

Understanding the Costs of Hiring Women in Pakistan *

Sakina Shibuya

Zunia Tirmazee

Hana Zahir

October 30, 2024

Abstract

74% of the employers in our survey think women are at least as good as, if not better than, men at on-site production jobs in the garment manufacturing industry in Pakistan. Yet, most factories still do not hire many women on factory floors. While much research has been conducted into the constraints women face in obtaining out-of-home employment, this paper instead investigates: What deters firms from hiring women? We utilized unique firm survey data, complemented by a hypothetical-choice methodology and a behavioral game, to examine the economic and non-economic costs of hiring women. Our study innovatively observes how monetary incentives can influence compliance with social norms related to gender segregation. Preliminary results suggest that the willingness to comply with these norms are not as relevant as expected, and costs of hiring women, especially the cost of providing safe transportation to women may be more salient to firms' decision to hire women.

The data collection for this project is still on-going. The current results are preliminary.

*Shibuya: University of Wisconsin-Madison. Corresponding author. sshibuya2@wisc.edu. Tirmazee: Lahore School of Economics. zuniatirmazee@gmail.com. Zahir: Lahore University of Management Science. hanazahir@gmail.com. We thank Laura Schechter, Jeremy Foltz, and Naved Hamid for their continuous feedback, suggestion, and fieldwork support. We also thank Jessica Goldberg, Nancy Quian and other participants at the graduate student workshop at the NBER Summer Institute on gender for their comments. We appreciate research assistance by Ahmed Hassan, Fatima Jamil, Ahmed Nazif, and Mrijan Rimal, and survey support from Kashif Abid and his team at RCons. We thank International Growth Centre (PAK-22075) and Private Enterprise Development in Low-Income Countries (PEDL) for their generous financial support. This document is an output from the research initiative, PEDL, a programme funded jointly by the Centre for Economic Policy Research (CEPR) and the Foreign, Commonwealth and Development Office (FCDO), contract reference PEDL.LOA.8897.Shibuya. The views expressed are not necessarily those of CEPR or FCDO. Shibuya also thanks the Center for Research in Economics and Business (CREB), the UW-Madison Graduate School, and SurveyCTO for their financial and in-kind support. We received IRB approval from UW-Madison (2022-0635). AEA Registry-0009222. All errors are our own.

1 Introduction

Pakistan’s female labor force participation (FLFP) substantially lags behind the rest of the world and its neighboring countries in South Asia. In 2019, the country’s FLFP rate for women ages 15 to 64 was 22.6% which is significantly lower than the world average of 52.6% and even low compared to the South Asian average of 25.2% (World Bank Data Bank, 2021). The social norms of physical gender segregation, prevalent in some Islamic countries, makes entering the formal labor market difficult for Pakistani women, but they can also affect Pakistani employers’ decisions to hire women, as demonstrated in the recent studies in Saudi Arabia by Miller et al. (2019) and Eger et al. (2022).

Even in this social norm environment unfriendly to formal female employment, some Pakistani employers are seemingly willing to hire women, and think women are viable job candidates. Our survey of garment manufacturing firms in the Punjab province in Pakistan shows 74% of the interviewed employers think women are at least as good as, if not better than, men at on-site production jobs (Figure 8 in the appendix).

This project, therefore, asks: **What deters firms from hiring women in Pakistan?** We examine whether there are economic and non-economic costs associated with employing women. The main goals of this paper are to provide evidence on 1) the presence of economic and non-economic costs in hiring women, 2) whether alleviating the burden of economic integration costs can motivate firms to hire women, and 3) the potential for monetary incentives to influence the willingness to comply with social norms.

To answer our research question and fulfil our research objectives, we collect original firm survey data on constraints on hiring women from garment manufacturing companies in the Province of Punjab, Pakistan. At each firm, we identified and interviewed a top manager who held decision making power over the company’s hiring policy. We used two experimental methods to collect unique data: a hypothetical-choice methodology and a behavioral game.

The hypothetical-choice experiment was designed to record firms’ response to difference types of costs associated with hiring women, and collected their reported probabilities to hire women given a hypothetical set of economic costs associated with hiring women. Through our previous work and existing studies, we have identified five types of economic costs: 1) investment in providing safe physical workspace; 2) cost of providing additional training due to the underlying education gender gap; 3) cost of additional management burden to mitigate and solve harassment and disputes; 4) cost of providing safe transportation; 5) compensating for household and care work. We exogenously varied which of these five costs are to be burdened by respondents’ firms, and recorded their reported probabilities of employing women under different cost environments. The exogenous variation in cost

burden allows us to causally estimate what kind of economic costs of hiring women are more restrictive to employers.

Additionally, we asked all respondents to think of themselves as the manager of a hypothetical company, A, with the aim to minimize omitted variable bias in the subsequent statistical analysis. In this experiment, respondents were first shown a profile of the hypothetical Company A, and asked to assume the role of its top manager. They were then assigned to the task of filling ten sawing positions under 16 hypothetical situations, each differing by the types of costs associated with hiring women that their hypothetical company would bear.

The behavioral game is aimed to measure top managers' willingness to comply with social norms that dictate string physical gender segregation. The key feature of this game is that top managers were made to face a trade-off between earnings and complying with social norms of gender segregation. This game asked top managers to sort marbles for one minute for a monetary prize which was determined by the quantity they sorted multiplied by the per 10 gram rate of 60 Pakistani Rupees (PKR) or about 20 cents. They were also offered a helper who can help him sort more marbles. Randomly chosen 90 % of the sample was offered a female helper, while the rest was offered a male helper. The male helper group was intended to serve as a comparison group in the later analysis.

The trade-off was designed such that when a helper joined the game, their entire prize earnings were transferred to the top manager, significantly increasing the manager's game earnings. If a female was assigned as a helper, the top manager needed to compare the expected benefit in prize earning of working with a woman in the same physical space, to the expected benefit of complying with the norms of gender segregation. The top manager was then asked to name the lowest "price" at which he is willing to work with the assigned helper, using the Becker-DeGroot-Marschak (BDM) mechanism (Becker et al., 1964). This price was to be used as the per 10 gram rate to calculate the helper's prize earning. This reported price is what we call top managers' willingness to comply with social norms (WTC).

The interpretation of WTC is as follows. Because WTC is the *lowest* willing price of working the helper, the higher WTC is, the more adherent to social norms of gender segregation the top manager is. In an extreme example where the top manager is extremely willing to comply with the social norms, he would report a very high WTC. On the contrary, if the top manager is extremely eager to work with the helper, he would report a very low non-zero WTC value.

The findings of this study indicate that among the surveyed firms, the adherence of top managers to social norms of physical gender segregation is not a significant factor. The research also reveals that top managers generally perceive their workers as more open to

collaborating with female counterparts, and their willingness to comply (WTC) for a typical male worker does not exceed their own WTC. This suggests that, in the given context, traditional norms of physical gender segregation are not deemed crucial. Moreover, the study highlights that the visibility of a typical male worker's actions is a critical consideration for top managers. Concerns about male workers' behavior towards females escalate when their actions are not observable.

The study also points out that other factors, beyond social norms of gender segregation, play a more dominant role in influencing the hiring of women. Specifically, the cost associated with providing safe transportation for female workers emerges as a key determinant in the decision-making process of firms regarding the employment of women.

This paper investigates the constraints firms face in hiring women and makes three contributions. First, our paper adds a new perspective to the existing literature on FLFP, which primarily focuses on women's constraints, by instead investigating firms' disincentives in hiring female workers. While women's constraints in the labor market have received much attention, the general literature on FLFP lacks an understanding of why firms may also be constrained to hire women, especially in a context like Pakistan, in which strict norms dictate the behaviors of men and women. The extensive literature on FLFP has identified several determining factors including gender differentials in human capital accumulation, and preference-based and statistical discrimination (Altonji and Blank, 1999); behavioral and psychological gender differences and identity (Fortin, 2005; Fernández, 2008; Fernández and Fogli, 2009; Bertrand, 2011; Fernández, 2013; Campos-Vazquez and Velez-Grajales, 2014; Olivetti et al., 2020; Bursztyn et al., 2020); and technological change (Albanesi and Olivetti, 2007). Furthermore, recent Pakistan-specific studies on this topic have identified 1) the burden of household chores and care work (Amir et al., 2018); 2) severely restricted mobility (Field et al., 2020; Cheema et al., 2020); and 3) fear of harassment at work (Makino, 2019) as salient women's constraints in participating in the formal labor market. Since the labor market consists of people who want jobs and others who want workers, a closer look at why employers cannot hire a large share of the country's working-age population may help solve the low FLFP problem.

Second, it complements recent studies by Miller et al. (2019) and Eger et al. (2022) which demonstrate that firms can be deterred from hiring women by fixed integration costs in the context of Saudi Arabia, where cultural norms idealize physical gender segregation, similarly to Pakistan. The latter study, in particular, provides a more granular understanding on different types of costs associated with hiring women with correlational evidence on the relationship between various costs and female hiring. Our paper complements the previous two studies, while it substantially differs in two critical way. First, our study provides be-

havioral evidence on the role of monetary incentives in changing social norm compliance. To our knowledge, this is the first paper to provide such evidence, particularly among employers. Second, this paper provide more causal evidence of the effects of different types of women-related hiring costs.

Lastly, we believe this study’s results hold significant implications for policy development. It demonstrates that in settings where gender-based social norms are prevalent, such norms may be less influential for specific demographics, like well-educated managers at larger firms. Yet, these companies continue to employ a relatively small proportion of women. The study suggests that aiding firms with the costs of providing safe transportation might boost female employment. However, further research is needed to determine the most effective methods for offering this support. Additionally, since the study primarily focused on larger firms and the Pakistani industry is mainly composed of smaller enterprises, further research to understand the hiring constraints for women in these smaller firms would be valuable.

The rest of this paper is structured as follows. First, we explain the context of Pakistan’s garment manufacturing industry in which the current study takes place. Second, we provide detailed information about the survey data we have collected, and describe the data. Third, we describe our estimation strategy. Fourth, we present our estimation results. Lastly, we conclude by interpreting the estimation results.

