Missing Women, Integration Costs, and Big Push Policies in the Saudi Labor Market

Conrad Miller[1], Jennifer Peck[2], and Mehmet Seflek[*1]

[1]Haas School of Business, UC Berkeley
[2]Department of Economics, Swarthmore College

July 2020

Abstract

We study firm demand for female labor in Saudi Arabia, where norms favor gender segregation and make employing both men and women costly to firms. Motivated by a simple model of firm hiring with fixed integration costs, we develop a methodology that uses the distribution of female employment across firms to estimate counterfactual female employment at all-male firms where integration costs bind. We validate our approach using policy variation from Nitaqat, a gender-neutral quota program that incentivized the hiring of Saudi nationals at private sector firms. We argue that Nitaqat dramatically increased Saudi female employment by inducing firms to integrate. *JEL* Codes: J16, J23, J71, O53.

1   Introduction

Women's employment rates are particularly low in the Middle East and North Africa (MENA) and South Asia.[1] These low employment rates are attributed in part to social norms regarding gender roles and their effects on labor supply decisions (e.g., Fernandez and Fogli, 2009; Alesina et al., 2013). In these regions, strong preferences for gender segregation are common (Jayachandran, 2015). Saudi Arabia is a particularly extreme example: women made up only 15.9 percent of the Saudi labor force in 2018 (GaStat, 2011). Saudi cultural norms around gender segregation are particularly strict and, until recently, labor regulations explicitly mandated gender-segregated workplace facilities. Increasing women's employment, however, has recently been a key economic and social goal for the Kingdom: since 2011, a series of ambitious labor reforms has corresponded with rapid growth of women's employment in the Saudi private sector. Some of these policies have been consequential for women's employment even when not gender-specific. The most significant of these reforms was Nitaqat, a gender-neutral quota program designed to increase number of Saudi nationals working in the private sector. The program, launched in 2011, has corresponded with a dramatic increase in the number of Saudi women working in the private sector: from just 56,000 in 2010 to 606,000 by the end of 2017, increasing the female share of the Saudi private sector workforce from 8% to 32%.

We argue that this dramatic transformation occurred in part because gender-neutral Nitaqat quotas led firms to overcome firm-level barriers to gender integration, where integration is defined here as employing both men and women. These barriers are the largely fixed costs of accommodating social norms and regulations that require a physically and socially segregated workplace. Workers, customers, and regulators may expect firms to establish gender-segregated facilities, including restrooms, entrances, and workspaces. Firms may also segregate tasks or teams to limit interactions between male and female employees. For male-dominated firms, hiring women may necessitate changes in their workplace culture. These social expectations constrain the production process for integrated firms. We study the consequences of these integration costs for Saudi women's employment and argue that Nitaqat pushed firms to integrate by incentivizing them to hire more Saudis, reducing the per-worker burden of the fixed costs. Paying these fixed integration costs may be difficult to justify when hiring only a small number of workers, but become less onerous when hiring a large number of Saudis to meet hiring quotas.

The distribution of women's employment across firms provides prima facie evidence that integration is costly. In January 2009, our first month of data, 73% of private sector firms with at least 5 Saudi employees employ no Saudi women.[2] This is substantially larger than what one would expect by chance, even with the low female share of employment.[3] This strongly suggests that

---

[1]The female share of the labor force ranges from a high of 46.4% in sub-Saharan Africa to 20.3% and 24% in the Middle East and North Africa (MENA) and South Asia (World Bank, 2018b).

[2]We describe these data and sample restrictions we apply in more detail in Section 3.4.

[3]In January 2009, 8% of Saudis in the private sector are women. If the gender of each employee were independent draws from a binomial distribution where the probability an employee is female is 0.08, the share of all-male firms that would occur by random chance is 34%.

firms face an extensive margin decision of whether to integrate their workforce.

Motivated by a simple model of firm hiring, we develop a methodology that uses the distribution of women's employment across firms to assess whether and how integration costs constrain women's employment at firms. We apply and validate the methodology using administrative data on Saudi citizens' employment in the private sector. We find that the majority of Saudi firms employ only men because they face binding integration costs. We then show that Nitaqat led to a dramatic increase in female share of the Saudi workforce at least in part by inducing firms to integrate.

We first build a partial equilibrium model of firm hiring based on Kuhn and Shen (2013). A firm posts an exogenously determined number of vacancies and receives a random draw of candidates for each vacancy. The firm hires their most preferred candidate for each vacancy from the candidate pools: they can choose to hire from a pool of only male candidates or both male and female candidates for all vacancies. To hire from both pools, the firm must pay a fixed integration cost to accommodate social preferences for gender segregation. In the Saudi context, these costs include establishing legally required gender-segregated facilities, for example. This framework generates a threshold rule: firms pay for the ability to hire both male and female candidates ("ex-ante integrate") if their expected number of female hires under integration is sufficiently large.

Guided by the model, we develop a joint test for whether all firms are ex-ante integrated (i.e., whether integration costs affect hiring at any firm). We assume that, for ex-ante integrated firms, the probability of a female hire for a vacancy i is a function, $\theta(\cdot)$, of observable job characteristics, X_i. We apply our test to social security data covering Saudi nationals in the private sector from 2009 to 2015.[4] In our data, X_i includes occupation, industry, and the location of the job. To test the null hypothesis that all firms are ex-ante integrated, we estimate $\theta(X_i)$ using data on employees at all firms, simulate the distribution of women's employment across firms using this estimate, and compare the simulated and observed distributions. Under the null, we would expect some firms to have zero female employees by chance alone. However, if integration costs bind for *some* firms, we show that we should see an excess mass or "bunching" of firms with zero female employees. We find exactly this pattern and reject the null hypothesis of no binding integration costs. In January 2009, our first month of data, 8% of Saudis in the private sector are women. We simulate that 43% of firms in our sample should employ only men under the null hypothesis. In practice, 73% of firms in our sample employ only men.

We next estimate $\theta(X_i)$ using data from integrated firms only. For each segregated firm, we use the firm's job mix and our estimate for $\theta(X_i)$ to predict what its female employment would be if it were to integrate, holding the behavior of other firms fixed. Our estimates imply that about 65% of Saudi firms face binding integration costs. We find that ex-ante integration rates are increasing in a firm's expected number of female employees if integrated, consistent with largely fixed integration

---

[4]A limitation of these data is that they do not cover non-Saudi workers, who make up a substantial share of the private sector workforce. In particular, some firms that we identify as all male might in fact employ non-Saudi women. However, non-Saudi women are rare in the private sector, and firms that employ both Saudi nationals (a requirement for inclusion in our data) and non-Saudi women tend to also employ Saudi women. We discuss this issue in more detail in Section 3.4.

costs.

A key concern with our approach is that an estimate of $\theta(X_i)$ based on integrated firms may not produce accurate estimates for counterfactual employment of women at segregated firms. With sufficient unobserved heterogeneity in candidate pools or preferences across firms, the observed distribution of female employment—including a mass of all-male firms—can be rationalized in the absence of any integration costs at all. Moreover, it is not clear that the relationship between a firm's ex-ante integration status and its expected number of female employees under integration is *causal*, as our model implies. We address these concerns using two features of the Saudi data and context.

The first feature is the panel structure of the data. We conduct two tests that exploit this feature. In the first test, we examine firm transitions from segregated to integrated using an event study framework. Our theory predicts "lumpy" transition dynamics; once a firm integrates, it hires women at rates similar to that of incumbent integrated firms with a similar mix of jobs. This prediction is borne out: six months after a newly integrated firm's first female hire, 26% of their hires are women. This matches the female share of hires for similar incumbent integrated firms. Moreover, our estimate of $\theta(X_i)$, derived from incumbent integrated firms, predicts the female share of hires across newly integrated firms with little bias. We also test for *state dependence*, comparing hiring behavior at previously segregated and previously integrated firms. We expect previously integrated firms to tend to remain integrated, either because employee turnover is low, integration costs are sunk, or firm conditions that make integration appealing in the first place are persistent. Consistent with state dependence, we find strong evidence of bunching at previously segregated firms but not at previously integrated firms.

The second feature we use is the exogenous variation in Saudi employment across firms generated by the Nitaqat nationalization quotas. As shown in Peck (2017), firms generally responded to Nitaqat quotas by employing more Saudis. Firms vary in their distance from their quota at baseline, generating exogenous variation in Saudi hiring across firms. We use this variation in hiring incentives to further test the model. We find that firms that are above and below their Nitaqat quotas at the time the policy is implemented have similar observable characteristics and are on similar pre-trends. Following implementation, firms that are below their quotas experience a larger increase in Saudi hiring. Consistent with the model, we also find that (previously all-male) below quota firms integrate at higher rates and have a larger female share of hires than above quota firms. Moreover, the magnitude of the increase in female share of hires is in line with what we would predict based on our estimate of $\theta(X_i)$ derived from incumbent integrated firms. We conclude that integration costs are an important driver of firm behavior in Saudi Arabia and our stylized model fits the data well.

While integration costs reduce women's employment at individual firms, it is not clear what implications integration costs have for female labor market outcomes in the aggregate. As in Becker (1957), integrated firms may be sufficiently numerous or large to absorb female labor so that the existence of constrained male-only firms has no bearing on women's wages and employment.

We interpret the aggregate effects of Nitaqat as evidence that integration costs reduce aggregate women's employment in Saudi Arabia. In particular, the Nitaqat quota policy nearly tripled the female share of Saudis working in the private sector within four years, from 10% in 2011 to 27% in 2015. This increase is concentrated at firms that were previously all-male and were induced to integrate by the policy. This occurs despite a *decrease* in the gender wage gap over this period, which is in part driven by the introduction of a de facto minimum wage for Saudi workers in the private sector. Together, these findings suggest that Nitaqat increased relative aggregate demand for women by inducing more firms to integrate.

The workplace segregation we document is consistent with employees preferring to work with coworkers of the same gender as in Becker (1957). In our framework, firms can pay a fixed cost to accommodate these preferences and employ both men and women, e.g. by establishing gender-segregated facilities or teams. A distinguishing prediction of our model is that ex-ante integration rates are increasing in a firm's total hires. This prediction is borne out in the data. Moreover, following a firm's first female hire, we do not see the wages or separation rates of incumbent male employees increase as the logic of the Becker (1957) model of coworker discrimination would suggest if workplace segregation were driven male preferences for male coworkers. Women's preferences for female coworkers could generate our findings if their utility is highly nonlinear in their number of female coworkers and, for example, women particularly dislike working in firms where they are the sole female employee.[5]

The notion that gender integration involves substantial, largely fixed costs has important implications for policy. In particular, our results suggest that "big push" demand-side policies like Nitaqat that incentivize firms to integrate can substantially change firm hiring preferences at the margin.[6] These policies can also have the potential for feedback effects by attracting more women to the labor market, which could in turn induce more firms to integrate. Though we cannot test this directly here, our results also suggest that one-time incentives to integrate may have long-lasting effects on women's employment. This is because the types of costs we believe are associated with gender integration in this context—physical investment in new or restructured workspaces and facilities, change in organizational structure or culture—have a significant sunk component.[7]

Related Literature.—Methodologically, our approach to inferring integration costs is similar in spirit to bunching estimators (Kleven 2016). Canonical bunching estimators exploit bunching in observed income distributions around discontinuities in tax rates to measure behavioral responses to taxes and transfers. In an application more closely related to ours, Garicano et al. (2016) and Gourio and Roys (2014) examine bunching in the firm size distribution to study the costs of labor regulations that apply to firms above a known size threshold. By contrast, we infer the existence of integration costs based on observed bunching at zero in the distribution of female employment

---

[5]In this case, one can think of the cost of a "cluster hire" as an approximately fixed integration cost.

[6]In the literature on racial discrimination in the US, there is evidence that affirmative action policies and other shocks to minority hiring can have long-term effects on minority employment even after the policies end Miller and Segal, 2012; Miller, 2017; Whatley, 1990.

[7]Policies that increase exposure to integrated workplaces may also change gender attitudes, which could in turn affect integration costs (Dahl et al., 2018).

across firms. While in traditional approaches the presence of bunching is often visually clear from a density plot alone, a key challenge in our setting is that identifying bunching requires a model of counterfactual female employment at firms. Our setting requires this structure because fixed integration costs imply excess mass in the number of firms with zero female employees, a corner solution where we may expect a mass of firms even in the absence of fixed integration costs. Our simulation-based approach to constructing a counterfactual is similar to Ellison and Glaeser (1997) and Augereau et al. (2006). Our methodology can be applied to measure integration costs on the basis of other worker characteristics including disability (Acemoglu and Angrist, 2001), language (Lang, 1986; Hellerstein and Neumark, 2008), ethnicity (Hjort, 2014; Glover et al., 2017), and religion.

We contribute to a large literature on how social and cultural norms affect women's labor market outcomes. This literature primarily focuses on how social norms influence labor supply decisions and how policy interacts with social norms (e.g., Fernandez, 2013; Ashraf et al., 2019). Most closely related is Bursztyn et al. (2018), who study social norms over women's labor supply in Saudi Arabia. They show that husbands underestimate the share of their peers who support wives participating in the labor market, and they provide evidence that correcting those misperceptions increases husbands' willingness to support their wives joining the labor force. By contrast, we focus on how norms constrain labor demand and how firms respond to those constraints.

We also contribute to the literature on workplace segregation and its implications for labor market inequality. As some firms pay more than others, this segregation can have important implications for gender earnings inequality (Groshen, 1991; Bayard et al., 2003; Card et al., 2016). While prior research has shown that skill differences, occupational preferences (Goldin, 1986), and gender-based perceptions of prestige (Pan, 2015) can explain between-establishment segregation to some degree, at least along the intensive margin, there is little research explaining why some firms employ no women at all.

Finally, we build on a literature that studies dynamics and adjustment costs in firm-level labor demand, primarily as an input for understanding macroeconomic fluctuations. A series of papers document that firms tend to change employment in a manner consistent with nonconvex adjustment costs: adjustment tends to be lumpy, with extended periods of inactivity and sharp, large changes (see, e.g., Varejão and Portugal, 2007; Hamermesh and Pfann, 1996). We study a different type of adjustment, moving from an all-male to an integrated workforce, and document that the pattern of adjustment within and across firms is consistent with largely fixed, potentially one-time adjustment costs.

2 A Model of Firm Hiring

In this section we describe a simple, partial equilibrium model of an individual firm's hiring strategy to study the implications of integration costs for women's employment. The model is a modified version of Kuhn and Shen (2013). We assume wages and a firm's candidate pool of potential hires

are fixed. The firm must decide which pool of candidates to hire from.

A firm must fill n vacancies, where n is set exogenously. For each vacancy, the firm receives a fixed number of applications from two types of candidates: type F and type M. Let the net value to the firm of an individual candidate, j, be

$$U_j = v^G + \epsilon_j, \ G \in (M, F),$$

where ϵ_j is an independent draw from a type I extreme value distribution with scale parameter β.[8] β indexes how much candidate quality varies within group. The difference $v^F - v^M$ embodies between-group differences in expected revenue productivity, wage costs, and turnover. This difference may also reflect employer tastes.

The firm will choose the best worker among candidates it can hire. The question is, which candidate pool will it hire from? At no additional cost, the firm can hire from either the type F or type M pool, but not both. To hire from *both* pools for all vacancies, the firm must pay fixed integration cost c, e.g., the cost of establishing gender-segregated facilities. We assume the firm must choose their hiring strategy prior to observing their candidates.

Define U_*^M, U_*^F, and U_*^I as the expected value of the highest U_j value among group M candidates, group F candidates, and the combined pool, respectively. The firm's problem of choosing what pools to hire from is equivalent to choosing the maximum of nU_*^F (only type F), nU_*^M (only type M), and $nU_*^I - c$ (both types).

We first consider the choice between hiring only type M candidates and hiring from both types. The firm will pay the fixed integration cost and hire from both types if

$$U_*^I - U_*^M > \frac{c}{n}. \tag{1}$$

As we show in Appendix B, the left hand side of this expression can be expressed as

$$U_*^I - U_*^M = -\beta \log[1 - \theta]$$
$$\approx \beta\theta. \tag{2}$$

where θ denotes the probability that the firm's preferred candidate from the combined pool is type F. Combining (2) with (1), an approximate condition for the firm to pay the fixed integration cost and hire from the combined pool is

$$n\theta > \frac{c}{\beta}. \tag{3}$$

The left-hand side of (3) is the firm's expected number of type F hires if it were to integrate. Hence, the firm's integration decision follows a threshold rule. If $n\theta$ exceeds integration costs (rescaled by β), the firm integrates and hires from both pools. Intuitively, θ is increasing in female labor supply (δ) and $v^F - v^M$, which embodies net productivity of, and employer tastes for, women

---

[8] The CDF is $F(\epsilon_j) = \exp(\exp(-\epsilon_j/\beta))$. It follows that $Var(\epsilon_j) = \frac{\beta^2 \pi^2}{6}$ and $E(\epsilon_j) = \beta\gamma$, where γ is Euler's constant.

relative to men.

Next we consider the choice of hiring only from the type F candidate pool. Symmetrically, we have $U_*^I - U_*^F = \beta \log[\theta]$. Hence, if $\theta < \frac{1}{2}$, then $U_*^F < U_*^M$ and no firm will hire from only the type F candidate pool. We find below that in contexts where integration costs are relevant, θ is generally below $\frac{1}{2}$. This is consistent with the fact that all-female firms are rare in our data.

