into one of two transport treatments. Employers in the *Transport* treatment are informed that the job provides all workers free and safe transport home after the shift. Employers in the *No Transport* treatment are not informed about the transport. Second, we randomize employers into one of four subsidy treatments. Employers in the *No Subsidy* treatment are informed that neither employers nor workers receive a cash subsidy. Employers in the *Employer Subsidy* treatment are informed that every time they hire a female worker, the employer receives a cash subsidy. Employers in the *Male Subsidy* and *Female Subsidy* treatments are informed that every time they hire a male or female worker, the worker receives a surprise cash subsidy.

The experiment allows us to detect paternalistic discrimination against women for a one-time night-shift job in Bangladesh and demonstrate a dynamic that may also be important in other contexts. To control selection and construct a natural hiring environment (see List (2020)), we recruit real employers with recent hiring experience across a range of industries to hire from a pool of real job seekers. However, some key features of our environment may not generalize to other hiring settings. In jobs with repeated interactions, attrition can change the composition of employers and workers (e.g., employers with other-regarding preferences may go out of business). Over time, paternalistic discrimination may also decrease if employers learn more about applicants' true job costs or increase if employers develop stronger feelings of responsibility.[30]

[30] We can only speculate about whether paternalism drives discrimination in other markets (e.g., in high-income countries) or with respect to characteristics other than gender (e.g., race, pregnancy, or parental status). Discrimination in hiring is likely more muted in markets where differential treatment is banned by law. However, paternalistic discrimination on the basis of age is sufficiently problematic even in the US to prompt the Equal Employment Opportunity Commission to issue specific guidelines: "The ADEA [Age Discrimination in Employment Act] would

4.2 Hiring Analysis: Empirical Specification

We identify *within-applicant* differences in hiring across treatments, allowing us to rule out a myriad of endogeneity concerns. The transport treatment was stratified by applicant and by employer industry.[31] We restrict the sample to the 460 employers who answer all understanding questions correctly (94%). In this sample, employer characteristics are balanced across treatment arms (appendix tables D.1). Employers are more likely to report basing their hiring decisions on safety but not taste or statistical concerns without than with transport.[32]

We estimate the following equation among all female applicants:[33]

$$\begin{aligned} H_{ki} = &\alpha + \beta_1 NT_k + \beta_2 MS_k + \beta_3 FS_k + \beta_4 ES_k + \beta_5 (NT_k \times MS_k) \\ &+ \beta_6 (NT_k \times FS_k) + \beta_7 (NT_k \times ES_k) + \mu_i + \mu_j + \epsilon_{ki} \end{aligned} \tag{3}$$

where H_{ki} is an indicator that is 1 if employer k hires female applicant i. NT_k, MS_k, FS_k, and ES_k are indicators that are 1 if employer k is assigned to the *No Transport*, the *Male Subsidy*, the *Female Subsidy* or the *Employer Subsidy* treatment, respectively. μ_i and μ_j are strata fixed effects, i.e., female applicant and employer industry fixed effects.[34] We estimate Huber–White robust standard errors clustered at the employer level (the level of randomization).

This specification allows us to test whether employers from the same industry hire the same woman compared to the same man differentially across treatment arms. Specifically, we test:

- Prediction 1: Demand for female labor is lower without safe transport than with: $\beta_1 < 0$.
- Prediction 2: Demand for female labor is higher with safe transport than with subsidies paid to female workers: $\beta_1 + \beta_3 + \beta_6 < 0$.
- Prediction 3: Labor demand is increasing more in subsidies paid to employers than workers: $\beta_2 < \beta_4$, $\beta_3 < \beta_4$.

The first prediction implies that employers are other-regarding. Without taste, profit, or reputation concerns, employers do not have *self*-regarding motives to hire fewer women without transport. The second prediction implies that employers are paternalistic. A deferential employer's utility from hiring women relative to men is strictly greater in the *No Transport*+*Female*

prohibit a covered employer from excluding an individual involuntarily from the workplace based on being older, even if the employer acted for benevolent reasons such as protecting the employee due to higher risk of severe illness from COVID-19" (U.S. EEOC, 2022).

[31] As the subsidy treatments were drawn on-the-spot, they were not stratified.

[32] We observe small differences in perceived revenues across treatments unlikely to drive hiring choices: Employers in the *Female Subsidy* expect women to generate slightly lower revenues: BDT 29, USD 0.3 ($p = 0.04$, appendix table D.2).

[33] We exclude seven female applicants from the application experiment (used to incentivize beliefs, see section 4.1, stages 2 and 4) shown to only one employer. By design, all applicants recruited in the hiring experiment are shown to multiple employers.

[34] We also control for a vector of male applicant characteristics (Excel screening score, education, work experience, and marriage status) in the prediction pair from the application experiment, in which women were not always compared to the same man.

Subsidy treatment than the *Transport+No Subsidy* treatment. Independent of women's and men's valuation of transport, women are strictly better off (they receive a subsidy larger than the cost of transport) while men are strictly worse off (they do not receive transport). Employers may only hire fewer women with the subsidy than the transport if they expect to earn less other-regarding utility because they (i) perceive the subsidy's value to be lower than that of the transport, and (ii) believe women may not purchase the transport because they undervalue it. The third prediction implies that labor demand is locally downward sloping in wages.

Finally, we assess heterogeneity in hiring by employer characteristics elicited from employers *after* making their hiring choices:

- Prediction 4: Employers with larger concerns for women's welfare respond more to safe transport and subsidies paid to female workers.

We test whether the response to the transport and female worker subsidies is larger among:

1. Employers who reported above-median agreement with paternalistic laws in India that restrict women from working at night (on a 0–10 scale with a median response of 8).

2. Employers who reported above-median agreement with the statement that women should not work at night, even if they want to (on a 0–10 scale with a median response of 6).

3. Employers who transferred above-median amounts to the female worker in a three-way dictator game between themselves and two workers recruited separately from the hiring experiment (to ensure that employers did not simply try to compensate workers for not hiring them; BDT 0–100 with a median transfer of BDT 30 or USD 0.3 to male and female workers).[35]

4. Employers who reported maximum agreement with the statement that women should be protected from harmful jobs, even against their will (on a 0–10 scale with a median response of 10).

We also test whether the response to the transport and subsidies is larger among employers with an above-median Kling Mean effects index, i.e., average of the four measures, each standardized by their control mean (Kling et al., 2007).[36] If treatment effects are larger among employers with more other-regarding attitudes towards women, then the observed behavior likely also reflects true underlying other-regarding preferences rather than, for example, experimenter demand effects. Furthermore, if experimenter demand effects drive answers to these questions and observed behavior, employers likely perceive protecting women to be the norm, which is consistent with our interpretation of the findings.[37]

[35] Dictator game transfers are not a direct measure of α_{kg}. In our model, employers should keep the entire amount whenever $\alpha_{kg} < 1$. Instead, we suppose that dictator game transfers suggest individual-level variation in other-regarding preferences.

[36] We find similar results when using a correlation-adjusted index (Anderson, 2008).

[37] We do not differentiate between employers who hire women out of genuine concern, to avoid feeling guilt, or to receive "warm glow" (Andreoni, 1990), consciously or sub-consciously (e.g., through motivated beliefs Bénabou and Tirole (2005)). We consider all these mechanisms as forms of other-regarding preferences.

4.3 Results: Job Costs and Labor Demand

This subsection presents the results of our experimental tests of predictions 1 to 4. We first test whether information about the transport changed employers' beliefs about job costs, and then test whether exogenously changing perceived job costs or payments received by the workers or employers changes hiring decisions.

Not informing employers about the transport increases their perceived job costs (including commute, section 4.1, stage 4) by 1.6 points ($p < 0.01$) from a baseline of 0.9 for male applicants and by 3.1 points ($p < 0.01$) from a baseline of 3.3 for female applicants (appendix table D.2).

Consistent with prediction 1, not informing employers about the transport reduces the share of hired female applicants by 21% (−10ppts, $p < 0.01$) from a baseline of 45% (figure 3, bars 1 and 2).The reduction in demand for female labor seems to be driven by changes on the intensive rather than the extensive margin: Employers in the *No Transport* treatment hire about one fewer woman, on average ($p < 0.01$, appendix table A.2).

As employers only make hiring choices about applicants willing to take the job, these results imply that employers restrict women's employment opportunities when employers consider the opportunities unsafe. Moreover, we find that this employer behavior varies with both employer and applicant characteristics: Employers respond more strongly to the transport information when they are male (even though the difference is not statistically significant given the small number of female employers), or when the female applicant has less experience than the male applicant (appendix figure D.2). We do not find significant differences based on other employer characteristics (appendix figure D.1).[38]

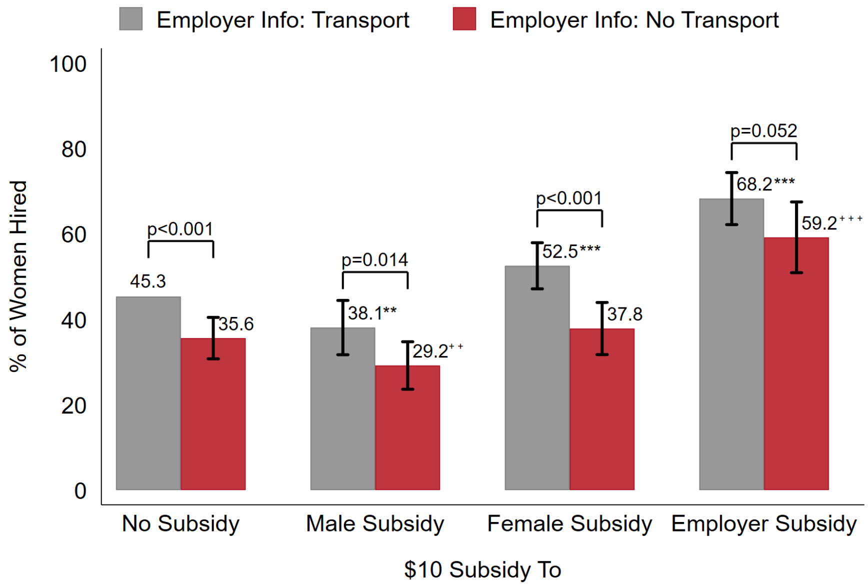
Consistent with prediction 2, employers behave paternalistically rather than deferentially: They hire women more under the *Transport+No Subsidy* condition than under the *No Transport+Female Subsidy* condition (45% versus 38%, bars 1 and 6, $p = 0.02$), i.e., when they know that women receive an additional BDT 1,000 (USD 10). In other words, employers do not hire women without transport—even when women can afford safe transport themselves. This is consistent with the finding that 93% of these employers agree that women should be protected even against their will (choosing 6–10 on the Likert scale, see section 4.2). By contrast, the male subsidy increases male hiring with and without transport (+13%, $p = 0.03$, and +10%, $p = 0.04$). In addition, consistent with prediction 3, the employer subsidy increases hiring more than the worker subsidies with and without transport (+51%, $p < 0.01$, and +66%, $p < 0.01$).

Using the change in hiring in response to applicant characteristics and the subsidy treatments, we estimate employers' valuation of the transport in terms of worker qualifications and payments to female workers or employers themselves. The coefficients on the Excel screening score, the

[38] We do not find evidence that information about the transport affects employers' hiring decisions between male-male applicant pairs, suggesting that paternalistic discrimination is a function of gender rather than characteristics (such as inexperience) associated with gender. We also construct causal trees following Athey and Imbens (2016) and find suggestive evidence that employers between the ages of 26 and 32 who have few female employees respond most to the transport treatment. However, these results do not appear to be robust (results not shown; available on request).

Female Subsidy, and the *Employer Subsidy* imply that employers value the safe transport as much as a seven percentage-point (0.5SD) increase in the Excel score, BDT 1,351 (USD 14) to the worker, or BDT 424 (USD 4) to the employer (appendix table D.4, columns (1) and (2)).

Figure 3: Hiring by Transport Information and Subsidy Assignment



Notes: The graph shows results from equation 3, i.e., the share of women hired by whether the employer knows about the transport or was offered no subsidy, a male or female worker subsidy, or an employer subsidy for hiring women. Each bar is the sum of the control mean and the relevant regression coefficients. We show 95% confidence intervals based on the estimated standard errors of the linear combinations of the regression coefficients. Asterisks are from p-values from Wald tests comparing hiring rates between *No Subsidy* and each of the subsidies with transport, $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$ (on the gray *Transport* bars, only), and pluses from comparing *No Subsidy* and each of the subsidies without transport, $p < 0.10^{+}$, $p < 0.05^{++}$, $p < 0.01^{+++}$ (on the red *No Transport* bars, only). P-values between bars compare hiring rates with and without transport within subsidies.[39]

Finally, consistent with prediction 4, employers with stronger other-regarding preferences respond more strongly to the *No Transport* treatment (figure 4, driven largely by employers with strong paternalistic preferences, appendix figure A.1) and directionally more to the *Female Subsidy* for three out of four measures (appendix figure A.1). Overall, we do not find substantial heterogeneity for two measures, for which we also observe little heterogeneity in the underlying responses, making meaningful employer classification difficult: Consistent with the 50–50 norm (Andreoni and Bernheim, 2009), 71% of employers gave the same dictator game transfers to men and women (85% within BDT 10), and 59% fully agreed (10 out of 10) that women should be protected against their will (only 4% disagreed, 0–4 out of 10).

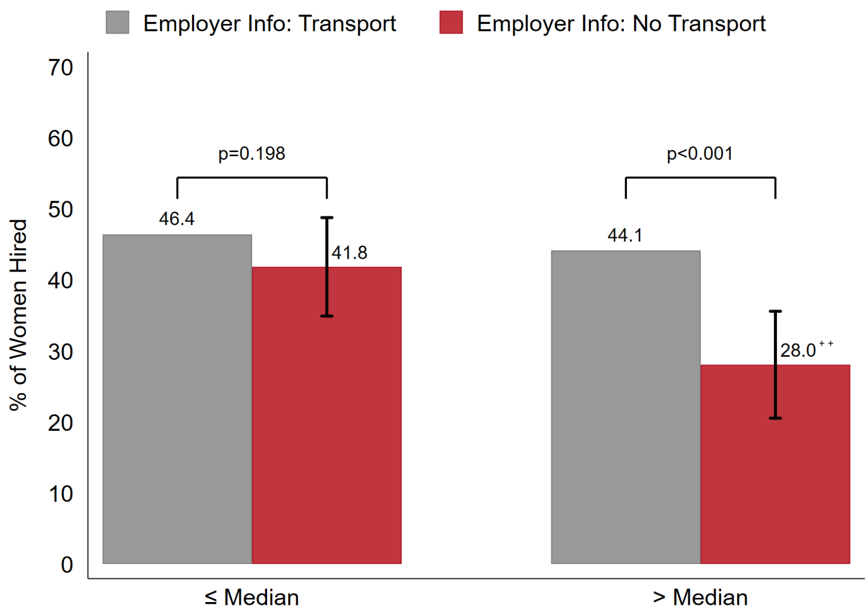
Results are robust to a series of different regression specifications (appendix table D.4). They are robust to (i) removing applicant fixed effects or (ii) all covariates, and (iii) selecting covariates using Belloni et al. (2014)'s post-double selection Lasso method. They are also robust to

[39] The p-values comparing the effect of the female with the male subsidy, with and without transport, are $p = 0.99$ and $p = 0.49$, respectively (not shown). The p-values comparing the effect of the female with the employer subsidy, with and without transport, are $p < 0.01$ and $p < 0.01$, respectively (not shown).

including (iv) employers who answer understanding questions incorrectly, (v) only employers who report that women in the *Transport* treatment will get home using provided transport and women in the *No Transport* will not, and (vi) only employers surveyed before the first night shift (for whom spillovers are impossible). They are also robust to (vii) excluding the applicants from the application experiment, (viii) clustering standard errors at both the employer and applicant level (Cameron et al., 2011), and (ix) using a Logit specification.

We also find that employers respond more to the transport information in a robustness check in a small sample of 41 employers, in which we reduced the salience of gender by presenting the subsidies as random payments to Candidate 1 or Candidate 2 (see appendix figure C.3 for the experimental interface).[40] This result is consistent with enumerator reports that employers made a conscious effort to reduce their biases against women as much as possible in the main study when the experiment's relationship to gender was more salient. Thus, reducing the salience of gender in the experiment *increases* paternalistic discrimination. By contrast, if the results were driven by experimenter demand effects, reducing the salience of gender in the experiment should *reduce* paternalistic discrimination.

Figure 4: Hiring by Transport and Other-Regarding (Deferential or Paternalistic) Preferences



Notes: We show results from regression 3 run separately among different subsets of employers (see section 4.2). Asterisks are from p-values from Wald tests comparing hiring rates of employers with $\leq$ median and above median other-regarding index with transport, $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$ (on the gray *Transport* bars, only), and pluses from Wald tests comparing hiring rates of employers with $\leq$ median and above median other-regarding index without transport, $p < 0.10^{+}$, $p < 0.05^{++}$, $p < 0.01^{+++}$ (on the red *No Transport* bars, only)

[40] The higher female hiring rate with transport is explained by women's Excel screening score being 1.8 points higher than men's in this subsample.

4.4 Mechanisms

We assess the role of three potential drivers of paternalistic discrimination: employers' beliefs about the value of job costs, employers' attitudes toward risk, and selection neglect (that is, failure to condition on applicants' selection into the applicant pool).

1. **Value of Job Costs**: As we discussed in section 4.3, employers assess the value of job costs (on a 0–10 scale, see section 4.1, stage 4) to be significantly higher without transport than with transport, and more so for women than men.

   Employers also believe that women underestimate the job costs more than men, both on the extensive and intensive margin. We elicited first- and second-order beliefs about job costs from 80 *Beliefs-Elicitation* employers, who guessed productivity and costs for eight applicants from the application experiment (second-order beliefs were incentivized with BDT 5, or USD 0.05, per correct answer).[41] Employers believe that both men and women underestimate the costs ($p < 0.01$). In addition, controlling for other characteristics, employers believe that 56% of women and 31% of men underestimate the value of job costs ($p < 0.01$) and that the average conditional mistake (the difference between first- and second-order beliefs) is 1.8 for women versus 1.4 for men ($p < 0.01$). Female employers do not believe that women make larger conditional mistakes than men, which is consistent with the absence of a transport treatment effect for them.

   In addition, we find that employers *overestimate* the frequency of negative events experienced on the night shift ($p < 0.01$ for all comparisons), and more so for female than male workers. We incentivized employers to guess the results of a small survey with 20 male and 20 female night-shift workers (BDT 5, or USD 0.05, per correct answer). Employers believe that (i) 3.3 men and 4.1 women were in a car accident ($p < 0.01$), with the true numbers being 5 and 2, (ii) 4.3 men and 6.3 women were robbed ($p < 0.01$), with the true numbers being 2 and 4, and (iii) 3.2 men and 8.8 women were attacked or assaulted ($p < 0.01$), with the true numbers being 1 and 3.

2. **Risk Preferences**: Employers who believe *women should be* relatively risk-averse (above-median in our sample of employers) but not employers who believe *women are* relatively risk-averse reduce hiring significantly more without transport (see appendix figure D.3). We measure both employers' risk preferences for women and perceptions about women's risk preferences by adapting a question from the Global Preference Survey (Falk et al., 2018, 2023): "In your opinion, on a scale of 0–10, how willing to take risks should women be [are women]?"[42] These results offer additional evidence that employers are paternalistic rather than deferential.[43]

[41] These employers are different from the 80 *Prediction-Only* employers who made predictions about applicants from the hiring experiment.

[42] We opted not to elicit incentivized risk preferences because gambling is illegal in Bangladesh.

[43] We do not believe that low reported risk-preferences for women simply proxy low perceived costs for women or

3. **Selection Neglect**: We find no evidence that selection neglect drives paternalistic discrimination in the experiment. We test for selection neglect by eliciting employers' perceptions of differences in reported job costs between applicants willing and unwilling to take the job at BDT 1,500 (USD 15) in the application experiment (see section 4.1, stage 4). If selection neglect drives discrimination (e.g., by causing employers to evaluate the selected pool of willing applicants as if they were a random draw from the general population), we would expect employers who underestimate the cost differences to respond more strongly to treatment (see, for example, Exley and Nielsen (2024)). However, we find that employers overestimate the reported cost differences between willing and unwilling applicants (2.3 for women and 1 for men, $p < 0.01$, compared to the true values of 0.8 and 0.5) and that hiring behavior does not vary with the perceived difference (see appendix figure D.3).