2 Background

This section provides necessary contextual information about the garment manufacturing industry in Punjab, Pakistan, and justifies the focus of this paper on the costs of hiring women. Firstly, we describe the current state of female labor force participation in the industry and identify contributing factors to the low female labor force participation. These factors are related to prevalent social norms that idealize the physical separation of the sexes. We highlight that the role of such social norms on employers is an understudied area. Secondly, we elaborate on the concept of costs of hiring women and explore how they may affect employers’ hiring decisions. Lastly, we provide a detailed description of our data collection process through experiments to gather data that can help us understand the relationship between costs of hiring women and the decision to hire women.

2.1 Setting

The female labor force participation in Punjab’s garment industry is low, reflecting its general pattern at the national level. Table 1 shows the share of firms with female employees

surveyed in the 2010-2011 Census of the Manufacturing Industry (CMI). The share of garment manufacturing companies with any female employees was 18.1% and 12.1% in 2006 and 2011, respectively. These values are on par the share of female hiring companies among all manufacturing firms in Punjab. Among the garment producing firms with female employees, the average share of female employees in their workforce is 2.4% and 1.6% in 2006 and 2011, respectively. The mean share of female workers in the garment manufacturing industry is much lower than the average in the whole private section, which was 7.5%, according to the 2013 Enterprise Survey (World Bank, 2015). Moreover, the garment industry’s mean share of female employees are driven by high outliers. The industry’s median values show that the majority of the garment manufacturers do not hire women.

Table 1: Female Employees in Punjab’s Garment Manufacturing Industry

	Share of Firms with Female Employees (%)	Share of Female Employees within Firm (%)					
		Mean	Median	SD	Min	Max	N
2006	18.1	2.42	0	10.3	0	100	1,385
2011	12.4	1.5	0	7.27	0	100	4,499

Source: Census of Manufacturing Industries (Pakistan Bureau of Statistics, 2006, 2011)

What can explain such low female labor force participation and female hiring in Punjab’s garment manufacturing industry? The rich literature on female labor force participation provides an extensive list of possible reasons for women’s low labor participation. Such factors include gender differentials in human capital accumulation, preference-based and statistical discrimination (Altonji and Blank, 1999); behavioral and psychological gender differences (Fortin, 2005; Fernández, 2008; Fernández and Fogli, 2009; Bertrand, 2011; Fernández, 2013; Campos-Vazquez and Velez-Grajales, 2014; Olivetti et al., 2020; Bursztyn et al., 2020); and technological change (Albanesi and Olivetti, 2007). Furthermore, recent Pakistan-specific studies on this topic have identified household chores and care work (Amir et al., 2018), severely restricted mobility (Field et al., 2020; Cheema et al., 2020), and fear of workplace harassment (Makino, 2019) as salient women’s constraints.

Many of the above-mentioned potential reasons are a manifestation of social norms around how women and men¹ are to behave and interact. In this study, we define social norms as “motivational mechanism,” shared by members of a community, which shapes individual behaviors by “the feelings of embarrassment, anxiety, guilt and shame”, and, in some cases, by social ostracism, borrowing from the characterization by Elster (1991).

¹While we recognize that gender and sex are non-binary, we focus on these two categories as gender is largely considered to be binary in the context we study in this paper.

In the context of Pakistan, it is a collection of social norms that determines the interaction between women and men. Two such norms that are particularly relevant to the focus of this paper are *purdah* and *izzat*.² *Purdah*, which translates to curtain, is a set of rules that physically and figuratively separate women and men in order to maintain purity, while *izzat*, which means honor or dignity, is a sense of family pride to be protected (Papanek, 1973). These two concepts are heavily intertwined. For example, *izzat* can be damaged by a failure to comply with *purdah* by a female family member who obtain out-of-home employment, or come to a close asexual contact with a man outside of her immediate family even in a professional setting (Syed, 2010; Masood, 2019).

Social norms like *purdah* and *izzat* can also affect employers. While the rich literature on FLFP discovered many women’s constraints in the labor market, it lacks an understanding of how social norms around gender also affects employers and their hiring behaviors. The recent studies demonstrated theoretically and empirically the existence of fixed costs to integrate female workers in an completely male work environment due to social norms for gender segregation (Miller et al., 2022a,b). In the next subsection, we discuss how social norms manifest themselves costs of hiring women faced by employers.

2.2 Costs of Hiring Women

Our initial fieldwork, conducted in the summer of 2021, and existing studies suggest that seven types of costs can potentially be incurred when hiring women. These are 1) investment in providing safe physical workspace (Makino, 2019); 2) cost of providing additional training due to the underlying education gender gap (Minardi et al., 2021) ; 3) cost of additional management burden to mitigate and solve harassment and disputes (fieldwork); 4) cost of providing safe transportation (Field et al., 2020; Cheema et al., 2020); 5) compensating for household and care work (Amir et al., 2018); 6) internal mental cost of breaking away from social norms that promote women and men to be physically segregated (fieldwork); and 7) reputation cost of overstepping such norms (Bernhardt et al., 2018).

There is a meaningful distinction between cost types 1-5 and cost types 6 and 7 as they differ in how they affect firms’ behaviors. We consider the first category the “economic” costs of hiring women.³ All of the economic costs, if incurred, directly enter a firm’s profit function and decrease its profit assuming perfect competition, profit maximization, and

²These concepts are multifaceted. We recognize that they are both social norms as well as religious and cultural values. However, we consider the functioning of these concepts as social norms in this paper so as not to deviate from its focus.

³Our characterization of potential economic costs also overlap with that of the paper by Eger et al. (2022) that studies barriers to hiring women in Saudi Arabia where essentially the same norms exist.

constant marginal revenue.⁴ For example, building a separate women-only bathroom to accommodate female workers requires paying a one-time investment, which increases the firm’s average total cost.

Meanwhile, we call the second category “non-economic” costs, which may not affect the firm’s profit function as straightforwardly as the economic costs. In addition, however, these non-economic costs possibly affect the firm’s top manager, whose norm adherence intentionally or unintentionally may permeate how he manages it. For instance, the top manager may believe that his reputation among his business peers, family members, and friends will be tarnished by deviating from norms and hiring women at his firm. If so, he may conclude that the expected reputation cost is higher than the expected benefit of hiring women at his firm, hence deciding to hire only men.

Furthermore, this distinction is salient because of its implication for future interventions. Following the earlier example, if the firm is constrained by its inability to absorb the cost of building a women-only bathroom, it may be incentivized to hire women if the cost is borne by someone else. However, the firm may still decide not to hire women even after the bathroom subsidy if its top manager adheres to norms for fear of damaging his reputation. To relax the latter constraint requires an entirely different set of interventions.

3 Data Collection Methodology

3.1 Survey Data

We surveyed 238 garment manufacturing firms in Punjab, Pakistan, in collaboration with trade associations in the sector. We partnered with the Pakistan Knitwear Training Institute (PKTI), run by the Pakistan Hosiery Manufacturers and Exporters Association (PHMA), and the Pakistan Readymade Garments Manufacturers & Exporters Association (PRGMEA). Of the total number of interviews, 197 have completed all components of the survey and constitute our analysis sample. Our respondents were top managers of member companies. We define a top manager as someone who holds i) a managerial position within a company and ii) a decision-making power that can set the company’s hiring policy. A top manager, for instance, can be a company owner, executive, or hiring manager. Our survey instrument had six modules and two experiments. Table 9 provides an overview of the question modules.

Modules 1 and 2 collected information about the basic characteristics of a company and

⁴Some types of costs can be more variable while others are fixed. While this distinction is important, this project currently focuses on understanding if these costs affects firms’ behaviors at all.

its top manager, such as years in operation, the number of employees, whether located in an industrial park, top manager's age, gender, education level, and other traits. Additionally, we asked top managers about their beliefs on women's ability relative to men's by occupations and skill levels by asking *In general, how do female workers perform compared to male workers on average?*. We also asked about their wives' employment status, and daughters' education, to get a sense of the top managers' attitude toward female empowerment. As the survey clearly focuses on female employment which can be a culturally controversial topic, we added a few questions from the social desirability measurement method developed originally by Crowne and Marlowe (1960). To minimize the length of the data collection, we were not able to include all 13 questions recommended by Reynolds (1982) who modified the original method later. Nonetheless, the responses to these questions will be useful for us to understand the respondents' social desirability bias.

Module 3 asked questions about the type of products the firm manufactures, revenues, labor costs, and whether the company exports abroad. The questions about exports and export destination is particularly important for this paper since these factors may be related to firms' female hiring.

Module 4 included questions about how the company finds, acquires, and maintains workers, and challenges in hiring. We additionally asked questions about preferred skills and characteristics that top managers seek in male or female workers.

Module 5 asked a series of questions on hiring women, depending on whether firms had have hired women in the past. We inquired reasons for hiring or not hiring women, and expected and actual consequences on hiring women on both firm and individual reputation. We also directly asked if firms face constraints in hiring women.

3.2 Challenges in Collecting Data on the Role of Social Norms

If one is to measure the effects of social norms on hiring decisions and top managers' willingness to comply with social norms, there are three main ways. The first and perhaps most straightforward method is to ask respondents direct questions about whether they factor social norms in hiring decisions and how willing they are to obey or disobey social norms. Unfortunately, while this method is advantageous in its relative ease of execution, a few problems can lead to lower measurement quality.

First, the method of direct questioning is prone to self-reporting bias. In our context, we are particularly concerned about the potential influence that the gender of an enumerator may have on the respondent's responses. The direction of this so-called social desirability bias in our context can depend on the respondent's belief about social norms. For instance,

the respondent may be compelled to show a higher willingness to comply with social norms because he wants to signal to a male enumerator that he is a norm-abiding member of society. Therefore, the direct questioning method does not offer an intuitive way to extract the “truth” away from this bias, especially when questions regard rather sensitive social and cultural norms.

Second, direct questioning also suffers from the potential disconnect between self-reported responses and actual behaviors. What respondents verbally report may not align with their actual behavior, which we cannot observe in this method. Thus, again, this method does not offer a way to reveal respondents’ behaviors.

The direct questioning method is not our preferred method of collecting data on top managers’ behaviors regarding social norms, although we still implement it in module 6 so that we can compare the direct questioning data to the data from the other methods.