The model is one period but can be readily extended to multiple periods. In a dynamic setting, where n or θ is varying over time, we must distinguish between ongoing and one-time sunk integration costs. Integration decisions now depend on the future path of n and θ and whether integration costs are ongoing or one-time sunk costs. For example, if θ is increasing over time, firms have more incentive to wait to integrate if integration costs are on-going rather than one-time costs.

While framed in terms of hiring, the model also has straightforward implications for women's employment. If turnover rates are similar for men and women—as we show they are in the Saudi private sector—then a firm's female share of hires will equal its female share of employees. Otherwise, the female share of employees will equal the duration-weighted female share of hires.

3   Saudi Arabia Context and Data

We apply the model to data to (1) develop a joint test for whether all firms are ex-ante integrated and (2) to estimate the counterfactual employment of women at segregated firms. We apply our methodology using administrative data from Saudi Arabia. In this section, we describe the Saudi context and data.

3.1   Women in the Saudi Workforce

There are several reasons to think that integration costs may be particularly important for Saudi firms. First, Saudi Arabia has extremely low female employment rates by international standards but also has high female unemployment rates. In 2008, before the start of our sample period, the employment rate for women was 8.4%, and for men was 56.8% (World Bank, 2018b); official unemployment rates were 26.9% for women and 6.8% for men (GaStat, 2011). These patterns are even more pronounced in the private sector, as Saudi women have typically relied on the public sector for work.[9] These disparities are not driven by differences in skill: education levels are also comparable for men and women, and women are more educated among private sector workers and the unemployed.[10] Wages also tend to be significantly lower for Saudi women as compared to Saudi men: in January 2009, our first month of administrative data, the average monthly full-time wage for women at baseline is about half the wage for men (Table 2). Even when controls are added for education, location, and occupation, women earn about 40% less than men in January 2009.[11]

---

[9]Even by 2014, women overwhelmingly worked in the public sector, with 74% of employed women working in girls' schools in 2014 (Evidence for Policy Design, 2015).

[10]Unemployed women with college degrees outnumbered men by almost four to one in 2008.

[11]A Mincer regression of the log of private sector wages at baseline on employee characteristics indicates that Saudi women earn 40% less than men within occupations after controlling for educational attainment, years of potential

The Saudi private sector is composed primarily of male, non-Saudi expatriate workers (see Tables 1 and 2). Among Saudis, women form a small share of private sector employment, with Saudi women making up 8.5% of Saudi employees in January 2009. Low employment of women in the private sector is likely attributable to a variety of factors on both sides of the market. Female employment in the public sector in part likely reflects women's work preferences: jobs in education are widely seen as culturally appropriate for women, and completely segregated gender environments are also seen as highly desirable (Evidence for Policy Design, 2015). As we will argue, low female employment in the private sector also reflects significant additional firm-level costs to employing women. At the same time, women's employment has become a priority for the Saudi government. The Kingdom's Vision 2030 economic strategy has an explicit goal of increasing women's labor force participation to 30% by 2030.

[Table 1 about here.]

[Table 2 about here.]

3.2 Firm-Level Costs of Employing Women in Saudi Arabia

There are a variety of features of the Saudi labor market that may create additional costs for firms as they begin to hire women. Many of these costs are fixed in the sense that they do not depend on the number of female workers that firms employ. These include one-time switching costs as well as ongoing costs that apply to integrated firms. Firms may also face differential per worker, or variable, costs in employing women instead of men.[12]

In particular, it may have been costly for firms to comply with government regulations regarding gender segregation in the workplace. During the study period, the government required that a firm employing women (whether Saudi or non-Saudi) provide them separate workstations, a private space to pray and take breaks, convenient restroom access, and a separate entrance to the building or workplace. Meeting rooms also had to be adjusted to accommodate mixed-gender meetings: firms were initially required to hold these meetings only in private and later to make them fully visible to the rest of the office. Employing women exposes firms to inspections and potential fines through the Ministry of Labor and Social Development (MLSD) and the Ministry of Municipal and Rural Affairs (Khoja and Thomas, 2018). In addition to the explicit integration costs associated with making a workplace compliant with segregation regulations, the cost of learning how to comply with these rules may also present a barrier to hiring women.

Even outside of legal requirements, Saudi firms may incur similar costs to accommodate social preferences for gender segregation. Moreover, integration may be costly if employees prefer to work

---

experience, and location (all with indicator variables).

[12]These integration costs are sometimes explicitly cited when discussing obstacles to women's employment. One business owner told the *New York Times*, "If they hire women to work, they need another office, with electricity, a dedicated security guard, computers... This is a major cost, especially for small, local companies." (*New York Times*, 2012) Lubna Olayan, a female Saudi CEO, describes integration obstacles, such as difficulties navigating labor law and social customs, when providing the required segregation for her company's male and female employees (*Fortune*, 2015).

with coworkers of the same gender regardless of accommodations as in Becker (1957). In that case, integrated firms may have to increase compensation to offset the disutility of working with members of a different gender. A key prediction of the model outlined in Section 2 that is not present in the Becker (1957) model of coworker discrimination is that ex-ante integration rates are increasing in a firm's total hires. Moreover, if segregation is driven by male preferences for male coworkers, we should see the wages or separation rates of incumbent male employees increase following a firm's first female hire. We test this prediction is Section 4.3. Women's preferences for female coworkers could generate predictions similar to our model if their utility is highly nonlinear in their number of female coworkers. For example, if women particularly dislike working in firms where they are the sole female employee, firms may integrate by hiring a 'cluster' of women.

Low historical female employment may also lead to high search costs on both sides of the market: firms may have limited access to hiring and referral networks with female employees, and women may have little information about opportunities for private sector employment. Furthermore, Saudi firms must also develop a strategy for navigating the relationship with male guardians: this is no longer explicitly required by the government, but many firms do ask for guardian permission when recruiting female workers.[13] More broadly, firms may also need to develop different types of HR policies to attract and to retain female employees, such as offering parental leave, facilitating childcare, and addressing workplace harassment.[14,15] Addressing these HR issues involves learning by doing, and these costs will be higher for firms that have never recruited women than for firms that already have female employees.

Firms may also need to restructure their task allocations or working hours to accommodate female employees. This type of reassessment can similarly present a one-time hurdle to overcome before hiring women. For example, firms may have a narrow view of the qualifications they require (e.g., certain types of degrees)[16] or years of experience, which disqualify many female applicants. Overcoming these barriers may require firms to think flexibly about how they structure their tasks across occupations within the firm. This might include restructuring shifts and working hours, as Saudi Arabia is among the 44 countries that restrict the working hours of women. Firms may also face costs due to the lower mobility of female employees.[17] Some firms address this by providing group transportation for their employees, a lumpy, ongoing cost.

---

[13] The guardianship requirement was lifted by the Ministry of Labor in 2008. There are still 18 countries where women must have a (male) guardian's permission to get a job (World Bank, 2018a).

[14] Some of these adjustments are mandated for firms above a particular size: Saudi labor law requires firms that employ more than 50 women with at least ten children under age six must provide childcare access, and firms with more than 100 women must provide a childcare center.

[15] Addressing workplace harassment is an ongoing issue in labor markets with high levels of gender integration as well as an important barrier to further integration in markets with low integration. The current literature suggests that there is significant firm-level heterogeneity in the prevalence of harassment, and firms can take steps to create an organizational culture that prevents harassment (Willness et al., 2007; Hart et al., 2018). Developing the procedures for reporting and responding to complaints and creating credible messaging from leadership about organizational values can be thought of as an investment in both prevention and firm reputation. This likely affects recruiting as well as retention.

[16] Engineering, for example, was not offered to Saudi women as an undergraduate degree program until 2005.

[17] Women were not permitted to drive in Saudi Arabia until June 2018.

3.3 Nitaqat Nationalization Quotas

We analyze the Nitaqat quota policy, which generated exogenous variation in Saudi employment across firms, through the lens of the model. The Nitaqat program is an ongoing gender-neutral nationalization quota policy first instituted in 2011.[18,19] The policy was designed to address growing national unemployment, which in 2011 had reached 40% for Saudis in the 20–25 age group, in the context of the low participation of nationals in the private sector. At the time, foreign guest workers made up 90% of non-oil private sector employment, with the majority of Saudis employed in the public sector. Under Nitaqat, the Saudi government began requiring private sector firms to attain set nationalization quotas for their employees. These quotas varied by firm industry and size and assigned firms to four color bands according to their level of compliance: firms in the Green and Platinum bands were in-compliance with nationalization quotas, while firms in the Red and Yellow bands were required to increase their share of Saudi workers.[20] For our purposes, we categorize firms by their Nitaqat status at the start of the program: "above quota" firms are those in the Green and Platinum bands in July 2011, and "below quota" are those in the Red and Yellow bands. Compliance was monitored using a system integration social security data for Saudi workers and visas for expatriate workers from the National Information center. Firms in the Red and Yellow bands faced restrictions on their ability to renew existing visas, obtain new visas, and access the MLSD's foreign recruitment services; Green and Platinum firms were given access to a streamlined visa renewal service. The MLSD first announced plans for Nitaqat in early 2011, with detailed information about the program structure, targets, and penalties released to firms in June 2011. Sanctions for noncompliance were phased in, starting just three months later in September 2011.

Overall, Nitaqat quotas were effective at increasing Saudi employment in the private sector (Peck, 2017), with firms complying with the program by increasing their Saudi employment. Nitaqat quotas also served as a way for the government to introduce a de facto minimum wage for Saudis. In September 2012 the government announced that only Saudis paid at least 3,000 SAR per month would count as a full Saudi employee; those paid 1,500 SAR would count as half an employee for Nitaqat purposes, and those between 1,500 and 3,000 SAR would be linearly prorated. This restriction was applied to firms beginning in February 2013.[21]

3.4 Data

We test for integration costs in the Saudi context using administrative social security data from the General Organization for Social Insurance (GOSI). These data contain information on all Saudis employed in the private sector between January 2009 and June 2015. They record worker characteristics such as gender, age, education level, and marital status; job characteristics such as occupation, work location, full-time status, and wages; and firm information such as administrative

---

[18]See Peck (2017) for a more detailed description of the Nitaqat program and its effects.

[19]Other contemporary labor policies are described in Appendix C.

[20]Cutoffs for each band were set based on pre-Nitaqat Saudization rates so that slightly less than half of firms in each industry by size group would be classified as Green or Platinum.

[21]*Saudi Gazette*, September 9, 2012, "Nitaqat percentage linked to salary."

identifiers and industries. While we cannot identify establishments in the data, the definition of the firm we use in this paper can be thought to be a legal commercial organization within a particular province or major city. Our definition of firms is described in more detail in Appendix D. In total, the GOSI data set contains information on approximately 2.8 million unique individuals and 430,000 firms. We restructure this dataset into an unbalanced monthly panel for each full-time Saudi employee and standardize the occupations using two-digit codes from the International Labour Organization's ISCO-08 codes.

We also test our model using firm responses to Nitaqat Saudi employment quotas. We use the Nitaqat data to obtain a list of firms and their quota compliance status for the second week of June 2011, when the program began assessing quotas and began reporting status to firms. This gives us a sample of approximately 1.07 million firms at our baseline, over 990,000 of which were originally exempt from the program for having fewer than ten employees. Approximately 113,000 of these firms appear in the GOSI data.[22] The details of merging the two data sets are described in more detail in Appendix D.

There is a potential concern that GOSI records may not accurately reflect real employment if firms falsify their employee records with GOSI to meet their Nitaqat quotas. This may be a particular concern for female employment if firms are more likely to fraudulently register women's ID numbers. We discuss this possibility in Appendix E by examining the share of workers in the GOSI data with "active" subsequent career trajectories by month of hire. We find that women hired after Nitaqat are no less likely to have active careers than those hired in the pre-period, particularly when compared to men and when controlling for observable worker characteristics.

Unfortunately, our GOSI data do not include information on non-Saudi workers, so our references to the composition of workers throughout the paper refer only to Saudi employees. This means that firms that we identify as "all-male" may in fact employ non-Saudi women. Other government data suggest that this group is small: of the 11.5 million people employed in 2015 only 773,000, or 7 percent, were non-Saudi women. Of these, 94 percent worked in private households as domestic workers or in public sector jobs in education and health. Less than one percent of workers outside households, health, and education were non-Saudi women (GaStat, 2015). In our analysis we consider a firm segregated if it reports employing Saudi men but not Saudi women in the GOSI dataset. Across 2012 and 2013 there are 2.1 million such firms in the Nitaqat data, only 1.5 percent of which reported employing non-Saudi women. Firms that employ non-Saudi women are also very likely to also employ Saudi women: among all firms that employed both non-Saudi women and Saudis in either 2012 or 2013, 75 percent employed Saudi women. Among firms that employed both non-Saudis (women or otherwise) and Saudis, only 39 percent employed Saudi women. This suggests that the integration costs associated with employing either Saudi or non-Saudi women

---

[22]The big drop in the number of baseline firms between the two data sets is primarily due to the fact that many firms in the white color band do not need to hire any Saudi employees, and therefore they do not appear in the GOSI data since it only contains information on firms that have hired at least one Saudi between 2009 to 2015. Additionally, some firms exit the market before hiring any Saudis, as Peck (2017) documents, so they again would not appear in our GOSI data.

likely have substantial overlap.

4 Empirical Strategy and Results

The central ideas of the model are (1) firms face an extensive margin integration decision and (2) integration costs are largely fixed, so firms integrate only if they anticipate employing enough women to justify the costs. To take the model to the data, the central assumption we make is that the probability that the top candidate for position i is female is a function, $\theta(\cdot)$, of *observable* job characteristics, X_i. We assume other factors that determine this probability are *uncorrelated* with the identity of the firm.

Building on this assumption, we develop a joint test for whether integration costs are nonbinding at all firms so that all firms are ex-ante integrated. We then show how to estimate $\theta(X_i)$ when some firms are ex-ante segregated and use this estimate to measure ex-ante integration rates as a function of expected female employment under integration. We also test our central assumption in several ways described below.

4.1 Testing the Null of No Binding Integration Costs

We first test the null hypothesis that no firm faces binding integration costs and all firms are ex-ante integrated.[23] Under the null, the distribution of female hires across firms should be consistent with $\theta(X_i)$, the probability that the top candidate for position i is female given job characteristics X_i. In other words, conditional on job characteristics, different firms should hire women at similar rates, and any variation across firms is due to chance alone. Our procedure for testing the null hypothesis is as follows: (1) estimate $\theta(X_i)$, (2) simulate the implied distribution of female hires across firms, and (3) compare that to the distribution we observe in practice. We describe each step in more detail below.[24]

While the model is framed in terms of firm hiring, the test we first develop uses cross-sectional, firm-level data on women's *employment*. If turnover rates are similar for men and women—as they are in Saudi Arabia—then a firm's female share of hires will equal its female share of employees.[25] Otherwise, the female share of employees will equal the duration-weighted female share of hires. We conduct a similar test in Section 4.3 that examines hiring rather than employment.

4.1.1 Estimating $\theta(X_i)$

We estimate $\theta(X_i)$ with a job-level regression model using jobs at all firms meeting our sample criteria. We use cross-sectional data from January 2009, the first month of our data. We limit to

---

[23]A special case would be that integration costs are zero (i.e., do not exist).

[24]By contrast, if some firms *do* face binding integration costs, we show in Appendix F that the simulation will generally *underpredict* the number of firms with zero female hires. The intuition is that when some firms are in fact ex-ante segregated, female hires are more concentrated across firms than the simulation predicts.

[25]In Saudi Arabia, turnover rates are similar for Saudi men and women. The monthly turnover rates in the GOSI data are 3.5 and 4.2 percentage points for men and women. Adjusting for job characteristics (occupation, industry, and location) and month, turnover rates are 5% lower for women.

firms with at least five Saudi employees to reduce the degree of "chance" segregation.[26] While firms with fewer than five Saudi employees account for the majority of firms, they account for less than 10% of Saudi private sector employment. Table 3 compares the characteristics of all firms and those with at least five Saudi employees. The two sets of firms have comparable industry compositions.

[Table 3 about here.]

We estimate a logistic regression model of the form

$$P(\text{Worker } i \text{ is female}) = \Lambda(X_i\beta),$$

where X_i includes fixed effects for job location, two-digit occupation, and one-digit industry. We label the function we estimate as $\hat{\theta}^0(X_i)$.

In Table 4 we summarize $\hat{\theta}^0(X_i)$ across all jobs and the explanatory power of location, occupation, and industry fixed effects for these estimates. The mean is 0.08, the median is 0.027, and the standard deviation is 0.155. Across one-digit occupations, $\hat{\theta}^0(X_i)$ is largest among professionals at 0.23 and lowest among plant and machine operators at 0.007. Across industries, $\hat{\theta}^0(X_i)$ is largest in community and social services at 0.43 and lowest in electricity, gas, and water at 0.008. In separate linear regression models, occupation and industry explain 73% and 62% of the variance in $\hat{\theta}^0(X_i)$, while location explains only 6%. $\hat{\theta}^0(X_i)$ explains 31% of variation in worker gender across positions.