Overall, our experimental results in the hiring experiment suggest that employers engage in substantial paternalistic discrimination, which is primarily driven by employers beliefs that 1) job costs are high and women underestimate them, and 2) women should be more risk-averse. In the next section, we investigate how applicants respond to changing worker job costs in the application experiment.

5 The Application Experiment: Job Costs and Labor Supply

To measure the labor supply response to variations in gender-specific costs, we conduct an "application experiment" with *applicants* for the Excel workshop and job on the night shift. We recruit 391 men and 379 women aged 18 to 60 through in-person recruitment drives on 11 university campuses in March and April 2023 in Dhaka, Bangladesh.

Applicants are similar to those in the hiring experiment (see table 2). The male applicants are, on average, 26 years old, and the female applicants are, on average, 23 years old. Around one quarter are married (26% of men and 24% of women), and less than one fifth have children (18% of men and 14% of women). Female applicants are less experienced than male applicants (89% have have up to three years of experience versus 72% of male applicants) but have similar education (9% have a Bachelor's degree versus 14% of male applicants) and Excel screening scores (the average score is 26% versus 25% among male applicants).

5.1 Application Experiment Design

Applicants make real application decisions and are randomized into different treatment conditions that experimentally vary the perceived job costs for workers. The experiment is conducted in four stages described below and takes an average of 63 minutes.

paternalistic attitudes. We observe a very low correlation between risk-preferences and perceived costs (r = -0.02) and low correlations between risk-preferences and agreement with paternalism laws in India (r = -0.12), the statement that women should be protected even against their will (r = -0.08), and the statement that women should not work at night (r = -0.18).

1. **Applicant screening:** In the first stage, applicants take two 12-minute back-to-back Excel screening tests incentivized with BDT 2 (USD 0.02) per correct answer for a total compensation of up to BDT 40 (USD 0.4). After completing the tests, applicants are informed that the workshop and job will be from 7:00 P.M. to midnight and all hired workers will receive an Excel certificate of completion.

2. **Transport information randomization:** In the second stage, we randomize applicants into one of two transport treatments that experimentally vary the perceived job costs:[44]

   (a) *Transport:* Applicants are informed about the safe transport home.

   (b) *No transport:* Applicants are not informed about the safe transport home.

3. **Cost beliefs elicitation:** In the third stage, we elicit applicants' unincentivized beliefs about job costs (see section 4.1).

4. **Reservation wage elicitation:** In the fourth stage, we elicit applicants' reservation wages using the Becker–DeGroot–Marschak mechanism (Becker et al. (1964), see appendix figure C.4 for the experimental interface and appendix table C.1 for the wage distribution). Applicants then draw a random wage between BDT 100 (USD 1) and BDT 5,000 (USD 50) by selecting a slip of paper.[45] Applicants are hired if the random wage is at least as high as their reported reservation wage. In total, 231 men and 183 women are hired as part of the application experiment.

5.2 Application Analysis: Empirical Specification

We next test whether the design successfully varied perceived job costs and present the estimating equation that allows us to estimate applicants' valuation of safe transport. We restrict the sample to applicants who answer all understanding questions correctly (91% of male and female applicants) and winsorize reservation wages at the 95th percentile (results are robust to not winsorizing). Applicant characteristics are balanced across treatments (appendix table D.5).

We estimate the following equation separately among male and female applicants:

$$\overline{w}_i = \alpha + \beta_1 NT_i + \beta_2' X_i + \epsilon_i \tag{4}$$

[44] In addition, we experimentally vary the perceived non-wage benefits through two treatments: (i) *High Promotion:* Applicants are informed that 90% of workers hired for the job are promoted. (ii) *Low Promotion:* Applicants are informed that 10% of workers are promoted. In the *Low (High) Promotion* arm, promotions are conducted automatically, selecting the 10% (90%) highest-scoring workers. Applicants determine their promotion treatment assignments by drawing a piece of paper, signaling to them that all promotion assignments are due to chance.

[45] Applicants are informed about the wages in the distribution but not the probability of drawing each wage. We noticed a correlation between the random lottery wage and applicant characteristics mid-survey. In particular, women, educated applicants, and married applicants without children drew higher random wages on average. We were concerned that enumerators might be redrawing the wages to draw higher wages for applicants with higher opportunity costs. We discussed our concerns with the survey firm and started closely supervising the surveys. Enumerators never redrew a wage while we were watching, and we do not observe any correlation between the stated reservation wage and the randomly drawn reservation wage.

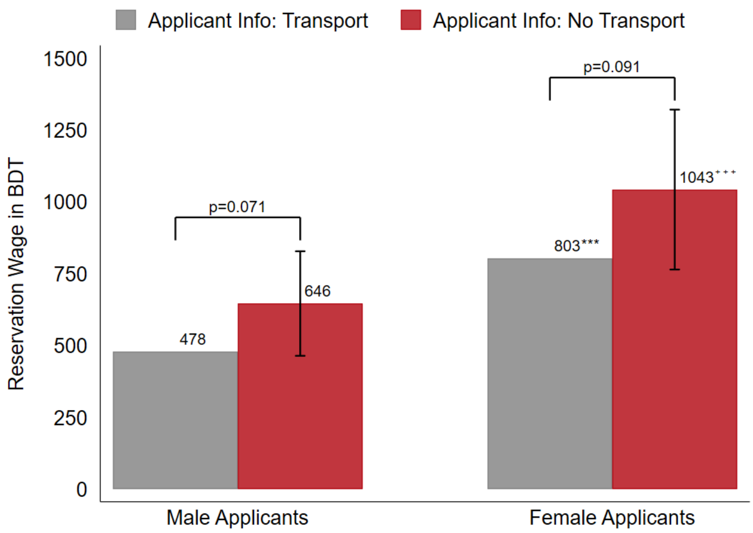
where $\overline{w}_i$ is the stated reservation wage of applicant i, and NT_i is an indicator that is 1 if applicant i is assigned to the *No Transport* treatment. X_i is a vector of applicant controls, including the applicant's age, Excel screening score, education, years of experience, and marital status.[46] ϵ_i are Huber–White robust standard errors.

5.3 Results: Job Costs and Labor Supply

We first test whether information about the transport changed applicants' beliefs about job costs and then whether exogenously changing perceived job costs changes application decisions.

Not informing applicants about the transport increases their perceived job costs by 0.4 points ($p > 0.1$) from a baseline of 2.3 among male applicants and by 0.8 points ($p = 0.03$) from a baseline of 5.9 among female applicants (on a scale from 0–10, appendix table D.5).

Figure 5: Reservation Wages by Applicant Gender and Transport Assignments



Notes: The graph shows results from equation 4 within gender (winsorized at the 95th percentile). Each bar is the sum of the control mean and the relevant regression coefficients. We show 95% confidence intervals based on the estimated standard errors of the regression coefficients. Asterisks are from p-values from Wald tests comparing reservation rates across genders with transport, $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$ (on the gray *Transport* bars, only), and pluses from reservation wages across genders without transport, $p < 0.10^{+}$, $p < 0.05^{++}$, $p < 0.01^{+++}$ (on the red *No Transport* bars, only). Within-gender comparisons of reservation wages with and without transport are given by p-values above bars.

Not informing applicants about the transport increases the reservation wage of male applicants by 35% (BDT 168, USD 2, $p = 0.07$) from a baseline of BDT 478 (USD 5) and that of female applicants by 30% (BDT 240, USD 2, $p = 0.09$) from a baseline of BDT 803 (USD 8, figure 5). Women's significantly higher reservation wage with transport ($p < 0.01$) is consistent with women's lower female labor force participation in this setting and their higher perceived costs even with transport. Interestingly, the valuation of the transport by both male and female applicants is similar to the expected transport price in Dhaka, suggesting that the applicants con-

[46] We also control for indicators for whether the applicant was assigned to the *High Promotion* rate and its interaction with the *No Transport* treatment (see footnote 44).

sidered safe transport as a means of reducing their transport costs. By contrast, employers value the transport significantly more for female workers, at BDT 1,351 (USD 14, $p < 0.01$, section 4.3).[47] Starting from a baseline wage of BDT 1,500 (USD 15), the wage paid in the experiment, male labor supply decreases by 5% (4ppts, p>0.01) without transport and female labor supply by 15% (13ppts, $p = 0.06$, appendix table D.6).

Results are robust to a series of different regression specifications (appendix table D.6). They are robust to (i) truncating and (ii) not winsorizing reservation wages, (iii) including or (iv) excluding outliers, (v) removing covariates or (vi) selecting them using Belloni et al. (2014)'s post-double selection Lasso method, and (vii) including applicants who answer understanding questions incorrectly.

6 Structural Estimation

To quantify the effect of paternalistic discrimination on equilibrium employment and wages and explore prediction 5, we combine the results from the labor demand and supply experiments described in the previous two sections in an equilibrium model. First, we construct the labor demand function using employers' decisions in the hiring experiment, allowing worker selection and expected productivity to vary endogenously with changes in perceived job costs. Second, we construct the labor supply function using applicants' reservation wages in the application experiment. Third, we combine the demand and supply functions to identify equilibrium employment and wages for men and women. Finally, we benchmark the importance of paternalistic discrimination in our setting and assess the effectiveness of safe transport and subsidy interventions.

6.1 Labor Demand

We construct the labor demand in three steps. First, we estimate employers' preference parameters, i.e., how employers trade off taste, profit, and other-regarding concerns in hiring. Second, we estimate how employers' beliefs about worker productivity and welfare respond to changes in wages and transport. Third, we use the estimated preference parameters and predicted beliefs to simulate labor demand in counterfactual markets for male and female workers.

6.1.1 Parametrization

We model the preferences of paternalistic employers from equation 2. We allow employers' profit and welfare beliefs to vary with wages and transport. We treat labor markets for each industry separately, and assume that employers and workers do not move between industries. Within each industry, the markets for male and female labor clear simultaneously. We set d_{jm} to

[47] This is not driven by the different selection of applicants in the two experiments: (i) women with any reservation wage in the application experiment versus (ii) women with a reservation wage of $\leq$ BDT $1,500$ in the hiring experiment, where women knew ex-ante they would be paid BDT 1,500. In fact, we find that women with a reservation wage of $\leq$ BDT $1,500$ in the application experiment react even less to the transport treatment than women with a reservation wage of $>$ BDT $1,500$ (appendix table D.6).

zero and drop the gender subscript such that d_j is the employer's taste for working with women relative to men.

Employer k's expected utility from hiring applicant i of gender $g \in \{m, f\}$ in industry j at wage w_{jg} and without transport $NT \in \{0, 1\}$ is given by the following equation:

$$
\begin{aligned}
u_{ki} &= \underbrace{d_j}_{\text{Taste utility}} + \underbrace{\beta_j \Pi_{ki}(w_{jg}, NT_{jg})}_{\text{Profit utility}} + \underbrace{\alpha_{jg} \mathcal{W}^E_{ki}(w_{jg}, NT_{jg})}_{\text{Other-regarding utility}} + \varepsilon_{ki} \qquad (5) \\
&= v_{ki} + \varepsilon_{ki}
\end{aligned}
$$

where v_{ki} is the utility that varies according to the applicant's gender, expected profit, and other-regarding utility, and ε_{ki} is an unobserved demand shock. The employer's *preference parameters* are given by taste parameter d_j, preference for profits β_j, and other-regarding utility weights α_{jg}. The employer's *beliefs* about the worker's profitability and welfare are given by $\Pi_{ki}(w_{jg}, NT_{jg}) = Y^E(\hat{Y}_{kf}(w_{jf}, NT_{jf}), \hat{Y}_{km}(w_{jm}, NT_{jm})) - L_{kf}w_{jf} - L_{km}w_{jm}$ and $\mathcal{W}^E_{ki}(w_{jg}, NT_{jg}) = w_{jg} - c^E_{kg}(w_{jg}, NT_{jg})$, where Y^E are the employer's expected earnings as a function of perceived female ($\hat{Y}_{kf}(w_{jf}, NT_{jf})$) and male ($\hat{Y}_{km}(w_{jm}, NT_{jm})$) productivity and $c^E_{kg}(w_{jg}, NT_{jg})$ are costs of gender g applicants.

6.1.2 Estimating Employers' Preference Parameters

We estimate employers' preference parameters using a binary choice model. The probability that employer k from industry j chooses to hire applicant i over applicant i' is determined by the relative utility of hiring each applicant:

$$
P_{kii'} = \Pr(u_{ki} > u_{ki'}) = \frac{exp(v_{ki})}{exp(v_{ki}) + exp(v_{ki'})} \qquad (6)
$$

where v_{ki} is the non-random utility of employer k from hiring applicant i in equation 5. Assuming an Extreme Value Type 1 distribution for the unobserved demand shock ε_{ki} in equation 5, we estimate preference parameters $(d_j, \beta_j, \alpha_{jm}, \alpha_{jf})$ in a logit model via maximum likelihood (McFadden, 1974) using the binary hiring decisions of each employer in the experiment.[48] We present results in money-metric utility (to the employer) by dividing d_j, α_{jm} and α_{jf} by β_j. We estimate standard errors via bootstrap.

We estimate equation 5 using the variation in employers' expected profits, Π_{ki}, and welfare, $\mathcal{W}^E_{ki}$. We calculate employers' profit and welfare expectations using the predictions from the *Hiring* employers for the four applicants for whom employers made both predictions and hiring decisions (section 4.1, steps 2 and 4).

First, we calculate employers' expected profits Π_{ki} in the experiment as the difference between the expected revenue generated by the worker and the wage paid. The expected revenue is

[48] We describe the estimation strategy and alternative specifications in appendix C.6. We show robustness to alternative modeling approaches in appendix D.4.

the sum of the employers' base pay of BDT 500 (USD 5) and a piece rate of BDT 5 (USD 0.05) multiplied with the predicted number of tasks completed (the incentivized expected attendance rate multiplied by the incentivized expected conditional number of tasks completed; see section 4.1, step 2). The wage paid is BDT 0 for male workers and $-$BDT 1,000 ($-$USD 10) for female workers in the *Employer Subsidy* treatment and BDT 0 for female workers in all other treatments.

Second, we calculate $\mathcal{W}_{ki}^E$ as the difference between the wage offered to the worker and the expected job costs, $\mathbb{E}_k[u_{ki}^E(c_{ki}^E)]$. Male and female workers in the *No Subsidy* and *Employer Subsidy* treatments receive BDT 1,500 (USD 15) and male workers in the *Male Subsidy* treatment and female workers in the *Female Subsidy* treatment receive BDT 2,500 (USD 25). The expected job costs are the predicted job costs on a scale of 0–10 (see section 4.1, step 4) converted to money-metric using conversion rates calculated from employers' hiring responses to increases in costs and worker wages (described in appendix section C.6.1). We thus assume that Likert scale costs reflect both beliefs about job costs and the employer's preferences over these costs.[49]

Parameter Estimates Employers internalize 11% of payments to male workers and 17% of payments to female workers (table 3).[50] Employers place similar weights on the welfare of male and female workers. We also observe a statistically insignificant distaste for hiring women relative to men (d_j). Holding productivity and worker welfare constant, employers are willing to forego BDT 115 (USD 1) to hire a man rather than a woman.

Table 3: Employer Preferences: Parameter Estimates

	Pooled	**Manufacturing**	**Services**	**Education**
d	-0.115	-0.024	-0.078	-0.190
	(0.085)	(0.139)	(0.528)	(0.120)
α_m	0.111*	0.011	0.256	0.134
	(0.058)	(0.089)	(0.771)	(0.089)
α_f	0.174***	0.175**	0.253	0.149*
	(0.045)	(0.074)	(0.617)	(0.081)
p-val ($\alpha_m = \alpha_f$)	0.393	0.155	0.997	0.901
Observations	1,816	606	588	622

Notes: The table presents parameter estimates from a logit model (equation 6). The sample includes the four applicants with predictions for the 460 employers who answer all understanding questions correctly and are not assigned to the *Employer Subsidy for Hiring Men*. We exclude two observations in a hiring pair with miscoded genders and ten observations in hiring pairs in which at least one applicant revoked their consent. We control for Excel screening score, education, work experience, and marriage status. All estimates in money metric. d in '000 BDT. Standard errors estimated via 1,500 bootstrap samples clustered at the employer level. p-values from testing whether α_m is statistically different from α_f. $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Results are robust to (i) omitting control variables, (ii) using probit rather than logit, (iii) estimating heterogeneous preferences using mixed logit, (iv) implementing a control function

[49] Neither predicted productivity nor costs differ between *Hiring* and *Prediction-Only* employers (see section 4.1, step 2, and appendix table D.3), alleviating concerns that employers may try to hide taste-based behind statistical or paternalistic concerns by understating productivity or overstating costs.

[50] Our estimated welfare weights are slightly lower than those estimated by Chen and Li (2009) (0.32–0.47).

approach (described in appendix C.6.2), and (v) using only random employer and worker subsidies while controlling for the expected revenue and job costs (appendix figure D.4).

6.1.3 Predicting Employer Beliefs as a Function of Wage and Transport

Employers choose to hire workers based on their perceptions of the workers' productivity and on-the-job welfare. We model a hiring process where employers hire randomly from the pool of applicants willing to work at a given wage, but cannot select particular applicants. We estimate the beliefs of employers in industry j about productivity, $\hat{Y}_{jg}(w_{jg}, NT_{jg})$ and costs, $c^E_{jg}(w_{jg}, NT_{jg})$, of gender g workers as a function of wage and transport, i.e., allowing beliefs to vary in applicant selection and productivity. We assume homogeneous beliefs for all employers in a given industry (and thus drop subscript k).

To identify the relationship between beliefs, wages, and transport, we estimate a random forest model in the sample of *Beliefs-Elicitation* employers who each predicted the productivity and costs of eight applicants from the application experiment (see section 4.4). *Beliefs-Elicitation* employers were provided with the applicants' wage and transport conditions in addition to the information provided to *Hiring* employers (gender, Excel screening score, education, work experience, and marital status). We then use the trained random forest to predict the productivity and costs beliefs of *Hiring* employers in each industry at every wage with and without transport based on the characteristics of the workers willing to work at each given wage and transport condition in the application experiment (see appendix section C.6.3 for additional detail). That is, we answer the question "What would the *Hiring* employers have thought about the applicant if we had allowed wages and information about the transport to affect selection and productivity?"

6.1.4 Constructing the Labor Demand Curve

Finally, we use the estimated preference parameters and predicted beliefs to simulate labor demand $\widehat{L^D_g}(w_{jg}, NT_{jg})$ as the fraction of gender g workers demanded at wage w_{jg} with and without transport $NT_{jg} \in \{0, 1\}$ in each industry. As in the experimental set-up, each market consists of 495 employers, 495 male applicants, and 495 female applicants. Each employer chooses how many male and female workers to hire and receives zero taste, profit, or other-regarding utility from any unhired applicant. We assume that the employers' expected revenues follow a constant-elasticity-of-substitution (CES) earnings function that takes as input the expected female ($\hat{Y}_{jf}$) and male ($\hat{Y}_{jm}$) productivity and derive the male and female labor demanded from the employers' first-order conditions (see appendix C.6.4 for the derivation and parameter selection).

6.2 Labor Supply

We estimate labor supply $\widehat{L^S_g}(w_{jg}, NT_{jg})$ non-parametrically as the fraction of gender g workers with reservation wages below every wage w_{jg} with and without transport $NT_{jg} \in \{0, 1\}$.

6.3 Counterfactuals

We conduct three sets of counterfactual analyses. First, we estimate the equilibria with and without transport. Second, we evaluate the importance of paternalistic discrimination relative to other drivers of the gender gaps in employment and wages in the experimental setting, such as taste-based and statistical discrimination or differences in labor supply. Third, we evaluate two counterfactual policies: safe transport and female worker subsidies. We consider three outcomes: worker welfare according to applicants, worker welfare according to employers, and profits.[51]

6.3.1 Counterfactual I: Baseline Equilibrium

Not offering transport to applicants reduces equilibrium female employment by 41% (from 75% to 44%) and female wages by 52% (from BDT 1085 to BDT 516; figure 6; we present results by industry in appendix figure D.5).[52] In addition, not offering transport reduces male employment by 9% and male wages by 24%. Note that the large reductions in both female and male employment and wages reflect both profit and other-regarding concerns, as employers believe that male and female workers are less productive without transport (both in terms of show-up rates and number of tasks completed). However, consistent with prediction 5, as the demand response to transport is larger than the supply response, not offering transport still reduces equilibrium female employment and wages by 30% even when holding constant selection and productivity across transport conditions (appendix figure A.2; male employment and wages do not respond to transport when holding selection and productivity constant).