The second method is to observe whether each company has hired women workers. Such data certainly tell us whether its top manager has broken away from norms; however, this cannot tell us whether there is a trade-off between the benefits attached to hiring women and its costs. Additionally, the circumstance under which each company and its top manager decide to hire women varies. This variation makes it challenging to use this information to measure willingness to comply with social norms. To mitigate this concern, we use a hypothetical choice method to observe top managers’ hiring decisions under hypothetical situations.

The third method is to create a lab-like environment in the field and use a game to record respondents’ behavior rather than self-reported responses. This is our most preferred method of measuring the willingness to comply with social norms for a few reasons. First, the lab game method measures respondents’ willingness by revealing their behaviors rather than verbal reporting. We believe that this reduces the risk of social desirability bias, mainly because misreporting through behavior is more costly for respondents than verbal misreporting. In other words, if a respondent wants to misreport his true willingness, he must act on it, which can potentially be unpleasant. Meanwhile, lying is much easier and requires no action. Secondly, this method allows us to link behavior and willingness much more concretely. Third, this method buys us more control over the environment where respondents’ behaviors are revealed.

This behavioral game method is certainly not without weakness. Namely, there is a trade-off between the control granted to the research team and the extent to which the lab environment mimics the reality of top managers. Since we are interested in understanding the association between top managers’ norm compliance and their hiring of women, the lab environment we create must not be entirely far off from their daily reality. To mitigate this

trade-off, we consider some variations of the basic game, which further try to replicate the real world.

3.3 Experiments

In order to understand the roles of social norms and related costs of hiring women on hiring decisions, we conducted the following experiments with employees in the garment manufacturing industry in Punjab, Pakistan. To measure employers' willingness to comply with social norms of gender segregation that can severely discourage female hiring, we have implemented a field experiment which we call the *marble sorting game*. This experiment also allows us to see if monetary incentives can change norm compliance. The second experiment, *hypothetical choice method experiment* was aimed to collect data on employers' reaction to the economic costs of hiring women described above when making hiring decisions.

3.3.1 Marble Sorting Game Experiment

The marble sorting game was implemented to elicit top managers' willingness to comply with social norms (WTC) that dictated strict physical separation of the sexes, using the Becker-DeGroot-Marschak (BDM) mechanism (Becker et al., 1964).

The key feature of this game is that top managers were made to face a trade-off between earnings and complying with social norms of gender segregation. This game asked top managers to sort marbles for a short period of time for a monetary prize. The prize was determined by the quantity they sorted in the given time multiplied by the per 10 gram rate of 60 Pakistani Rupees (PKR) or about 20 cents. They were also offered a helper who can help him sort more marbles, thus increasing his prize earnings. Randomly chosen, 90 % of the sample was offered a female helper. If the female helper joined him to play the game in the same physical space, the quantity sorted by her times a different fixed rate would be given to him. In short, his prize earnings, if he played with the female helper, was the sum of the contributions by him and the female helper. The remaining 10% was offered a male helper, which produces a comparison group that enables us to learn about the roles of potential attributes, other than the helper's gender, that may affect the norm compliance price point. Bot the female and male helper groups went through exactly the same game procedures as described below.

Box 2 below presents the flow of the marble sorting game flow. The game started with explaining to the top manager the flow of the game which included two rounds of sorting. In round 1, the top manager was asked to sort a bowl of mixed marbles for one minute alone. The bow of marbles contained five different colored marbles, which the top managers was to

sort into five bowls by color. The first round was intended as a practice run, and to give the respondent a sense of a baseline prize earning he could achieve alone.

In round 2, the top manager was first randomly offered a female or male helper, and told that whether he would sort marbles along or with the helper depended on 1) his reported lowest rate at which he was willing to work with the helper, and 2) the rate that he would randomly draw. We asked “*What is the lowest per gram rate at which you are willing to sort marbles with a [randomly assigned gender] helper?*” Before being asked this question, the respondent was explained that the rate they would report for this question was meant for the helper, as his per 10 gram rate was always 60 rupees. Thus, the reported lowest willing rate implied the additional prize earning he could get by working with the helper. This lowest reported rate is what we call the willingness to comply with social norms.

After the lowest willing rate was reported, the top manager was asked to draw a random number by pushing a button on the enumerator’s tablet screen. If the random draw was greater than his reported WTC, the respondent were to sort marbles *with* the helper. Meanwhile, if the random draw was less than or equal to his reported WTC, the respondent were to sort marbles *alone*. This format follows the key feature of the BDM mechanism. By detaching the actual outcome from the respondent’s decision with randomness, the mechanism makes truth-telling incentive compatible for the respondent. The exact script that the enumerators was instructed to follow during the game is presented in Appendix C.

Ensuring respondents’ comprehension of the game rule and what they were asked to report as the lowest willing rate is extremely important. If they misunderstood, for instance, the process in which whether they would play alone or with the helper in round 2, the effectiveness of the truth telling mechanism could be compromised. To avoid such mistake, we imbedded a sequence of comprehension questions, and without answering them correctly respondents were not allowed to proceed in round 2.

Box 2: Marble Sorting Game Flow

Introduction: Game overview

Round 1: The top manager sorts alone at the per 10 gram rate of PKR 60.

Round 2: WTC revelation and random draw

- If random draw $>$ revealed WTC, play with the helper
- If random draw \leq revealed WTC, play alone

This experiment produces data on each individual manager's lowest rate at which he is willing to play this game with the female helper through the BDM mechanism. This rate may be used to convert the female helper's contribution into a monetary prize. However, a randomly drawn rate determines whether the top manager plays with the female helper and the rate used to convert her sorted quantity to a monetary prize. This rate is drawn after the top manager reveals his lowest willing rate. If the randomly drawn rate is greater than or equal to his willing rate, he plays the game with the female helper and uses the randomly drawn rate for her. If it is less than his willing rate, he plays alone. This random drawing is the key aspect of the BDM mechanism; it induces truth-telling because the top manager has no control over if he gets to play the game with the female helper.

3.3.2 Interpretation of WTC

Because WTC is the *lowest* willing price of working the helper, the higher WTC is, the more adherent to social norms of gender segregation the top manager is. In an extreme example where the top manager is extremely willing to comply with the social norms, he would report a very high WTC. On the contrary, if the top manager is extremely eager to work with the helper, he would report a very low non-zero WTC value.

3.3.3 Hypothetical Choice Method Experiment

The second experiment used the hypothetical choice method which collected firms' reported probabilities to hire women given the economic costs. The basic concept of this method is the following. First, all respondents were shown a profile of a hypothetical company, A. Throughout this section, they were asked to think of themselves as the top manager of this hypothetical firm. As the top manager of Company A, they were tasked to fill ten sewing positions. They were then shown various scenarios in which some of the five economic costs associated with hiring women are paid for by a particular lottery and do not have to be paid for by Company A. Finally, under each cost environment, they were asked to report the probability of filling the ten sewing positions with women instead of men.

More specifically, each respondent was first given a prompt in which the respondent was asked to pretend as the top manager of a hypothetical company who needs to fill ten sewing positions. The English translation of the actual prompt given to the respondents are shown in the textbox below. The respondent is asked to consider each of 16 situations with various hiring costs environments. Each situation shows the five economic costs of hiring women, and it differs from one another in which costs are burdened by the hypothetical company. The respondent is told that costs that are not borne by the company is paid. for by a special

lottery for an indefinite amount of time.

We asked the respondents to report their *probability of hiring 10 women to fill the 10 sewing positions*, and the vacancies must be filled. This indicates that the reported probabilities are to be interpreted as the probabilities of hiring women against hiring men. Therefore, the reported probability of 0% indicates that the respondent's probability of hiring 10 men to fill the positions is 100%.

Box 1: Hypothetical Choice Method Prompt

As the top manager of Company A, one of your main responsibilities is to hire workers for Company A's factory. You are currently tasked to fill 10 sewing positions. In the following, we will ask you to provide the probabilities of hiring 10 people under different scenarios. Please think of yourself strictly as the top manager of company A during this exercise.

You will encounter 8 scenarios.

In each scenario, you will see 2 situations in which you are asked to hire 10 people for sewing positions. Your company really needs to fill 10 sewing positions. Therefore, you must hire 10 men or women to fill these positions.

You cannot choose not to hire.

We ask you to tell us the percent chance (or chances out of 100) of hiring women in each of the situations. The chance of each situation should be a number between 0 and 100.

This means if you say your likelihood of hiring women in a certain situation is 0, your likelihood of hiring men is 100.

In each situation, types of costs of hiring workers that your company must pay varies. In some cases, your company wins a special lottery that covers a particular type of cost indefinitely.

After the prompt was given, the respondent was shown the profile of the hypothetical company, Company A (Figure 1). We constructed profile partially based on real-world statistics to capture the profile of an "average" garment manufacturing firm in Punjab. Thus, we used the average revenue, profit, and mean daily persons engaged during a year in

2021 (Bureau of Statistics, 2022).

The major advantage of providing this hypothetical company profile is its ability to minimize omitted variable bias. By providing hypothetical scenarios to homogenize firm characteristics correlated with constraints to hire women. However, this identification strategy relies on respondents fully “buying in” the hypothetical situations and not bringing in their real-world considerations. To mitigate this risk, we thoroughly piloted our data collection and trained our enumerators.

Figure 1: Profile of Hypothetical Company A

Company A Profile	
Revenue	6,280 lakh Rs.
Profit	1,667 lakh Rs.
Average daily persons engaged during a year	286
Number of female employees	None
Main product	Garments for export
Factory environment	<ul style="list-style-type: none"> - Currently no separate workspace for women - no transportation service for women - no child care facility - no maternity leave
Position you are hiring for	Sewing positions

Notes: The values of revenue, profit, and average daily persons engaged during a year are the industry means obtained from the 2022 Punjab Development Statistics (Bureau of Statistics, 2022).

An example of the situations that the respondent had been given is shown in Figure 2. The costs that needed to be borne by Company A were highlighted in black so that the respondent easily distinguish what costs he needs to take seriously, when considering hiring women. In each situation, the respondent is asked ”How likely are you to hire the group of 10 women in each of the following situation?” The 16 situations were created using the *dcreate* command in Stata to (Hole, 2017) such that the variation across the situations over Company A’s cost burden is most suitable to estimate the parameters of interest, effects of each of the five economic costs associated with hiring women.