[Table 4 about here.]

4.1.2 Simulation Results

Next, we simulate the distribution of women's employment across firms using our estimate, $\hat{\theta}^0(X_i)$, and compare the result to the distribution we observe.[27]

We plot the simulated and observed distributions of women's employment in Figure 1. We plot the share of firms with zero female employees separately due to the difference in scale. We also plot the share of firms with 1, 2, 3, 4, 5, 6–10, 11–25, and > 25 female employees. The error bars represent the 5th and 95th percentiles across simulations for the share of firms with a given number of female employees.

[Figure 1 about here.]

We substantially underpredict the number of firms with zero female employees. While we predict that 43% of firms will have zero female employees, on average, across simulations, in fact, 73% of firms have zero female employees. We also overpredict the number of firms with few female

---

[26]In other words, this restriction reduces the number of firms that are potentially ex-ante integrated but ex-post segregated.

[27]In each simulation, we take a random draw from a uniform distribution for each position i. If that draw is below $\hat{\theta}^0(X_i)$, the worker in that position is labeled as female; if not, the worker is labeled as male. We then sum up to the firm level to get the simulated total of female employees at each firm. We repeat this procedure 1,000 times.

employees, particularly in the one to four range. For all simulations, we reject equality of the distributions in a Kolmogorov-Smirnov test at the 1% significance level.[28] Overall, the pattern is consistent with binding integration costs at many firms.

However, our simulated distribution may also fail to match the observed distribution because we have misspecified $\theta(X_i)$. Relative to traditional bunching approaches, misspecification is a particularly important concern in our context because, even in the absence of fixed integration costs, we expect a mass of all-male firms. Hence, the presence of a mass of all-male firms alone does not imply an 'excess' mass. $\theta(X_i)$ is misspecified if there may be job characteristics that are not included in X_i that (1) help to explain the probability that the top candidate for a position is female and (2) vary systematically across firms, conditional on X_i.

One concern is that the occupation and industry classifications in our data may be too coarse, as there may be systematic variation in gender composition between subcategories of jobs. For example, for the same occupation, commerce firms that sell men's clothing may skew male compared to commerce firms that sell women's clothing. Under this misspecification, our simulation may underpredict the number of firms with zero female employees, not because some firms have not ex-ante integrated but because some firms in fact have smaller θ values than we estimated. In other words, we may underestimate the number of firms that are all male simply because those firms employ workers in job types that few women work in.

There is reason to think misspecification is not a first-order issue. Generating the number of all-male firms we observe would require a substantial role for unobservable job characteristics relative to observable characteristics in determining θ. In Panel A of Appendix Table A.1 we compare the observed distribution to the simulated distribution for several specifications of $\theta(X_i)$, where we vary the set of job characteristics we include in X_i. If we had estimated θ using no covariates so that $\hat{\theta}^0 = 0.08$ for all jobs, we would simulate that 34% of firms would be all male. While adding controls for job location makes little difference, adding controls for 2-digit occupation increases this value to 39%. Adding 1-digit industry increases this value further to 43%. This value is unchanged if we replace occupation and industry fixed effects with 1-digit occupation by 1-digit industry interactions. Hence, while the simulated percentage of firms that are all male depends somewhat on the job characteristics we include in constructing $\theta(X_i)$, including all of our observable job characteristics brings the value nowhere close to the observed value, 73%. Unobservable job characteristics would need to be highly influential relative to observable job characteristics to match the distribution of female employment in the data.[29]

We conduct more direct tests of our specification of $\theta(X_i)$ in Sections 4.3 and 4.4.

4.2   Estimating $\theta(X_i)$ When Integration Costs Bind

If some firms are ex-ante segregated, $\hat{\theta}^0(X_i)$ will underestimate $\theta(X_i)$ because we include these firms in its estimation. To correctly estimate $\theta(X_i)$, we must limit the data to ex-ante integrated

---

[28]We compare the distributions of the number of female employees, not the binned data presented in Figure 1.

[29]The reasoning behind this argument is similar to that of Altonji et al. (2005) and Oster (2019).

firms. A key problem with executing this is that we *do not observe* whether firms are ex-ante integrated. Instead, we observe whether they are "*ex-post*" integrated—whether they employ both men and women in practice. We take two approaches to address this issue. First, we ignore the distinction and limit the data to ex-post integrated firms. This will lead to an upward bias in our estimate for $\theta(X_i)$. The bias is small in practice because the distinction between ex-ante and ex-post integrated firms is only relevant for small firms, which account for a small share of employment. Second, we estimate a more parametric model that accommodates ex-ante integration as a potentially unobserved firm state.

We first limit the data to ex-post integrated firms when estimating $\theta(X_i)$. Table 3 summarizes the characteristics of ex-post integrated firms, using the same sample restrictions described in Section 4.1.1. Overall, the female share of employment is 8.2%, while the female share of employment at integrated firms is 12.5%. We label our function estimated using only ex-post integrated as $\hat{\theta}^{EP}(X_i)$. Column (2) of Table 4 summarizes $\hat{\theta}^{EP}(X_i)$ for *all* jobs, not just those at ex-post integrated firms. In separate linear regression models, occupation and industry explain 69% and 60% of the variance in $\theta(X_i)$, while location explains only 9%.

The second approach we take is to directly model the distinction between ex-ante and ex-post integrated firms and to structurally estimate $\theta(X_i)$. We jointly estimate $\theta(X_i)$ and the probability that a firm is ex-ante integrated as a function of firm characteristics. Estimation details are provided in Appendix G. Column (3) of Table 4 summarizes the estimates and how they vary across jobs. The average value of $\hat{\theta}^S(X_i)$ is 0.123. These estimates are similar to those using only ex-post integrated firms; the correlation between $\hat{\theta}^S(X_i)$ and $\hat{\theta}^{EP}(X_i)$ is 0.82. The average value of π_j is 0.65, indicating 65% of firms are ex-ante segregated.

4.2.1 Applying Estimates of $\theta(X_i)$

We next examine how integration rates relate to a firm's expected number of female employees if ex-ante integrated. Equipped with an estimate of $\theta(X_i)$, we can use the following to estimate counterfactual female employment for firms that did not integrate:

$$\sum_{i \in \text{ firm } j} \theta(X_i) n_{ij} = \bar{\theta}_j n_j,$$

where n_{ij} is the number of type i jobs at firm j, n_j is the number of jobs at firm j, and $\bar{\theta}_j$ is average value of $\theta(X_i)$ at firm j given its job composition. In other words, once we know the probability that the top candidate for a given job is female, we can predict female employment for each firm under integration given its job mix.[30] We can also test whether firm ex-ante integration rates are increasing in $\bar{\theta}_j n_j$, as our assumption that integration costs are largely fixed would suggest.

Panel A of Figure 2 plots ex-post integration rates as a function of $\bar{\theta}_j n_j$. It also plots simulated ex-post integration rates under the counterfactual that all firms are ex-ante integrated. The distinction between ex-ante and ex-post integration is that firms that pay their integration costs may

---

[30]For this construction, we assume that a firm's job mix does not depend on its integration status.

still hire only men by chance alone. The simulated ex-post integration rate in Panel A of Figure 2 estimates this chance factor.

[Figure 2 about here.]

We use the relationships illustrated in Panel A of Figure 2 to estimate *ex-ante* integration rates as a function of $\bar{\theta}_j n_j$. The ratio of the actual and simulated ex-post integration rates provides an estimate of the ex-ante integration rate, as a function of $\bar{\theta}_j n_j$.[31] Panel B of Figure 2 plots this estimate of ex-ante integration rates as a function of $\bar{\theta}_j^{EP} n_j$. We also plot our structural estimates of ex-ante integration rates, π_j, as a function of $\bar{\theta}_j^S n_j$, where $\bar{\theta}_j^S$ is the estimate for $\bar{\theta}_j$ constructed using $\hat{\theta}^S(X_i)$. For both estimates, we find that ex-ante integration rates are increasing in $\bar{\theta}_j n_j$. This pattern is consistent with firms facing an integration threshold rule with respect to $\bar{\theta}_j n_j$.

As an additional test for whether $\theta(X_i)$ is well specified, we evaluate whether a simulation of the distribution of female employment across firms that allows for integration rates to vary by $\bar{\theta}_j n_j$ fits the observed distribution. These simulations are described in more detail in Appendix G. Panel B of Appendix Table A.1 compares the simulated distribution of female employment to the observed distribution for various specifications of $\theta(X_i)$. Our baseline (and most saturated) specification, where X_i includes job location, two-digit occupation, and one-digit industry fixed effects, matches the observed distribution. Across all simulations, we fail to reject equality of the simulated and observed distributions in a Kolmogorov-Smirnov test at the 1% significance level. The average p-value is 0.75. This suggests that we have included the most relevant job characteristics in X_i (or that other relevant characteristics are not concentrated within firms) and have mapped them appropriately to hiring probabilities, at least among ex-ante integrated firms.

4.3 Using Panel Structure to Validate the Model

The panel structure of the Saudi data permits additional tests of the model. In particular, panel data allow us to further probe our key assumption that $\theta(X_i)$ dictates counterfactual employment of women for ex-ante segregated firms. In Appendix G we use the panel structure to conduct two tests. First, we test whether our estimate of $\theta(X_i)$ provides unbiased predictions for the female share of hires at *newly integrated* firms. With panel data, we can observe firm transitions from segregated to integrated. The model predicts an *extensive margin* adjustment: abrupt changes occur in the gender composition of hires for these firms as they move from hiring no women to

---

[31]To see how, let I_j be an indicator for whether firm j has ex-ante integrated. Then

$$P(K_j > 0) = P(K_j > 0 | I_j = 1) \times P(I_j = 1),$$

where $P(K_j > 0)$ is the probability that firm j is *ex-post* integrated. Grouping firms by their value of $\bar{\theta}_j n_j$, we get

$$P(K_j > 0 | \bar{\theta}_j n_j) = E[P(K_j > 0 | I_j = 1) \times P(I_j = 1) | \bar{\theta}_j n_j]$$
$$\approx P(K_j > 0 | I_j = 1, \bar{\theta}_j n_j) \times P(I_j = 1 | \bar{\theta}_j n_j),$$

where the approximation holds because conditional on $\bar{\theta}_j n_j$, $P(K_j > 0 | I_j = 1)$ varies little across firms.

The two relationships depicted in Panel A of Figure 2 correspond to $P(K_j > 0 | \bar{\theta}_j^{EP} n_j)$ and $P(K_j > 0 | I_j = 1, \bar{\theta}_j^{EP} n_j)$, where $\bar{\theta}_j^{EP}$ is the natural estimate for $\bar{\theta}_j$ constructed using $\hat{\theta}^{EP}(X_i)$.

hiring women at a rate dictated by their job composition. By contrast, if the bunching at zero we observe in Figure 1 is driven by unobserved heterogeneity in job characteristics, we would expect these transitions to reflect *intensive margin* changes in θ or chance variation in the candidate pool. In this case, we expect transitions to be smooth and the female share of hires at newly integrated firms to be low relative to observably comparable incumbent integrated firms.

We find that firm transitions from segregated to integrated are consistent with the model. Prior to integrating, we observe transitioning firms in the GOSI data for an average of 39 months. Six months after a newly integrated firm's first female hire, 26% of their hires are women. This share matches the female share of hires for similar incumbent integrated firms. Moreover, our estimate of $\theta(X_i)$, derived from incumbent integrated firms, predicts the female share of hires across newly integrated firms with little bias.

Second, we test for *state dependence*: the hiring behavior of firms that have already paid their integration costs should differ from the behavior of firms that have not. In particular, we should *not* observe bunching for the former set of firms. While we cannot observe each firm's current state, we can proxy for their current state using their baseline ex-post segregation status. This proxy should closely correlate with a firm's current state if integration costs are sunk or if the conditions that led the firm to integrate are highly persistent over time. We test the null hypothesis of no binding integration costs but conduct separate tests for firms that are ex-post integrated and ex-post segregated as of January 2009.

For baseline all-male firms, the contrast between the simulated and observed distributions of female employment is similar to that shown in Figure 1. By contrast, the simulated distribution for baseline integrated firms matches the observed distribution relatively well. Consistent with our interpretation of bunching as evidence for the presence of ex-ante segregated firms, there is little evidence of bunching at firms that are likely ex-ante integrated.

We also examine how incumbent male employees respond to their employer's first female hire. If between-firm gender segregation were purely driven by employee preferences for same-gender coworkers as in Becker (1957), then we should observe an increase in separation rates or in wages for these employees to compensate for the disutility they incur when working in an integrated firm. In Appendix Figure A.1 we plot average monthly separation rates and wages for male employees that joined transitioning firms more than 18 months prior to their first female hire. Neither outcome appears to meaningfully respond to a firm's first female hire. This suggests that gender segregation is not driven by the preferences of male employees, or that firms can accommodate these preferences while employing both men and women, e.g. by establishing gender-segregated facilities or teams.

4.4 Using Policy Variation to Verify the Threshold Rule

We next test whether the Nitaqat employment quotas induce firm integration and increase hiring of women in a manner consistent with the model. In particular, Nitaqat provides a direct test for the model prediction that firm integration decisions follow a threshold rule in $\bar{\theta}_j n_j$. As we will demonstrate, Nitaqat incentivizes some firms to increase their number of Saudi hires (n) and

incentivizes larger increases at some firms than others. An implication of the model is that by increasing n, Nitaqat will induce some firms to integrate and will increase their female share of hires by a magnitude predicted by $\bar{\theta}_j$.

We study the causal effects of Nitaqat using a difference-in-difference research design, comparing above and below quota firms before and after the policy is implemented. We first show that Nitaqat increases Saudi employment at private sector firms, with larger increases at firms that needed to increase their Saudi share of employees to satisfy their quota.

Panel A of Figure 3 plots the average number of Saudi hires each half-year, separately for below and above quota firms. Prior to Nitaqat, hiring at below and above quota firms move in tandem. Following the implementation of Nitaqat, hiring at below quota firms increases sharply relative to above quota firms. A gap of five Saudi hires per half-year emerges in the second half of 2011, which drops to about two hires by the second half of 2012 and stagnates thereafter. We use this policy-induced variation in hiring (n) to test the predictions of the model.

[Figure 3 about here.]

In particular, we test whether Nitaqat increases (1) integration rates and (2) the female share of hires at below quota firms relative to above quota firms. We also evaluate whether the female share of hires increases by a degree consistent with $\theta(X_i)$. For the analysis below, we limit to Above and Below firms that were ex-post segregated and employed at least five Saudis in January 2009. Each plot as an average across firms, where each firm that is present in that period is weighted equally.

Appendix Table A.2 compares descriptive statistics for these two sets of firms as of June 2011. There are 2,224 below quota firms and 1,559 above quota firms satisfying our sample criteria. The firms are generally similar except, as expected, above quota firms have more Saudi employees (an average of 45.5) than below quota firms (33.7). Above quota firms also pay higher average wages.

First, we look at integration rates. Our model predicts that by increasing n at below quota firms relative to above quota firms, Nitaqat will increase relative integration rates at below quota firms too, as more firms will cross their integration threshold. We plot integration rates in Panel B of Figure 3. As the model predicts, the share of below quota firms that are integrated increases relative to the same share of above quota firms following the implementation of Nitaqat. An immediate difference of about 8 percentage points emerges by the second half of 2011. The gap fluctuates between 8 and 14 percentage points thereafter.

Third, we look at the female share of hires. With constant underlying rates of female hiring, Nitaqat could increase integration rates at below quota versus above quota firms by chance alone. By contrast, the model predicts an increase in the female share of hires at below quota versus above quota firms, pooling both firms that do and do not integrate. There is no mechanical reason that an increase in total hires would increase the female share of hires. We plot the female share of hires in Panel B of Figure 3. As the model predicts, the female share of hires at below quota firms increases relative to above quota firms following the implementation of Nitaqat. The magnitude of this relative increase—2–4 percentage points—is in line with what we would predict given the

19

differential in integration rates (about 11 percentage points) and our $\hat{\theta}(X_i)$ estimates from Section 4.3. Averaging across post-Nitaqat hires within firms, the average estimated value of $\bar{\theta}_j$ is 0.25.

Finally, we present corresponding difference-in-difference estimates in table form. We estimate models of the form:

$$Y_{jt} = \alpha_j + \tau_t + \beta \text{Post}_t \times \text{Below}_j + \epsilon_{jt}, \tag{4}$$

where α_j are firm fixed effects, τ_t are half-year fixed effects, Post_t is an indicator for post-Nitaqat implementation, and Below_i is an indicator for a below quota firm. The coefficient β is the post-Nitaqat differential change in the outcome for below quota firms relative to above quota firms. We estimate equation (4) for the same three outcomes: total hires, integration status, and female share of hires.

The DD estimates are overlayed on Figure 3. Over the full period, Saudi hires increase at below quota firms relative to above quota firms by about three per half-year. The integration rate increases by about 11 percentage points, and the female share of hires increases by 2.26 percentage points.

5   Aggregate Effects of Integration Costs: Evidence from Nitaqat

The results above do not consider what would happen in the aggregate if integration costs were eliminated in the labor market. As in Becker (1957), integrated firms may be sufficiently numerous or large to absorb female labor so that the existence of constrained male-only firms has no bearing on women's wages and employment. On the other hand, in the presence of search frictions or insufficient entry or growth of integrated firms, integration costs will reduce aggregate demand for female labor. We discuss these aggregate effects in more detail in Appendix H.