6.3.2 Counterfactual II: Benchmarking Paternalistic Discrimination

We benchmark the importance of paternalistic discrimination in explaining the gender employment and wage gaps observed without transport in the experiment. We consider a series of counterfactuals that one-by-one eliminate the following sources of gender disparities:

1. Paternalistic discrimination: We equalize other-regarding utility by setting female (i) welfare weights to that of men ($\alpha_{jf} = \alpha_{jm} = \widehat{\alpha_{jm}}$), (ii) expected welfare ($\mathcal{W}^E_{jf}(w) = \mathcal{W}^E_{jm}(w) = \widehat{\mathcal{W}^E}_{jm}(w)$), or (iii) both welfare weights and welfare ($\alpha_{jf}\mathcal{W}^E_{jf}(w) = \alpha_{jm}\mathcal{W}^E_{jm}(w) = \widehat{\alpha_{jm}}\widehat{\mathcal{W}^E}_{jm}(w)$). We present welfare results under employers' status quo beliefs.[53]

[51] We do not consider employer utility as an outcome since it may include employers' discriminatory preferences. The debate over utilitarianism as a guiding principle in economics has a long history (see, for example, Sen (1979)) and is beyond the scope of this paper. See Bernheim and Taubinsky (2018) for a broader discussion of various approaches to welfare analysis with behavioral agents. We use employers' beliefs about welfare and profits instead of workers' true welfare and profits because (i) true welfare is not observed, and (ii) paternalistic employers may not update their beliefs if they rarely hire women.

[52] We do not model firms' entry and exit in a competitive equilibrium. Whether paternalistic firms will exit the market over time depends on a multitude of factors, including the degree of heterogeneity in preferences, the substitutability of male and female labor, and the accuracy of employers' and workers' beliefs.

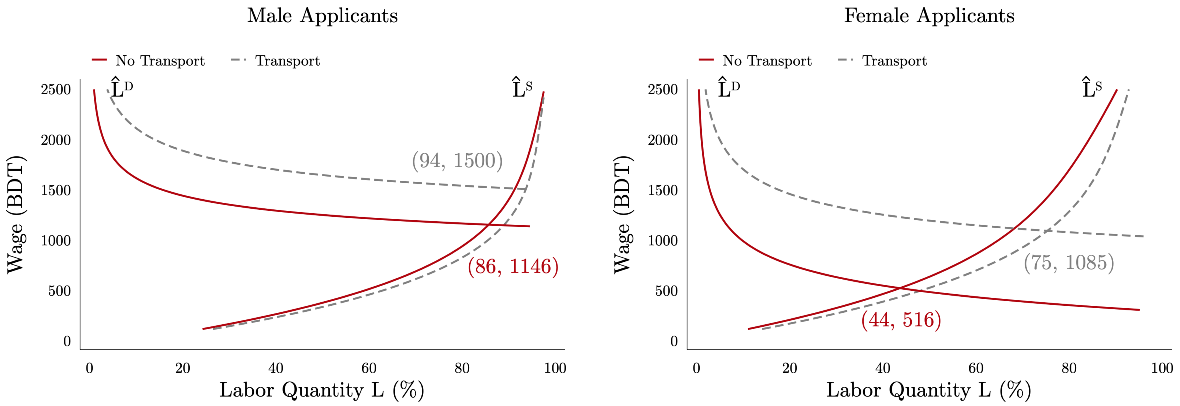
[53] That is, employers in the simulation behave as if they have the same beliefs and preferences toward men and women, but we evaluate employers' perceptions of worker welfare using employers' original beliefs.

2. Taste-based discrimination: We set non-pecuniary returns to zero ($d_j = 0$).

3. Statistical discrimination: We equalize the expected profit at every wage by setting expected female profits to that of men $\left(\Pi^E_{jf}(w) = \Pi^E_{jm}(w) = \widehat{\Pi^E}_{jm}(w)\right)$.

4. Differences in labor supply: We equalize labor supply at every wage by setting female labor supply to that of men $(L^S_f(w) = L^S_m(w) = \widehat{L^S}_m(w))$.[54]

Figure 6: Equilibria in the Male and Female Labor Markets



Notes: The graph shows the share of workers demanded and supplied in male and female labor markets at each wage with and without transport. We use predicted productivity and cost beliefs from the *Beliefs-Elicitation* employers (see section 6.1.3) and calculate profits using the CES production function described in appendix C.6.4. Numbers in parentheses in the graph give (L^*_g, w^*_g). Numbers in gray on top are the equilibrium with transport and numbers in red in the bottom are the equilibrium without transport.

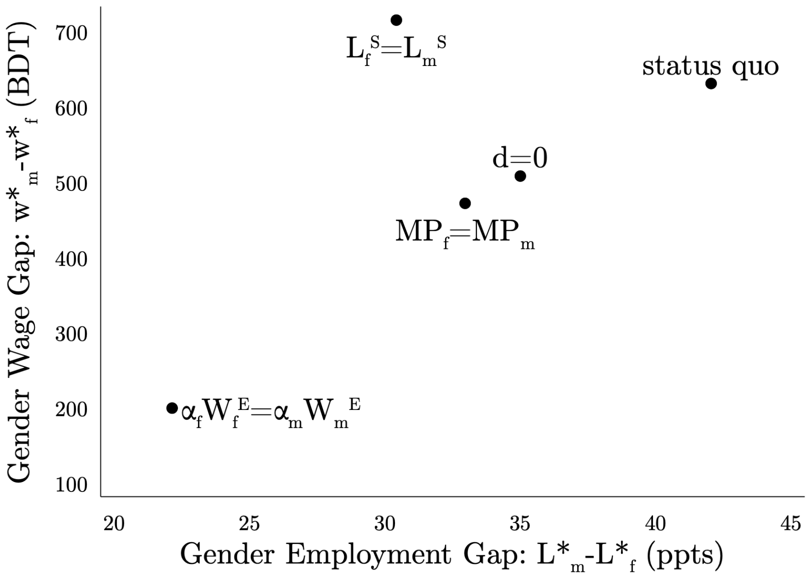
Benchmarking Results Paternalistic discrimination drives a larger share of the gender and wage gaps than statistical and taste-based discrimination in these counterfactual markets (figure 7). Eliminating paternalistic discrimination reduces the gender employment gap by 48% (20 ppts) and the gender wage gap by 68% (BDT 431 or USD 4.3, appendix table D.9). In addition, eliminating paternalistic discrimination reduces worker welfare by 16% using employers' perceptions of worker welfare but increases welfare by 35% using workers' own perceptions of worker welfare. Effects are larger when eliminating differences in perceived welfare rather than differences in the welfare weights placed on men and women.

By contrast, eliminating taste-based and statistical discrimination reduces the gender employment gap by 17% (7 ppts) and 21% (9 ppts) and the gender wage gap by 19% (BDT 123 or USD 1.2) and 25% (BDT 160 or USD 1.6), respectively, while eliminating differences in labor supply reduces the gender employment gap by 29% (12 ppts) but increases the gender wage gap (as the increase in female labor supply puts downward pressure on female wages). Note that paternalistic discrimination may be particularly prevalent in our experimental setting since

[54] We rank male and female applicants by their reservation wages and equate each female applicant's reservation wage with that of her male counterpart. We then recompute female applicants' perceived welfare using these updated wages. We also update demand estimates to account for the changes in selection and its effects on employers' productivity and welfare beliefs (see also section 6.1.3). After equalizing labor supply, we calculate worker welfare based on workers' simulated perceptions at the new levels of supply. This ensures that simulated workers only choose to work when their perceived welfare is positive.

the night shift is highly salient and taste-based and statistical discrimination might be relatively small as employers receive a highly informative signal of applicant quality (the Excel screening score) and do not directly interact with applicants.

Figure 7: Benchmarking Paternalistic Discrimination



Notes: The graph shows the gender employment gap ($L^*_m - L^*_f$) and the gender wage gap ($w^*_m - w^*_f$) of the status quo (the equilibrium in figure 6) as well as in four different counterfactuals that eliminate one-by-one: 1) paternalistic discrimination $\left(\alpha_f \mathcal{W}^E_f = \alpha_m \mathcal{W}^E_m\right)$, 2) taste-based discrimination ($d = 0$), 3) statistical discrimination $\left(\Pi^E_f = \Pi^E_m\right)$, and 4) differences in labor supply $\left(L^S_f = L^S_m\right)$.

We also find that if employers made deferential hiring decisions ($\mathcal{W}^E_g = \mathcal{W}^{E:A}_g$) or used workers' perception of welfare ($\mathcal{W}^E_g = \mathcal{W}^A_g$), total experienced worker welfare would increase by 12–32% and 66–180% using employers' or applicants' perception of worker welfare (appendix table D.9).[55] That is, paradoxically, even employers would consider worker welfare larger in both scenarios as female and male employment and wages would increase.

6.3.3 Counterfactual III: Policy Interventions

Finally, we consider the welfare effects and cost effectiveness of two counterfactual policies in our setting: safe transport or hiring subsidies for female workers.

Safe Transport for Female Workers We estimate the welfare effects and financial cost of providing safe transport to female workers at a cost to the policymaker of BDT 800 (USD 8) for each woman hired. This expense reflects the cost of a private car service in our setting.

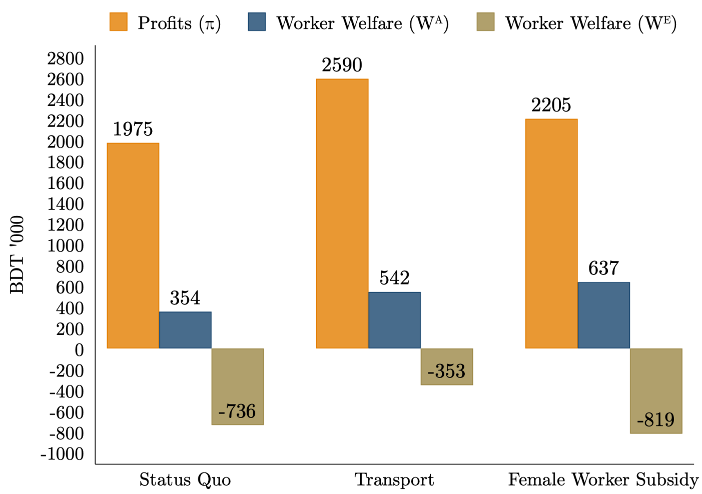
Female Subsidy We estimate the effects of providing female workers a subsidy s. The labor supply at each wage w^E is given by all female workers willing to work at the wage $w^A = w^E + s$.

[55] We use the *Beliefs-Elicitation* employers' reported second-order beliefs to simulate the behavior of deferential employers.

The labor demand at each wage w^E is given by employers' demand for workers willing to work at wage w^A when paying wage w^E. The equilibrium wage equalizes supply and demand. We calculate the subsidy that equalizes the policymaker's expenditures of the transport and subsidy interventions, amounting to BDT 791 (USD 8) per woman hired.[56]

Policy Intervention Results Both the transport and subsidy interventions increase employers' expected profits and workers' assessment of worker welfare; the transport intervention (but not the subsidy) increases employers' assessment of worker welfare (figure 8). The female hiring subsidy reduces employers' assessment of worker welfare by inducing more women to work at a welfare loss that is not sufficiently countered by the increase in wages. Transport and hiring subsidies are about equally effective in reducing the gender employment gap (reducing the gap from 42ppts to 12 and 11ppts, respectively; appendix table D.10). The transport reduces the gender wage gap (from BDT 631, USD 6 to BDT 120, USD 1) while the subsidy reverses the gender gap (–BDT 258, –USD 3). Overall, employers believe that the transport intervention results in a larger increase in expected profits (BDT 615k, USD 6,150) and worker welfare (BDT 383k, USD 3,830) than the subsidy intervention (BDT 230k, USD 2,300 and –BDT 83k, –USD 830) while workers believe that the subsidy intervention results in a larger increase in worker welfare (BDT 283k, USD 2,830) than the transport intervention (BDT 189k, USD 1,890).

Figure 8: Welfare Effects of Transport and Subsidy Interventions



Notes: The figure shows total expected profits and total worker welfare (male + female worker welfare) using applicants' perceptions of worker welfare ($\mathcal{W}^A$) and employers' perceptions of worker welfare ($\mathcal{W}^E$) in three different equilibria: the status quo, a counterfactual equilibrium in which female workers receive free transport and a counterfactual equilibrium in which female workers receive a subsidy of BDT 791 (USD 8). Results in BDT $'000$.

[56] The expense per female worker is higher under the subsidy policy as fewer women get hired than under the transport policy.

7 Conclusion

This paper considers paternalism as a source of labor market discrimination. Combining a labor market model with data from two parallel field experiments, we document a high degree of paternalistic discrimination. An equilibrium model predicts that eliminating paternalistic discrimination reduces the gender employment gap by 48% and the gender wage gap by 68% (USD 4.3) in our experimental setting.

The magnitude of paternalism in our setting is likely similar to other settings in which safety is a salient factor. However, in other hiring settings, employers directly interact with workers, which may increase paternalism if employers feel a greater responsibility for their workers or decrease paternalism if employers update their beliefs about job costs to workers. In addition, paternalism is likely a larger factor in hiring in settings where the information gap between employers and applicants is large.

Studying paternalistic discrimination offers valuable insights for policymakers aiming to affect labor market outcomes. For one, decreasing workers' job costs, both in the workplace or during the commute, or increasing workers' benefits may induce increases in both the supply of and demand for labor. Meanwhile, programs targeting women's labor supplymay be more effective if employers are informed of them. Fundamentally, paternalistic discrimination is driven by the perception that one group faces larger costs from employment than another. If minority status in the workforce or in society itself generates costs to minorities, paternalistic discrimination may lead to a "minority trap" (Cabral, 2022). That is, a disadvantaged group may not be hired because of the very costs related to being disadvantaged, reinforcing the disadvantaged status.[57]

Future research should examine the long-run dynamics of paternalistic discrimination, for both firms and workers. Under what conditions can paternalistic firms remain in business? To what extent does paternalism affect women's career trajectories or preferences, thus contributing to systemic discrimination? Our data suggest that those who suffer the most from paternalistic discrimination are women with little experience. Obstacles to early-career employment may keep these applicants off the career ladder, slowing human capital accumulation and eliminating some future opportunities. While we focus on hiring decisions, other-regarding preferences may also lead to differential treatment in task assignments, promotions, or layoff decisions. Moreover, paternalistic discrimination might occur not only in the labor market but also inside the household (towards daughters) or in school (towards female students), thus differentially shaping the preferences of girls and boys during their most formative stages. Studying these issues can enhance our understanding of gender differences in and outside the labor market and our analysis of available policies.

[57] For example, employers may not hire minority workers if they believe the workers will suffer ostracism.

References

Aachol Foundation (2022, June). Harassment in Public Transport in Dhaka City; Its Impact on Mental Health of Adolescent and Young Women. Technical report.

Adamovic, M. and A. Leibbrandt (2023). A large-scale field experiment on occupational gender segregation and hiring discrimination. *Industrial Relations: A Journal of Economy and Society 62*(1), 34–59.

Aguilar, A., E. Gutiérrez, and P. S. Villagrán (2021). Benefits and unintended consequences of gender segregation in public transportation: Evidence from Mexico City's subway system. *Economic Development and Cultural Change 69*(4), 1379–1410.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics 97*(4), 543–569.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Al-Mamun, M., M. J. Hossain, M. Alam, M. S. Parvez, B. K. Dhar, and M. R. Islam (2022, October). Discrimination and social exclusion of third-gender population (Hijra) in Bangladesh: A brief review. *Heliyon 8*(10).

Allcott, H., J. Kim, D. Taubinsky, and J. Zinman (2021, 09). Are High-Interest Loans Predatory? Theory and Evidence from Payday Lending. *The Review of Economic Studies 89*(3), 1041–1084.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019a). Regressive Sin Taxes, with an Application to the Optimal Soda Tax. *Quarterly Journal of Economics 134*(3).

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019b, Summer). Should We Tax Sugar-Sweetened Beverages? An Overview of Theory and Evidence. *Journal of Economic Perspectives 33*(3), 202–227.

Allcott, H. and D. Taubinsky (2015). Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market. *American Economic Review 105*(8), 2501–38.

Ambuehl, S., B. D. Bernheim, and A. Ockenfels (2021, March). What Motivates Paternalism? An Experimental Study. *American Economic Review 111*(3), 787–830.

Ambuehl, S., A. Ockenfels, and C. Stewart (2022, 11). Who Opts In? Composition Effects and Disappointment from Participation Payments. *The Review of Economics and Statistics*, 1–45.

Anand, B. and S. Kaur (2022, March). State of Discrimination Report: Sub-national comparison of legal barriers to women's right to choose work in India.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal 100*(401), 464–477.

Andreoni, J. and B. D. Bernheim (2009, September). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica 77*(5), 1607–1636.

Arrow, K. J. (1973). The Theory of Discrimination. In O. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets*. Princeton University Press.

Asad, S. A., R. Banerjee, and J. Bhattacharya (2023). Do workers discriminate against their out-group employers? evidence from an online platform economy. *Journal of Economic Behavior Organization 216*, 221–242.

Athey, S. and G. Imbens (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences 113*(27), 7353–7360.

Baert, S. (2018). *Hiring Discrimination: An Overview of (Almost) All Correspondence Experiments Since 2005*, pp. 63–77. Cham: Springer International Publishing.

Bajtelsmit, V. L. and A. Bernasek (1996). Why Do Women Invest Differently than Men? *Financial Counseling and Planning 7*. Available at SSRN: https://ssrn.com/abstract=2238 or http://dx.doi.org/10.2139/ssrn.2238.

Bandiera, O., I. Barankay, and I. Rasul (2005, 08). Social Preferences and the Response to Incentives: Evidence from Personnel Data. *The Quarterly Journal of Economics 120*(3), 917–962.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica 77*(4), 1047–1094.

Banerjee, A., M. Bertrand, S. Datta, and S. Mullainathan (2009). Labor market discrimination in Delhi: Evidence from a field experiment. *Journal of Comparative Economics 37*(1), 14–27. Symposium: Labor Regulation In Developing Countries.

Bangladesh Labour Act (2006). Bangladesh Labour Act, 2006 (XLII of 2006). https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/76402/110637/F-1265526237/BGD76402%20Eng.pdf.

Bartling, B., A. W. Cappelen, H. Hermes, and B. Tungodden (2023, May). Free to Fail? Paternalistic Preferences in the United States. Available at SSRN: https://ssrn.com/abstract=4454985.

BBS (2018, January). Report on Labour Force Survey (LFS) 2016–17. https://mccibd.org/wp-content/uploads/2021/09/Labour-Force-Survey-2016-17.pdf.

BBS (2021). Bangladesh Statistics 2020. https://bbs.portal.gov.bd/sites/default/files/files/bbs.portal.gov.bd/page/a1d32f13_8553_44f1_92e6_8ff80a4ff82e/2021-05-14-06-22-47723b0e1476ed905d1c121f8f07d935.pdf.

BDHS (2016). Bangladesh Demographic and Health Survey 2014.

Becerra, O. and J.-A. Guerra (2023). Personal safety first: Do workers value safer jobs? *Journal of Economic Behavior Organization 212*, 996–1016.

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

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies 81*(2), 608–650.

Bénabou, R. and J. Tirole (2005). Self-confidence and personal motivation. *Psychology, Rationality and Economic Behaviour: Challenging Standard Assumptions*, 19–57.

Bernhardt, A., E. Field, R. Pande, and N. Rigol (2019). Household matters: Revisiting the returns to capital among female microentrepreneurs. *American Economic Review: Insights 1*(2), 141–60.

Bernhardt, A., E. Field, R. Pande, N. Rigol, S. Schaner, and C. Troyer-Moore (2018). Male social status and women's work. In *AEA Papers and Proceedings*, Volume 108, pp. 363–367. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Bernheim, B. D. and D. Taubinsky (2018). Behavioral public economics. In B. D. Bernheim, S. DellaVigna, and D. Laibson (Eds.), *Handbook of Behavioral Economics — Foundations and Applications 1*, Volume 1 of *Handbook of Behavioral Economics: Applications and Foundations 1*, Chapter 5, pp. 381–516. North-Holland.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics 130*(2), 571–614.

Bertrand, M. and S. Mullainathan (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review 94*(4), 991–1013.

Bezirganian, S. and P. Cohen (1992). Sex differences in the interaction between temperament and parenting. *Journal of the American Academy of Child & Adolescent Psychiatry 31*(5), 790–801.