Figure 2: Example Situations

	Situation 1	Situation 2
Cost of building a wall/partition to create separate space	Your company pays this cost.	Your company does NOT pay this cost.
Cost of training	Your company pays this cost.	Your company does NOT pay this cost.
Cost of hiring an assistant to deal with management issues	Your company does NOT pay this cost.	Your company pays this cost.
Cost of providing transportation to and from factory	Your company does NOT pay this cost.	Your company pays this cost.
Cost of providing a child care service at the factory	Your company pays this cost.	Your company does NOT pay this cost.

3.4 Characterization of the Study Sample

3.4.1 Firm Characteristics

Table 2 describes the garment manufacturing firms in our analysis sample. The first observation to note is that the firms in our sample tend to be well-established. The majority are located in non-residential areas, have production space larger than several floors of a building, and hire a large number of employees at a given time. The average number of on-premise employees in the past 30 days was 386 workers, while the median was 200 workers. While the distribution of the employee size is quite spread out, the high value of the median tells that the sampled companies are quite large. Additionally, the average and median number of years in operations are 26 and 22 years, respectively, which further indicates that the firms in the sample are fairly well-established.

Second, 95% of our sampled firms are exporters, as we used the membership lists of the trade associations dedicated to apparel exporters. Garment exporters everywhere tend to be subject to different criterion for quality control and labor and environmental rules, often set by international clients based in advanced economies. Our data also reflect this: 90% of the firms has at least one social, audit, environmental, or product international certification.

Third, the most common form of employment among the sampled firms is on-premise contractual production positions. While the average share of this employment type is 88%, the median is 100%, which implies that the majority hires factory workers on contractual basis. This is consistent with the labor intensive nature of garment manufacturing and with hiring practices elsewhere including India, a leading garment exporter (International Labour Office, 2015).

Forth, most of the sampled firms hire men to fill on-premise positions. The average shares of male workers in full-time and contractual on-premise positions in the past 30 days

are 82%, while the median values are 90% and 95%. These numbers starkly contrast the average and median shares of male workers on off-premise contractual positions which is 6% and 0%, respectively. This implies that off-premise jobs, which normally involves take-home small tasks such as cutting lose threads off finished products, are held by women.

These statistics are suggestive of costs associated with hiring women. Most of firms are clearly willing to hire women off the premise of their companies or factories, where they do not have to deal with the possibility of their female and male employees sharing the same physical space. To be clear, this paper investigates the costs of hiring women on-premise.

Lastly, 63% of the companies firms reported that their factory space required women and men to share space, although, surprisingly, over 95% have separate bathrooms and prayer rooms for women in their factories.

Table 2: Basic Company Characteristics

	Mean	Median	SD	Min	Max	Obs
<i>Company location</i>						
Residential area	0.30			0	1	197
Commercial area	0.17			0	1	197
Industrial area	0.51			0	1	197
Industrial park	0.02			0	1	197
<i>Production space</i>						
One room/ shop	0.01			0	1	197
One floor in a building	0.08			0	1	197
Several floors in a building	0.48			0	1	197
An entire building	0.24			0	1	197
Several buildings	0.20			0	1	197
Part of a larger company	0.06			0	1	197
<i>International certificate</i>						
Has social certification	0.78			0	1	197
Has audit certification	0.83			0	1	197
Has environmental certification	0.79			0	1	197
Has product certification	0.85			0	1	197
At least one int'l certificate	0.89			0	1	197
Years in operation	26.99	23.00	16.86	3	99	197
# on-premise employees in last 30 days	529.53	200.00	1914.81	6	26,000	197
<i>Employee shares by contract types (%)</i>						
Full-time on-premise production workers	65.56	70.00	24.12	0	100	197
Full-time on-premise non-production workers	34.44	30.00	24.12	0	100	197
Contractual on-premise production workers	86.84	100.00	32.25	0	100	197
Contractual on-premise non-production workers	2.50	0.00	11.67	0	100	197
Contractual off-premise production workers	12.49	0.00	32.88	0	100	197
Contractual off-premise non-production workers	1.73	0.00	12.37	0	100	197
<i>Employee shares by gender (%)</i>						
Male in full-time on-premise jobs	82.92	90.00	24.88	0	100	197
Male in contractual on-premise jobs	80.20	95.00	32.78	0	100	197
Male in contractual off-premise jobs	6.39	0.00	22.99	0	100	197
Exporter	0.94			0	1	197

3.4.2 Top Manager Characteristics

Table 3 presents the characteristics of the top managers of the sampled firms we interviewed. The average top manager in our analysis is 41 years old, has the industry experience of 14 years, and has worked at the current firm for about 10 years. The top managers in our sample are highly education: 87% have at least a Bachelor degree or equivalent, over half of them have a Master's degree or equivalent. Most of them are also married and living together with their spouses.

As mentioned earlier, we identified and interviewed top managers who had control over their firms' hiring policy. Therefore, 77% of the top managers were human resource (HR)

managers, and 73% reported their primary responsibility to be HR management.

Table 3: Basic Top Manager Characteristics

	Mean	Median	SD	Min	Max	Obs
If male	0.98			0	1	197
Age	39.33	38.00	10.54	22	73	197
Industry experience (years)	13.41	11.00	9.05	1	50	197
Firm experience (years)	9.14			1	45	197
<i>Education</i>						
Primary school	0.01			0	1	197
Middle school	0.01			0	1	197
Matriculation	0.03			0	1	197
Intermediate	0.08			0	1	197
Bachelor degree or equivalent	0.42			0	1	197
Masters degree or more	0.46			0	1	197
<i>Marital status</i>						
Married (Living together)	0.86			0	1	197
Single	0.12			0	1	197
Engaged	0.02			0	1	197
<i>Position within firm</i>						
Owner	0.10			0	1	197
HR manager	0.79			0	1	197
Factory manager	0.05			0	1	197
Other	0.06			0	1	197
<i>Primary responsibility within firm</i>						
Financial management	0.16			0	1	197
HR	0.77			0	1	197
Sales	0.02			0	1	197
Production management	0.01			0	1	197
Other	0.06			0	1	197

Table 4 describes the top managers' views on women and employment. The top managers in our sample were generally open to the idea of women's out-of-home employment. Eighty five percent believed that it was acceptable for women to work outside the home, and for garment manufacturing firms to hire women for both managerial and non-managerial positions. When asked about women's competency in common garment manufacturing tasks against men's, nearly 50% answered women always or sometimes outperforms men in cutting, sewing, and finishing. Only 23-25% thought that women underperforms in these tasks compared to men.

Table 4: Managers' Views on Women and Employment

	Share
<i>Acceptable for women to work outside the home</i>	
Strongly agree	0.57
Somewhat agree	0.29
Neutral	0.04
Somewhat disagree	0.04
Strongly disagree	0.06
<i>Views on performance on cutting and sewing fabric</i>	
Female workers always outperform male workers.	0.28
Female workers sometimes outperform male workers.	0.23
Female workers are comparable to male workers.	0.27
Female workers seldom outperform male workers.	0.15
Female workers never outperform male workers	0.07
<i>Views on performance on finishing ^a</i>	
Female workers always outperform male workers.	0.25
Female workers sometimes outperform male workers.	0.23
Female workers are comparable to male workers.	0.27
Female workers seldom outperform male workers.	0.16
Female workers never outperform male workers	0.09
<i>Acceptable for a firm in the garment industry to hire women for non-managerial positions</i>	
Strongly agree	0.37
Somewhat agree	0.44
Neutral	0.09
Somewhat disagree	0.04
Strongly disagree	0.07
<i>Acceptable for a firm in the garment industry to hire women for managerial positions</i>	
Strongly agree	0.41
Somewhat agree	0.44
Neutral	0.08
Somewhat disagree	0.04
Strongly disagree	0.04
Obs	197

^a Finishing in the garment manufacturing involves cutting loose threads, washing, ironing, and packing of products.

Table 5 reports top managers' perceived constraints in hiring women by occupation types. For non-managerial positions, women as disruption at work was the most reported constraint in hiring women, followed by women's family c commitment. For managerial jobs, the lack of women with management experience was the most reported constraint.

Table 5: Perceived Constraints in Hiring Women

	Share
Feel constrained in hiring women for production	0.17
<i>For non-managerial jobs</i>	
Women's family commitments	0.26
Gov't regulations	0.16
Disruption at work	0.30
Expenses related to hiring women	0.13
Other	0.43
<i>For managerial jobs</i>	
Unavailability of women with mgnt experience	0.21
Women's family commitments	0.19
Gov't regulations	0.11
Disruption at work	0.21
Expenses related to hiring women	0.09
Other	0.47
Obs	197

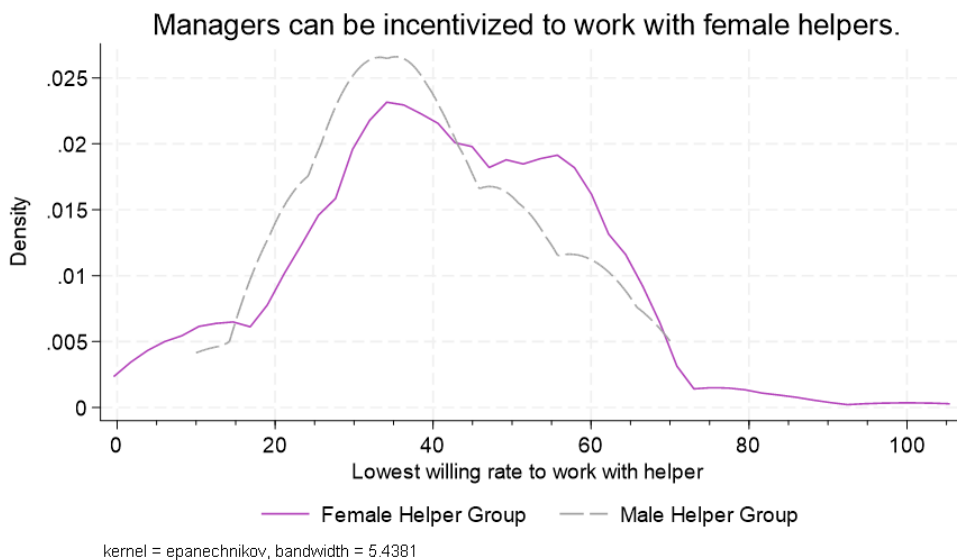
4 Experiment Results

4.1 Monetary Incentives in Changing Top Managers' WTC

Figure 3 presents the results from the marble sorting game experiment, and plots the willingness to comply with social norms of gender segregation (WTC) for the group of top managers who were assigned a female helper. The mean WTC is 41 Pakistani rupees, which is close to the median.