To assess the aggregate consequences of integration costs, we would ideally use exogenous variation in integration costs across labor markets. Lacking such variation, we examine the labor market response to Nitaqat, which induced many firms to integrate and hire women (see Appendix Figure H.1). Following the introduction of Nitaqat in 2011, the female share of the Saudi private sector workforce increased from 10% to 27% in 2015. This increase is concentrated in firms that were previously all-male. We also find that this increase in the female share of Saudis in the private sector is not offset by a decrease in the public sector: the female share of Saudis working in the public sector instead increases over this period, from 33% in 2011 to 40% in 2015.

While this increase is striking, it does not necessarily indicate an important role for integration costs. Nitaqat may also increase the female share of employment through a *price effect*. If the increase in demand for labor bids up wages for males, then we may expect a demand increase for relatively cheaper female labor. We find, however, that the gender wage gap *decreases* following Nitaqat. Moreover, the establishment of the effective minimum wage in 2013 reduces the wage gap even further, and the female share of the workforce remains elevated (see Appendix Figure H.2). The fact that both women's relative wages and employment increase is difficult to reconcile

with a price-based explanation. Instead, the evidence is consistent with Nitaqat increasing relative demand for female labor by increasing the set of firms that integrate.

Lastly, it is also possible that Nitaqat led to a shift in women's labor *supply*. Female labor force participation in Saudi Arabia is among the lowest in the world, at 17.8% in 2011 (GaStat, 2011). One reason for this low rate may be that households perceive that few firms are willing to hire women in the first place. Since we do not have data on labor supply decisions, we look at the response to Nitaqat for firms that had integrated prior to the policy's implementation. These firms are already employing a mix of men and women and face an increase in the relative price of female labor due to Nitaqat. In the absence of a supply response, we would expect to see the female share of employment at these firms weakly decreasing. Instead, for *both* sets of firms, there is a marked increase in the female share of employment beginning with Nitaqat's integration, implying that there may have been an outward shift in women's labor supply.[32]

6   Conclusion

We posit that where there are social norms for gender segregation, firms face costs to employing both men and women that are largely fixed. Motivated by a simple model of firm hiring, we develop a joint test for whether integration costs bind for any firm and a methodology for evaluating the firm-level consequences of those costs. We validate our approach using administrative employer-employee data and unique policy variation from Saudi Arabia, a country that strictly regulates between-gender interactions in the workplace during our period of study. We also find evidence that integration costs depress aggregate demand for female labor. In particular, we document that Nitaqat—a gender-neutral policy in Saudi Arabia that had the unintended consequence of inducing many firms to integrate—increased women's employment and wages.

Integration costs may seem particularly likely to exist and bind in Saudi Arabia, a country with uniquely explicit restrictions on between-gender interactions. However, there is suggestive evidence that integration costs also bind in other countries with strong social norms for gender segregation. In particular, in recent World Bank survey data on manufacturing firms, 50% of medium firms and 25% of large firms in MENA and South Asia are all male, a far higher share of firms than one would expect by chance alone (see Appendix Table A.3). In World Bank surveys from 2013 and 2014, 29% of South Asian firms claim that hiring women "could cause disruption in the working environment" and cite this as a constraint to hiring women.[33] Integration costs may therefore constrain women's employment more broadly across the world, providing an important barrier to growth as well as an important driver of global employment dynamics.

Integration costs also have the potential to generate a coordination problem: firms may not integrate unless enough women enter the labor market, and women may not enter the labor market

---

[32]Appendix Figure H.3 plots the female share of employment over time in firms that employed Saudis in January 2009, split by the firm's ex-post integration status in that month.

[33]These data are from World Bank Enterprise Surveys in Afghanistan (2014), Bangladesh (2013), India (2014), Nepal (2014), and Pakistan (2013).

<div align="center">21</div>

unless enough firms have integrated. This interaction between the two sides of the market may generate a feedback loop: for example, a firm's decision to integrate may increase the supply of women searching in the labor market, which in turn induces other firms to integrate. Unfortunately, we are unable to quantify the potential magnitude of these spillovers because we do not have data on labor supply and how female labor supply responds to the integration of local firms. However, such a coordination problem could be solved by policy: big-push policies like Nitaqat could have large equilibrium-switching effects, and complementary policies that address labor supply may help magnify these impacts.

Our results suggest that integration costs prevent some firms from hiring superior female candidates. Though beyond the scope of this paper, a natural question for further research is: how do integration costs affect productivity, both for firms and in the aggregate? Investments in overcoming integration costs may have longer-term implications for productivity in addition to employment.

References

Acemoglu, Daron and Joshua D. Angrist, "Consequences of Employment Protection? The Case of the Americans with Disabilities Act," *Journal of Political Economy*, October 2001, *109* (5), 915–957.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, May 2013, *128* (2), 469–530.

Altonji, Joseph G., Todd E. Elder, and Christoper R. Taber, "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling," *Journal of Human Resources*, 2005, *40* (4), 791–821.

Arab News, "Women's rights gain focus in the Kingdom," *Arab News*, 2010. June 15, 2010.

_ , "Saudi Women No Longer Need Guardians' Consent to Receive Services," *Arab News*, 2017. May 5, 2017.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena, "Bride Price and Female Education," *Journal of Political Economy*, 2019.

Augereau, Angelique, Shane Greenstein, and Marc Rysman, "Coordination versus Differentiation in a Standards War: 56K Modems," *RAND Journal of Economics*, Winter 2006, *37* (4), 887–909.

Bayard, Kimberly, Judith Hellerstein, David Neumark, and Kenneth Troske, "New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data," *Journal of Labor Economics*, October 2003, *21* (4), 887–922.

Becker, Gary, *The Economics of Discrimination*, University of Chicago Press, 1957.

Bursztyn, Leonardo, Alessandra L. Gonzalez, and David Yanagizawa-Drott, "Misperceived Social Norms: Female Labor Force Participation in Saudi Arabia," June 2018.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *Quarterly Journal of Economics*, May 2016, *131* (2), 633–686.

Dahl, Gordon B., Andreas Kotsadam, and Dan-Olof Rooth, "Does Integration Change Gender Attitudes? The Effect of Randomly Assigning Women to Traditionally Male Teams," February 2018. Unpublished manuscript.

Ellison, Glenn and Edward L. Glaeser, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 1997, *105* (5), 889–927.

Evidence for Policy Design, "Back to Work in a New Economy: Background Paper on the Saudi Labor Market," Harvard Kennedy School, Harvard University April 2015.

_ , "The JPCs seem to be working. Why not use them to increase female employment?," EPoD Policy Insight, Harvard Kennedy School, Harvard University May 2017. https://epod.cid.harvard.edu/sites/default/files/2018-06/JPC.pdf.

Fernandez, Raquel, "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century," *American Economic Review*, 2013, *103* (1), 472–500.

_ **and Alessandra Fogli**, "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, January 2009, *1* (1), 146–177.

Fry, Erika, "This CEO is helping Saudi women break a gender barrier," *Fortune*, 2015. September 14, 2015.

Garicano, Luis, Claire Lelarge, and John Van Reenan, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, November 2016, *106* (11), 3439–79.

GaStat, *Labor Force Survey 2011*, General Authority for Statistics, Kingdom of Saudi Arabia, 2011.

_ , *Labor Force Survey 2015*, General Authority for Statistics, Kingdom of Saudi Arabia, 2015.

Glover, Dylan, Amanda Pallais, and William Pariente, "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores," *Quarterly Journal of Economics*, August 2017, *132* (3), 1219–1260.

Goldin, Claudia, "Monitoring Costs and Occupational Segregation by Sex: A Historical Analysis," *Journal of Labor Economics*, January 1986, *4* (1), 1–27.

Gourio, François and Nicolas Roys, "Size-dependent regulations, firm size distribution, and reallocation," *Quantitative Economics*, July 2014, *5* (2), 377–416.

Groshen, Erica L., "The Structure of the Female/Male Wage Differential: Is It Who You Are, What You Do, or Where You Work?," *Journal of Human Resources*, Summer 1991, *26* (3), 457–472.

Hamermesh, Daniel S. and Gerard A. Pfann, "Adjustment Costs in Factor Demand," *Journal of Economic Literature*, September 1996, *34* (3), 1264–1292.

Hart, Chloe Grace, Alison Dahl Crossley, and Shelley J. Correll, "Leader Messaging and Attitudes toward Sexual Violence," *Socius*, 2018, *4*.

Hellerstein, Judith K. and David Neumark, "Workplace Segregation in the United States: Race, Ethnicity, and Skill," *The Review of Economics and Statistics*, 2008, *90* (3), 459–477.

Hjort, Jonas, "Ethnic Divisions and Production in Firms," *Quarterly Journal of Economics*, November 2014, *129* (4), 1899–1946.

Jayachandran, Seema, "The Roots of Gender Inequality in Developing Countries," *Annual Re-*

view of Economics, August 2015, *7*, 63–88.

Khoja, Sara and Sarit Thomas, "Labour and Employment Compliance in Saudi Arabia," in del Rey Salvador and Robert J. Mignin, eds., *International Labour and Employment Compliance Handbook*, Kluwer Law International BV, Netherlands, 2018.

Kleven, Henrik, "Bunching," *Annual Review of Economics*, October 2016, *8*, 435–464.

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

Lang, Kevin, "A Language Theory of Discrimination," *Quarterly Journal of Economics*, May 1986, *101* (2), 363–382.

Miller, Amalia R. and Carmit Segal, "Does Temporary Affirmative Action Produce Persistent Effects? A Study of Black and Female Employment in Law Enforcement," *Review of Economics and Statistics*, 2012, *94* (4), 1107–1125.

Miller, Conrad, "The Persistent Effect of Temporary Affirmative Action," *American Economic Journal: Applied Economics*, July 2017, *9* (3), 152–90.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, 2019, *37* (2), 187–204.

Pan, Jessica, "Gender Segregation in Occupations: The Role of Tipping and Social Interactions," *Journal of Labor Economics*, April 2015, *33* (2), 365–408.

Peck, Jennifer R., "Can Hiring Quotas Work? The Effect of the Nitaqat Program on the Saudi Private Sector," *American Economic Journal: Economic Policy*, May 2017, *9* (2), 316–347.

Sara Hamdan, "Saudi Arabia Signals Openness to Women Seeking Work," *New York Times*, 2012. September 5, 2012.

Saudi Gazette, "Nitaqat percentage linked to salary," *Saudi Gazette*, 2012. September 9, 2012.

Varejão, José and Pedro Portugal, "Employment Dynamics and the Structure of Labor Adjustment Costs," *Journal of Labor Economics*, January 2007, *25* (1), 137–165.

Whatley, Warren C., "Getting a Foot in the Door: "Learning," State Dependence, and the Racial Integration of Firms," *Journal of Economic History*, March 1990, *50* (1), 43–66.

Willness, Chelsea R., Piers Steel, and Kibeom Lee, "A Meta-Analysis of the Antecedents and Consequences of Workplace Sexual Harassment," *Personnel Psychology*, 2007, *60* (1), 127–162.

World Bank, *Women, Business and the Law 2018*, The World Bank, 2018.

_ , *World Development Indicators 2018*, The World Bank, 2018.

FIGURE 1

DISTRIBUTION OF FEMALE EMPLOYMENT ACROSS FIRMS
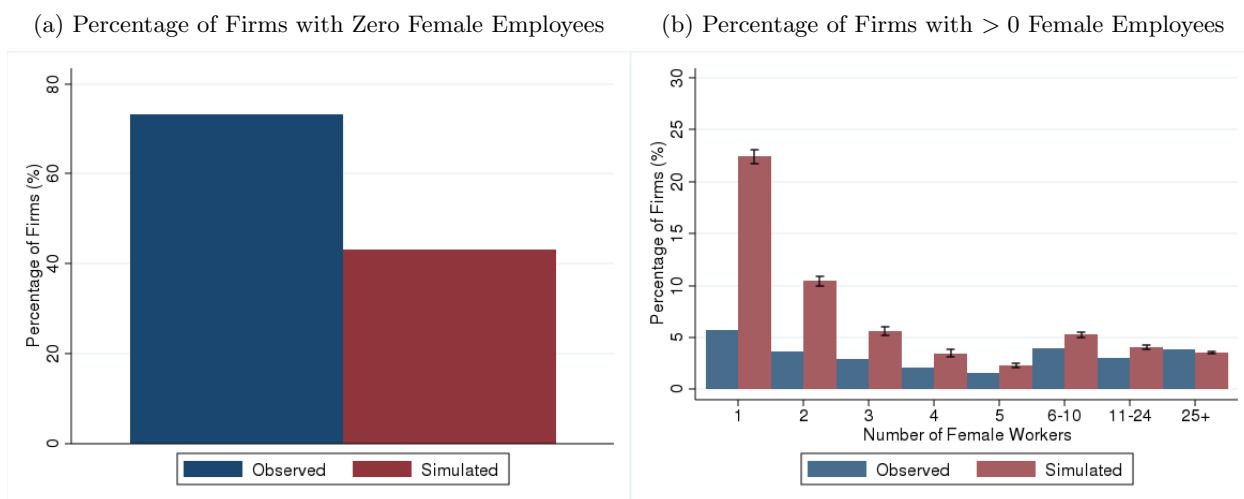
(a) Percentage of Firms with Zero Female Employees (b) Percentage of Firms with > 0 Female Employees
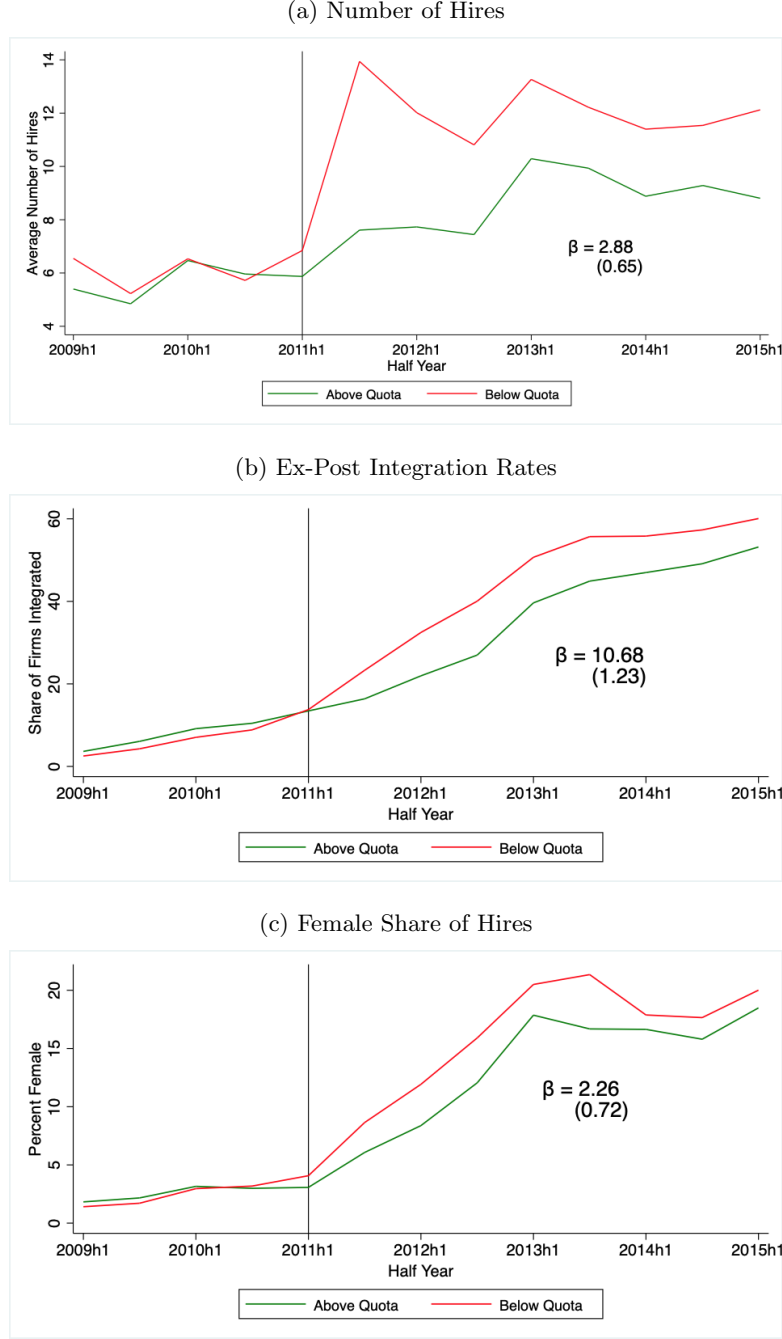

Note: This set of figures compares the observed distribution of female employment across firms in January 2009 to distributions simulated under the null hypothesis that no firm faces binding integration costs. Sample selection and simulation details are described in Sections 4.1.1 and 4.1.2. Panel A plots the share of firms with zero female employees in both the observed and simulation distributions. Panel B plots the share of firms with various nonzero totals of female employees in both the observed and simulated distributions. For all simulations, a Kolmogorov-Smirnov test rejects equality of the observed and simulated distributions at the 1% significance level.