Bohren, J. A., P. Hull, and A. Imas (2022, March). Systemic discrimination: Theory and measurement. Working Paper.

Bohren, J. A., A. Imas, and M. Rosenberg (2019, October). The Dynamics of Discrimination: Theory and Evidence. *American Economic Review 109*(10), 3395–3436.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016, 07). Stereotypes. *The Quarterly Journal of Economics 131*(4), 1753–1794.

Borker, G., G. Kreindler, and D. Patel (2022). Women's Urban Mobility Barriers: Evidence from Delhi's Free Public Transport Policy. Research ongoing.

Boudet, A. M. M. (2013). *On norms and agency: Conversations about gender equality with women and men in 20 countries*. World Bank Publications.

Bursztyn, L., A. W. Cappelen, B. Tungodden, A. Voena, and D. H. Yanagizawa-Drott (2023, March). How are gender norms perceived? Working Paper 31049, National Bureau of Economic Research.

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

Cabral, L. (2022, July). Minority traps. Technical report. Working Paper. http://luiscabral.net/economics/workingpapers/minority.pdf.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Carlana, M. (2019, 03). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *The Quarterly Journal of Economics 134*(3), 1163–1224.

Chan, A. (2022, October). Discrimination against doctors: A field experiment. Unpublished manuscript.

Chaudhary, R. et al. (2021). India's emerging gig economy: shaping the future of work for women.

Cheema, A., A. I. Khwaja, M. F. Naseer, and J. N. Shapiro (2022). Glass walls: Experimental evidence on access constraints faced by women.

Chen, Y. and S. X. Li (2009). Group identity and social preferences. *American Economic Review 99*(1), 431–457.

Choudhury, P. (2015, 07). Explaining Gender Discrimination in the Employment and Earnings of Engineering Graduates in India. *Journal of Educational Planning and Administration 29*, 225–246.

Christensen, P. and A. Osman (2023). The Demand for Mobility: Evidence from an Experiment with Uber Riders. Technical report, National Bureau of Economic Research.

Coutts, A., B. H. Koh, and Z. Murad (2024, January). The signals we give: Gender, feedback, and competition. Available at SSRN: `https://ssrn.com/abstract=4635599` or `http://dx.doi.org/10.2139/ssrn.4635599`.

DeCicca, P., D. Kenkel, and M. F. Lovenheim (2022, September). The economics of tobacco regulation: A comprehensive review. *Journal of Economic Literature 60*(3), 883–970.

Dhillon, A., R. Peeters, O. Bartrum, and A. M. Yüksel (2020). Hiring an employee's friends is good for business: Overcoming moral hazard with social networks. *Labour Economics 67*, 101928.

Dianat, A., F. Echenique, and L. Yariv (2022). Statistical discrimination and affirmative action in the lab. *Games and Economic Behavior 132*, 41–58.

Elías, J. J., N. Lacetera, and M. Macis (2023, April). Is the Price Right? The Role of Economic Tradeoffs in Explaining Reactions to Price Surges. Working Paper 29963, National Bureau of Economic Research.

Erkal, N., L. Gangadharan, and B. H. Koh (2023). Do women receive less blame than men? Attribution of outcomes in a prosocial setting. *Journal of Economic Behavior & Organization 210*, 441–452.

Esponda, I., R. Oprea, and S. Yuksel (2023, 05). Seeing What is Representative. *The Quarterly Journal of Economics 138*(4), 2607–2657.

Evans, G. (1974). Benign discrimination and the right to equality. *Federal Law Review 6*(1), 26–83.

Exley, C. L. and K. Nielsen (2024, March). The Gender Gap in Confidence: Expected but Not Accounted For. *American Economic Review 114*(3), 851–85.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics 133*(4), 1645–1692.

Falk, A., A. Becker, T. Dohmen, D. Huffman, and U. Sunde (2023). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Management Science 69*(4), 1935–1950.

Farole, T., Y. Cho, L. Bossavie, and R. Aterido (2017, October). Bangladesh jobs diagnostic. *2*(9).

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review 103*(1), 472–500.

Field, E., R. Pande, N. Rigol, S. Schaner, and C. Troyer Moore (2021, July). On her own account: How strengthening women's financial control impacts labor supply and gender norms. *American Economic Review 111*(7), 2342–75.

Field, E. and K. Vyborny (2022). Women's Mobility and Labor Supply: Experimental Evidence from Pakistan. *Asian Development Bank Economics Working Paper Series* (655).

Fitriani, R. G. Cooper, and R. Matthews (2016). Women in ground close combat. *The RUSI Journal 161*(1), 14–24.

Fraser, C. (2015). From ladies first to asking for it: Benevolent sexism in the maintenance of rape culture. *California Law Review 103*, 141.

Gallen, Y. and M. Wasserman (2021). Informed choices: Gender gaps in career advice. IZA Discussion Paper 14072. Available at SSRN: `https://ssrn.com/abstract=3775977` or `http://dx.doi.org/10.2139/ssrn.3775977`.

Glick, P. and S. T. Fiske (1997). Hostile and Benevolent Sexism: Measuring Ambivalent Sexist Attitudes Toward Women. *Psychology of Women Quarterly 21*, 119–135.

Glick, P., S. T. Fiske, A. Mladinic, J. L. Saiz, D. Abrams, B. Masser, B. Adetoun, J. E. Osagie, A. Akande, A. Alao, A. Brunner, T. M. Willemsen, K. Chipeta, B. Dardenne, A. Dijksterhuis, D. Wigboldus, T. Eckes, I. Six-Materna, F. Expósito, M. Moya, M. Foddy, H.-J. Kim, M. Lameiras, M. J. Sotelo, A. Mucchi-Faina, M. Romani, N. Sakalh, B. Udegbe, M. Yamamoto, M. Ui, M. C. Ferreira, and W. L. López (2000, Nov). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psycholology 79*(5), 763–75.

Glick, P. and L. Raberg (2018). Benevolent sexism and the status of women. In C. B. Travis, J. W. White, A. Rutherford, W. S. Williams, S. L. Cook, and K. F. Wyche (Eds.), *APA Handbook of the Psychology of Women: History, Theory, and Battlegrounds*, Chapter 18, pp. 363–380. American Psychological Association.

Glover, D., A. Pallais, and W. Pariente (2017). Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores. *The Quarterly Journal of Economics 132*(3), 1219–1260.

Gneezy, U., J. List, and M. K. Price (2012, February). Toward an understanding of why people discriminate: Evidence from a series of natural field experiments. Working Paper 17855, National Bureau of Economic Research.

Goldin, C. and C. Rouse (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review 90*(4), 715–741.

Grosset, F. (2024, March). Complementarities in labor supply. Working paper.

Guenzel, M., C. Hamilton, and U. Malmendier (2023, August). CEO Social Preferences and Layoffs. Working paper.

Hahn, J. and G. Ridder (2017). Instrumental variable estimation of nonlinear models with non-classical measurement error using control variables. *Journal of Econometrics 200*(2), 238–250.

Hjort, J. (2014, 10). Ethnic Divisions and Production in Firms. *The Quarterly Journal of Economics 129*(4), 1899–1946.

Ho, L., S. Jalota, and A. Karandikar (2024, April). Bringing work home: Flexible work arrangements as gateway jobs for women in west bengal. Working paper.

Hossain, T. and R. Okui (2013). The binarized scoring rule. *Review of Economic Studies 80*(3), 984–1001.

International Labour Organization (2018). Global Wage Report 2018/19: What lies behind gender pay gaps. https://www.ilo.org/wcmsp5/groups/public/—dgreports/—dcomm/—publ/documents/publication/wcms_650553.pdf.

Islam, A., D. Pakrashi, S. Sahoo, L. C. Wang, and Y. Zenou (2021). Gender inequality and caste: Field experimental evidence from India. *Journal of Economic Behavior & Organization 190*, 111–124.

Ito, Takahiro (2009). Caste discrimination and transaction costs in the labor market: Evidence from rural North India. *Journal of Development Economics 88*(2), 292–300.

Jacobsson, F., M. Johannesson, and L. Borgquist (2007, 06). Is Altruism Paternalistic? *The Economic Journal 117*(520), 761–781.

Jalota, S. and L. Ho (2024, January). What works for her? how work-from-home jobs affect female labor force participation in urban india. Working paper.

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

Kabir, H. and S. Islam (2023). Sexual Harassment in Public Transport in Dhaka City: A Socio-legal Assessment. *Asian Journal of Sciences and Legal Studies 5*(2), 31–42.

Kessler, J. B., C. Low, and C. D. Sullivan (2019, November). Incentivized resume rating: Eliciting employer preferences without deception. *American Economic Review 109*(11), 3713–44.

Kline, P., E. K. Rose, and C. R. Walters (2022, 06). Systemic Discrimination Among Large U.S. Employers. *The Quarterly Journal of Economics 137*(4), 1963–2036.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Kondylis, F., A. Legovini, K. Vyborny, A. M. T. Zwager, and L. Cardoso De Andrade (2020). Demand for safe spaces: Avoiding harassment and stigma. *World Bank Policy Research Working Paper* (9269).

Kuhle, B. X., D. K. Melzer, C. A. Cooper, A. J. Merkle, N. A. Pepe, A. Ribanovic, A. L. Verdesco, and T. L. Wettstein (2015). The "birds and the bees" differ for boys and girls: Sex differences in the nature of sex talks. *Evolutionary Behavioral Sciences 9*(2), 107–115.

Lata, L. N., P. Walters, and S. Roitman (2021). The politics of gendered space: Social norms and purdah affecting female informal work in Dhaka, Bangladesh. *Gender, Work & Organization 28*(1), 318–336.

Leider, S. and A. E. Roth (2010). Kidneys for sale: Who disapproves, and why? *American Journal of Transplantation 10*(5), 1221–1227.

Levitt, S. D. (2004). Testing theories of discrimination: Evidence from Weakest Link. *The Journal of Law and Economics 47*(2), 431–452.

List, J. A. (2020, July). Non est Disputandum de Generalizability? A Glimpse into The External Validity Trial. Working Paper 27535, National Bureau of Economic Research.

Macchi, E. (2023). Worth your weight: experimental evidence on the benefits of obesity in low-income countries. *American Economic Review 113*(9), 2287–2322.

Mamun, M. Z., M. Ahmed, and S. M. Jahan (2019). Bangladesh's readiness for the international call center industry. *Journal of Bangladesh Studies 21*(2).

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Fontiers in Econometrics*, pp. 105–142. New York: Academic press.

McKelway, M. (2023, August). Information, Norms, and Female Employment: An Experiment in India. Working paper available at https://drive.google.com/file/d/1G_ZI0ug2Zzy47YYXiWhXyJ_WX1ZkzBZx/view.

Naher, K. (2022, October). Female employment is rising rapidly. Then why are women-only buses 'unprofitable'? *The Business Standard*.

Niederle, M. and E. Vespa (2023, February). Cognitive limitations: Failures of contingent thinking. https://drive.google.com/file/d/104jv_CrDqzVbmajXCxrwR2dR-NSaDTre/view. Unpublished manuscript.

Patty, W. F. (1989). Constitutional Law—Affirmative Action—The Supreme Court Decides on a Standard of Review for Benign Discrimination. *Cumberland Law Review 20*, 849.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research 47*(1), 3–13.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review 62*(4), 659–661.

Quayyum, N. (2019). Women workers in Bangladesh's ready-made garment industry: Building an infrastructure of dissent. *Journal of Labor and Society 22*(4), 835–852.

Rabin, M. (2013). Risk aversion and expected-utility theory: A calibration theorem. In L. C. MacLean and W. T. Ziemba (Eds.), *Handbook of the Fundamentals of Financial Decision Making: Part I*, pp. 241–252. World Scientific.

Rahman, M. and M. Al-Hasan (2022). The Reverse Gender Wage Gap in Bangladesh: Demystifying the Counterintuitive. *Indian Journal of Labour Economics 65*, 929–950.

Rahman, M. S.-U. (2010, 12). Bus Service for 'Women Only' in Dhaka City: An Investigation. *Journal of Bangladesh Institute of Planners 3*, 17–32.

Rivers, D. and Q. H. Vuong (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics 39*(3), 347–366.

Sen, A. (1979). Utilitarianism and welfarism. *The Journal of Philosophy 76*(9), 463–489.

Shnabel, N., Y. Bar-Anan, A. Kende, O. Bareket, and Y. Lazar (2016). Help to perpetuate traditional gender roles: Benevolent sexism increases engagement in dependency-oriented cross-gender helping. *Journal of Personality and Social Psychology 110*(1), 55–75.

Siddique, Z. (2011). Evidence on caste based discrimination. *Labour Economics 18*, S146–S159. Labour markets in developing countries.

Siddique, Z. (2022). Media-reported violence and female labor supply. *Economic Development and Cultural Change 70*(4), 1337–1365.

The Labour Protection Act B.E. 2541 (2014). The Labour Protection Act B.E. 2541.

Thorat, S. and P. Attewell (2007). The legacy of social exclusion: A correspondence study of job discrimination in India. *Economic and Political Weekly 42*(41), 4141–4145.

Uhl, M. (2011). Do self-committers mind other-imposed commitment? An experiment on weak paternalism. Unpublished manuscript. Available at https://jlupub.ub.uni-giessen.de/handle/jlupub/447.

US Department of State (2022a). Argentina Human Rights Report. https://www.state.gov/reports/2022-country-reports-on-human-rights-practices/argentina/.

US Department of State (2022b). Korea Human Rights Report. https://www.state.gov/reports/2022-country-reports-on-human-rights-practices/south-korea/.

U.S. EEOC (2007, May). Enforcement guidance: Unlawful disparate treatment of workers with caregiving responsibilities. https://www.eeoc.gov/laws/guidance/enforcement-guidance-unlawful-disparate-treatment-workers-caregiving-responsibilities#benevolent.

U.S. EEOC (2022, July). What You Should Know About COVID-19 and the ADA, the Rehabilitation Act, and Other EEO Laws. https://www.eeoc.gov/wysk/what-you-should-know-about-covid-19-and-ada-rehabilitation-act-and-other-eeo-laws. Technical Assistance Questions and Answers.

Villas-Boas, J. M. and R. S. Winer (1999). Endogeneity in brand choice models. *Management Science 45*(10), 1324–1338.

White, S. C. (2017). Patriarchal Investments: Marriage, Dowry and the Political Economy of Development in Bangladesh. *Journal of Contemporary Asia 47*(2), 247–272.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources 50*(2), 420–445.

World Bank (2023a). Women, Business and the Law 1971–2023 Dataset. Technical report, World Bank. https://wbl.worldbank.org/content/dam/sites/wbl/documents/2023/WBL-1971-2023-Dataset.xlsx.

World Bank (2023b). Women, Business and the Law 2023. Technical report, World Bank.

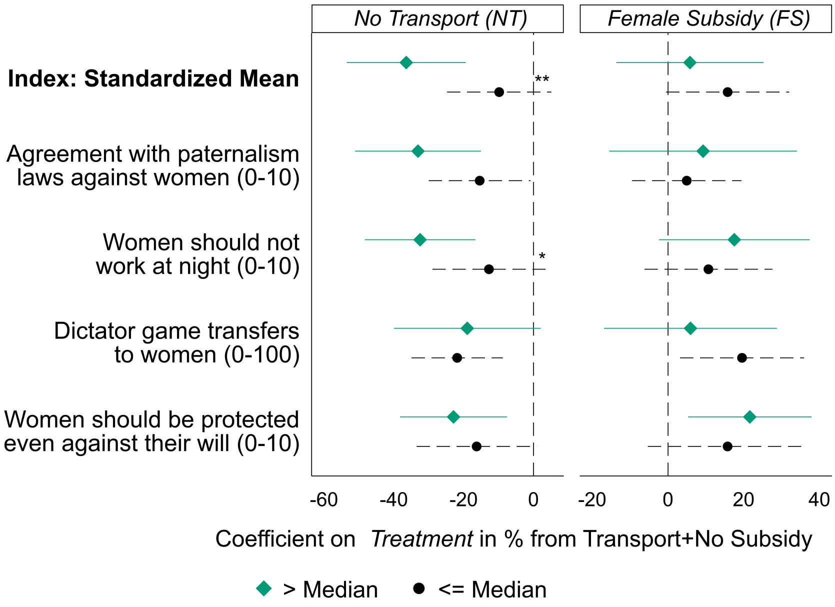
World Bank DataBank (2023). World development indicators. Technical report, World Bank. https://databank.worldbank.org/source/world-development-indicators.

World Values Survey (2018). World Values Survey (WVS) 2018. https://www.worldvaluessurvey.org/WVSOnline.jsp.

A Appendix

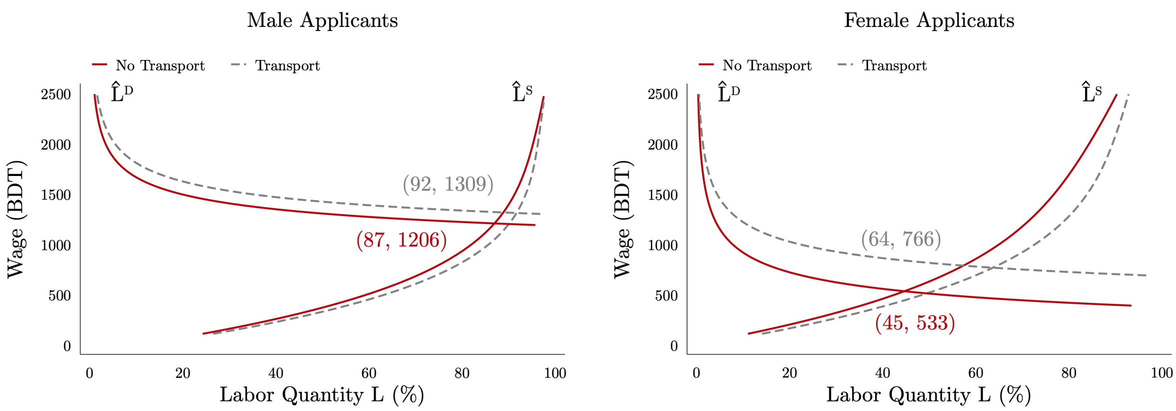
A.1 Figures

Figure A.1: Hiring by Transport Information, Female Subsidy and Other-Regarding Preferences



Notes: The graph shows the coefficients on the *No Transport* and *Female Subsidy* indicators from regression 3, respectively. Regressions are run separately among different subsets of employers (see section 4.2). That is, each coefficient shows how much employers in that group reduce female hiring when they do not know about the safe transport or increase female hiring when they know about the female subsidy. The index is formed as the mean of the standardized continuous and not binary variables. Thus, the treatment effects do not need to be the averages of the treatment effects of the binary measures. Asterisks from comparing the coefficients across subsamples. $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Figure A.2: Equilibria in the Male and Female Labor Markets, Holding Selection and Productivity Constant Across Wages and Transport Conditions



Notes: The graph shows the share of workers demanded and supplied in male and female labor markets at each wage with and without transport. We use predicted productivity and cost beliefs from the *Beliefs-Elicitation* employers (see section 6.1.3) and calculate profits using the CES production function described in appendix C.6.4. Numbers in parentheses in the graph give (L_g^*, w_g^*). Numbers in gray on top are the equilibrium with transport and numbers in red in the bottom are the equilibrium without transport.

A.2 Tables

Table A.1: Employer Characteristics, by Industry

	Manufacturing		Services		Education	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Female (%)	3.9	19.5	1.3	11.5	13.3	34.1
Age (Years)	32.4	7.8	32.2	7.8	29.9	7.7
Married (%)	72.5	44.8	62.4	48.6	41.1	49.4
Children (%)	58.2	49.5	49.7	50.2	28.5	45.3
Bachelor's (%)	11.8	32.4	32.0	46.8	81.0	39.3
# Male Employees	11.0	37.6	3.2	4.0	12.4	17.0
# Female Employees	11.5	71.0	0.2	0.9	6.3	9.6
# Hiring Decisions Last 3 Years	53.2	405.3	10.3	25.5	17.8	37.1

Notes: The table shows the means and standard deviations of characteristics of employers by industry in the analysis sample of the hiring experiment. *Children* is an indicator equal to 1 if the applicant has at least one child. *Bachelor's* is an indicator equal to 1 if the applicant has at least a Bachelor's degree.