While managers have different thresholds at which they are willing to work with the helper, the gender of the helper does not seem to affect their decisions. Comparing the distributions of managers' reported lowest willing rate to work with the helper across the groups of managers who were assigned the female or male helper, we learn that there is not statistically meaningful difference between the means of these groups.

Figure 3: Willingness to Comply with Social Norms



	Mean	Median	SD	Min	Max	Obs
Assigned female helper	40.99	40.00	16.94	5	100	173
Assigned male helper [†]	38.33	37.50	15.00	10	70	24
Difference	2.66					
p-value	0.47					

[†] 10% of the respondents were assigned to the male helper for comparison.

4.1.1 Why social norms are seemingly unimportant?

The experimental results presented above suggest that the social norms around physical gender segregation may not be important for the samples managers. In the following we consider

Managers’ preference for keep decisions private and work alone.

We asked managers to report what was the most important determining factor when reporting the lowest willing rate (Table 6)⁵. About 40% reported enumerators’ observance of their decisions, and possible visibility from other people at the interview sites on their company premises affected their decisions. Additionally, 15% reported that they preferred working alone.

Table 6: Managers’ Decision Making Process during the Sorting Game

	Share	Obs
<i>Most important factor in determining lowest rate</i>		
Concerned about COVID19	0.04	172
Prefer to work alone	0.15	172
The helper’s gender	0.02	172
Enumerator (male) knows one’s decision	0.17	172
Game is played at one’s workplace	0.05	172
Other people may observe one’s decision	0.40	172
Other	0.17	172
<i>Why concern about the helper’s gender ^a</i>		
Concerned about helper’s ability due to gender	1.00	3

^a Responses of those who reported that the helper’s gender as the most important factor

Managers have different WTC for their male workers

It is possible that managers in our sample simply do not care about complying with social norms of gender segregation but they do for their male workers. For instance, highly educated managers in our sample may feel that someone like their male factory workers with less education may be more likely to misbehave toward women, and believe that they should be physically separated from women. If this is the case, we expect that managers’ WTC for the hypothetical male helper to be higher than their own. We now compare managers’ WTC for

⁵The number of observations is lower than in the rest of the study on this table, because some interviewees were not asked these questions due to a coding mistake in the survey instrument.

themselves to their WTC for their workers as well as their guess of the hypothetical male worker's WTC.

Figure 4 compare these different WTC across the four different hypothetical scenarios. First, we learn that managers' WTC for the male worker and their guessed worker WTC is generally lower than the WTC for themselves. Second, across the hypothetical scenarios, the distributions of managers' WTC and guessed worker WTC are very similar.

Figure 4: Managers' WTC for Male Worker and Male Worker's WTC Guessed by Managers

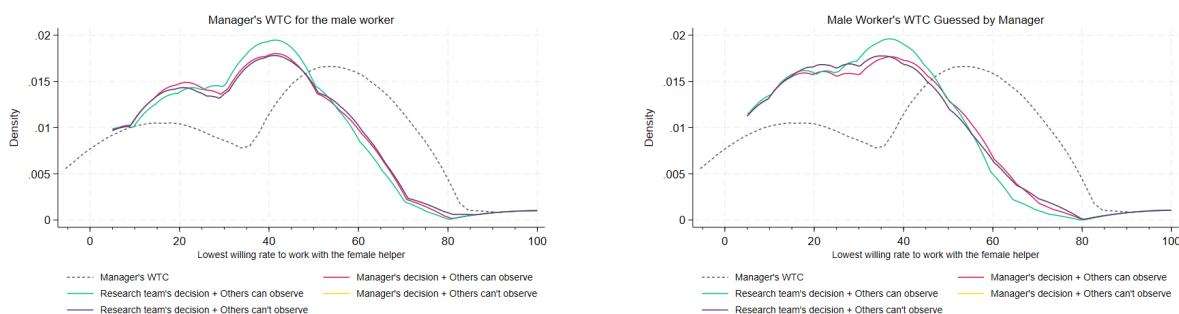
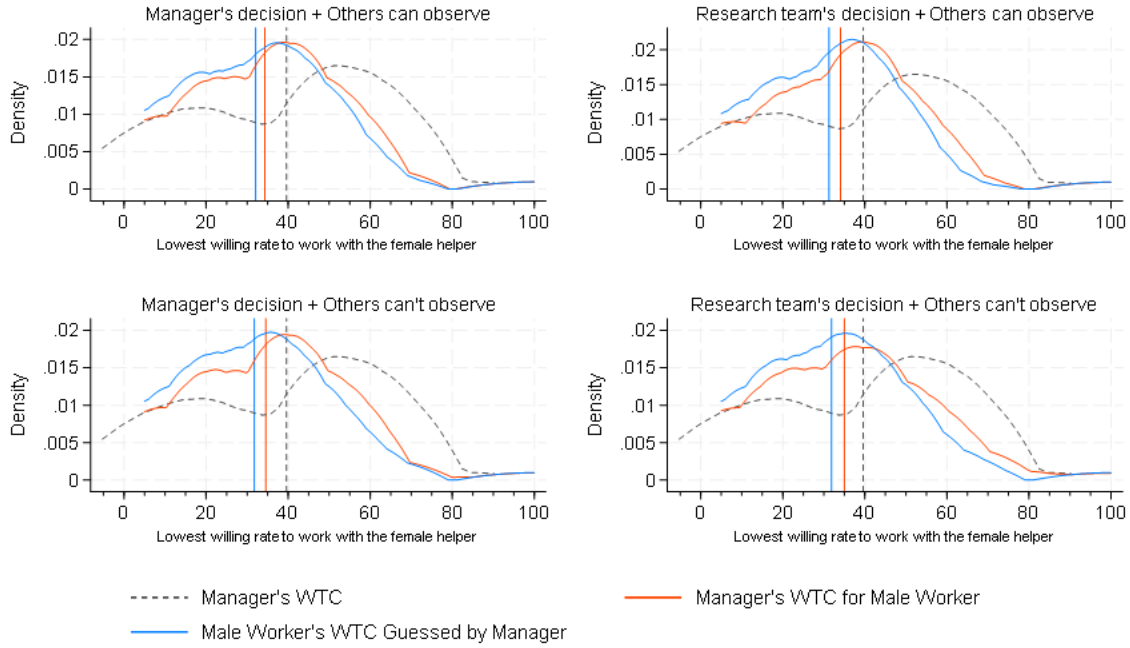


Figure 5 shows the distribution of managers' WTC for the male worker (orange line), and the male worker's WTC guessed by managers (blue line), compared against managers' WTC obtained in the marble sorting game (gray dotted line). We first note that managers' own WTC is the highest on average under all four situations, which implies that they are more willing to comply with norms for themselves than for their male factory workers.

Second, managers' WTC for the male worker is lower than their WTC but higher than their guess of workers' WTC. This may indicate either managers think that workers have lower norm cost or they think that workers have higher opportunity cost on forgone prize earning.

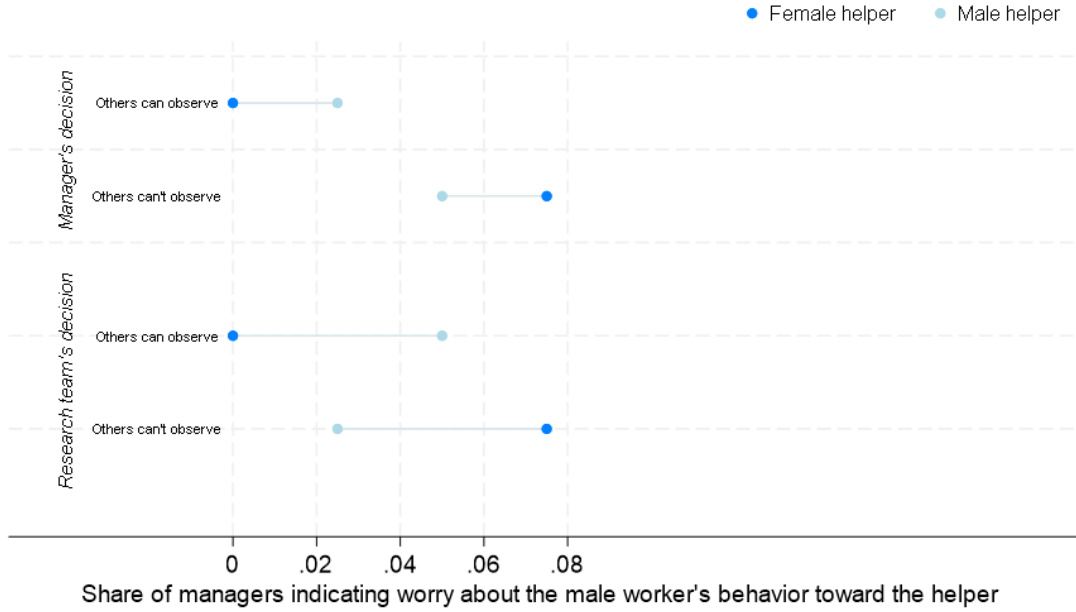
Figure 5: Comparing WTC under Hypothetical Situations



Note: The number of observations is 40. *Manager's WTC* was recorded in the marble sorting game. *Manager's WTC for Male Workers* and *Male Worker's WTC Guessed by Manager* were reported by top managers under the four different hypothetical situations.

We also learn that observability of the male worker's behavior is very important for managers. Figure 6 plots the difference in the share of managers who expressed concern for the male worker's behavior toward the female and male helper under the four situations. Surprisingly, when *others can observe* the male worker's game behavior, no managers expressed concern for the worker's behavior toward the female helper, although the share of concerned managers jumps nearly 8 percentage points when there is no observability of the worker's behavior.