FIGURE 2

INTEGRATION RATES BY $\bar{\theta}_j n_j$

(a) Integration Status

(b) Ex-Ante Integration Rates
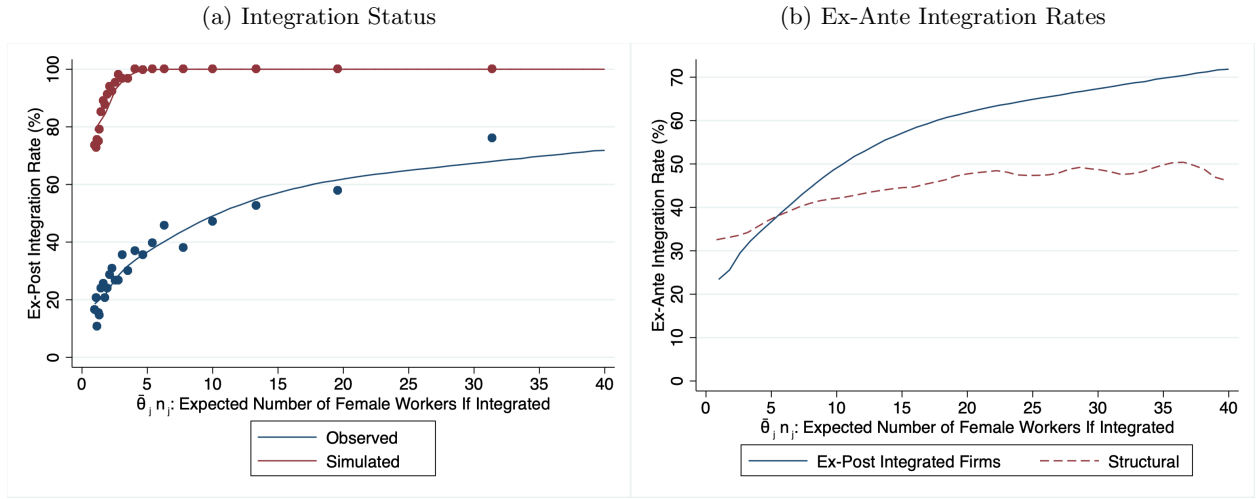
Note: This set of figures depicts the relationship between ex-post and ex-ante integration rates and $\bar{\theta}_j n_j$, a firm's expected number of female employees if ex-ante integrated. We construct $\bar{\theta}_j n_j$ for each firm j using an estimate of $\theta(X_i)$—either $\hat{\theta}^{EP}(X_i)$ or $\hat{\theta}^S(X_i)$—and the job composition of firm j. Panel A plots both the observed ex-post integration rate and the simulated ex-post integration rate, where the latter is simulated under the null hypothesis that all firms are ex-ante integrated. In Panel A, $\bar{\theta}_j$ is constructed using $\hat{\theta}^{EP}(X_i)$ and the job mix in firm j. Panel B plots ex-ante integration rates. "Ex-Post Integrated Firms" is constructed as described in Section 4.2.1. "Structural" plots the average estimated values of π_j (described in Section G.1) as a function of $\bar{\theta}_j n_j$, where $\bar{\theta}_j$ is constructed using $\hat{\theta}^S(X_i)$ and the job mix in firm j. The sample includes a cross-section of firms from 2009 with at least five Saudi employees. Sample selection and simulation details are described in Sections 4.1.1 and 4.1.2.

26

FIGURE 3
INTEGRATION RATES AND FEMALE SHARE OF HIRES OVER TIME

(a) Number of Hires

(b) Ex-Post Integration Rates

(c) Female Share of Hires

Note: This set of figures compares the number of hires, ex-post integration rates, and female share of hires for above quota and below quota firms that were ex-post segregated in January 2009. Above quota firms are Green and Platinum firms, and below quota firms are Yellow and Red firms. Color refers to firm quota status in June 2011. We restrict to firms that had at least five Saudi employees in January 2009. There are 2,224 below quota firms and 1,559 above quota firms satisfying our sample criteria. The vertical line marks the first half of 2011. Nitaqat is implemented in June 2011. Panel A plots the average number of hires by half-year. Panel B plots the share of firms that are ex-post integrated by half-year. Panel C plots the female share of hires at each firm, averaged across firms. Firms that do not make any hires in a given half-year are not included in the calculation of the female share of hires for that period. Each panel includes OLS coefficient estimates for equation (4), a firm-level difference-in-difference model for number of hires, integration rates, and the female share of hires. Each observation reflects a firm by half-year pair. Robust standard errors are clustered at the firm level.

TABLE 1
COMPOSITION OF PRIVATE SECTOR

	Saudi		Non-Saudi	
	Male	Female	Male	Female
Share of workforce				
2005	11.0%	0.6%	86.9%	1.5%
2010	9.6%	0.8%	88.4%	1.3%
2015	11.7%	4.7%	79.2%	2.2%
Occupational distribution among group in 2015				
Managers	8.0	8.1	0.7	0.5
Professionals	5.6	6.3	7.7	13.6
Technicians	8.0	12.5	7.7	29.8
Clerical	23.8	39.8	0.6	1.3
Sales	9.8	20.5	5.5	0.8
Service	24.1	7.5	29.3	44.8
Agriculture	0.2	0.1	6.6	0.1
Industrial, Chemical, and Food Industries	1.8	1.6	2.5	8.4
Engineering Support	16.0	3.3	39.3	0.4
Armed Forces and Security	2.6	0.2	0.1	0.1
Other	0.0	0.0	0.0	0.3

Note: The top half of this table tabulates the distribution of private sector workers by year. The second half tabulate the occupational distribution of each subgroup of private sector workers in 2015. Numbers exclude domestic workers. Source: Saudi Ministry of Labor and Social Development (MLSD) via Saudi Arabian Monetary Agency (SAMA).

TABLE 2

SAUDI WORKERS SUMMARY STATISTICS,
JANUARY 2009

	Men	Women
Share of employment (%)	91.6	8.4
Age	32.1	30.3
	(10.1)	(7.6)
Married	24.6	32.9
Education level (%)		
Less than Secondary	5.5	4.1
Secondary	40.0	42.9
University	5.6	32.6
Missing	48.9	20.4
Monthly Wage	7206	3308
(Riyals)	(8016)	(4178)

Source: General Organization for Social Insurance (GOSI) administrative data. Data include only Saudi nationals in the private sector. These statistics describe 1,396,962 Saudis employed in the private sector in January 2009. Section 3.4 provides more details about the dataset used for this table.

Table 3

TABLE 3

FIRMS WITH SAUDI EMPLOYEES IN JANUARY 2009

	All	≥ 5 Saudi employees	
		All	Ex-post integrated
# of firms	27,294	7,943	2,123
Number of Saudi employees			
Mean	16.3	52.0	118.2
Median	2	12	20
SD	353	654	1254.3
Female share of employees (%)			
Mean	3.6	9.1	34.0
Median	0	0	25
SD	13.7	21.1	28.8
Avg. monthly wage (Riyals)	3,058	3,971	4,112
Industry (%):			
Agriculture and fishing	0.9	1.0	0.4
Commerce	32.5	28.1	23.1
Community/social services	9.4	13.9	35.0
Construction	28.5	21.6	12.2
Electricity, gas, and water	0.7	1.4	1.0
FIRE	10.8	11.1	14.7
Manufacturing	13.3	17.1	10.7
Mining	1.0	1.6	0.8
Telecommunications	3.1	4.2	2.1

Notes: This table presents descriptive statistics for firms with any Saudi employee in January 2009. The second column limits to firms with at least five Saudi employees. We limit the analysis to firms with at least five Saudi employees throughout Section 4.1. The third column further limits to firms that are employment both men and women. The average wage at a firm is measured in nominal Saudi Riyals in January 2009.

TABLE 4

SUMMARY OF θ ESTIMATES, JANUARY 2009

	Naive ($\hat{\theta}^0$)	Ex-post integrated ($\hat{\theta}^{EP}$)	Structural ($\hat{\theta}^S$)
Mean	0.082	0.125	0.123
Median	0.027	0.063	0.061
Standard deviation	0.155	0.180	0.177
Pairwise R^2:			
Location	0.06	0.09	0.09
Occupation	0.73	0.69	0.43
Industry	0.62	0.60	0.62

Notes: This table summarizes three estimates for $\theta(X_i)$: (1) the "naive" estimate ($\hat{\theta}^0$), described in Section 4.1.1, which estimated using data from all firms; (2) the estimate using ex-post integrated firms ($\hat{\theta}^{EP}$); and (3) the structural estimates ($\hat{\theta}^S$), where the model and estimation are described in Section G.1. Each estimated function is applied to all jobs in firms meeting the sample criteria described in Section 4.1. The *Pairwise R^2* values are the R^2 values from separate linear regressions of the $\theta(X_i)$ estimates on location fixed effects, two-digit occupation fixed effects, and one-digit industry fixed effects.

For Online Publication

A   Appendix: Additional Tables and Figures

FIGURE A.1

SEPARATION RATES AND WAGES FOR INCUMBENT MEN AT NEWLY INTEGRATED FIRMS

(a) Separation Rates

(b) Log Wages

Note: This set of figures describes the separation rates and wages of incumbent male employees at newly integrated firms in the GOSI data. Incumbent male employees are defined as male employees that joined the firm more than 18 months prior to the firm's first female hire. Panel A plots the average monthly separation rate of incumbent male employees at integrating firms in six-month increments relative to a firm's first observed female hire, averaged across firms. Panel B plots the average log monthly wage of incumbent male employees at integrating firms in six-month increments relative to a firm's first observed female hire, averaged across firms. For both panels, we restrict to firms with at least five Saudi employees in the month prior to integration.

TABLE A.1

OBSERVED AND SIMULATED DISTRIBUTION OF FEMALE
EMPLOYMENT ACROSS FIRMS

Panel A: θ Estimated Under Null

# of Female Employees	Observed	(1)	(2)	(3)	(4)	(5)	(6)
0	73.27	33.37	32.73	34.37	38.66	42.92	42.56
1	5.70	26.97	25.27	25.03	25.05	22.35	22.07
2	3.68	13.79	14.08	13.29	12.01	10.48	10.32
3	2.88	7.13	7.61	7.15	6.01	5.59	5.73
4	2.13	4.19	4.45	4.24	3.56	3.50	3.60
5	1.54	2.63	2.93	2.81	2.27	2.70	2.41
6-10	3.93	6.09	6.60	6.48	5.32	5.26	5.59
11-24	3.03	3.89	4.17	4.26	4.10	4.11	4.34
25+	3.83	1.94	2.17	2.37	3.01	3.53	3.37
Location			✓	✓	✓	✓	✓
1-Digit Occ.				✓			
2-Digit Occ.					✓	✓	
1-Digit Ind.						✓	
1-Digit Ind. × 1-Digit Occ.							✓

Panel B: θ Estimated Using Ex-Post Integrated

# of Female Employees	Observed	(7)	(8)	(9)	(10)	(11)	(12)
0	73.27	73.48	73.36	73.46	73.40	73.39	73.29
1	5.70	6.97	6.09	6.33	6.59	5.34	5.34
2	3.68	4.77	4.32	4.15	4.19	3.83	3.75
3	2.88	2.92	2.90	2.65	2.56	2.62	2.58
4	2.13	1.94	1.98	1.80	1.73	1.83	1.94
5	1.54	1.38	1.41	1.36	1.24	1.40	1.40
6-10	3.93	3.54	3.82	3.67	3.28	3.81	3.93
11-24	3.03	2.90	3.43	3.60	3.46	3.62	3.88
25+	3.83	2.10	2.70	2.98	3.54	4.15	3.90
Location			✓	✓	✓	✓	✓
1-Digit Occ.				✓			
2-Digit Occ.					✓	✓	
1-Digit Ind.						✓	
1-Digit Ind. × 1-Digit Occ.							✓

Source: This table compares the observed distribution of female employment across firms in January 2009 to various simulated distributions. Sample selection and simulation details are described in Sections 4.1.1 and 4.1.2. In Panel A, distributions are simulated under the null hypothesis that no firm faces binding integration costs. We use estimates of $\theta(X_i)$ for varying sets of observable job characteristics. In Panel B, we simulate the distribution of female employment while allow ex-ante integration rates to vary by $\bar{\theta}_j n_j$. The details of this simulation exercise are described in Section G.1.1.

TABLE A.2
BASELINE-SEGREGATED FIRMS IN JUNE 2011

	Below Quota	Above Quota
# of firms	2,224	1,559
Number of Saudi employees		
Mean	33.7	45.5
Median	12	16
SD	101	130
Female share of employees (%)		
Mean	2.2	1.7
Median	0	0
SD	8.6	7.0
Avg. monthly wage (Riyals)	3,680	4,898
Industry (%):		
Agriculture and fishing	1.4	0.8
Commerce	25.4	25.1
Community/social services	7.0	4.6
Construction	30.3	26.2
Electricity, gas, and water	1.2	1.6
FIRE	8.0	10.3
Manufacturing	20.1	22.4
Mining	1.7	3.7
Telecommunications	5.0	5.4

Notes: This table presents descriptive statistics as measured in June 2011 for firms with at least five Saudi employees in January 2009. The first column limits to Below Quota firms, those with Yellow and Red color statuses in June 2011. The second column limits to Above Quota firms, those with Green and Yellow color statuses in June 2011.

Manufacturing Firms with Zero Female Employees and Workforce Composition, by Region

	All-male share of firms (%), by size		Female share (%)	
	Medium (20–99)	Large (100+)	Surveyed firms	Labor force
Sub-Saharan Africa	10.5	2.3	27.0	47.5
East Asia and Pacific	1.8	0.5	41.2	42.8
Eastern and Central Europe	2.5	0.7	38.4	43.9
Latin America and Caribbean	3.0	0.8	32.8	41.1
Middle East and North Africa	48.1	22.7	16.9	21.1
South Asia	49.9	28.6	14.5	23.5

Source: World Bank Enterprise Survey, 2006–2018. Survey data cover manufacturing firms in 65 countries. Statistics are calculated using survey weights within each country and year, then averaged across years within a country, then averaged across countries within a region, weighting by 2018 population. Female share of labor force is derived from 2018 World Bank Development Indicators for the same countries, is also a population-weighted average, and is not restricted to manufacturing.

B   Appendix: Model

For each vacancy, the firm receives k applications from two types of candidates: type F and type M. Share δ of candidates are type F, and share $(1 - \delta)$ are type M.

The expected value of the highest U_j in a sample of size $s \in \{\delta k, (1 - \delta)k\}$ drawn from a single group, F or M, is

$$U_*^G = \mu^G + \beta \log(s), \ G \in (M, F),$$

where $\mu^G \equiv v^G + \beta\gamma$ is the expected net value of a single candidate from group G.

The expected value of the highest U_j drawn from a combined sample of all candidates is

$$U_*^I = \beta \log \left[\delta \exp \left(\frac{\mu^F}{\beta} \right) + (1 - \delta) \exp \left(\frac{\mu^M}{\beta} \right) \right] + \beta \log k.$$

The firm's problem of choosing what pools to hire from is equivalent to choosing the maximum of nU_*^F (only type F), nU_*^M (only type M), and $nU_*^I - c$ (both types).

We first consider the choice between hiring only type M candidates and hiring from both types. The firm will pay the fixed integration cost and hire from both types if

$$U_*^I - U_*^M > \frac{c}{n}. \tag{5}$$

The left-hand side of this expression can be expressed as

$$U_*^I - U_*^M = \beta \log \left[\frac{\delta}{1 - \delta} \exp \left(\frac{v^F - v^M}{\beta} \right) + 1 \right]. \tag{6}$$

Let θ denote the probability that the firm's preferred candidate from the combined pool is type F, where

$$\theta = \frac{\delta \exp \left(\frac{v^F}{\beta} \right)}{\delta \exp \left(\frac{v^F}{\beta} \right) + (1 - \delta) \exp \left(\frac{v^M}{\beta} \right)}.$$

Rearranging, we get

$$\frac{1}{1 - \theta} = \frac{\delta}{1 - \delta} \exp \left(\frac{v^F - v^M}{\beta} \right) + 1. \tag{7}$$

Combining (6) and (7), we have

$$\begin{aligned}
U_*^I - U_*^M &= \beta \log \left[\frac{\delta}{1 - \delta} \exp \left(\frac{v^F - v^M}{\beta} \right) + 1 \right] \\
&= -\beta \log[1 - \theta] \\
&\approx \beta\theta.
\end{aligned}$$

Combining the expression above with (5), an *approximate* condition for the firm to pay the fixed integration cost and hire from the combined pool is

$$n\theta > \frac{c}{\beta}. \tag{8}$$

C   Appendix: Saudi Female Employment Policies

In addition to Nitaqat, the Saudi government also pursued a slate of practical measures designed to increase women's employment over the study period, including the Retail Employment Decree, the Hafiz program, and updates to the guardianship system. The King issued a royal decree in 2011 mandating that shops selling lingerie and cosmetics employ only Saudi women as salesclerks beginning in August 2012. The decree was expanded to also cover stores selling women's clothing and accessories beginning in January 2014. There were recently plans to further expand the decree to cover all stores selling goods of primary interest to women, such as pharmacies with cosmetics sections and fabric stores (Evidence for Policy Design, 2015).