Table A.2: Treatment Effect on Share of Employers Hiring Women and Number of Women Hired

	Hired Woman (%)	# Women
	(1)	(2)
No transport (NT)	-0.017	-0.920***
	(0.024)	(0.236)
Male subsidy (MS)	-0.033	-0.619**
	(0.041)	(0.286)
Female subsidy (FS)	0.001	0.531**
	(0.026)	(0.247)
Employer subsidy (ES)	0.022	1.974***
	(0.015)	(0.300)
NT*MS	-0.003	0.030
	(0.063)	(0.409)
NT*FS	0.015	-0.343
	(0.039)	(0.389)
NT*ES	-0.012	0.374
	(0.037)	(0.510)
Control Mean	0.979	4.587
Observations	460	446

Notes: The table shows results from OLS regressions with Huber–White robust SEs, controlling for industry fixed effects. The unit of observation is the employer. Column (1) keeps all employers who correctly answer the understanding questions. The outcome is whether the employer hires at least one woman. Column (2) keeps all employers who answer the understanding questions correctly and hire at least one woman. The outcome is the number of women hired by the employer. $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

ONLINE APPENDIX
FOR ONLINE PUBLICATION ONLY

B Theory Appendix

B.1 Production Function

We make several assumptions about the production function to ensure a unique solution to the employer's problem:

1. $Y^E(L_{kf}, L_{km})$ is a non-negative, continuously differentiable function with existing second derivatives.

2. $\lim_{L_{kg}\to 0^+} \dfrac{\partial Y^E(L_{kf}, L_{km})}{\partial L_{kg}} \to \infty$.

3. $\frac{\partial^2 Y^E(L_{kf},L_{km})}{\partial L_{kf}^2} < 0$ and $\frac{\partial^2 Y^E(L_{kf},L_{km})}{\partial L_{km}^2} < 0$ for all L_{kf}, L_{km}.

4. $\frac{\partial^2 Y^E(L_{kf},L_{km})}{\partial L_{kf}^2}\frac{\partial^2 Y^E(L_{kf},L_{km})}{\partial L_{km}^2} > \left(\frac{\partial^2 Y^E(L_{kf},L_{km})}{\partial L_{km}\partial L_{kf}}\right)^2$ for all L_{kf}, L_{km}.

Assumption 1 ensures tractability. Assumption 2 ensures that each employer hires both men and women. Assumptions 3 and 4 ensure that the production function is concave. For example, the CES production function satisfies these assumptions.

B.2 Derivation of Prediction 1

We derive prediction 1 in three steps: First, we describe the employers' expected change in worker welfare in response to increases in job costs c_g. Second, we derive the first-order conditions that describe the employers' problems. Third, we use the employers' expected change in worker welfare as well as the first-order conditions to derive the demand response to changes in job costs c_g.

First, we describe the change in welfare in response to increases in gender-specific costs c_g.

$$\frac{\partial \mathcal{W}_{kg}}{\partial c_g} = -\frac{\partial}{\partial c_g}\mathbb{E}_k[u_{ki}(c_g + c_i + c_{kg})\,|\mathbb{E}_i[u_i^{E:A}(c_g + c_i + c_{kg})] \leq w_g], \tag{7}$$

for $u_{ki} \in \{u_{ki}^E, u_i^{E:A}\}$.

A change in job costs has two effects: (i) *direct*: job costs increase, thereby reducing the employer's perception of applicant utility, and (ii) *selection*: workers with smaller individual job costs self-select into the job, thereby partially offsetting the increase in perceived job costs. Without selection effect, for example, when employers engage in *selection neglect*, i.e., they do not consider how changing job costs changes the selection of workers, $\frac{\partial \mathcal{W}_{kg}}{\partial c_g} < 0$. As we make predictions for the experiment, we assume that selection is fixed going forward, i.e., that $\frac{\partial \mathcal{W}_{kg}}{\partial c_g} < 0$ and equal to the direct effect.

Second, we pin down the labor demand using the first-order conditions implied by the employer's problem (equation 2).

$$\begin{aligned}
\underset{L_{kf}}{FOC} \quad & d_{kf} + \frac{\partial Y^E(L_{kf}, L_{km})}{\partial L_{kf}} + \alpha_{kf}\mathcal{W}_{kf} - w_f = 0 \\
\underset{L_{km}}{FOC} \quad & d_{km} + \frac{\partial Y^E(L_{kf}, L_{km})}{\partial L_{km}} + \alpha_{km}\mathcal{W}_{km} - w_m = 0.
\end{aligned} \tag{8}$$

Given assumptions 1–4 about the shape of the production function, the above system of equations has a unique maximum. Note that the employer hires until the utility contributed by the marginal worker is equal to the wage.

Third, implicit differentiation of the first-order conditions 8 yields the following comparative static:

$$\frac{\partial L_{kf}}{\partial c_f} = -\frac{\alpha_{kf} \overbrace{\frac{\partial \mathcal{W}_{kf}}{\partial c_f}}^{<0} \overbrace{\frac{\partial^2 Y^E}{\partial L_{km}^2}}^{<0}}{\underbrace{\frac{\partial^2 Y^E}{\partial L_{kf}^2}\frac{\partial^2 Y^E}{\partial L_{km}^2} - \left(\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}}\right)^2}_{>0}} \tag{9}$$

This is $= 0$ if and only if $\alpha_{kf} = 0$, > 0 if and only if $\alpha_{kf} < 0$ and < 0 if and only if $\alpha_{kf} > 0$. The results are equivalent when considering $\frac{\partial L_{km}}{\partial c_m}$.

B.3 Derivation of Prediction 2

Prediction 2 states that holding constant selection and productivity, the labor demand of deferential employers is increasing weakly more in cash payments to workers than in equally (or lower) priced worker amenities.

To see this, consider a worker who does not receive a given amenity as part of her job. If the amenity is sufficiently valuable and within the worker's budget, the worker will purchase the amenity. As a result, the worker is at least as well off with a cash payment that allows her to purchase the amenity than with receiving the amenity directly. That is, an in-kind amenity is (weakly) less valuable than a cost-equivalent cash payment.

Knowing this, a deferential employer understands that a subsidy to women is at least as valuable to female workers as a transport home, as long as the value of the subsidy exceeds the cost of the transport. Thus, a deferential employer's other-regarding utility is larger from hiring a worker who has the option to either purchase the amenity or keep the extra cash compared to hiring a worker who can only receive the amenity.

B.4 Derivation of Prediction 3

We derive prediction 3 in two steps: First, we describe the employers' expected change in worker welfare in response to increases in wages w_g. Second, we use the employers' expected change in worker welfare as well as the first-order conditions to derive the demand response to changes in wages w_g.

First, we describe the change in welfare in response to increases in gender-specific wages w_g.

$$\frac{\partial \mathcal{W}_{kg}}{\partial w_g} = 1 - \frac{\partial}{\partial w_g}\mathbb{E}_k[u_{ki}(c_i + c_{kg} + c_g)|\mathbb{E}_i[u_i^{E:A}(c_i + c_{kg} + c_g)] \leq w_g], \tag{10}$$

for $u_{ki} \in \{u_{ki}^E, u_i^{E:A}\}$.

Wage affects the employer's view of worker welfare through two channels. First, a wage increase directly contributes to worker welfare; higher wages are more desirable. Second, a selection effect changes the composition of workers. In particular, when the wage increases, the higher wage attracts workers with higher worker-specific costs, resulting in a decrease in worker welfare. The relative size of the direct and selection effects depend on the levels of cost as well as the utility functions. Welfare is unambiguously increasing when holding selection fixed, or when employers engage in selection neglect.

As we make predictions for the experiment, we assume that selection is fixed going forward, i.e., that $\frac{\partial \mathcal{W}_{kg}}{\partial w_g} = 1$.

Second, implicit differentiation of the first-order conditions 8 yields the following comparative static:

$$\frac{\partial L_{kf}}{\partial w_f} = \frac{\overbrace{\frac{\partial^2 Y^E}{\partial L_{km}^2}}^{<0}(1-\alpha_{kf})}{\underbrace{\frac{\partial^2 Y^E}{\partial L_{kf}^2}\frac{\partial^2 Y^E}{\partial L_{km}^2} - \left(\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}}\right)^2}_{>0}} \tag{11}$$

The above is ≥ 0 if and only if $\alpha_{kf} \geq 1$, i.e., when employers do not place a higher weight on their own welfare than that of the worker.

B.5 Derivation of Prediction 4

Note that, if $\alpha_{kf} > 0$, then equations 9 and 11 are decreasing in absolute value in α_{kf}.

B.6 Derivation of Prediction 5

We derive prediction 5 in four steps. First, we assess how the demand for workers of the opposite gender is changing in gender-specific costs. Second, we set up the system of equations describing the equilibrium. Third, we show that this system of equations has a solution. Fourth, we show how the equilibrium labor quantity and wages respond to changes in gender-specific costs c_g.

First, to assess under which conditions male and female workers are substitutes, complements, or neither, we calculate the cross-wage elasticity of demand of male labor with respect to female wages. Male and female workers are substitutes if the cross-wage elasticity is positive, i.e., an increase in female wages increases the demand for male workers, complements if the cross-wage elasticity is negative, i.e., an increase in female wages decreases the demand for male workers, and neither substitutes nor complements if the cross-wage elasticity is 0.

$$\epsilon_{w_f,w_m} = \frac{w_f}{L_{km}}\frac{\partial L_{km}}{\partial w_f} = -\underbrace{\frac{w_f}{L_{km}}}_{<0}\frac{\overbrace{\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}}}^{?}\overbrace{(1-\alpha_{kf})}^{>0}}{\underbrace{\frac{\partial^2 Y^E}{\partial L_{kf}^2}\frac{\partial^2 Y^E}{\partial L_{km}^2} - \left(\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}}\right)^2}_{>0}} \tag{12}$$

Thus, $\epsilon_{w_f,w_m} > 0$ if and only if $\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}} < 0$, $\epsilon_{w_f,w_m} < 0$ if and only if $\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}} > 0$ and $\epsilon_{w_f,w_m} = 0$ if and only if $\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}} = 0$. That is, male and female workers are substitutes if and only if $\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}} < 0$, complements if and only if $\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}} > 0$, and neither substitutes nor complements if and only if $\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}} = 0$. Therefore, the demand for male workers is increasing in female wages if male and female workers are substitutes, decreasing if they are complements, and constant if they are neither substitutes nor complements. The change in the demand for male workers is decreasing in α_{kf}.

Next, we derive how labor demand is changing in increases in gender-specific costs of the opposite gender using implicit differentiation of the first-order conditions 8:

$$\frac{\partial L_{km}}{\partial c_f} = \frac{\alpha_{kf}\overbrace{\frac{\partial \mathcal{W}_{kf}}{\partial c_f}}^{<0}\overbrace{\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}}}^{?}}{\underbrace{\frac{\partial^2 Y^E}{\partial L_{kf}^2}\frac{\partial^2 Y^E}{\partial L_{km}^2} - \left(\frac{\partial^2 Y^E}{\partial L_{km}\partial L_{kf}}\right)^2}_{>0}}. \tag{13}$$

Assuming $\alpha_{kf} \geq 0$, the above is 0 if $\alpha_{kf} = 0$ or if $Y^E_{L_{km},L_{kf}} = 0$, i.e., male and female workers are neither substitutes nor complements. For $\alpha_{kf} > 0$, it is < 0 if $Y^E_{L_{km},L_{kf}} > 0$, i.e., male and female workers are complements, and > 0 if $Y^E_{L_{km},L_{kf}} < 0$, i.e., male and female workers are substitutes. That is, as the costs to female workers increase, other-regarding employers hire more male workers if male and female workers are substitutes, and fewer male workers if they are complements. Note that the change in male hiring is increasing in α_{kf}. Note from equation 9 that an increase in costs has the smallest effect on demand for female labor when male and female labor are neither substitutes nor complements. Intuitively, if male and female labor are substitutes, employers substitute towards male labor when the perceived costs to female labor increase. On the other hand, if male and female labor are complements, the demand for male labor decreases, further reducing the demand for female labor.

Second, we set up the system of equations describing the equilibrium.

Let c_i follow distribution h^I_g, which is a continuously differentiable density function with no mass points. The labor supply of gender g labor is then given by the following equation:

$$L_g = L_g^S \equiv \int_i \mathbb{1}\left(\mathbb{E}_i[u^A(c_i + c_{kg} + c_g)] \leq w_g\right) h_g^I(c_i) dc_i \tag{14}$$

As $\frac{\partial Y^E(L_{kf}, L_{km})}{\partial L_{kg}}$ is a continuous function that increases monotonically in L_{kg} for $g \in \{f, m\}$, the inverse function $\left(\frac{\partial Y^E(L_{kf}, L_{km})}{\partial L_{kg}}\right)^{-1}$ exists. The demand for gender g labor is then given by the following equation:

$$L_g = L_g^D \equiv \int_k \left(\frac{\partial Y^E(w_g - d_{kg} - \alpha_{kg}\mathcal{W}_{kg}, L_{kg'})}{\partial L_{kg}}\right)^{-1} dk, \tag{15}$$

where g' is the opposite gender.

The system of equations given by 14 and 15 then describes the equilibrium.

Third, we show that the system of equations describing the equilibrium has a solution. It has a local solution if it has continuous partial derivatives with respect to all endogenous and exogenous variables and the determinant of the Jacobian of the system of equations is non-zero. This Jacobian is given by the matrix on the left of the following equation:

$$\begin{bmatrix} 1 & 0 & -\frac{\partial L_f^S}{\partial w_f} & 0 \\ 0 & 1 & 0 & -\frac{\partial L_m^S}{\partial w_m} \\ 1 & 0 & -\frac{\partial L_f^D}{\partial w_f} & -\frac{\partial L_f^D}{\partial w_m} \\ 0 & 1 & -\frac{\partial L_m^D}{\partial w_f} & -\frac{\partial L_m^D}{\partial w_m} \end{bmatrix} \begin{bmatrix} \frac{\partial L_f^*}{\partial c_f} \\ \frac{\partial L_m^*}{\partial c_f} \\ \frac{\partial w_f^*}{\partial c_f} \\ \frac{\partial w_m^*}{\partial c_f} \end{bmatrix} = \begin{bmatrix} \frac{\partial L_f^S}{\partial c_f} \\ 0 \\ \frac{\partial L_f^D}{\partial c_f} \\ \frac{\partial L_m^D}{\partial c_f} \end{bmatrix} \tag{16}$$

The following equation gives the determinant of the Jacobian:

$$|J| = \underbrace{\left(\frac{\partial L_f^D}{\partial w_f}\frac{\partial L_m^D}{\partial w_m} - \frac{\partial L_f^D}{\partial w_m}\frac{\partial L_m^D}{\partial w_f}\right)}_{=(1-\alpha_{kf})(1-\alpha_{km})>0} \underbrace{+\frac{\partial L_f^S}{\partial w_f}\left(\frac{\partial L_m^S}{\partial w_m} - \frac{\partial L_m^D}{\partial w_m}\right)}_{>0} \underbrace{-\frac{\partial L_f^D}{\partial w_f}\frac{\partial L_m^S}{\partial w_m}}_{>0}$$

Thus, as the Jacobian is positive, the system of equation has a solution.

Next, we show how the equilibrium labor quantity and wages respond to changes in gender-specific costs c_g. By Cramer's rule, the aggregate solution can be expressed as

$$\frac{\partial L_f^*}{\partial c_f} = \frac{|J_1|}{|J|} \qquad \frac{\partial L_m^*}{\partial c_f} = \frac{|J_2|}{|J|}$$

$$\frac{\partial w_f^*}{\partial c_f} = \frac{|J_3|}{|J|} \qquad \frac{\partial w_m^*}{\partial c_f} = \frac{|J_4|}{|J|}.$$

Here $|J_j|$ is the matrix resulting from replacing the jth column of the Jacobian matrix with the solution to the system of equations. As we already know that $|J| > 0$, we only have to sign $|J_j|$.

To evaluate the sign of $\frac{\partial L_f^*}{\partial c_f} = \frac{|J_1|}{|J|}$, we calculate $|J_1|$ and re-arrange:

$$|J_1| = \underbrace{\frac{\partial L_f^S}{\partial c_f}}_{<0} \left(\underbrace{\frac{\partial L_f^D}{\partial w_f}\frac{\partial L_m^D}{\partial w_m} - \frac{\partial L_f^D}{\partial w_m}\frac{\partial L_m^D}{\partial w_f}}_{(1-\alpha_{kf})(1-\alpha_{km})>0} \underbrace{- \frac{\partial L_f^D}{\partial w_f}\frac{\partial L_m^S}{\partial w_m}}_{>0} \right)$$
$$\underbrace{-\frac{\partial L_f^S}{\partial w_f}}_{<0} \left(\underbrace{\frac{\partial L_f^D}{\partial c_f}\frac{\partial L_m^D}{\partial w_m} - \frac{\partial L_f^D}{\partial w_m}\frac{\partial L_m^D}{\partial c_f}}_{-(1-\alpha_{km})\alpha_{kf}\frac{\partial \mathcal{W}_f}{\partial c_f}>0} \underbrace{- \frac{\partial L_f^D}{\partial c_f}\frac{\partial L_m^S}{\partial w_m}}_{>0} \right) < 0 \tag{17}$$

An increase in costs to female labor reduces the equilibrium quantity of women hired by reducing both female labor supply and demand for female labor. The magnitude of the shift depends on both the responsiveness of demand and supply to costs and wages. Note that the above is true for any $\alpha_f \in [0,1]$, implying that the equilibrium female labor quantity is decreasing in costs to female labor in a model with and without other-regarding preferences.

The results are equivalent when considering $\frac{\partial L_m^*}{\partial c_m}$.

To evaluate the sign of $\frac{\partial w_f^*}{\partial c_f} = \frac{|J_3|}{|J|}$, we calculate $|J_3|$ and re-arrange:

$$|J_3| = \underbrace{\frac{\partial L_f^D}{\partial w_m}\frac{\partial L_m^D}{\partial c_f}}_{>0} + \underbrace{\left(\frac{\partial L_m^D}{\partial w_m} - \frac{\partial L_m^S}{\partial w_m}\right)}_{<0} \underbrace{\left(\frac{\partial L_f^S}{\partial c_f} - \frac{\partial L_f^D}{\partial c_f}\right)}_{?} \tag{18}$$

We define $r_{w_m} \equiv \frac{\frac{\partial L_m^S}{\partial w_m}}{\frac{\partial L_m^D}{\partial w_m}} < 0$ as the ratio of the supply and demand elasticities with respect to male wages and $r_{w_m}^D \equiv \frac{\frac{\partial L_f^D}{\partial w_m}}{\frac{\partial L_m^D}{\partial w_m}}$ and $r_{c_f}^D \equiv \frac{\frac{\partial L_m^D}{\partial c_f}}{\frac{\partial L_f^D}{\partial c_f}}$ as the ratio of the demand elasticities with respect to male wages and female costs. Rearranging equation 18, we find that equilibrium wages *decrease* if and only if:

$$\left|\frac{\partial L_f^S}{\partial c_f}\right| < \delta \left|\frac{\partial L_f^D}{\partial c_f}\right|. \tag{19}$$

Here,

$$\delta \equiv 1 - \frac{r_{w_m}^D r_{c_f}^D}{1 - r_{w_m}} \in (0,1].$$

This implies that equilibrium wages decrease if the labor demand shift is proportionally larger than the labor supply shift. The ratios $r_{w_m}^D$ and $r_{c_f}^D$ estimate the labor demand response of the other gender to changes in wages or costs relative to the labor demand response of the same gender. $r_{w_m}^D r_{c_f}^D = 0$ if male and female labor are neither substitutes nor complements, and $r_{w_m}^D r_{c_f}^D > 0$ if male and female labor are either substitutes or complements. Given assumption 4 and substituting from equations 9, 11, 12 and 13, $r_{w_m}^D r_{c_f}^D < 1$.

If male and female labor are neither substitutes nor complements, in which case $\delta = 1$, then the equilibrium wage decreases whenever the demand elasticity in costs is larger than the supply elasticity in costs.

The cutoff value δ is increasing in absolute value in r_{w_m}, the ratio of the supply and demand elasticities with respect to male wages. Intuitively, if an increase in the wage for male labor increases male labor supply by a relatively large amount, employers will be able to replace female labor with male labor without having to pay drastically higher wages for all of their male labor. This makes substituting more attractive, putting downward pressure on female wages.

The cutoff δ is decreasing in the relative curvature of the production function across versus within gender. We consider the two cases of female and male workers being substitutes or complements:

- Complements, $Y^E_{L_f L_m} > 0$: An increase in gender-specific costs for female labor reduces the female labor hired and the marginal productivity of male labor. Intuitively, strong complementarity between men and women means that each woman has a large effect on the productivity of male workers. Since women are necessary for production, this increases the demand for female workers and puts upward pressure on female equilibrium wages.