Figure 6: Managers' Concerns for Male Workers' Behavior toward Women



Note: The number of observations for this analysis is 40 as the relevant questions were asked for only this subset of the whole sample. Manager's worries for misbehavior by the hypothetical male worker toward the fe/male helper are managers' responses to the question, *Do you worry if your male factory worker misbehaves toward the fe/male helper?* under four hypothetical situations where a hypothetical male factory worker who works for the manager would sort marbles with the fe/male helper. The situations vary in two dimensions: 1) if the decision to sort marbles with the helper has been made by the manager or the research team; and 2) if other people can observe the male worker is working with the helper.

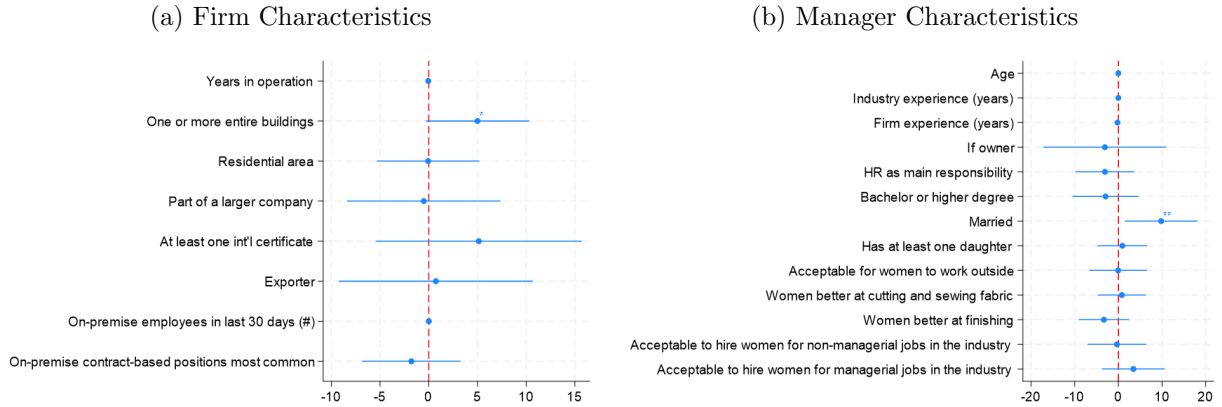
In sum, it seems that managers are higher willingness to let the hypothetical male subordinate work with the helpers; however, they worry about his behavior against the female helper when his behavior is not visible. This analysis produced a rather unintuitive result where managers' WTC is higher than their WTC for the male worker, indicating that managers show higher willingness for the male worker to work with the female helper. What this result seems to imply is that the gender of the helper is not as important in manager's mind.

4.2 Correlates of WTC

Figure 7 investigates the correlation between various firm and top manager characteristics and WTC. We observe that WTC is not correlated with any observable characteristics of firms, except for the size of production facilities (*one or more entire buildings*) (Panel A). We also find that top managers' being married was positively correlated with WTC, implying being married is associated with higher willingness to abide by the social norms of physical

gender segregation (Panel B).

Figure 7: Correlates of Willingness to Comply with Social Norms



Note: Robust standard errors

4.3 Relationship between WTC and Female Hiring

Table 7 presents the correlational relationship between WTC and firms' female hiring and related hiring practices. More specifically, we investigate outcomes indicating if a company has ever hired women for non-managerial or managerial positions, if it considers whether female job candidates live in the same city and has a means of transportation, and the male shares in on-premise and off-premise contractual production positions. We observe no statistically meaningful relationship between WTC and these outcomes.

Table 7: Relationship of WTC to Female Hiring
 Indep. Var: Willingness to Comply

Dependent Variables	Coeff	R-squared	Obs	D.V. Mean (%)
If ever hired women: non-managerial production	0.098 (0.189)	0.211	197	81.73
If ever hired women: managerial production	0.044 (0.229)	0.132	197	53.81
If consider if female living in the same city	0.056 (0.092)	0.071	197	4.06
If consider female's transportations	0.029 (0.049)	0.118	197	4.57
If team has mixed gender	-0.284 (0.188)	0.080	197	25.89
Male share in on-premise full-time production jobs	0.133 (0.112)	0.125	197	82.92
Male share in on-premise contractual production jobs	0.069 (0.144)	0.195	197	80.20
Male share in off-premise contractual production jobs	-0.017 (0.107)	0.079	197	6.39

Robust standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Covariates included are managers' characteristics including age, education, marital status, if they have daughters, experience, position within the firm, and main responsibility, and firm characteristics including years in operations, size of production workforce, premise size, location, if part of a larger company, international certification, commonly used contract type, and exporter status. The binary dependent variables are scaled by 100 for easier interpretation.

4.4 Other Mechanisms

Table 8 shows the estimated effect of each of the five economic costs of hiring women (Column 1). All five costs decreased the likelihood of hiring women. Of the five costs, the cost of transportation has the largest magnitude, while the cost of additional assistance to deal with management issues has the least impact. This is perhaps because providing transportation is not a fixed cost that requires one-time investment, and possibly perceived as more costly by top managers.

Table 8: Effects of Economic Costs of Hiring Women

	Dependent var.: Reported probability of hiring 10 women					
	(1)	(2)	(3)	(4)	(5)	(6)
Cost of additional assistant	-5.854*** (0.642)					-5.854*** (0.643)
Cost of wall/partition		-6.655*** (0.773)				-6.655*** (0.774)
Cost of child care			-8.118*** (0.853)			-8.118*** (0.853)
Cost of training				-7.208*** (0.768)		-7.208*** (0.769)
Cost of transportation					-9.007*** (0.830)	-9.007*** (0.831)
Obs.	2768	2768	2768	2768	2768	2768
Manager FE	✓	✓	✓	✓	✓	✓
Dep var mean	60.34	60.34	60.34	60.34	60.34	60.34

Note: Robust standard errors in parentheses clustered at the firm level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5 Interpretation and Conclusion

The results from this study have shown that, at least among the firms sampled in this study, top managers' willingness to comply with the social norms of physical gender segregation are not relevant. The study has also shown that top managers seem to believe that their workers are more willing to work with female, and that their WTC for a typical male worker is not higher than their WTC for themselves, suggesting that, at least in the current context, the social norms of physical gender segregation seems unimportant.

However, this study has also produced some evidence that the observability of a typical male worker's behavior is an important factor for top managers. When the male worker's behavior cannot be observed, top managers are more worried about his conduct against female.

Other constraints to hiring women may dominate over the social norms of gender segregation. In particular, the cost of providing save transportation to female workers seems to be the most salient in determining firms' likelihood of hiring women.

We believe that the finding from this study is relevant to policy making. The study shows that even in a context where social norms around gender are a dominant part of life, norms may not matter so much for certain groups such as well-educated managers of relatively large firms. However, these firms still hire women for only a small fraction of their total workforce. While the finding from the current study suggests that assisting firms

on the cost of providing safe transportation may encourage firms to hire more women, we believe that a further study that can identify an efficient way to provide such support is necessary. Moreover, the current study sample is limited to relatively large companies, while the industry in Pakistan predominantly consists of much smaller firms. An additional study that can understand constraints of hiring women among these firms will also be relevant.

References

- Albanesi, S. and Olivetti, C. (2007). Gender roles and technological progress NBER working paper series 13179.
- Altonji, J. G. and Blank, R. M. (1999). Race and gender in the labour market. *Handbook of Labor Economics, Volume 3*, 3:3143–3259.
- Amir, S., Kotikula, A., Pande, R. P., Bossavie, L. L. Y., and Khadka, U. (2018). Female labor force participation in pakistan. Technical report, World Bank.
- Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behav. Sci.*, 9(3):226–232.
- Bernhardt, A., Field, E., Pande, R., Rigol, N., Schaner, S., and Troyer-Moore, C. (2018). Male social status and women’s work. *American Economic Association (AEA) Papers and Proceedings*, 108:363–367.
- Bertrand, M. (2011). New perspectives on gender. In *Handbook of Labor Economics*, volume 4, pages 1543–1590. Elsevier B.V.
- Bureau of Statistics (2022). Punjab development statistics 2022. Technical report, Bureau of Statistics Planning & Development Board Government of the Punjab.
- Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in saudi arabia. *Am. Econ. Rev.*, 110(10):2997–3029.
- Campos-Vazquez, R. M. and Velez-Grajales, R. (2014). Female labour supply and inter-generational preference formation: Evidence for mexico. *Oxford Development Studies*, 42(4):553–569.
- Cheema, A., Khwaja, A. I., Naseer, M. F., and Shapiro, J. N. (2020). Glass walls: Experimental evidence on access constraints faced by women (unpublished working paper).
- Crowne, D. P. and Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 24:349–354.
- Eger, C., Fetzer, T., Peck, J., and Alodayni, S. (2022). Organizational, economic or cultural? firm-side barriers to employing women in saudi arabia. *World Dev.*, 160:106058.
- Elster, J. (1991). Rationality and social norms. *Arch. Eur. Sociol.*, 32(1):109–129.

Fernández, R. (2008). Alfred marshall lecture: Women, work, and culture. *J. Eur. Econ. Assoc.*, 6(2-3):295–320.

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

Fernández, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177.

Field, E., Junaid, S. U., Majid, H., Malik, A., Shahid, A., and Vyborny, K. (2020). Transport and urban labour market integration: Evidence on travel time and congestion from a mass transit quasi-experimental evaluation and evidence on firms from a randomised control trial in pakistan. Technical report, International Initiative for Impact Evaluation (3ie).

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, 21(3):416–438.

Hole, A. R. (2017). DCREATE: Stata module to create efficient designs for discrete choice experiments. *Statistical Software Components*.

International Labour Office (2015). Insights into working conditions in india’s garment industry. Technical report, International Labour Organization.

Makino, M. (2019). Labor market information and parental attitudes toward the labor force participation of their daughters: Experimental evidence from rural pakistan IDE-JETRO discussion paper 826.

Masood, A. (2019). Doing gender, modestly: Conceptualizing workplace experiences of pakistani women doctors. *Gend. Work Organ.*, 26(2):214–228.

Miller, C., Peck, J., and Seflek, M. (2019). Integration costs and missing women in firms.

Miller, C., Peck, J., and Seflek, M. (2022a). Integration costs and missing women in firms around the world. *AEA Papers and Proceedings*, 112:578–582.

Miller, C., Peck, J., and Seflek, M. (2022b). Missing women, integration costs, and big push policies in the saudi labor market. *Am. Econ. J. Appl. Econ.*, 14(2):51–77.