Though not gender-specific, the Hafiz unemployment assistance program has also drawn women into the workforce and supported their private sector job search. Hafiz provides a monthly financial stipend to unemployed Saudis who make weekly check-ins to a government-sponsored online job search portal (Taqat Online). More than 90% of Hafiz beneficiaries have been women (Evidence for Policy Design, 2017). The MLSD removed regulations requiring women to obtain permission from a male guardian to apply for private sector jobs.[34] Many firms still require a guardian's approval, though the Ministry recently forbade this practice among government employers.[35]

D   Appendix: Data

D.1   Matching GOSI and Nitaqat Firms

Administrative data from the Nitaqat program is used to identify the Nitaqat compliance status of firms. As described by Peck (2017), the Nitaqat database is used to track compliance with national quotas on Saudi employment in the private sector. The database collects information on whether a given firm was subject to quotas during a given week, and, if so, whether it met the quotas for that particular week. These data provide weekly quota compliance information from June, 2011 (the start of the Nitaqat program) until December, 2013.

Firms are defined differently between the Nitaqat and GOSI data sets. In the latter, firms are defined by their legal status as a commercial organization operating in potentially multiple industries. In the Nitaqat data, however, the operations of such firms are further classified into entities, which are subject to different quotas depending on the industry category each entity operates in and, as described in the main text, the size group based on the total number of employees. For example, a firm operating a bakery and a jewelry store would be considered two separate entities facing different quotas (and would therefore contain two entries in the data for each time period)[36]. In the GOSI data, however, such a firm would be considered a single firm. Firms with multiple entities can also list as a single entity (in the "Multiple Economic Activities" industry) but would be subject to the most stringent quota they face based on the entities under their umbrella. To harmonize the definition of the firm between the two data sets, firms with multiple entities in the Nitaqat data were aggregated together by summing their employee counts, and assigning the color and size status by the most binding entity quota (as measured by the number of Saudis required to fulfill it) the firm faces. The number of Saudis the firm needs to hire, however, was summed across all entities to create a single metric for the distance of the firm to the quota. This transformation

---

[34]Jafar AlShayeb, *Arab News* June 15, 2010 "Women's rights gain focus in the Kingdom"

[35]Lulwa Shalhoub, *Arab News* May 5, 2017 "Saudi women no longer need guardians' consent to receive services" http://www.arabnews.com/node/1094681/saudi-arabia

[36]An entity consisting of multiple branches (e.g., a national franchise) are counted as a single entity for each branch of the MLSD labor office they are linked to.

only affects 58,000 of the approximately 1.07 million firms in the Nitaqat data.

In addition to the distinction between entities and firms, it should be noted that the firm identifiers used by both GOSI and the Nitaqat data define firms with a national or multicity presence as separate commercial organizations depending on the geographic MLSD office they register with. For example, a firm with branches in Riyadh and Dammam would count as two firms, both of which are subject to separate quota calculations. The geographic scope of the MLSD offices is quite broad, and are typically at the provincial level. The definition of the firm we use in this paper therefore can be thought of as a legal commercial organization within a particular province.

D.2 Other Data Notes

We classify each occupation to the two-digit ISCO-08 group, reducing the number of occupations from 2,151 to 40. This significant drop in occupations is primarily due to inconsistent naming, misspellings, and changes to the GOSI classification scheme over time. Table D.1 lists the top ten most common ISCO-08 coded occupations in June 2011.

TABLE D.1

EMPLOYEES BY ISCO-08 OCCUPATION, JUNE, 2011

ISCO Code	ISCO Category	Frequency	Percent
96	Refuse workers and other elementary workers	104,744	14.4
41	General and keyboard clerks	84,406	11.6
54	Protective services workers	65,032	9.0
42	Customer services clerks	64,265	8.9
99	Unclassified	48,382	6.7
33	Business and administration associate professionals	36,547	5.0
52	Sales workers	32,943	4.5
74	Electrical and electronic trades workers	26,754	3.7
83	Drivers and mobile plant operators	25,296	3.5
21	Science and engineering professionals	23,465	3.2
	Total	511,834	70.5

Note: This table presents the number of Saudi employees in the ten most common ISCO-08 2-digit occupation group in the GOSI data. The large number of unclassified occupations is due to the significantly large number of cases where the GOSI occupation verification process was still processing or was incomplete.

There are 37 work locations provided in the data. We limit our analysis to locations with at least 50 firms with five or more Saudi employees in January 2009. This leaves us with 17 locations that account for 95% of firms and 98% of workers. In January 2009, 83% of workers are located in four cities: Riyadh, Jeddah, Dammam, and Khobar.

To clean up potentially erroneous observations, we drop individuals with ages below 10 or above 100 in the GOSI data. We also drop entries for part-time work, which only affects about 47,000 of the 2.8 million employees in the data. If an individual has more than one full-time job in a given month, we keep only the observation for the job with the highest wage.

E   Appendix: Ghost Employment

The main text mentions the concern that firms may falsify their employee records with GOSI to meet their quotas after Nitaqat, so reported employment numbers may not reflect real employment, particularly for women. Private sector firms are required to register their employees with GOSI and to pay a fraction of the reported wage into the employee's social security account. Nationals may not be registered as full-time employees for more than one firm at the same time. Workers have some incentive to make sure these records are filed accurately so that their eventual retirement payments are accurate. The Nitaqat enforcement system draws directly on these GOSI records to monitor the number of Saudi workers registered as employees at each firm. "Ghost employment" is used to refer to a variety of situations in which the worker is not doing the job as reported to GOSI. This can range from cases of outright fraud (e.g., where a worker's National ID Number is used without the worker's knowledge or permission) to cases where the worker draws the reported salary but does not perform meaningful work at the firm.[37] This ghost employment would cause our analysis to overstate the degree to which firms hire Saudi women in response to employment quotas. In this analysis we investigate whether this phenomenon becomes more common after the start of Nitaqat and whether it appears to be more common for women than for men.

To do this, we examine the share of workers hired in each month who appear to have "active" career trajectories. We define a worker as being active if their job history shows that they switch firms, receive wage increases, change occupations, or make above minimum wage. We can be reasonably confident that workers that experience these events are "real" employees: firms have no incentive to report paying fake workers above minimum wage (as this simply increases their GOSI payments without providing Nitaqat benefits), and there is similarly no reason to promote them, give them raises, or move their IDs to other firms. We construct an indicator equal to 1 if the worker experiences any of these actions (change wage or occupation, switch firms, or make above minimum wage) within 24 months of their first appearance in the GOSI system.[38]

In addition to capturing ghost employment, GOSI records may be inaccurate for several other reasons. First, firms may register artificially low wages in order to minimize their social security payments on behalf of their employees. This can in principle be checked by the worker, but there are some accounts of workers being surprised by their wage records upon retirement. Firms may also neglect to record promotions in the GOSI system, so recorded wages may lag actual wages. Movements across firms seem likely to be accurate, as a prior employer will not want to make payments for people who are no longer employees, and new firms will want to have the worker's national ID number released so they can register a new hire. These will bias the measure toward under-counting active employees, so the count of "inactive" workers should be assumed to include not only ghost employees, but also employees whose records are not updated promptly as well as workers who simply do not experience job status changes over the period.[39]

---

[37]There may also be cases in between, for example where workers collect a one-time payment or ongoing small payment from the firm to use their ID numbers.

[38]One potential issue is the de facto increase in the minimum wage in 2013. GOSI had previously required firms to enter a minimum wage of 1500SAR per month. In January 2013 firms were only given pro-rated Nitaqat credit for Saudi employees paid less than 3000SAR a month (e.g., a worker being paid the previous minimum of 1500SAR would count as 0.5 Saudis for Nitaqat purposes). Because of this, we do not consider increases from 1500 to 3000SAR that occur after January 2013 to be wage increases.

[39]Firms may also retain previous workers who have exited the labor market on their GOSI employment rolls. These workers will mistakenly appear to be active. Because we focus on workers hired between 2009 and 2013 we expect that this will comprise a only a very small part of the workforce, as these workers would need to enter the labor force after 2009, experience a change in wage, occupation, or firm, and then leave the private sector workforce without retiring and drawing their GOSI pension.

Figure E.1 shows a plot of the share of workers hired in each month that experience at least one of these events within 24 months of being hired. The share of workers who change job status is relatively steady for both genders at about 58% for men and 47% for women. As discussed before, there are a variety of reasons (aside from ghost employment) why this might only apply to half of workers. First, workers may simply not be promoted within 24 months of their first entry into the private sector. Second, they may be promoted but not have the promotions recorded in GOSI. Although only about half of workers experience official status changes within two years of hire, the patterns are similar across genders and relatively stable over time. There is a slight decrease in the share of workers promoted for those hired after Nitaqat.

FIGURE E.1
SHARE HIRED IN MONTH WHO CHANGE STATUS
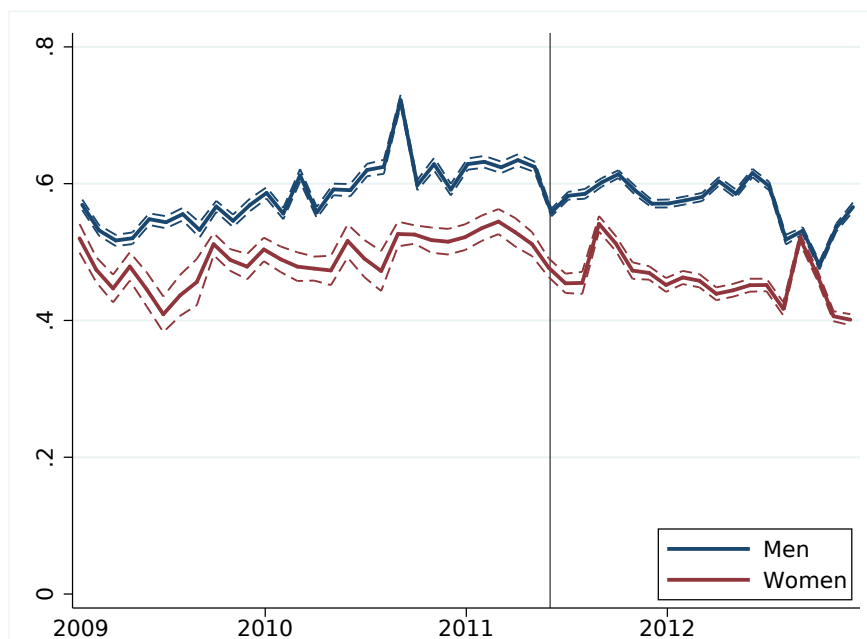
Note: This figure plots the share of Saudi employees in the GOSI matched employee-employer data who are first hired in each month who change wage or occupation, switch firms, or earn above minimum wage within two years of hire. Dashed lines show the 95% confidence interval for month indicator variables.

Within these series we may be concerned also about compositional changes in the types of workers that are being hired before and after Nitaqat as well as the types of firms that hire Saudis before and after the policy change. There is ample evidence that Saudis hired after Nitaqat are different from those hired before: more are women, more are hired with lower skill levels, and married women are more likely to join the labor force. Red and Yellow firms, which were most incentivized to increase Saudi hiring, were also potentially less desirable places for Saudis to work and may be less likely to keep their GOSI records up to date and to promote their employees over time. Figure E.2 shows the plot of these shares controlling for some worker characteristics: age, education, and marital status of the new hires.

Women are more likely to be active workers when controlling for observable worker characteristics, and the likelihood of promotion appears to be steadily increasing over time for women. We therefore conclude that even if ghost employment is captured by the GOSI data it does not appear to worsen after Nitaqat, and does not worsen for women in particular.

SHARE HIRED IN MONTH WHO CHANGE STATUS (WITH WORKER-LEVEL CONTROLS)

Note: This figure plots the share of Saudi employees in the GOSI matched employee-employer data who are first hired in each month who change wage or occupation, switch firms, or earn above minimum wage within two years of hire when controlling for employee characteristics. Indicator variables are used to flexibly control for age, education, and marital status of new hires. Dashed lines show the 95% confidence interval for month indicator variables.

F   Appendix: Testing and Estimation Details

F.1   Simulating Bunching at Zero When Integration Costs Bind

In this section we demonstrate the rationale for our test of the null hypothesis that no firms face binding integration costs, developed in Section 4.1. In this test, we simulate the distribution of female hires across firms under the null hypothesis, and compare this to the distribution we observe in practice. We show here that if some firms are in fact ex-ante segregated, the distribution we simulate will generally *underpredict* the number of firms with zero female hires. We demonstrate this point using simulation.

Our simulation exercise builds on the model above by positing that some exogeneously determined γ share of firms are integrated and that, under the null hypothesis, θ_0 is the probability that each hire is female. Under both hypotheses, θ_0 is the expected female share of employees pooled across all firms. Firms are characterized by their number of hires, n.

Under the null hypothesis (H_0), all firms are integrated ($\gamma = 1$). In this case, firms which do not hire any women do so by chance alone. Alternatively (H_a), if $\gamma < 1$, then some firms do not hire women because they are ex-ante segregated. We show via simulation that under H_a there are generally a greater share of firms with zero female employees.

We consider two scenarios: one where the probability of integration is constant across firms and a second where integration rates are increasing in firm size.

F.1.1   Constant Integration Probability

First, we assume that the probability of integration is constant across firms, and given by γ. In this case, under H_a, the probability that a hire is female at an ex-ante integrated firm is $\theta_a = \theta_0/\gamma$. Our simulation is structured as follows. We first set a value of γ, the share of integrated firms, and θ_a, the probability a hire is female in an ex-ante integrated firm under H_a. Then, for each run of the simulation, we:

1. sample firm sizes (ie. the total number of employees) from a log-normal distribution with mean 50 and standard deviation 500, approximately matching the distribution of firm sizes we observe in our Saudi employment data (see Table 3);

2. determine whether a firm in our sample is integrated with probability $\gamma < 1$ for H_a; all firms are considered integrated under H_0 ($\gamma = 1$);

3. for H_a determine the gender of each hire via a binomial draw with probability θ_a that each hire is female. Sum these hires to determine the count of female employees for each firm under H_a;

4. set θ_0, the probability of a hire being female under H_0 using the overall female share of employment simulated in the prior step[40], then similarly determine the gender of each hire and count the number of female employees for each firm under H_0

After running the above simulation 1,000 times, we calculate the share of simulations where the number of firms with zero female employees under H_a (Z_a) exceeds the same value under H_0 (Z_0). We show in Figure F.1 what the distribution of female employee counts look like under

---

[40]This allows us to have approximately equal numbers of female employees under both hypotheses.

both hypotheses for $\gamma = 0.7$ and $\theta_a = 0.5$.[41] Each column represents the mean across simulations, whereas the error bars represent the 5th and 95th percentiles.

FIGURE F.1

SIMULATED DISTRIBUTION OF FEMALE EMPLOYMENT FOR $\gamma = 0.7$ AND $\theta_a = 0.5$
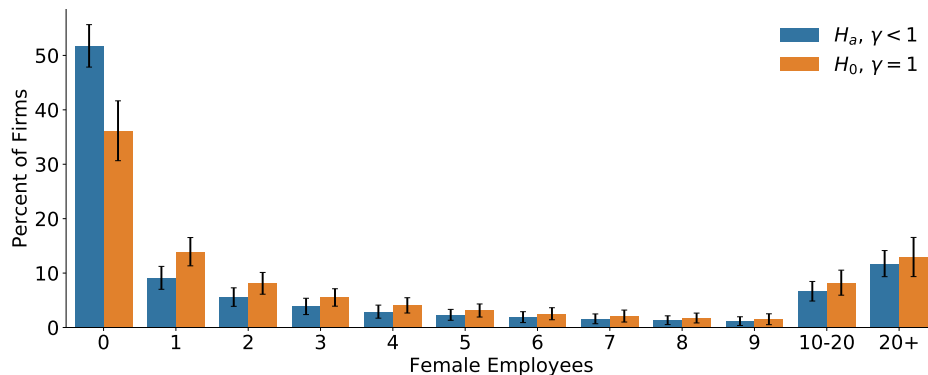
Note: This figure plots the distribution of the count of female employees across firms based on 1,000 simulations of firm sizes, integration probabilities (γ) and the share of female labor in the workforce (θ_a). The H_0 category supposes that all firms are integrated ($\gamma = 1$), and the H_a category supposes that some firms are ex-ante segregated ($\gamma < 1$). Sample selection and simulation details are described in Sections 4.1.1 and 4.1.2.

We then repeat this exercise by iterating over values of $\gamma \in (0, 1)$ and $\theta_a \in (0, 1)$. We plot the share of simulations where $Z_a > Z_0$ for each γ and θ_a value in Figure F.2 below.