- Substitutes, $Y^E_{L_f L_m} < 0$: An increase in gender-specific costs for female labor reduces the female labor hired and increases the marginal productivity of male labor.

B.7 Prediction 6

We next evaluate how the equilibrium quantity and wage of the other gender respond to an increase in gender-specific costs. As gender-specific costs increase, employers substitute toward labor of the other gender if male and female workers are substitutes, as labor of the other gender can generate similar revenues at larger other-regarding utility. Note that this implies that the gender employment gap is unambiguously increasing in gender-specific costs if male and female workers are substitutes. We formalize this in the following auxiliary prediction:

Prediction 6 (Substitutability)**.** *Holding selection and productivity constant, the demand for labor and wages are increasing in gender-specific costs to substitute labor and decreasing in gender-specific costs to complement labor.*

Derivation We have shown in equations 12 and 13 that labor demand is increasing in gender-specific costs and wages to substitute labor and decreasing in gender-specific costs and wages to complement labor. We now evaluate how the equilibrium labor quantity and wages are changing in gender-specific costs. To evaluate the sign of $\frac{\partial L^*_m}{\partial c_f} = \frac{|J_2|}{|J|}$, we calculate $|J_2|$ and re-arrange:

$$|J_2| = \underbrace{\frac{\partial L^S_m}{\partial w_m}}_{>0} \left(\underbrace{\frac{\partial L^D_m}{\partial c_f}}_{?} \underbrace{\frac{\partial L^S_f}{\partial w_f}}_{+} - \underbrace{\frac{\partial L^D_m}{\partial w_f}}_{?} \underbrace{\frac{\partial L^S_f}{\partial c_f}}_{-} \right) \tag{20}$$

The above results from the fact that $\frac{\partial L^D_m}{\partial w_f}\frac{\partial L^D_f}{\partial c_f} - \frac{\partial L^D_m}{\partial c_f}\frac{\partial L^D_f}{\partial w_f} = 0$. The effect of an increase in costs to female labor depends on the substitutability of male and female labor. $\frac{\partial L^*_m}{\partial c_f} = 0$

iff male and female workers are neither substitutes nor complements, $\frac{\partial L_m^*}{\partial c_f} > 0$ iff they are substitutes and $\frac{\partial L_m^*}{\partial c_f} < 0$ iff they are complements. This is true for any $\alpha_f \in [0, 1]$. The results are equivalent when considering $\frac{\partial L_f^*}{\partial c_m}$.

Finally, to evaluate the sign of $\frac{\partial w_m^*}{\partial c_f} = \frac{|J_4|}{|J|}$, we calculate $|J_4|$ and re-arrange:

$$|J_4| = \underbrace{\frac{\partial L_m^D}{\partial c_f}}_{?} \underbrace{\frac{\partial L_f^S}{\partial w_f}}_{+} - \underbrace{\frac{\partial L_m^D}{\partial w_f}}_{?} \underbrace{\frac{\partial L_f^S}{\partial c_f}}_{-} \tag{21}$$

The above results from the fact that $\frac{\partial L_m^D}{\partial w_f} \frac{\partial L_f^D}{\partial c_f} - \frac{\partial L_m^D}{\partial c_f} \frac{\partial L_f^D}{\partial w_f} = 0$.

The effect of an increase in costs to female labor depends on the substitutability of male and female labor. $\frac{\partial w_m^*}{\partial c_f} = 0$ iff male and female workers are neither substitutes nor complements, $\frac{\partial w_m^*}{\partial c_f} > 0$ iff they are substitutes as the demand for male labor increases and $\frac{\partial w_m^*}{\partial c_f} < 0$ iff they are complements as the demand for male labor decreases. This is again true for any $\alpha_f \in [0, 1]$. The results are equivalent when considering $\frac{\partial w_f^*}{\partial c_m}$.

C Experiment Appendix

C.1 Information about Sample Industries

We recruit employers from three industries: Manufacturing, Retail & Services, and Education. According to the Bangladesh Bureau of Labour Statistics' 2016—2017 Labour Force Survey, urban workers in Retail & Services are 77% male, Manufacturing workers 61% male, and Education workers 53% male. We calculate the Retail & Services employment rate combining wholesale and retail trade and repair of motor vehicles; accommodation and food service activities, activities of households as employers, and other service activities.

The gender wage gap is largest in Manufacturing, where men earn about BDT 4,200 more than women per month (USD 42; 14,570 for men versus 10,346 for women). Male Services & Retail workers in urban areas earn about BDT 3800 more than women (USD 38; BDT 14,131 for men versus BDT 10,313 for women). In Education, men earn about BDT 3,200 more than women (USD 32; BDT 26,790 for men versus 23,568 for women) BBS (2018).

C.2 Night-Shift Workshop and Job

Figure C.1: Night-Shift Workshop and Job

C.3 Matching of Applicant Pairs in the Hiring Experiment

To mimic a realistic hiring process in which similar applicants apply for the same job, we randomly matched applicants with similar scores to each other using the following procedure. First, we ordered the 14 male and ten female workers by score. Second, we randomly matched two men from the bottom half with each other and two men from the top half. Third, we randomly matched the remaining top five men with the top five women and the remaining bottom five men with the bottom five women.

C.4 Random Wage Distribution in the Application Experiment

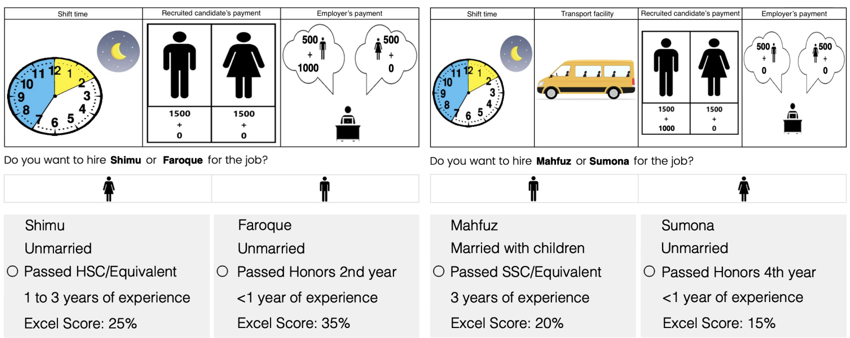
Table C.1: Random Wage Distribution in the Application Experiment

BDT	100	250	500	1,000	2,000	3,000	4,000	5,000
%	40%	40%	15%	1%	1%	1%	1%	1%

Notes: Table shows the wage distribution used to incentivize the reservation wage BDM in the application experiment.

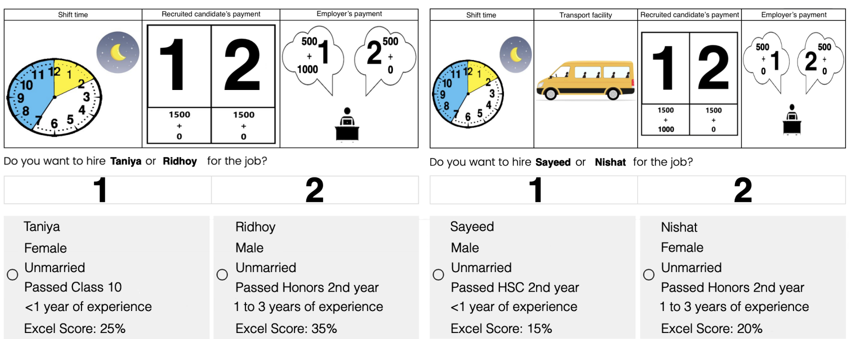
C.5 Experimental Interfaces[58]

Figure C.2: Experimental Interface for Hiring Decisions



a) No Transport + Male Employer Subsidy   b) Transport + Male Worker Subsidy

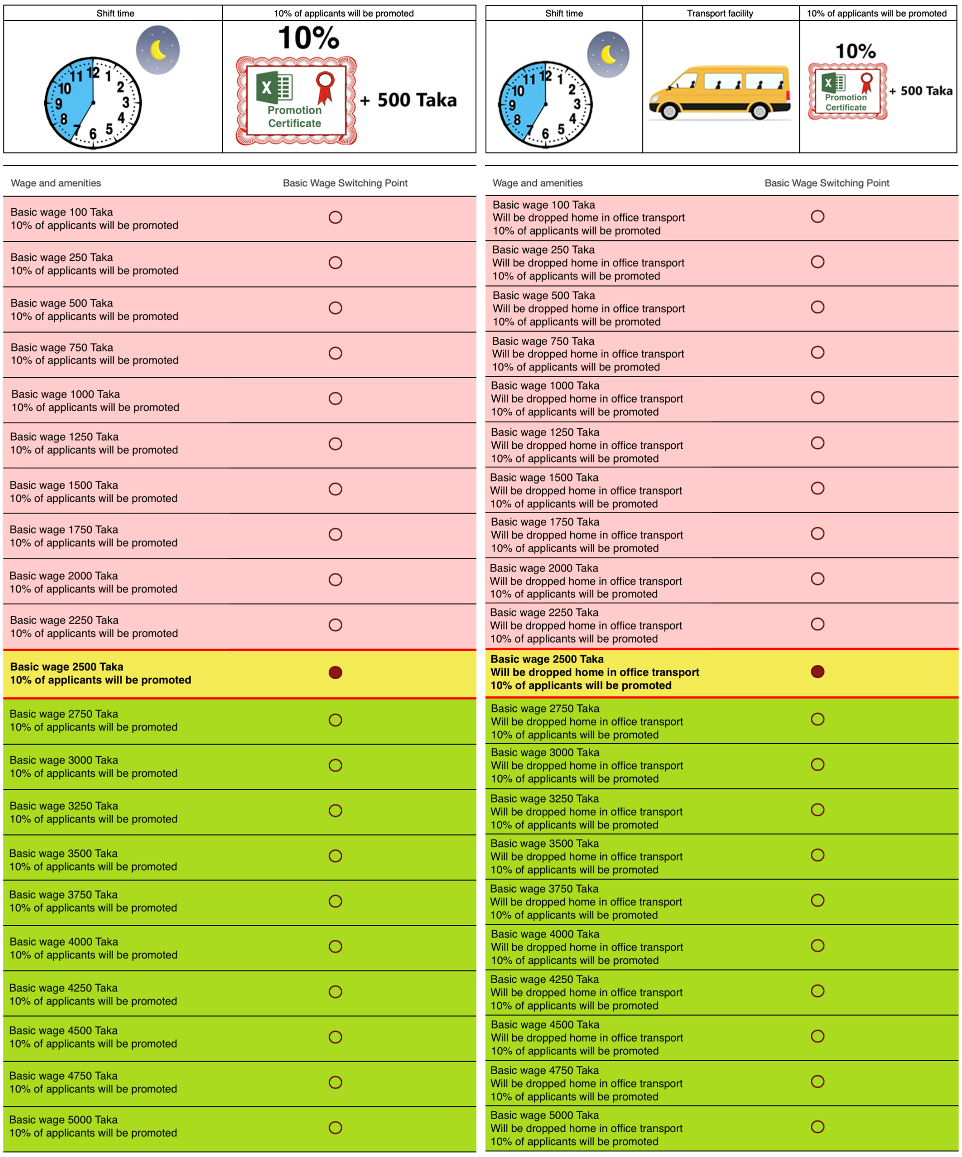
Figure C.3: Experimental Interface for Hiring Decisions; Candidate 1 vs. Candidate 2 Setup



a) No Transport + Male Employer Subsidy   b) Transport + No Subsidy

[58] Translated from Bangla to English.

Figure C.4: Experimental Interface to Elicit Reservation Wage Decisions



a) No Transport b) Transport

C.6 Structural Appendix

C.6.1 Calculating the Cost Conversion Rate

First, we estimate equation 3 replacing the *No Transport* indicator with the predicted costs for the female and male workers compared in each pair. That is, we include the costs reported for the female worker and the compared male worker as well as their interactions with all subsidy treatments. We also replace the applicant fixed effects with the worker characteristics shown to the employer as we only have two male and female cost predictions per employer. The coefficients on the male and female costs give us employers' reaction to male and female costs holding constant the costs of the other applicant. Second, we calculate the conversion factors as the coefficients on the costs for men or women divided by the coefficients on the male and female subsidies from equation 3, multiplied by $-1,000$. This gives us employers' evaluations of a one-Likert scale increase in terms of BDT paid to the worker.

C.6.2 Alternative Preference Estimation Specification: Control Function Approach

We test for robustness using a control function approach to address potential misreporting in welfare and profit predictions. We would like to estimate the following equation:

$$v_{ki} + \varepsilon_{ki} = d_j + \beta_j \Pi_{ki} + \alpha_{jg} \mathcal{W}_{ki} + \varepsilon_{ki}. \tag{22}$$

Assume we do not observe the true profit and welfare beliefs because of measurement error or misreporting. Instead, we observe $\Pi^*_{ki} = \Pi_{ki} + \varepsilon^{\Pi}_{ki}$ and $\mathcal{W}^*_{ki} = \mathcal{W}_{ki} + \varepsilon^{\mathcal{W}}_{ki}$ (for example, employers with high social image concerns might report low profits or welfare whenever they do not hire women to avoid appearing sexist). We can thus rewrite equation 22:

$$v_{ki} + \varepsilon_{ki} = d_j + \beta_j \Pi_{ki} + \alpha_{jg} \mathcal{W}_{ki} + \underbrace{\beta_j \varepsilon^{\Pi}_{ki} + \alpha_{jg} \varepsilon^{\mathcal{W}}_{ki}}_{\varepsilon^{end}_{ki}} + \varepsilon^{ex}_{ki}, \tag{23}$$

where ε^{end}_{ki} is potentially correlated with d_j, Π_{ki} and $\mathcal{W}_{ki}$ and ε^{ex}_{ki} is not correlated with d_j, $\mathcal{W}_{ki}$ nor Π_{ki}.

We adopt a two-step procedure similar to that developed by Rivers and Vuong (1988).[59] First, let

$$\Pi_{ki} = Z'_k \kappa^{\Pi}_j + X'_i \gamma^{\Pi}_j + \tilde{\varepsilon}^{\Pi}_{ki} \tag{24}$$

and

$$\mathcal{W}_{ki} = Z'_k \kappa^{\mathcal{W}}_j + X'_i \gamma^{\mathcal{W}}_j + \tilde{\varepsilon}^{\mathcal{W}}_{ki}, \tag{25}$$

where X_i is a vector of worker characteristics shown to the employer, i.e., the applicant's gender, Excel screening score, education, work experience, and marital status, and Z_k constitutes a vector of transport and subsidy treatment assignments, which are independent of X_i, ε^{Π}_{ki}, $\varepsilon^{\mathcal{W}}_{ki}$, ε^{end}_{ki}, and ε^{ex}_{ki}. $\tilde{\varepsilon}^{\mathcal{W}}_{ki}$, $\tilde{\varepsilon}^{\Pi}_{ki}$, and ε^{end}_{ki} are jointly normal. We estimate equations 24 and 25 using OLS separately by industry (and across industries, including industry fixed effects).

Second, we plug the fitted residuals $\widehat{\tilde{\varepsilon}}^{\Pi}_{ki}$ and $\widehat{\tilde{\varepsilon}}^{\mathcal{W}}_{ki}$ (i.e., the endogenous parts of Π_{ki} and $\mathcal{W}_{ki}$ not explained by the random treatment assignments Z_k or applicant characteristics X_i) into equation 22 and estimate the following probit model:

$$v_{ki} + \varepsilon_{ki} = d_j + \beta_j \Pi_{ki} + \alpha_{jg} \mathcal{W}_{ki} + X_i \gamma + \delta^{\Pi} \widehat{\tilde{\varepsilon}}^{\Pi}_{ki} + \delta^{\mathcal{W}} \widehat{\tilde{\varepsilon}}^{\mathcal{W}}_{ki} + \tilde{\varepsilon}^{ex}_{ki}, \tag{26}$$

where $\tilde{\varepsilon}^{ex}_{ki}$, the error term after controlling for the fitted residuals (we also include industry fixed effects if we estimate equation 26 across industries), is i.i.d. normal with zero mean.

As expected, the employer subsidy increases the expected profit by approximately BDT 1,000 (USD 10) for female applicants but not male applicants in equation 24. By contrast, the *No Transport* treatment reduces the expected welfare by BDT 924 (USD 9) and BDT 1,700 (USD 17) for male and female applicants, respectively. The male and female worker subsidies increase the expected welfare of male and female workers by approximately BDT 1,000 (USD 10) each. Results from equation 26 suggest no mismeasurement in reported welfare or profits in the pooled sample ($p > 0.1$ for both $\widehat{\varepsilon^{\Pi}}$ and $\widehat{\varepsilon^{\mathcal{W}}}$, results not shown).

[59] See also Villas-Boas and Winer (1999), Petrin and Train (2010), Wooldridge (2015) and Hahn and Ridder (2017).

C.6.3 Beliefs Predictions: Random Forest Algorithm

We use a random forest algorithm to predict out-of-sample employer beliefs about profits and welfare based on treatment assignment, employer characteristics, and applicant characteristics. To determine the number of variables to consider in each tree and the number of trees to estimate, we use a grid search and select the combination of parameters that creates the lowest mean out-of-sample error.

The main predictors of productivity are transport, the number of male employees the employer has, the employer's industry, and the worker's Excel screening score. The main predictors of perceived costs are transport, applicant gender, the employer's industry, and how many hiring choices the employer made in the last three years.

C.6.4 Simulating Labor Demand: CES Production Function

The production function of employers in industry j is:

$$Y^E(\hat{Y}_{jf}, \hat{Y}_{jm}) = p\left(a_{jf}(\widehat{Y}_{jf}L_{kf})^\rho + a_{jm}(\widehat{Y}_{jm}L_{km})^\rho\right)^{\frac{\upsilon}{\rho}},$$

where p is the piece rate, $\widehat{Y}_{jf}$ and $\widehat{Y}_{jm}$ are the employer's beliefs about the productivity of female and male workers in industry j, $\rho < 1$ is the substitution parameter, υ is the degree of homogeneity of the production function (where $\upsilon = 1$ is constant returns to scale, $\upsilon < 1$ is decreasing returns to scale, and $\upsilon > 1$ is increasing returns to scale) and a_{jf} and $a_{jm} = 1 - a_{jf}$ are the share parameters. We assume that employers have degenerate, point-beliefs. This assumption implies that the expected value of a function, $E[f(x)]$, equals the function of the expected value, $f(E[x])$, which allows us to derive the results in this section. The employer's utility from profits is β_j (we use this notation to match our structural analysis, in monetary terms $\beta_j = 1$). In addition, the employer receives non-pecuniary benefits d_{kg} from hiring a worker of gender g, and internalizes fraction α_{kg} of the applicant's expected net on-the-job utility $\mathcal{W}_{kg}$.

The first-order conditions are:

$$\underset{L_{kg}}{FOC} \quad d_{kg} + \beta_j p a_{jg}\widehat{Y}_{jg}^\rho L_{kg}^{\rho-1}\upsilon(a_{jf}(\widehat{Y}_{jf}L_{kf})^\rho + a_{jm}(\widehat{Y}_{jm}L_{km})^\rho)^{\frac{\upsilon-\rho}{\rho}} + \alpha_{kg}\mathcal{W}_{kg} - w_g = 0$$

Rearranging, we can solve for the labor g demand:

$$\widehat{L^D}_{kg} = \frac{(\beta_j p\upsilon)^{\frac{1}{1-\upsilon}}\left(\frac{a_{jg}\widehat{Y}_{jg}^\rho}{w_g - d_{kg} - \alpha_{kg}\mathcal{W}_{kg}}\right)^{\frac{1}{1-\rho}}}{\left(\sum_{g'\in\{f,m\}}\left(\frac{a_{jg'}\widehat{Y}_{jg'}^\rho}{(w_{g'} - d_{kg'} - \alpha_{kg'}\mathcal{W}_{kg'})^\rho}\right)^{\frac{1}{1-\rho}}\right)^{\frac{\rho-\upsilon}{\rho(1-\upsilon)}}}.$$

We simulate labor demand using a symmetric CES function, $a_{jf} = a_{jm} = \frac{1}{2}$, a substitution parameter of $\rho = 0.8$ (i.e., male and female workers are substitutes as in the experiment, results are qualitatively the same when using a substitution parameter of $\rho = 0.7$ or $\rho = 0.9$). We calibrate the returns-to-scale parameter υ such that the equilibrium wage for male workers with transport is BDT 1,500 (USD 15) as in the experiment (this results in $\upsilon \approx 0.896$). We calibrate the piece rate p paid to employers for each task to match the average payoffs in the experiment, resulting in BDT 124 (USD 1).