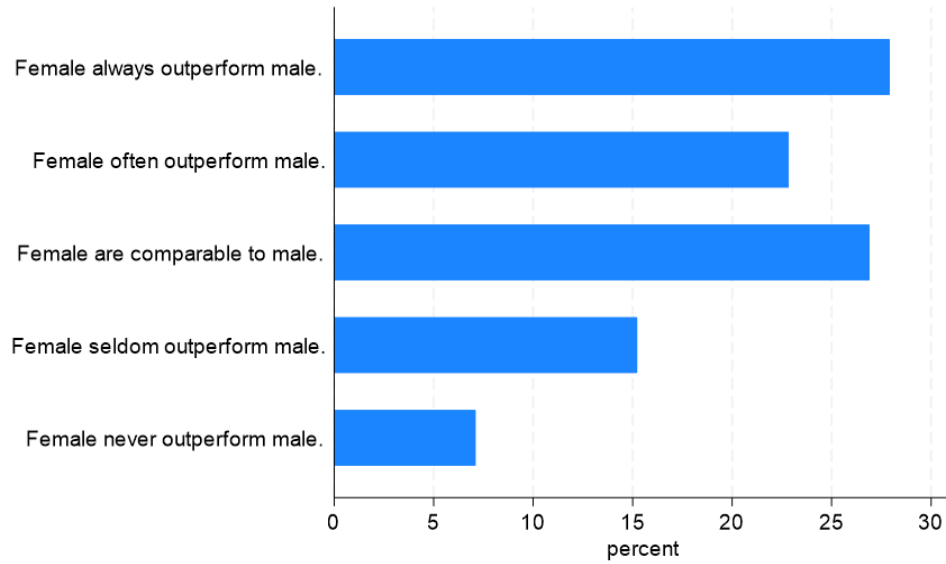
Minardi, A. L., Akmal, M., Crawford, L., and Hares, S. (2021). CGD blog: Promoting gender equality in pakistan means tackling both real and misperceived gender norms. <https://www.cgdev.org/blog/promoting-gender-equality-pakistan-means-tackling-both-real-and-misperceived-gender-n>

- Olivetti, C., Patacchini, E., and Zenou, Y. (2020). Mothers, peers, and gender-role identity. *J. Eur. Econ. Assoc.*, 18(1):266–301.
- Pakistan Bureau of Statistics (2006). Census of manufacturing industries 2006 punjab.
- Pakistan Bureau of Statistics (2011). Census of manufacturing industries 2011 punjab. <https://www.pbs.gov.pk/content/industry>.
- Papanek, H. (1973). Purdah: Separate worlds and symbolic shelter. *Comp. Stud. Soc. Hist.*, 15(3):289–325.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the marlowe-crowne social desirability scale. *J. Clin. Psychol.*, 38(1):119–125.
- Syed, J. (2010). An historical perspective on islamic modesty and its implications for female employment. *Equality, Diversity and Inclusion: An International Journal*, 29(2):150–166.
- World Bank (2015). Enterprise survey 2013 - pakistan.
- World Bank Data Bank (2021). Labor force participation rate, female (% of female population ages 15+) (modeled ILO estimate) - pakistan. <https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS?locations=PK>.

A Graphs

Figure 8: Employers' Perception of Relative Female Job Performance

With which of the statements do you most agree, with respect to non-finishing production positions?



Notes: Non-finishing production positions were defined in the survey as factory positions in which people work on tasks such as knitting, dyeing, cutting, embroidery, and stitching. There is no statistically significant correlation between employers' positive beliefs about women's relative performance and their social desirability bias.

Figure 9: Distribution of Share of Male Workers at Factory by Employment Types



B Summary of Questions

Table 9: Summary of the Questionnaire

Module	Description
1 Top manager characteristics	<ul style="list-style-type: none"> • Gender, age, education, and marital status • Origin and family's origin • Years of experience in the firm and garment industry • Beliefs about women's general ability compared to men • Beliefs about women's out-of-home employment • If his wife works • Daughters' education • Crowne-Marlowe social desirability questions
2 General firm characteristics	<ul style="list-style-type: none"> • Legal status and ownership by gender • Year of establishment and registration • Number of employees by occupation • Share of female employees by occupation • Internationally-recognized certification status • If located in an industrial park • If access to finance an obstacle
3 Production	<ul style="list-style-type: none"> • Type of manufacturing and products • Annual sales (export vs domestic) • Export destination country
4 General hiring practice	<ul style="list-style-type: none"> • Type of position most frequently hired for • Frequency of hiring • Mode of worker acquisition • Challenges in hiring • Preferred skills and characteristics of workers
5 Hiring women	<ul style="list-style-type: none"> • Experience hiring women (now or in the past) • Reasons for hiring/not hiring women • Consequences on hiring women on reputation • Constraints in hiring women

C Marble Sorting Game Script

Our enumerators were strictly instructed to read exactly from the script programmed in the tablets, and not to deviate from the script.

Introduction

READ:

- In this part, we want you to sort marbles as fast as you can for a monetary prize and take a very short survey at the end.
- There will be two rounds of the marble sorting game.
- You will be paid the prize after the short survey.

Round 1

READ:

- In your first round, you will sort a bowl of mixed marbles by yourself.
- You will sort the marbles for 1 minute.
- Your prize is determined by how much you sort in the given time.
- The per 10 gram rate is 60 rupees.
- So, for example, if you sort 100 grams in 1 minute, your prize is 600 rupees (100 grams divided by 10 grams times 60 rupees).

ENUMERATOR: After the respondent finishes his first round, please weigh the sorted marbles with your scale and input the sorted quantity in the survey form on your tablet. The survey form will calculate the respondent's prize.

READ: You sorted [Total Quantity Sorted] grams, and earned [Prize] rupees.

Round 2

Enumerator: Before you start this round, please make sure there is no one else other than you and your respondent in the room.

READ:

- Now, you will sort a bowl of mixed marbles for 1 minute again.
- However, this time, we want to offer you a [randomly assigned gender] helper to help you sort marbles and earn a larger prize.
- If you decide to work with the helper, you and the helper sort marbles for 1 minute together in this room. Otherwise, you sort alone.
- If you work with the helper, your prize is calculated based on the sum of the quantity you sort and the quantity the helper sorts.
- Your per 10 gram rate is 60 rupees, the same rate as your first round.
- The helper' per gram rate is not yet decided. It will be determined by chance after you tell us if you want to work with the helper. I'll explain how the rate is decided shortly in detail.
- The helper is someone that the research team has hired. They are paid an hourly wage to be part of this event and also are paid the game prize.

Enumerator: Please ask your respondent “Do you have any questions so far?”

READ:

- Now let me explain how the helper's per gram rate is decided for this round.
- I will ask you to tell me the lowest per gram rate at which you are willing to play with a [randomly assigned gender] helper. The per gram rate you name is the rate for the helper.
- In this bag, I have many different balls with different numbers on them.
- Each number represents a potential per gram rate for the helper.
- After you tell me your lowest willing per gram rate, I will ask you to pick a ball from the bag, and we will look at the number together.
- If the number you draw from the bag is greater than or equal to your willing rate that you name for the helper, you will sort marbles with a [randomly assigned gender] helper in this room. You will play at the per 10 gram rate of 60 rupees, the same rate as before. The helper will play at the randomly drawn rate.

- If the number you draw from the bag is less than your willing rate, you will sort marbles alone at the same rate at which you played in the first round.
- After the helper’s per gram rate for this round is decided, I will bring in the helper to this room.
- You cannot change your mind after you draw a number from the bag.
- It is best if you name the lowest per gram rate at which you are actually willing to play the marble sorting game with the helper.
- When you are deciding to accept or reject the helper offer, you can think about how much you were able to sort by yourself in the first round, and think of how much contribution another person can bring to your prize earning.

Enumerator: Please ask your respondent “Do you have any questions?”

ENUMERATOR: Remember that we want the respondent to state the **LOWEST** per gram rate at which they are **WILLING** to sort marbles with the helper.

READ:

- Let us begin the second round.
- What is the lowest per gram rate at which you are willing to sort marbles with a [randomly assigned gender] helper?

Note: The respondent’s lowest per gram rate is noted as X in the following.

★

- Now, if you draw a number from the bag that is greater than or equal to X , you will sort marbles with a [randomly assigned gender] helper. If you draw a number from the bag that is less than X , you will sort marbles alone, even if you are willing to sort at a lower rate. You cannot change your stated per gram rate for the helper after you draw a number from the bag. Do you understand?
- Please, tell me: if you draw [$X+1$ rupees] from the bag now, will you work alone or with the helper?

1. I will work alone.

2. I will work with the helper.
- Please tell me: if you draw $[X+1]$ rupees from the bag now, which per gram rate will be applied to you?
 1. The per gram rate used in the first round
 2. The per gram rate of $[X+1]$ rupees
 - Please tell me: if you draw $[X+1]$ rupees from the bag now, which per gram rate will be applied to the helper?
 1. The per gram rate used in the first round
 2. The per gram rate of $[X+1]$ rupees

ENUMERATOR: If the respondent does not give the correct answer, explain the rules again and then ask the question again.

- Please, tell me: if you draw $[X-1]$ rupees from the bag now, will you work alone or with the helper?
 1. I will work alone.
 2. I will work with the helper.
- Please tell me: if you draw $[X-1]$ rupees from the bag now, which per gram rate will be applied to you?
 1. The per gram rate used in the first round
 2. The per gram rate of $[X-1]$ rupees
- Please tell me: if you draw $[X-1]$ rupees from the bag now, which per gram rate will be applied to the helper?
 1. The per gram rate used in the first round
 2. The per gram rate of $[X-1]$ rupees

ENUMERATOR: If the respondent does not give the correct answer, explain the rules again and then ask the question again.

- If you draw $[X-1]$ from the bag, will you want to play with the helper for $[X-1]$?

1. Yes
 2. No
- Do you want to change your bid to [X-1]?
 1. If YES: Go to ★ and use [X-1] instead of X
 2. If NO: Go to the next question
 - Is X truly the lowest rate at which you are willing to sort marbles with the helper?
 1. If YES: Proceed
 2. If NO: Go back to ★

Note: The respondent's lowest per gram rate is noted as X and the randomly drawn rate is noted as Y in the following.