---

[41]These values are chosen primarily for testing purposes. Repeating the exercise for different values results in similar patterns as shown below.
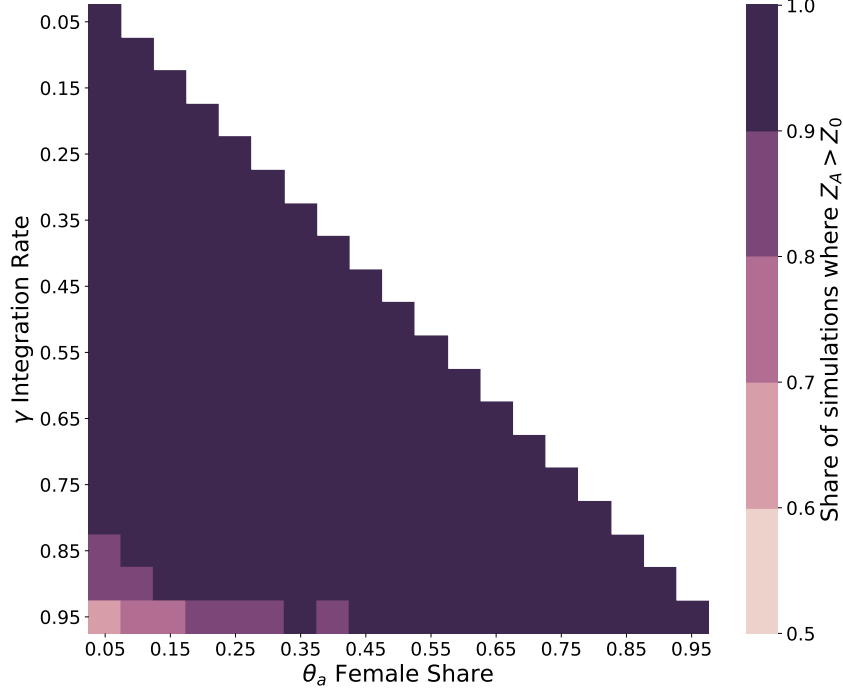
Note: This heatmap plots the share of simulations with $Z_a > Z_0$, or the share of simulations where there are more firms with no female employees under H_0 vs. H_a while varying values of θ and γ. Sample selection and simulation details are described in Sections 4.1.1 and 4.1.2.

Except the largest values of γ ($\gamma \geq 0.9$), $Z_a > Z_0$ for virtually all simulation draws. When γ is large, $Z_a > Z_0$ for the majority of simulation draws, but this share gets as low as the $0.6 - 0.7$ range (when $\gamma = 0.95$ and $\theta_a < 0.075$).

F.1.2 Integration Rates Increasing in n

If integrated costs are largely fixed, firms which have to hire more employees may be more likely to integrate. In this case, integration rates are increasing in n. To account for this, we again draw n from log-normal distribution, and also generate firm specific integration likelihoods $\gamma_i \sim Beta(\beta \frac{\gamma}{1-\gamma}, \beta)$ where $\gamma = \bar{\gamma}_i$.[42] To introduce the correlation between these two marginal distributions, we conduct a Cholesky decomposition to create a joint distribution of n_i and γ_i across such that the correlation between n and γ is positive.

We then continue the simulations as above, but iterate over values of $\bar{\gamma}_i$ and θ_a, and determine whether a firm is integrated according to its specific γ_i integration probability. We show in Figure

---

[42]We pick this particular form of the *Beta* distribution as its mean is γ. In other words, for a given share of firms integrated, we can generate a distribution of integration likelihoods for each firm such that the mean is equal to the overall share of firms integrated. In this case β acts as a scaling parameter but does not affect the mean.

F.3 the distribution of female employment for $\gamma = 0.7$ and $\theta_a = 0.5$ as above. We similarly plot the share of simulations where $Z_a > Z_0$ for each γ and θ_a value in Figure F.4.

FIGURE F.3

SIMULATED DISTRIBUTION OF FEMALE EMPLOYMENT FOR $\bar{\gamma}_i = 0.7$ AND $\theta_a = 0.5$ – INTEGRATION RATES INCREASING IN n
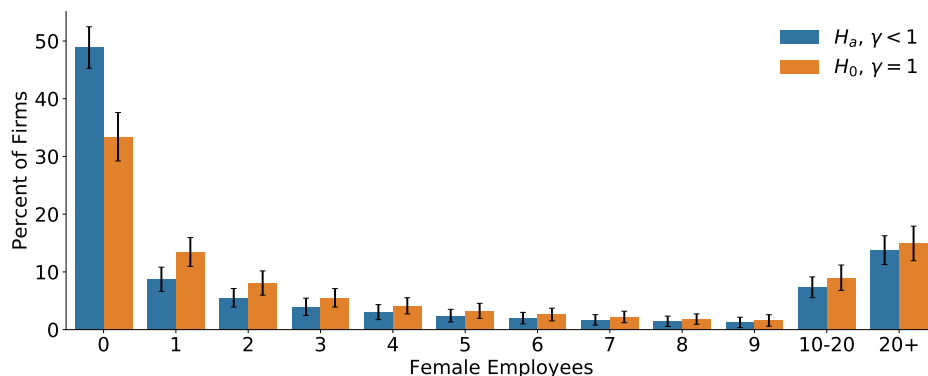
Note: This figure plots the distribution of the count of female employees across firms based on 1,000 simulations of firm sizes, integration probabilities (γ) and the share of female labor in the workforce (θ_a) when firm integration probabilities correlate positively with firm size. The H_0 category supposes that all firms are integrated ($\gamma = 1$), and the H_a category supposes that some firms may still be segregated ($\gamma < 1$). Sample selection and simulation details are described in Sections 4.1.1 and 4.1.2.
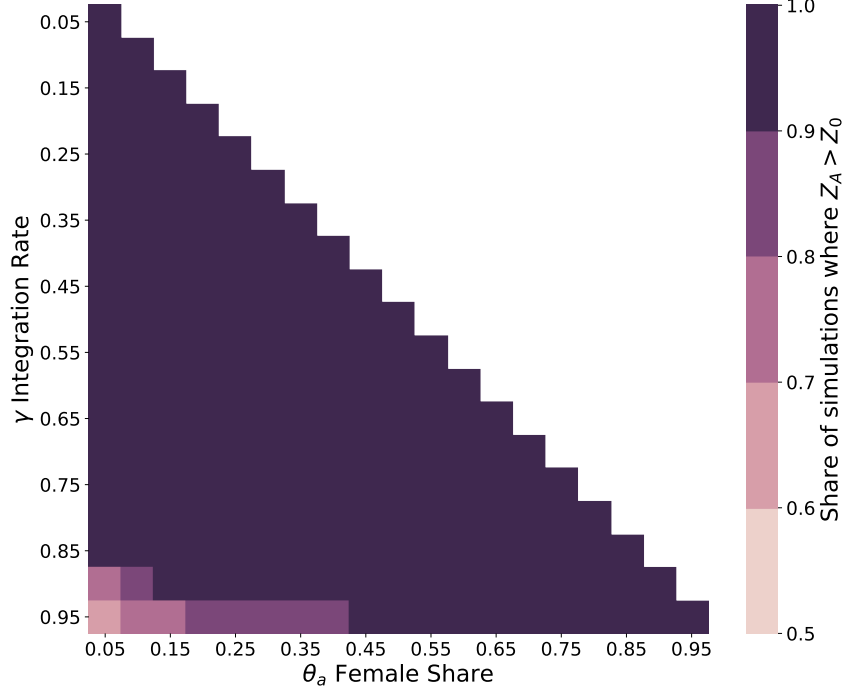
Note: This heatmap plots the share of simulations with $Z_a > Z_0$, or the share of simulations where there are more firms with no female employees under H_0 vs. H_a while varying values of θ and γ and when firm integration probabilities correlate positively with firm size. Sample selection and simulation details are described in Sections 4.1.1 and 4.1.2.

As above, except the largest values of γ ($\gamma \geq 0.9$), $Z_a > Z_0$ for virtually all simulation draws.

F.2   Structural Estimation of θ using Expectation-Maximization

In Section G.1 we modeled the distinction between ex-ante and ex-post integrated firms to structurally estimate $\theta(X_i)$. We use an expectation-maximization algorithm to estimate these parameters. Continuing from Section G.1, the likelihood function for firm j is

$$
P(Y_j = Y) = \begin{cases}
\pi_j + (1 - \pi_j) \prod_{i=1}^{N_j} (1 - \theta_{ij}) & \text{if } K_j = 0 \\
(1 - \pi_j) \prod_{i=1}^{K_j} \theta_{ij} \prod_{K_j+1}^{N_j} (1 - \theta_{ij}) & \text{if } 0 < K_j < N_j \\
(1 - \pi_j) \prod_{i=1}^{N_j} \theta_{ij} & \text{if } K_j = N_j.
\end{cases}
$$

We model both θ_{ij} and π_j in logistic regression models with explanatory variables X_{ij} and Z_j, respectively:

$$\theta_{ij} = \Lambda(X_{ij}\beta)$$
$$\pi_j = \Lambda(Z_j\gamma)$$

where Λ is the logistic function.

The log-likelihood for firm j is

$$\log(f_j) = \log(P(Y_j = Y)) = \begin{cases} -\log\left(1 + e^{Z_j\gamma}\right) + \log\left(e^{Z_j\gamma} + \prod_{i=1}^{N_j}\left(1 + e^{X_{ij}\beta}\right)^{-1}\right) & \text{if } K_j = 0 \\ -\log\left(1 + e^{Z_j\gamma}\right) - \sum_{i=1}^{N_j}\left(1 + e^{X_{ij}\beta}\right) + \sum_{i=1}^{K_j} X_{ij}\beta & \text{if } 0 < K_j < N_j \\ -\log\left(1 + e^{Z_j\gamma}\right) + \sum_{i=1}^{N_j}\left[X_{ij}\beta - \log\left(1 + e^{X_{ij}\beta}\right)\right] & \text{if } K_j = N_j. \end{cases}$$

Combining each firm's log likelihood, we write our log-likelihood function as:

$$l(\beta, \gamma; Y_j, X_{ij}, Z_j) = \sum_{j=1}^{J} \log(f_j)$$

We obtain maximum likelihood estimates of γ and β using the expectation-maximization (EM) algorithm. The EM algorithm is an iterative method to find maximum likelihood estimates, where the model depends on unobserved latent variables. The EM algorithm alternates between an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated at the current estimate for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found in the E step.

For each firm j, let the unobserved random variable I_j indicate whether a firm has ex-ante integrated. When $I_j = 0$, firm j is ex-ante segregated and Y_j is necessarily zero. When $I_j = 1$, firm j is ex-ante integrated. If we could observe I_j for every firm, then the log-likelihood for firm j given complete data (Y_j, I_j) would be

$$\log(f_j | I_j, X_{ij}, Z_j) = (1 - I_j)\underbrace{\left(Z_j\gamma - \log\left(1 + e^{Z_j\gamma}\right)\right)}_{\log_\gamma(f_j | I_j, X_{ij}, Z_j)} + I_j\underbrace{\left[-\sum_{i=1}^{N_j}\log\left(1 + e^{X_{ij}\beta}\right) + \mathbf{1}_{0 < K_j \le N_j}\sum_{i=1}^{K_j} X_{ij}\beta\right]}_{\log_\beta(f_j | I_j, X_{ij}, Z_j)}.$$

Therefore the complete data log-likelihood function is

$$l_c(\beta, \gamma | Y_j, I_j, X_{ij}, Z_j) = \sum_{j=1}^{J} \log(f_j | I_j, X_{ij}, Z_j)$$

$$= \sum_{j=1}^{J} [\log_{\gamma}(f_j | I_j, X_{ij}, Z_j) + \log_{\beta}(f_j | I_j, X_{ij}, Z_j)]$$

$$= l_c(\gamma | Y_j, I_j, X_{ij}, Z_j) + l_c(\beta | Y_j, I_j, X_{ij}, Z_j).$$

The EM algorithm begins with starting values $\omega^{(0)} = (\gamma^{(0)}, \beta^{(0)})$. Our starting value for $\beta^{(0)}$ is derived from estimating the linear regression $Y_j = X_{ij}\beta$ and setting $\beta^{(0)} = \hat{\beta}$. For $\gamma^{(0)}$, we estimate the regression $\mathbf{1}_{K_j > 0} = Z_j \gamma$ and similarly set $\gamma^{(0)} = \hat{\gamma}$.

From these initial values, we proceed iteratively, with (r) indexing the iteration:

- **E Step**: estimate I_j by its conditional mean $I_j^{(r)}$ given $\omega^{(r)} = (\gamma^{(r)}, \beta^{(r)})$:

$$\hat{I}_j^{(r)} = E[I_j | Y_j, X_{ij}, Z_j, \gamma^{(r)}, \beta^{(r)}]$$

$$= \frac{P(Y_j | I_j = 0)P(I_j = 0)}{P(Y_j | I_j = 0)P(I_j = 0) + P(Y_j | I_j = 1)P(I_j = 1)}$$

$$= \begin{cases} \left[\left[1 + e^{-G_j \gamma} \prod_{i=1}^{N_j} \left(1 + e^{X_{ij}\beta} \right)^{-1} \right] \right]^{-1} & \text{if } K_j = 0 \\ 0 & \text{if } K_j \neq 0 \end{cases}$$

- **M Step for** γ: we find $\gamma^{(r+1)}$ by maximizing $l_c(\gamma | Y_j, I_j, X_{ij}, Z_j)$. This can be accomplished by logistic regression of $I_j^{(r)}$ on Z_j. It is equivalent to solving the FOC of $l_c(\gamma | Y_j, I_j, X_{ij}, Z_j)$:

$$\sum_{j=1}^{J} \left(I_j^{(r)} - \frac{e^{Z_j \gamma}}{1 + e^{Z_j \gamma}} \right) Z_j = 0.$$

- **M Step for** γ: we find $\gamma^{(r+1)}$ by maximizing $l_c(\gamma | Y_j, I_j, X_{ij}, Z_j)$. This can be accomplished by logistic regression of $I_j^{(r)}$ on Z_j. It is equivalent to solving the FOC of $l_c(\gamma | Y_j, I_j, X_{ij}, Z_j)$:

$$\sum_{j=1}^{J} \left(I_j^{(r)} - \frac{e^{Z_j \gamma}}{1 + e^{Z_j \gamma}} \right) Z_j = 0.$$

From the above, we obtain estimates for β and γ for iteration (r) and repeat the exercise until $\left\| \beta^{(r+1)} - \beta^{(r)} \right\| + \left\| \gamma^{(r+1)} - \gamma^{(r)} \right\| < 0.0001$.

G   Appendix: Additional Analyses

G.1   Modeling Firm Integration States

The second approach we take is to directly model the distinction between ex-ante and ex-post integrated firms and to structurally estimate $\theta(X_i)$.

Let j index firms, and let N_j denote the number of positions at firm j. Let y_{ij} be an indicator that equals one if position i in firm j is filled by a female employee. Denote $K_j = \sum_i^{N_j} y_{ij}$ as the number of female employees at firm j.

Let π_j denote the probability that firm j has *not* paid its integration cost and so is not able to employ women. Hence, with probability $1 - \pi_j$, the firm is ex-ante integrated. We will model π_j as a function of observable firm characteristics. Finally, among ex-ante integrated firms, denote the probability that position i is filled by a female employee as θ_{ij}. As above, we model θ_{ij} as a function of observable job characteristics, X_{ij}.[43]

With these terms defined, we can define the likelihood function for each firm. Without loss of generality, we order each firm's workers such that the first K_j workers are female and the remaining $N_j - K_j$ are male. Denote $Y_j = (Y_{1j}, ..., Y_{N_j j})$ as the firm-specific vector of outcomes. The likelihood function for firm j is

$$
P(Y_j = Y) = \begin{cases} \pi_j + (1 - \pi_j) \prod_{i=1}^{N_j} (1 - \theta_{ij}) & \text{if } K_j = 0 \\ (1 - \pi_j) \prod_{i=1}^{K_j} \theta_{ij} \prod_{K_j+1}^{N_j} (1 - \theta_{ij}) & \text{if } 0 < K_j < N_j \\ (1 - \pi_j) \prod_{i=1}^{N_j} \theta_{ij} & \text{if } K_j = N_j. \end{cases}
$$

We model both θ_{ij} and π_j in logistic regression models with explanatory variables X_{ij} and Z_j, respectively:

$$
\theta_{ij} = \Lambda(X_{ij}\beta)
$$
$$
\pi_j = \Lambda(Z_j\gamma)
$$

where Λ is the logistic function. In the vector of firm characteristics, Z_j, we include fixed effects for location and industry and a cubic in log firm size.[44] For the vector of hire characteristics, X_{ij}, we include fixed effects for two-digit occupation codes, location, and one-digit industry. We estimate the model using an expectation-maximization (EM) algorithm. Estimation details are provided in Appendix F. We label these structural estimates for $\theta(X_i)$ as $\hat{\theta}^S(X_i)$.

Column (3) of Table 4 summarizes the estimates and how they vary across jobs. The average value of $\hat{\theta}^S(X_i)$ is 0.123. These estimates are similar to those from Section 4.2 using only ex-post integrated firms; the correlation between $\hat{\theta}^S(X_i)$ and $\hat{\theta}^{EP}(X_i)$ is 0.82. The average value of π_j is 0.65, indicating 65% of firms are ex-ante segregated.

---

[43]For ease of notation, in this section we index positions separately by firm.
[44]We measure firm size here using the firm's number of Saudi employees.

G.1.1 Can $\theta(X_i)$ Match the Distribution of Female Employment?