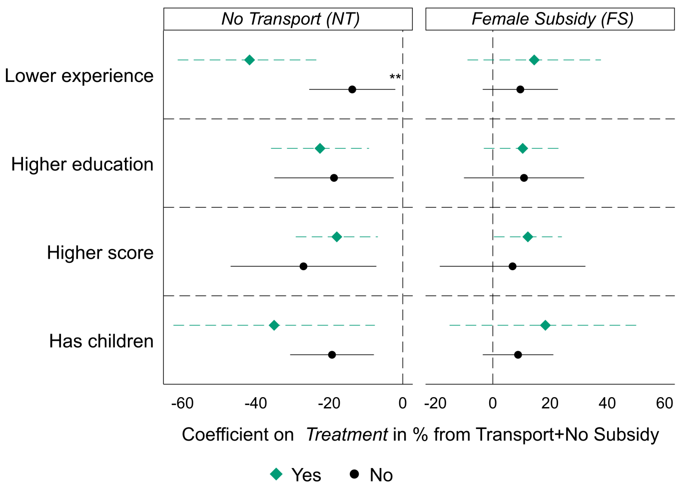
D Empirical Appendix

D.1 Figures

Figure D.1: Hiring by Transport and Female Subsidy Information and Employer Characteristics

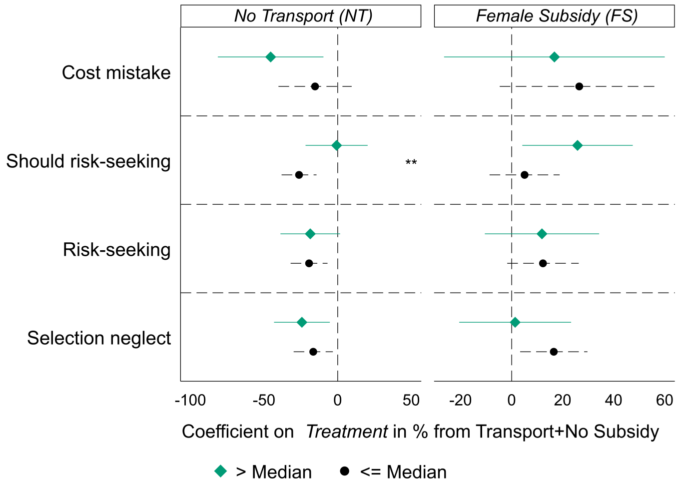
Notes: The graphs show the coefficients on the *No Transport* and the *Female Subsidy* indicators from regression 3 run separately within different employer subsamples. We do not include applicant fixed effects for insufficient observations and instead control for all characteristics of both applicants shown to the employer. We compare female versus male employers, employers who made up to seven hiring choices in the last three years (the median) versus employers who made more than seven hiring choices, employers age 30 or younger (the median) with employers that are older than 30, employers with a high school degree or less (the median) with employers with more than a high school degree, employers with or without daughters and with or without female employees. Asterisks indicate significantly different coefficients between subsamples. $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Figure D.2: Hiring by Transport and Female Subsidy Information and Applicant Characteristics



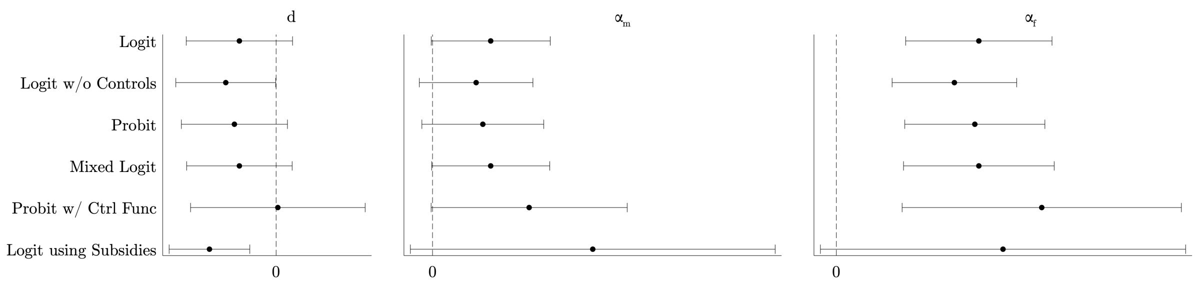
Notes: The graphs show the coefficients on the *No Transport* and the *Female Subsidy* indicators from regression 3 run separately within different pairs. We do not include applicant fixed effects for insufficient observations and instead control for all characteristics of both applicants shown to the employer. We compare pairs in which the female applicant has less work experience, higher education, or a higher Excel score than the male applicant versus pairs in which the woman has the same or more work experience, the same or less education, or the same or a lower Excel score as well as pairs in which the female applicant has children versus pairs in which she does not. Asterisks indicate significantly different coefficients between subsamples. $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Figure D.3: Hiring by Transport and Female Subsidy Information and Employer Characteristics



Notes: The graphs show the coefficient estimates and 95% confidence intervals for *No Transport* and the *Female Subsidy*. We run the regressions in different subsets of employers (see section 4.2). We do not include applicant fixed effects for insufficient observations and instead control for all characteristics of both applicants shown to the employer. Asterisks indicate significantly different coefficients between subsamples. $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

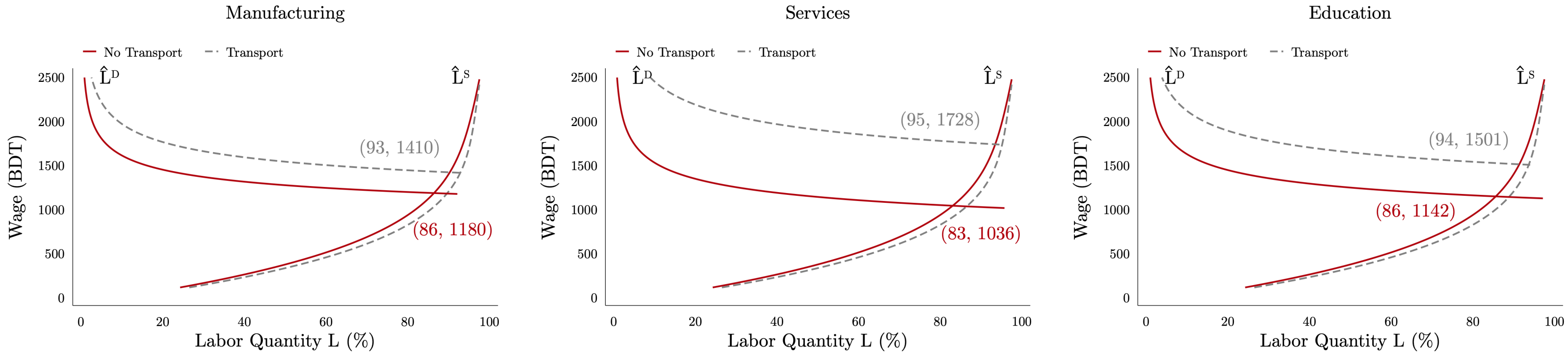
Figure D.4: Parameter Robustness



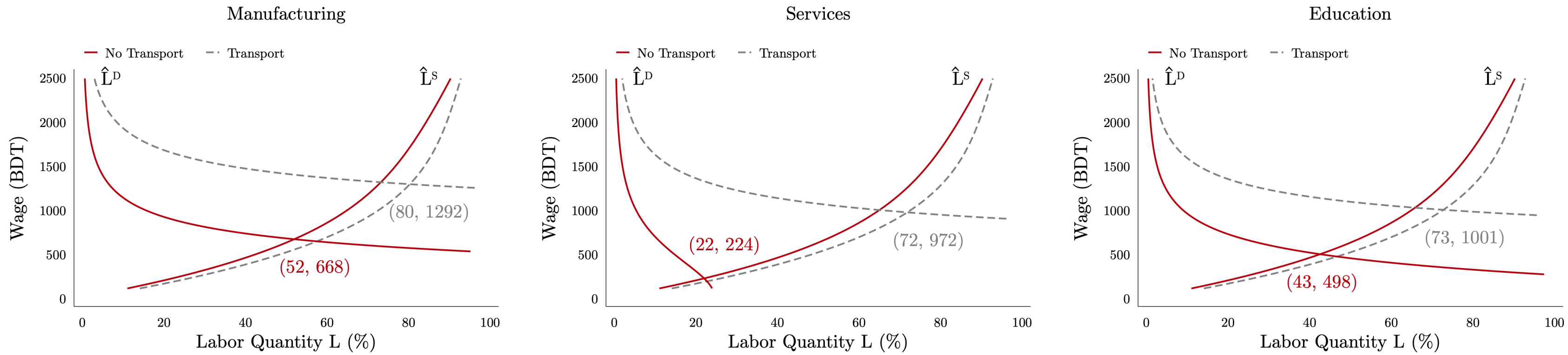
Notes: The graph shows the estimated coefficients and 95% confidence intervals of the preference parameters for a series of specifications. "Logit" estimates equation 5 controlling for a vector of applicant characteristics (Excel screening score, education, work experience, and marriage status). "Logit without Controls" estimates the same specification without controls. "Probit" estimates the same specification using probit instead of logit regression. "Mixed Logit" estimates heterogeneous preferences using mixed logit that takes advantage that we have four observations per employer. "Probit with Control Function" employs a control function approach to address potential endogeneity concerns of the reported productivity and cost beliefs (see also section C.6.2). "Logit using Subsidies" estimates α_m and α_f on only the exogenously varied wages paid by employers or received by workers. The confidence intervals are based on bootstrapped standard errors. Standard errors are based on 1,500 bootstrap samples clustered at the employer level. We retain only those bootstrap samples where the estimation routine converged within 50 iterations, resulting in 1,500, 1,500, 1,499, 1,082, 1,500, 1,500 bootstrap samples for the 6 datasets, respectively.

Figure D.5: Equilibria in the Male and Female Labor Markets

(D.5.1) Market for Male Workers



(D.5.2) Market for Female Workers



Notes: The graph shows the share of workers demanded and supplied in male and female labor markets at each wage with and without transport. We use predicted productivity and cost beliefs from the *Beliefs-Elicitation* employers (see section 6.1.3) and calculate profits using the CES production function described in appendix C.6.4. Numbers in parentheses in the graph give (L_g^*, w_g^*). Numbers in gray on top are the equilibrium with transport and numbers in red in the bottom are the equilibrium without transport.

D.2 Tables

Table D.1: Employer Characteristics in the Hiring Experiment, By Transport Information and Subsidy Assignment

	No Transport (NT)		Male Subsidy (MS)		Female Subsidy (FS)		Employer Subsidy (ES)		NT+MS		NT+FS		NT+ES		Transport
N	101		36		47		60		39		47		36		94
	Mean	β_{NT}	Mean	β_{MS}	Mean	β_{FS}	Mean	β_{ES}	Mean	β_{NT+MS}	Mean	β_{NT+FS}	Mean	β_{NT+ES}	Mean
	(SD)	(p-val)	(SD)	(p-val)	(SD)	(p-val)	(SD)	(p-val)	(SD)	(p-val)	(SD)	(p-val)	(SD)	(p-val)	(SD)
Manufacturing (%)	34.65	3.80	36.11	5.26	36.17	5.32	33.33	2.48	41.03	1.11	27.66	-12.31	27.78	-9.36	30.85
	(47.82)	(0.57)	(48.71)	(0.58)	(48.57)	(0.53)	(47.54)	(0.75)	(49.83)	(0.93)	(45.22)	(0.30)	(45.43)	(0.43)	(46.44)
Retail & Services (%)	37.62	0.39	33.33	-3.90	25.53	-11.70	30.00	-7.23	35.90	2.17	23.40	-2.52	25.00	-5.39	37.23
	(48.69)	(0.96)	(47.81)	(0.68)	(44.08)	(0.15)	(46.21)	(0.35)	(48.60)	(0.87)	(42.80)	(0.82)	(43.92)	(0.65)	(48.60)
Education (%)	27.72	-4.19	30.56	-1.36	38.30	6.38	36.67	4.75	23.08	-3.29	48.94	14.83	47.22	14.75	31.91
	(44.99)	(0.53)	(46.72)	(0.88)	(49.14)	(0.46)	(48.60)	(0.55)	(42.68)	(0.79)	(50.53)	(0.22)	(50.63)	(0.23)	(46.86)
Age	31.51	0.36	31.25	0.09	30.94	-0.22	32.27	1.11	32.62	1.01	29.23	-2.06	33.31	0.68	31.16
	(7.28)	(0.74)	(7.00)	(0.95)	(7.81)	(0.87)	(8.98)	(0.43)	(7.54)	(0.61)	(7.22)	(0.28)	(9.14)	(0.76)	(7.80)
Bachelor's (%)	41.58	-3.58	44.44	-0.72	42.55	-2.61	38.33	-6.83	32.43	-8.43	40.43	1.45	52.78	18.02	45.16
	(49.53)	(0.62)	(50.40)	(0.94)	(49.98)	(0.77)	(49.03)	(0.41)	(47.46)	(0.53)	(49.61)	(0.91)	(50.63)	(0.16)	(50.04)
Married (%)	66.34	15.27	55.56	4.49	57.45	6.38	58.33	7.27	71.79	0.97	48.94	-23.78	58.33	-15.27	51.06
	(47.49)	(0.03)	(50.40)	(0.65)	(49.98)	(0.48)	(49.72)	(0.38)	(45.59)	(0.94)	(50.53)	(0.06)	(50.00)	(0.23)	(50.26)
Children (%)	49.50	11.21	41.67	3.37	44.68	6.38	50.00	11.70	61.54	8.66	34.04	-21.85	44.44	-16.76	38.30
	(50.25)	(0.12)	(50.00)	(0.73)	(50.25)	(0.47)	(50.42)	(0.16)	(49.29)	(0.52)	(47.90)	(0.08)	(50.40)	(0.19)	(48.87)
# Daughters	0.36	0.05	0.36	0.05	0.34	0.03	0.48	0.17	0.44	0.03	0.34	-0.05	0.39	-0.14	0.31
	(0.61)	(0.59)	(0.64)	(0.67)	(0.64)	(0.78)	(0.77)	(0.14)	(0.72)	(0.88)	(0.67)	(0.77)	(0.69)	(0.42)	(0.64)
# Female Employees	13.99	11.34	2.69	0.05	2.51	-0.14	4.68	2.03	6.00	-8.04	6.40	-7.45	2.58	-13.44	2.65
	(86.55)	(0.19)	(4.90)	(0.97)	(6.27)	(0.91)	(11.26)	(0.22)	(15.20)	(0.38)	(12.36)	(0.40)	(5.65)	(0.13)	(7.59)
# Hiring Decisions Last 6 Months	66.75	55.78	11.19	0.23	12.04	1.07	14.10	3.13	29.41	-37.57	17.57	-50.25	25.83	-44.05	10.97
	(496.98)	(0.26)	(16.59)	(0.94)	(19.00)	(0.72)	(19.55)	(0.26)	(85.42)	(0.47)	(27.72)	(0.31)	(66.02)	(0.39)	(11.20)
All Understanding Questions Correct (%)	96.19	0.27	94.74	-1.18	95.92	0.00	98.36	2.44	86.67	-8.34	87.04	-9.15	92.31	-6.33	95.92
	(19.23)	(0.92)	(22.63)	(0.78)	(19.99)	(1.00)	(12.80)	(0.35)	(34.38)	(0.22)	(33.90)	(0.13)	(27.00)	(0.24)	(19.89)
Made Hiring Choices b/c of Taste (%)	4.95	2.82	5.56	3.43	0.00	-2.13	0.00	-2.13	7.69	-0.69	2.13	-0.70	0.00	-2.82	2.13
	(21.80)	(0.29)	(23.23)	(0.41)	(0.00)	(0.16)	(0.00)	(0.16)	(27.00)	(0.91)	(14.59)	(0.84)	(0.00)	(0.29)	(14.51)
Made Hiring Choices b/c of Productivity (%)	100.00	2.13	97.22	-0.65	100.00	2.13	100.00	2.13	100.00	0.65	100.00	-2.13	100.00	-2.13	97.87
	(0.00)	(0.16)	(16.67)	(0.84)	(0.00)	(0.16)	(0.00)	(0.16)	(0.00)	(0.84)	(0.00)	(0.16)	(0.00)	(0.16)	(14.51)
***p*-value from joint significance test**		0.47		0.99		0.72		0.29		0.96		0.47		0.19	
Made Hiring Choices b/c of Applicant Welfare (%)	82.37	25.72	80.88	24.23	65.66	9.01	73.38	16.73	67.23	-39.38	74.28	-17.11	58.29	-40.82	56.65
	(38.11)	(0.00)	(39.35)	(0.00)	(47.51)	(0.30)	(44.22)	(0.03)	(46.97)	(0.00)	(43.73)	(0.13)	(49.34)	(0.00)	(49.57)

Notes: The table shows characteristics by treatment arm of all employers in the analysis sample of the hiring experiment (except for "All Understanding Questions Correct (%)", for which we include all employers in each treatment). **No Transport (NT)** includes all employers in the *No Transport+No Subsidy* treatment; **Male Subsidy (MS)** includes all employers in the *Transport+Male Subsidy* treatment; **Female Subsidy (FS)** includes all employers in the *Transport+Female Subsidy* treatment; **Employer Subsidy (ES)** includes all employers in the *Transport+Employer Subsidy* treatment. "Made Hiring Choices b/c of Taste" is an indicator that is 1 for employers who reported that women belong at home. "Made Hiring Choices b/c of Productivity" is an indicator that is 1 for employers who report that they based their hiring choices based on absenteeism, performance, firm reputation, experience, education, or because women are hard to manage. "Made Hiring Choices b/c of Applicant Welfare" is an indicator that is 1 for employers who report that they based their hiring choices based on the applicants' safety, health, or marital status, or because they stated it would be inappropriate for women to work at night or that men would need money more than women. We show means and standard deviations within treatment arms as well as coefficients and p-values on the treatment indicators in OLS regressions with modified Huber-White robust SEs. ***p*-values from joint significance test** indicates the joint significance of all variables in predicting treatment assignment (excluding "Education (%)", which is perfectly collinear with "Manufacturing (%)" and "Retail & Services (%)").