READ:

- Now, you will randomly draw a number. You will do so by clicking a button on my tablet which starts a lottery.
- If you draw a number that is greater than or equal to X, you will sort marbles with a [randomly assigned gender] helper.
- If you draw a number that is less than X, you will sort marbles alone.
- Are you ready to draw?
- Please push the Next button to draw a number.
- Your drawn rate for this round is Y.

Case 1: If the randomly drawn rate (Y) > the respondent's lowest willing rate (X)

READ:

- Your drawn rate is Y, which is greater than or equal to X, the rate at which you said you would be willing to play with the helper.
- You will sort marbles with your helper.
- The helper's per 10 gram rate is Y.

ENUMERATOR: Please bring in the helper and set up the sorting game now. After sorting, please record the quantity sorted by the respondent and helper separately in your tablet.

Case 2: If the randomly drawn rate (Y) \leq the respondent's lowest willing rate (X)

READ:

- Your drawn rate is Y , which is less than X , the rate at which you said you would be willing to play with the helper.
- You will sort marbles alone.

ENUMERATOR: Please set up the sorting game now. After sorting, please record the quantity sorted by the respondent in your tablet.

Box 3: Post-game Hypothetical Question Prompt

I want to ask you some more hypothetical questions.

Recall Round 2 of the game you just played.

Imagine that one of your male factory workers who work on the factory floor will play the same game in your stead with the gender helper.

If your male factory worker sorts alone, he earns a prize calculated based on the quantity he sorts, and the per 10-gram rate of Rs. 60. And you will also earn the same prize.

For example, if your male factory worker sorts 100 grams. He receives Rs. 600 ($= 100/10 \cdot 60$). And you will also receive Rs. 600.

If your male factory worker sorts with the gender helper, he AND you each earn the sum of your male factory worker's contribution and the gender helper's contribution.

For example, suppose that your male factory worker and the gender helper each sort 100 grams. Assume for the sake of demonstration, that the gender helper's per 10-gram rate is also Rs. 60. Then, your male factory worker earns Rs 1,200 ($= 100/10 \cdot 60 + 100/10 \cdot 60$). You also earn Rs 1,200.

The gender helper is part of the research team and is paid hourly. Your male factory worker is told that the helper is paid, and you also receive a prize.

When your male factory worker sorts with the gender helper, each of them will be given a large bowl of mixed marbles, and five smaller bowls to sort out marbles by color. They will be sharing the same table.

Similarly to Round 2, we want you to tell us the lowest rate at which you are willing to allow your male factory worker to work with the gender helper under the following conditions.

As in Round 2, the lowest willing rate you tell us will be used to convert the helper's contribution to a monetary prize.

D Social Desirability Bias

Table 10: Summary Statistics: Top Managers' Social Desirability Bias

	Mean	Median	SD	Min	Max	Obs
<i>Most important factor in determining lowest rate</i>						
Social desirability index ^b	0.62	0.67	0.24	0	1	197
Individual questions						
If thinks himself as a good listener ^b	0.98					197
If never insisted on having things his own way ^c	0.28					197
If never irked on people with different ideas ^d	0.59					197

^a Social desirability index is the sum of the scores on the three social desirability questions below divided by three. Thus, the index ranges from 0 to 1 with 1 indicating the greatest sign of social desirability. The questions and scoring method are based on Crowne and Marlowe (1960). In each question, a top manager was asked if each statement was true or false. If the top manager's response on a particular question matches the one determined by the authors to show social disirability, a score of 1 is assigned.

^b *No matter who I'm talking to, I'm always a good listener. (True)*

^c *At times I have really insisted on having things my own way. (False)*

^d *I have never been irked when people expressed ideas very different from my own. (True)*

Table 11: Correlation: Managers' Social Desirability and Views on Women's Ability

Dependent Variables	Coeff	R-squared	Obs	D.V. Mean (%)
Female always outperform male.	-0.235* (0.124)	0.016	197	0.28
Female often outperform male.	0.108 (0.125)	0.004	197	0.23
Female are comparable to male.	0.113 (0.127)	0.004	197	0.27
Female seldom outperform male.	0.102 (0.119)	0.005	197	0.15
Female never outperform male.	-0.088 (0.088)	0.007	197	0.07
Female are at least as good if not better.	-0.014 (0.138)	0.000	197	0.78

Robust standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The dependent variables are top managers' binary responses to the question, "With which of the statements do you most agree, with respect to non-finishing production positions? 1) Female always outperform male. 2) Female often outperform male.; 3) Female are comparable to male. 4) Female seldom outperform male.; and 5) Female never outperform male. The last dependent variable is the combination of responses that agree with statement 1 to 3.

The independent variable, social desirability index is the sum of the scores on the three social desirability questions below divided by three. Thus, the index ranges from 0 to 1 with 1 indicating the greatest sign of social desirability. The questions and scoring method are based on Crowne and Marlowe (1960). In each question, a top manager was asked if each statement (shown below) was true or false. If the top manager's response on a particular question matches the one determined by the authors to show social disirability, a score of 1 is assigned.

Statements: 1) *No matter who I'm talking to, I'm always a good listener. (True)*; 2) *At times I have really insisted on having things my own way. (False)*; and 3) *I have never been irked when people expressed ideas very different from my own. (True)*

Table 12: Comparing Managers' Lowest Rates for the Male Factory Worker

	Mean	Median	SD	Min	Max	N
<i>Manager's decision + Others can observe</i>						
With the female helper	34.30	31	20.23	5	100	40
With the Male helper	34.92	34	20.94	5	100	40
<i>Research team's decision + Others can observe</i>						
With the female helper	34.05	34	19.81	5	100	40
With the Male helper	34.92	35	20.82	5	100	40
<i>Manager's decision + Others can't observe</i>						
With the female helper	34.55	31	20.46	5	100	40
With the Male helper	35.25	35	20.85	5	100	40
<i>Research team's decision + Others can't observe</i>						
With the female helper	35.00	32	21.27	5	100	40
With the Male helper	35.12	35	20.63	5	100	40

Note: Managers' lowest rates for the male factory worker are managers' responses to the question, *What is the lowest rate at which you are willing for your male factory worker to work with the fe/male helper?* under four hypothetical situations where a hypothetical male factory worker who works for the manager would sort marbles with the fe/male helper. The situations vary in two dimensions: 1) if the decision to sort marbles with the helper has been made by the manager or the research team; and 2) if other people can observe the male worker is working with the helper.

Table 13: Comparing Hypothetical Male Factory Worker’s Lowest Rates Guessed by Managers

	Mean	Median	SD	Min	Max	N
<i>Manager’s decision + Others can observe</i>						
With the female helper	32.08	30	19.84	5	100	40
With the Male helper	32.62	30	20.29	5	100	40
<i>Research team’s decision + Others can observe</i>						
With the female helper	31.25	30	19.04	5	100	40
With the Male helper	32.83	32	20.20	5	100	40
<i>Manager’s decision + Others can’t observe</i>						
With the female helper	31.75	30	19.77	5	100	40
With the Male helper	33.10	34	20.24	5	100	40
<i>Research team’s decision + Others can’t observe</i>						
With the female helper	31.88	30	19.83	5	100	40
With the Male helper	32.88	34	20.22	5	100	40

Note: The hypothetical male factory worker’s lowest rates guessed by managers are managers’ responses to the question, *What do you think is the lowest rate at which your male factory worker will be willing to work with the fe/male helper?* under four hypothetical situations where a hypothetical male factory worker who works for the manager would sort marbles with the fe/male helper. The situations vary in two dimensions: 1) if the decision to sort marbles with the helper has been made by the manager or the research team; and 2) if other people can observe the male worker is working with the helper.

Table 14: Managers' Worries for Misbehavior by the Hypothetical Male Worker toward the Helper

	Mean	Median	Min	Max	N
<i>Manager's decision + Others can observe</i>					
With the female helper	0.00	0	0	0	40
With the male helper	0.03	0	0	1	40
<i>Research team's decision + Others can observe</i>					
With the female helper	0.00	0	0	0	40
With the male helper	0.05	0	0	1	40
<i>Manager's decision + Others can't observe</i>					
With the female helper	0.07	0	0	1	40
With the male helper	0.05	0	0	1	40
<i>Research team's decision + Others can't observe</i>					
With the female helper	0.07	0	0	1	40
With the male helper	0.03	0	0	1	40

Note: Manager's worries for misbehavior by the hypothetical male worker toward the fe/male helper are managers' responses to the question, *Do you worry if your male factory worker misbehaves toward the fe/male helper?* under four hypothetical situations where a hypothetical male factory worker who works for the manager would sort marbles with the fe/male helper. The situations vary in two dimensions: 1) if the decision to sort marbles with the helper has been made by the manager or the research team; and 2) if other people can observe the male worker is working with the helper.

Table 15: Beliefs about Gender and Religion in Punjab, Pakistan

	Male		Female	
	Share	Obs.	Share	Obs.
<i>Beating wife justifiable</i>				
2012	.18	299	.13	305
2018	.21	569	.23	570
<i>Men make better politicians than women</i>				
1997	.72	372	.52	361
2001	.54	508	.58	652
2012	.74	299	.67	305
2018	.85	569	.78	570
<i>Coed classes are violation of Islam</i>				
2001	.58	508	.53	652
<i>Islam requires women dress modestly but not veil faces</i>				
2001	.5	508	.46	652
<i>Wife earning more than husband is problematic</i>				
1997	.63	372	.38	361
2012	.4	299	.31	305
2018	.79	569	.77	570
<i>Wife and husband should both contribute to income</i>				
1997	.44	372	.75	361
2001	.65	508	.66	652
<i>Men deserve jobs more than women when jobs are scarce</i>				
1997	.8	372	.73	361
2001	.64	508	.65	652
2012	.76	299	.69	305
2018	.9	569	.9	570
<i>Wife must obey</i>				
2001	.46	508	.44	652
<i>Religion is important in life</i>				
1997	.98	372	.89	361
2001	.92	508	.92	652
2012	.99	299	.95	305
2018	.99	569	.98	570
<i>Consider self as religious person</i>				
2001	.79	508	.81	652
2012	.99	299	1	305
2018	.94	569	.98	570

Data source: World Value Survey 3, 4, 6, and 7