Table D.2: Employer Beliefs about Applicants in the Hiring Experiment, by Transport Information and Subsidy Assignment

	No Transport (NT)		**Male Subsidy (MS)**		**Female Subsidy (FS)**		**Employer Subsidy (ES)**		**NT+MS**		**NT+FS**		**NT+ES**		**Transport**
	Mean (SD)	β_{NT} (p-val)	Mean (SD)	β_{MS} (p-val)	Mean (SD)	β_{FS} (p-val)	Mean (SD)	β_{ES} (p-val)	Mean (SD)	β_{NT+MS} (p-val)	Mean (SD)	β_{NT+FS} (p-val)	Mean (SD)	β_{NT+ES} (p-val)	Mean (SD)
Male Applicants: N	990		349		463		586		384		466		355		928
Productivity (P(Show-up) $\times$ E[Tasks|Show-up])	39.65	0.98	39.75	1.09	34.78	-3.88	38.05	-0.62	47.46	6.72	41.41	5.64	35.30	-3.73	38.67
	(21.17)	(0.70)	(21.04)	(0.76)	(18.56)	(0.19)	(21.32)	(0.84)	(22.25)	(0.18)	(21.95)	(0.20)	(19.69)	(0.40)	(20.17)
Predicted Revenue (BDT)	697.76	2.96	701.14	6.35	674.18	-20.61	690.32	-4.47	736.73	32.63	707.04	29.89	675.97	-17.32	694.80
	(106.25)	(0.82)	(106.36)	(0.73)	(93.26)	(0.17)	(107.07)	(0.77)	(111.87)	(0.20)	(109.75)	(0.18)	(97.49)	(0.44)	(100.85)
Actual Revenue (BDT)	585.27	0.57	585.45	0.75	586.02	1.32	583.98	-0.72	587.80	1.78	586.07	-0.52	583.99	-0.56	584.70
	(28.53)	(0.79)	(29.46)	(0.79)	(27.81)	(0.61)	(29.50)	(0.77)	(29.35)	(0.65)	(27.88)	(0.88)	(29.20)	(0.88)	(29.18)
Perceived Costs (0–10)	2.50	1.63	1.08	0.22	0.64	-0.22	0.72	-0.14	1.89	-0.82	2.16	-0.11	1.83	-0.52	0.87
	(2.18)	(0.00)	(1.10)	(0.29)	(1.25)	(0.24)	(1.21)	(0.43)	(1.82)	(0.04)	(1.87)	(0.78)	(2.04)	(0.22)	(1.20)
Perceived Costs (BDT)	1399.77	924.39	598.07	122.69	361.56	-113.82	406.43	-68.95	1048.39	-474.07	1206.23	-79.72	1050.66	-280.16	475.38
	(1215.21)	(0.00)	(614.75)	(0.29)	(702.31)	(0.29)	(678.73)	(0.49)	(1016.59)	(0.04)	(1040.09)	(0.71)	(1141.80)	(0.24)	(670.38)
Female Applicants: N	990		352		466		586		385		467		357		929
Productivity (P(Show-up) $\times$ E[Tasks|Show-up])	31.72	0.62	28.87	-2.23	25.37	-5.73	30.83	-0.27	36.60	7.12	32.23	6.25	27.35	-4.10	31.10
	(19.50)	(0.80)	(17.96)	(0.45)	(17.63)	(0.04)	(20.83)	(0.93)	(22.98)	(0.13)	(21.55)	(0.14)	(19.81)	(0.35)	(21.20)
Predicted Revenue (BDT)	657.81	2.01	644.95	-10.84	626.83	-28.97	654.51	-1.28	683.02	36.06	661.17	32.33	635.79	-20.74	655.80
	(97.07)	(0.87)	(90.27)	(0.47)	(88.17)	(0.04)	(104.53)	(0.93)	(114.88)	(0.13)	(107.75)	(0.12)	(99.39)	(0.35)	(106.19)
Actual Revenue (BDT)	588.47	5.48	583.50	0.51	587.25	4.25	593.32	10.33	588.73	-0.25	586.40	-6.32	591.10	-7.69	582.99
	(28.26)	(0.10)	(30.40)	(0.91)	(30.82)	(0.32)	(26.43)	(0.01)	(28.25)	(0.97)	(29.26)	(0.31)	(27.64)	(0.15)	(29.46)
Perceived Costs (0–10)	6.30	3.05	4.43	1.18	3.00	-0.25	3.07	-0.19	6.01	-1.46	5.84	-0.21	5.17	-0.95	3.25
	(2.50)	(0.00)	(2.02)	(0.00)	(2.20)	(0.48)	(2.18)	(0.59)	(2.17)	(0.01)	(2.31)	(0.70)	(2.92)	(0.13)	(2.20)
Perceived Costs (BDT)	3507.20	1702.56	2476.51	671.87	1669.29	-135.35	1702.30	-102.34	3345.81	-833.26	3248.83	-123.01	2860.52	-544.34	1804.64
	(1400.40)	(0.00)	(1127.82)	(0.00)	(1222.39)	(0.50)	(1217.97)	(0.60)	(1205.95)	(0.01)	(1284.91)	(0.68)	(1629.86)	(0.12)	(1225.17)

Notes: The table shows employers' beliefs about applicant characteristics by treatment arm of all applicants in the hiring experiment. N indicates observations of employer–applicant pairs. **No Transport (NT)** includes all employers in the *No Transport+No Subsidy* treatment; **Male Subsidy (MS)** includes all employers in the *Transport+Male Subsidy* treatment; **Female Subsidy (FS)** includes all employers in the *Transport+Female Subsidy* treatment; **Employer Subsidy (ES)** includes all employers in the *Transport+Employer Subsidy* treatment. $\beta_{[treatment]}$ indicates the coefficient estimates from a regression of the variable on an indicator for treatment status with modified Huber–White robust SEs. *Productivity* indicates employers' beliefs about the number of tasks (out of 100) that an applicant will complete conditional on showing up for the shift. *Predicted Revenue* indicates employers expected payoff from hiring a given applicant. *Actual Revenue* indicates the realized payoffs to employers from hired workers. *Perceived Costs (0–10)* indicates employers' beliefs about applicants' on-the-job costs on a 0–10 Likert scale. *Perceived Costs (BDT)* indicates employers' beliefs about on-the-job costs converted to money.

Table D.3: Productivity and Costs Predictions from *Hiring* and *Prediction-Only* Employers

	Employer Type		
	Hiring		*Prediction-Only*
	Mean (SD)	β_{Pred} (p-val)	Mean (SD)
Male Applicants: *N*	1,408		319
Productivity (P(Show-up) $\times$ E[Tasks|Show-up])	41.54	-3.85	37.69
	(21.93)	(0.11)	(21.21)
Perceived Costs (0–10)	2.28	0.15	2.43
	(2.06)	(0.59)	(2.22)
Female Applicants: *N*	1,406		320
Productivity (P(Show-up) $\times$ E[Tasks|Show-up])	32.27	-2.92	29.35
	(21.92)	(0.22)	(20.00)
Perceived Costs (0–10)	6.06	-0.41	5.64
	(2.44)	(0.18)	(2.27)

Notes: Table shows predictions of the *Hiring* and *Prediction-Only* employers recruited simultaneously. We exclude employers *Hiring* employers recruited before *Prediction-Only* recruitment began. N indicates the number of employer–applicant pairs. β_{Pred} shows the coefficient and p-values from an OLS regression of each variable on an indicator that is 1 for *Prediction-Only* employers (modified Huber–White standard errors not shown).

Table D.4: Hired by Transport Information and Subsidy Assignment, Robustness Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
No transport (NT)	-9.729***	-9.653***	-9.729***	-9.189***	-9.377***	-9.943***	-11.528***	-9.774***	-9.729***	-0.617***	-34.902***
	(2.486)	(2.444)	(2.487)	(2.465)	(2.443)	(2.469)	(3.548)	(2.590)	(2.630)	(0.150)	(6.021)
Male subsidy (MS)	-7.263**	-6.899**	-7.263**	-6.230*	-6.848**	-6.869**	-9.104**	-7.104**	-7.263**	-0.459**	
	(3.256)	(3.264)	(3.259)	(3.261)	(3.059)	(3.224)	(4.146)	(3.251)	(3.487)	(0.191)	
Female subsidy (FS)	7.203***	4.843*	7.203***	4.859*	7.685***	7.714***	8.285**	6.787**	7.203**	0.433***	
	(2.765)	(2.667)	(2.767)	(2.702)	(2.739)	(2.806)	(3.678)	(2.857)	(2.949)	(0.159)	
Employer subsidy (ES)	22.931***	22.729***	22.931***	22.855***	22.807***	23.100***	22.974***	22.330***	22.931***	1.385***	
	(3.122)	(3.063)	(3.124)	(3.089)	(3.084)	(3.138)	(4.347)	(3.223)	(3.297)	(0.191)	
NT*MS	0.851	-0.963	0.851	-1.126	0.968	0.876	3.093	1.414	0.851	0.054	
	(4.401)	(4.300)	(4.404)	(4.351)	(4.301)	(4.384)	(5.997)	(4.564)	(4.702)	(0.269)	
NT*FS	-4.969	-3.850	-4.969	-4.486	-6.894*	-4.996	-7.385	-4.923	-4.969	-0.271	
	(4.079)	(4.007)	(4.082)	(4.008)	(3.908)	(4.099)	(5.779)	(4.147)	(4.276)	(0.238)	
NT*ES	0.666	-0.335	0.666	-0.814	0.458	1.514	0.835	1.228	0.666	0.036	
	(5.339)	(5.268)	(5.343)	(5.307)	(5.119)	(5.372)	(7.800)	(5.441)	(5.553)	(0.308)	
Applicant: Excel score		1.338***		1.313***							
		(0.060)		(0.059)							
Applicant: Education		2.250***		1.832***							
		(0.310)		(0.303)							
Applicant: $\leq$ 3 yrs work experience		-10.727***									
		(2.334)									
Applicant: Married		-5.586***									
		(2.069)									
Applicant: Has children		-1.714									
		(2.805)									
Control Mean	45.318	45.269	45.269	45.318	45.971	45.158	48.092	45.998	45.318	45.318	55.000
Observations	4532	4539	4539	4532	4815	4482	2570	4080	4532	4173	241
Main	✓										
No fixed effects		✓									
No controls			✓								
Post-Double-Selection				✓							
Understanding					✓						
Correct commute						✓					
Before first shift							✓				
No prediction applicants								✓			
Two-way clustering									✓		
Logit										✓	
Candidate 1 versus 2											✓

Notes: The table shows results from OLS regressions with Huber–White robust SEs clustered at the employer level (see notes to figure 3). Column (2) excludes applicant fixed effects, column (3) excludes all covariates, and column (4) uses covariates selected using the post-double selection (PSD) Lasso method of Belloni et al. (2014). Column (5) includes employers who answer understanding questions incorrectly, column (6) includes only employers who report that women in the *Transport* treatment will get home using provided transport and that women in the *No Transport* treatment will not get home using provided transport, and column (7) includes only employers surveyed before the first night shift. Column (8) excludes the applicants from the application experiment. column (9) clusters standard errors both at the employer and the applicant level, column (10) uses a Logit specification, and column (11) includes hiring decisions over candidate 1 versus 2 (not disaggregated by subsidies and using the covariates from column (2) due to small sample size). $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table D.5: Applicant Characteristics in the Application Experiment, by Transport Information

	Male Applicants			**Female Applicants**		
	No Transport		*Transport*	*No Transport*		*Transport*
N	171		183	175		169
	Mean (SD)	β_{NT} (p-val)	Mean (SD)	Mean (SD)	β_{NT} (p-val)	Mean (SD)
Age	25.34	-1.07	26.49	23.01	0.76	23.05
	(7.26)	(0.36)	(8.56)	(6.37)	(0.49)	(6.64)
Education (Years)	14.43	-0.05	14.64	13.70	-0.29	13.86
	(2.36)	(0.89)	(2.29)	(2.21)	(0.43)	(2.30)
$\leq$ 3 Years Work Experience (%)	73.68	3.68	71.58	86.29	-9.62	90.53
	(44.16)	(0.59)	(45.22)	(34.50)	(0.08)	(29.36)
Excel Screening Score (%)	24.65	0.44	25.08	26.31	0.06	26.42
	(11.65)	(0.80)	(11.38)	(12.01)	(0.97)	(12.27)
Married (%)	23.98	-2.12	27.32	21.14	-6.38	27.22
	(42.82)	(0.74)	(44.68)	(40.95)	(0.37)	(44.64)
Children (%)	18.71	-0.58	18.03	10.29	-8.92	17.16
	(39.12)	(0.92)	(38.55)	(30.46)	(0.13)	(37.82)
All Understanding Questions Correct (%)	89.53	-4.15	91.50	93.58	6.43	88.02
	(30.70)	(0.30)	(27.96)	(24.57)	(0.10)	(32.56)
***p*-value from joint significance test**		0.80			0.14	
Reported Costs (0–10)	2.30	0.35	1.81	5.89	0.82	4.88
	(2.44)	(0.23)	(2.33)	(2.97)	(0.03)	(3.05)

Notes: The table shows characteristics by treatment arm of all female and male workers in the application experiment. We show means and standard deviations within treatment arms as well as coefficients and p-values on the treatment indicators from regression 4 without applicant controls. *p-value from joint significance test* indicates the results of a joint significance test of age, education, work experience, Excel screening score, marriage status, child status, and correct responses to understanding questions.

Table D.6: Reservation Wages in the Application Experiment by Transport Information, Robustness Analysis

	Male Workers								Female Workers							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
No transport (NT)	167.516*	-0.044	140.286*	217.414	179.465**	203.575**	105.113*	176.543**	239.841*	-0.130*	246.949*	215.525	238.095*	221.388	-16.016	162.023
	(92.523)	(0.048)	(76.212)	(146.505)	(88.303)	(89.773)	(59.888)	(85.941)	(141.490)	(0.069)	(128.997)	(173.132)	(137.016)	(136.263)	(64.003)	(135.064)
Control Mean	478.297	0.934	408.333	552.473	478.297	478.297	370.882	474.623	802.663	0.858	735.671	848.521	802.663	802.663	521.724	810.677
Observations	352	352	335	352	352	352	326	389	344	344	333	344	344	344	279	379
Main	✓								✓							
Applied		✓								✓						
Truncating			✓								✓					
Keep outliers				✓								✓				
No controls					✓								✓			
Post-Double-Selection						✓								✓		
Reservation wage $\leq 1,500$							✓								✓	
Understanding								✓								✓

Notes: The table shows results from OLS regressions with Huber–White robust SEs (see equation 4 and notes to figure 5). We always control for assignment to the *High promotion* treatment and its interaction with *No transport*. We winsorize the data at the 95th percentile and control for the worker's education, marriage status (unmarried, married without children, or married with children), work experience, Excel screening score, and age in the main specification in columns (1) and (9). We use a reservation wage of $\leq$ BDT 1,500 (the wage in the hiring experiment) as an outcome in columns (2) and (10). We truncate the data at the 95th percentile in columns (3) and (11) and do not exclude outliers in columns (4) and (12). We exclude all covariates in columns (5) and (13) and include covariates selected using the post-double selection (PSD) Lasso method of Belloni et al. (2014) in columns (6) and (14). We only keep applicants with a reservation wage of $\leq$ BDT 1,500 in columns (7) and (15) and include applicants with incorrect understanding questions in columns (8) and (16). $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table D.7: Employer Preference Parameter Estimates, Mixed Logit

	Pooled		**Manufacturing**		**Services**		**Education**	
	μ	σ	μ	σ	μ	σ	μ	σ
d	-0.115	0.000	-0.024	0.000	-0.073	0.000	-0.201	0.000
	(0.084)	(0.047)	(0.137)	(0.075)	(0.432)	(0.044)	(0.114)	(0.050)
α_m	0.111*	0.000	0.011	0.000	0.261	0.000	0.121	0.000
	(0.057)	(0.019)	(0.093)	(0.099)	(3.018)	(0.091)	(0.084)	(0.013)
α_f	0.174***	0.005	0.175***	0.000	0.257	0.117	0.155**	0.182
	(0.047)	(0.078)	(0.067)	(0.028)	(4.248)	(2.198)	(0.076)	(0.117)
p-val ($\alpha_m = \alpha_f$)	0.396	.	0.153	.	0.999	.	0.763	.
Observations	1,816	.	606	.	588	.	622	.

Notes: The table presents parameter estimates from a mixed logit model, estimated among mixed-gender hiring pairs, assuming normally distributed random parameters. We control for the applicant characteristics shown to employers, including Excel screening score, education, work experience, and marital status. All estimates in money metric. d in '000 BDT. Standard errors are based on 1,500 bootstrap samples clustered at the employer level. Standard errors are calculated using 1,500 bootstrap samples clustered at the employer level. We retain only those bootstrap samples where the estimation routine converged within 50 iterations, resulting in 1,082, 1,223, 1,177, and 1,290 samples for the pooled data and the three industry-specific datasets, respectively. $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table D.8: Employer Preferences: Heterogeneity by Other-Regarding Index

	$\leq$ Median	$>$ Median
d	-0.226	-0.058
	(0.125)	(0.141)
α_m	0.022	0.180
	(0.083)	(0.113)
α_f	0.064	0.268***
	(0.061)	(0.104)
p-val ($\alpha_m = \alpha_f$)	0.682	0.564
Observations	910	906

Notes: The table presents parameter estimates from a logit model, estimated among mixed-gender hiring pairs. We control for the applicant characteristics shown to employers, including Excel screening score, education, work experience, and marital status. We divided the sample by the median other-regarding index used to test prediction 4. All estimates in money metric. d in '000 BDT. Standard errors are based on 1,500 bootstrap samples clustered at the employer level. We retain only those bootstrap samples where the estimation routine converged within 50 iterations, resulting in 1,496 and 1,499 bootstrap samples for the below and above median datasets, respectively. $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table D.9: Counterfactuals: Benchmarking the Importance of Paternalistic Discrimination

	Status Quo	$\alpha_m \mathcal{W}_m^E = \alpha_f \mathcal{W}_f^E$	$\mathcal{W}_m^E = \mathcal{W}_f^E$	$\alpha_m = \alpha_f$	$d = 0$	$\Pi_m^E = \Pi_f^E$	$L_m^S = L_f^S$	$\mathcal{W}_g^E = \mathcal{W}_g^{E:A}$	$\mathcal{W}_g^E = \mathcal{W}_g^A$
L_m^* (%)	86	86	86	86	86	86	86	87	89
L_f^* (%)	44	64	63	54	51	53	55	50	69
$L_m^* - L_f^*$ (ppts)	42	22	23	31	35	33	30	37	20
w_m^* (BDT)	1146	1166	1165	1157	1154	1163	1157	1200	1304
w_f^* (BDT)	516	967	931	720	646	692	442	621	1137
$w_m^* - w_f^*$ (BDT)	631	200	234	437	508	471	715	579	167
$\mathcal{W}_m^E$ ('000 BDT)	-122	-115	-115	-118	-119	-116	-118	50	371
$\mathcal{W}_m^A$ ('000 BDT)	308	316	315	312	311	315	312	329	371
$\mathcal{W}_f^E$ ('000 BDT)	-614	-739	-736	-699	-675	-691	-808	-547	216
$\mathcal{W}_f^A$ ('000 BDT)	46	161	150	91	73	84	47	67	216
Π ('000 BDT)	2157	2305	2298	2247	2220	2237	2415	2209	2319

Notes: The table shows the results from the industry counterfactuals. We use both employers' and applicants' beliefs about the job costs and productivity in our experiment. We conduct the following counterfactual exercises: 1) eliminating paternalistic discrimination, either by equalizing male and female other-regarding utility, $\alpha_m \mathcal{W}_m^E = \alpha_f \mathcal{W}_f^E$, by equalizing male and female perceived welfare, $\mathcal{W}_m^E = \mathcal{W}_f^E$, or equalizing the welfare weights, $\alpha_m = \alpha_f$, 2) eliminating taste-based discrimination by setting $d = 0$, 3) eliminating statistical discrimination by equalizing male and female perceived profits, $\Pi_m^E = \Pi_f^E$, or 4) eliminating differences in labor supply by equalizing male and female labor supply, $L_m^S = L_f^S$. We present effects on the following outcomes: 1) male and female employment as well as the the gender employment gap, $L_m^*, L_f^*, L_m^* - L_f^*$ (in percentage points), 2) male and female wages as well as the gender wage gap, $w_m^*, w_f^*, w_m^* - w_f^*$ (in BDT), 3) total male welfare as perceived by employers, $\mathcal{W}_m^E$ (in '000 BDT), and applicants, $\mathcal{W}_m^A$ (in '000 BDT), 4) total female welfare as perceived by employers, $\mathcal{W}_f^E$ (in '000 BDT) and applicants, $\mathcal{W}_f^A$ (in '000 BDT).

Table D.10: Effects of Counterfactual Transport and Subsidy Interventions

	Status Quo	**Transport**	**Subsidies**
L_m^* (%)	86	87	86
L_f^* (%)	44	75	76
$L_m^* - L_f^*$ (ppts)	42	12	11
w_m^* (BDT)	1146	1186	1176
w_f^* (BDT)	516	1067	1434
$w_m^* - w_f^*$ (BDT)	631	120	-258
$\mathcal{W}_m^E$ ('000 BDT)	-122	-107	-111
$\mathcal{W}_m^A$ ('000 BDT)	308	324	320
$\mathcal{W}_f^E$ ('000 BDT)	-614	-246	-708
$\mathcal{W}_f^A$ ('000 BDT)	46	219	317
Π ('000 BDT)	1975	2590	2205
Gov't Cost ('000 BDT)	0	269	269

Notes: The table shows the results from evaluating the effectiveness of transport and subsidy interventions. We use both employers' and applicants' beliefs about the job costs and productivity of the three industries in our sample. We evaluate the following interventions: 1) female transport paid by the policymaker and 2) a BDT 900 subsidy for hiring female workers paid to the employer. We present effects on the following outcomes: We present effects on the following outcomes: 1) male and female employment as well as the the gender employment gap, $L_m^*, L_f^*, L_m^* - L_f^*$ (in percentage points), 2) male and female wages as well as the gender wage gap, $w_m^*, w_f^*, w_m^* - w_f^*$ (in BDT), 3) total male welfare as perceived by employers, $\mathcal{W}_m^E$ (in '000 BDT), and applicants, $\mathcal{W}_m^A$ (in '000 BDT), 4) total female welfare as perceived by employers, $\mathcal{W}_f^E$ (in '000 BDT) and applicants, $\mathcal{W}_f^A$ (in '000 BDT), 5) total profits (in '000 BDT), 6) total costs to the government (in '000 BDT).