As an additional test for whether $\theta(X_i)$ is well specified, we evaluate whether a simulation of the distribution of female employment across firms that allows for integration rates to vary by $\bar{\theta}_j n_j$ fits the observed distribution. For each firm, we take a uniform random draw and label the firm as integrated if the draw is below the corresponding values in Panel B of Figure 2 given the firm's value of $\bar{\theta}_j^{EP} n_j$. If the firm is not labeled as integrated, we assign it a value of zero for its female employment. For firms labeled as integrated, we simulate a value of female employment as above, this time using $\hat{\theta}^{EP}(X_i)$ to assign the gender for the employee in each position.

Panel B of Appendix Table A.1 compares the simulated distribution of female employment to the observed distribution for various specifications of $\theta(X_i)$. While, by construction, we will match the share of firms with zero female employees, the simulation is not guaranteed to match other parts of the distribution. Yet, our baseline specification, where X_i includes job location, two-digit occupation, and one-digit industry fixed effects, matches the observed distribution. Across all simulations, we fail to reject equality of the simulated and observed distributions in a Kolmogorov-Smirnov test at the 1% significance level. The average p-value is 0.75. This suggests that we have included the most relevant job characteristics in X_i (or that other relevant characteristics are not concentrated within firms) and have mapped them appropriately to hiring probabilities, at least among ex-ante integrated firms.

G.1.2 Checking Predictions for Counterfactual Female Employment

First, we test whether our estimate of $\theta(X_i)$ provides unbiased predictions for the female share of hires at newly integrated firms. This is a powerful out-of-sample test for whether our estimate of $\theta(X_i)$ predicts counterfactual female employment at segregated firms because we do not use this set of firms to estimate $\theta(X_i)$. We also examine the transition dynamics of these newly integrated firms.

We first examine hiring at newly integrated firms in an event study. We plot the female share of new hires at integrating firms in the months following a firm's first observed female hire.[45] We limit to firms with at least five Saudi employees in the month prior to integration. We observe 8,307 transitioning firms meeting this size threshold. Prior to integrating, we observe transitioning firms in the GOSI data for an average of 39 months. At each firm, we calculate the female share of all new hires made in six-month increments before and after a firm's first female hire.[46] We then take the average across all firms meeting the sample restrictions and exclude firms that do not make a hire in a given six-month increment from the calculation for that period.

The event study is shown in Panel A of Figure G.1. By construction, among hires made prior to integration, there are no women. Among firms that we observe integrating, we observe an average of 33 male hires made prior to a firm's first female hire. We see that the female share of hires changes abruptly at newly integrated firms, consistent with an extensive margin response.[47] Among hires made in the six months following integration, including the first female hire, 55% are female. This drops to about 26% in the following six-month period and remains relatively steady thereafter. By contrast, if the excess mass of firms with zero female workers we observed in Figure 1 was driven by unobserved heterogeneity in job characteristics, we would expect a gradual and potentially short-lived increase in female hiring rather than the discrete and sustained increase we observe.
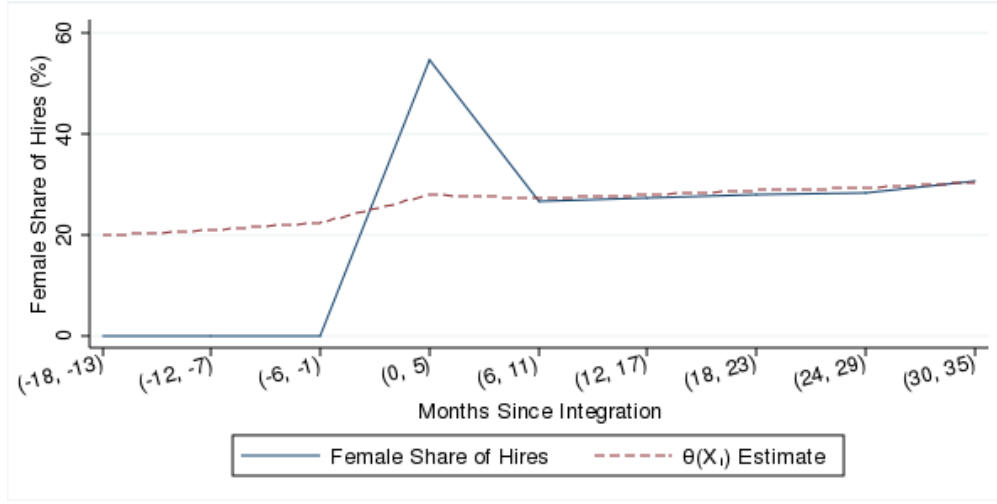
---

[45]In this exercise, we exclude firms that have female employees when they are first observed in the GOSI data.

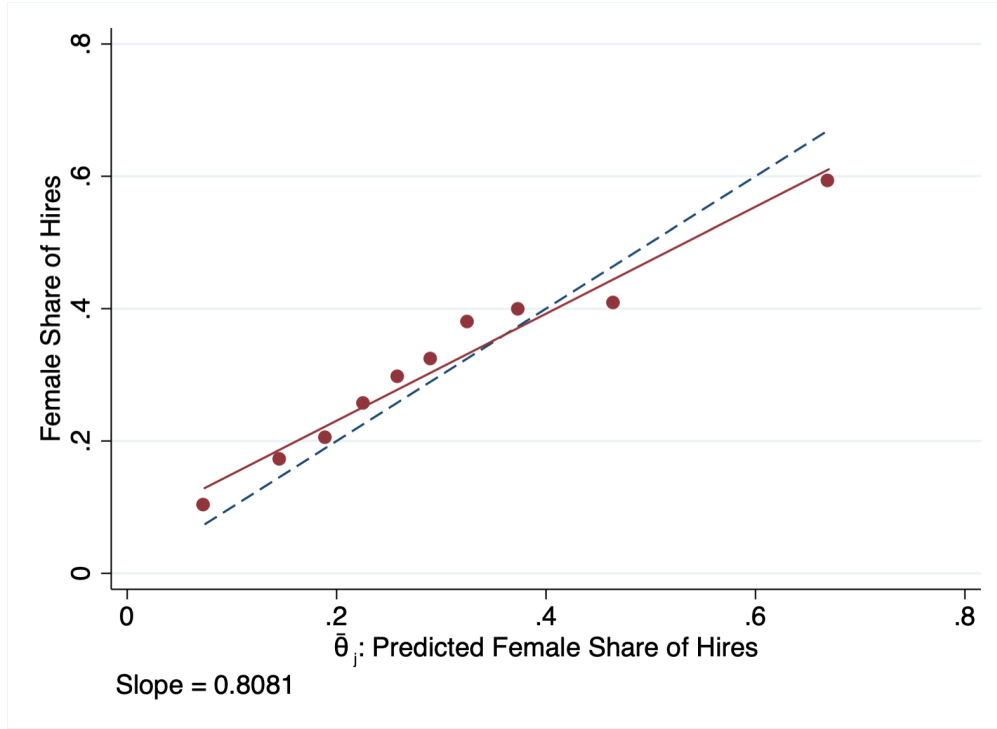[46]Hires include any employee that begins a new job spell at the firm in a given period.

[47]Note that the period labeled as "0 to 5" months includes the first female hire herself.

(a) Female Share of Hires

(b) Observed Versus Predicted Female Share of Hires

Slope = 0.8081

Note: This set of figures describes the gender composition of hires at newly integrated firms. Panel A plots the female share of hires made at integrating firms in six-month increments relative to a firm's first observed female hire, averaged across firms. Using the GOSI data, we restrict to firms with at least five Saudi employees in the month prior to integration. Panel B compares the female share of hires at newly integrated firms to their $\theta(X_i)$-based predicted values, where $\theta(X_i)$ is estimated using firms that are ex-post integrated in January 2009. $\theta(X_i)$ estimation details are provided in Section G.1.2. The vertical axis depicts the female share of hires that are made 12 or more months following a firm's first female hire. The horizontal axis depicts the $\theta(X_i)$-based prediction for this value.

We compare the observed increase to what we would predict using an estimate of $\theta(X_i)$ derived from hires at incumbent integrated firms. As in Section 4.2, we construct our predictions by estimating a logistic regression of the form:

$$P(\text{Worker } i \text{ is female}) = \Lambda(X_i\beta),$$

where i indexes the position for each hire. As above, X_i includes fixed effects for job location, two-digit occupation, and one-digit industry. In addition, we allow predictions to vary over time by including in X_i fixed effects for each half-year and interactions between the location, occupation, and industry controls with an indicator for hires made after June 2011, the month Nitaqat is implemented. We limit estimation to all firms that are ex-post integrated in January 2009, regardless of whether any of their subsequent hires are female. These firms should provide a valid estimate for $\theta(X_i)$ if integrated firms remain ex-ante integrated moving forward, an assumption we verify in the next section. We label this estimate $\hat{\theta}(X_i)$.

We include the $\hat{\theta}(X_i)$-based prediction for the female share of hires in Panel A of Figure G.1. We find that the magnitude of this change matches what we would predict using $\hat{\theta}(X_i)$, at least on average. Next, we check how well firm-specific predictions for the female share of hires of newly integrated firms matches the realized female share of hires. In Panel B of Figure G.1, we group newly integrated firms into deciles based on the predicted female share of hires and plot bin averages against their observed female share of hires 12 or more months following their first female hire. If the predictions are *unbiased*, the binned averages will fall on the 45-degree line. This is similar to the pattern we observe, though the observed female share of hires is slightly below the 45-degree line, with the gap increasing in the predicted female share of hires.

G.2 State Dependence

An immediate implication of the model is state dependence: the hiring behavior of firms that have already paid their integration costs will differ from the behavior of firms that have not. In particular, we should *not* observe bunching for the former set of firms. While we cannot observe each firm's current state, we proxy for their current state using their baseline ex-post segregation status. This proxy should closely correlate with a firm's current state if integration costs are sunk or if the conditions that led the firm to integrate are highly persistent over time. We test the null hypothesis of no binding integration costs but conduct separate tests for firms that are ex-post integrated and ex-post segregated as of January 2009.

We conduct a similar test to that described in Section 4.1, except we pool *hires* between February 2009 and June 2015. We limit to firms that have at least five Saudi hires over this period. To classify firms as ex-post integrated or segregated in January 2009, we also limit to firms that had Saudi employees in January 2009. We estimate $\theta(X_i)$ separately by baseline integration status and include the same job characteristics we use in Section G.1.2: fixed effects for each half-year and fixed effects for job location, two-digit occupation, and one-digit industry, all interacted with an indicator for hires made after June 2011.

Table G.1 compares the two sets of firms. There are 2,796 firms meeting the sample criteria that were ex-post integrated in January 2009 ("baseline integrated") and 12,617 firms that were ex-post segregated ("baseline all male"). Baseline integrated firms are larger, pay higher wages, and concentrated in community and Social services. For baseline all-male and integrated firms, the female share of recent hires is 19.2% and 48.4%. Figure G.2 plots the simulated and observed distribution of female employment for baseline all-male (Panels A and B) and integrated firms (Panels C and D).

FIRM DESCRIPTIVE STATISTICS, BY BASELINE INTEGRATION STATUS

	Baseline all male	Baseline integrated
# of firms	12,617	2,796
Number of Saudi hires		
Mean	61.9	219.7
Median	18	47
SD	194.6	930.9
Female share of hires (%)		
Mean	19.2	48.4
Median	10.0	46.3
SD	23.4	33.4
Avg. monthly wage (Riyals)	3,238	3,709
Industry (%):		
Agriculture and fishing	1.0	0.3
Commerce	29.7	20.6
Community/social services	6.9	41.6
Construction	30.0	11.4
Electricity, gas, and water	0.8	0.9
FIRE	9.9	12.5
Manufacturing	16.3	10.4
Mining	1.4	0.6
Telecommunications	3.9	1.7

Notes: This table presents descriptive statistics for firms with any Saudi employee in January 2009 that hire at least five Saudis between February 2009 and June 2015 in the GOSI data. The first column includes firms that are all-male in January 2009. The second column includes firms that are ex-post integrated in January 2009. The average wage at a firm is measured in nominal Saudi Riyals at the time of hiring.
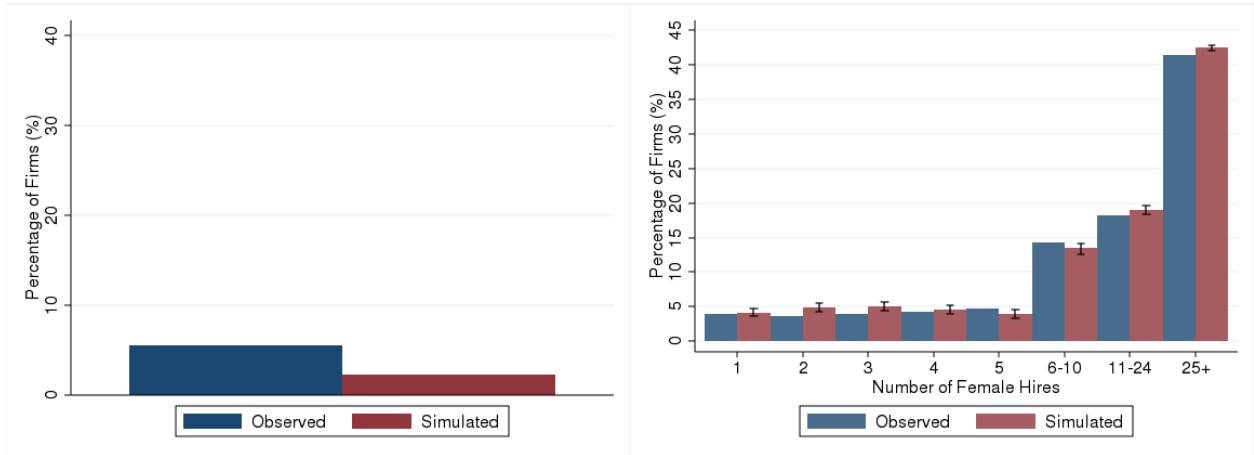
FIGURE G.2

DISTRIBUTION OF FEMALE HIRING, BASELINE ALL-MALE AND INTEGRATED FIRMS

(a) % of Baseline All-Male Firms with Zero Female Hires   (b) % of Baseline All-Male Firms with > 0 Female Hires
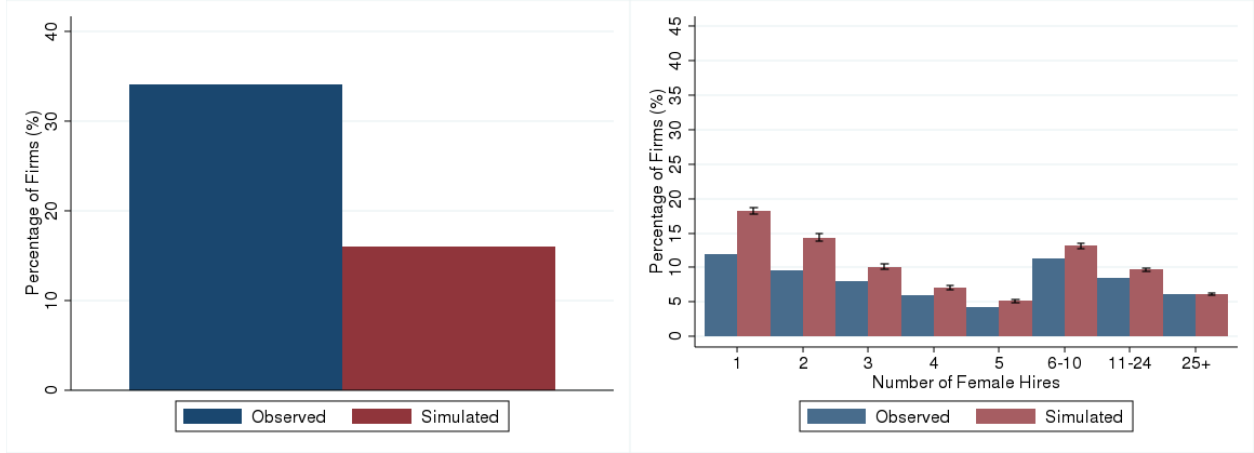
(c) % of Baseline Integrated Firms with Zero Female Employees   (d) % of Baseline Integrated Firms with > 0 Female Hires

Note: This set of figures compares the observed and simulated distributions of female hires across firms that are (1) ex-post segregated in January 2009 and (2) ex-post integrated in January 2009 for hires made between February 2009 and June 2015. The simulated distributions are simulated under the null hypothesis that no firm faces binding integration costs over the hiring period. Sample selection and simulation details are described in Sections G.2. Panels A and B plot simulation results for firms that are ex-post segregated in January 2009. Panels C and D plot simulation results for firms that are ex-post integrated in 2009. Panels A and C plots the share of firms with zero female hires in both the observed and simulation distributions. Panel B plots the share of firms with various nonzero totals of female hires in both the observed and simulated distributions. For all simulations for firms that are ex-post segregated in January 2009, a Kolmogorov-Smirnov test rejects equality of the observed and simulated distributions at the 1% significance level. For firms that are ex-post integrated in January 2009, across simulations we fail to reject equality of distributions with an average p-value of 0.10. Sample selection and simulation details are described in Sections 4.1.1 and 4.1.2.

For baseline all-male firms, the pattern is similar to that observed in Figure 1. The simulations underpredict the number of firms that employ zero female workers (16.1% versus 34.2%) and over-predict the number of firms that employ fewer than ten (68.1% versus 51.2%). For all simulations, a Kolmogorov-Smirnov test rejects equality of the observed and simulated distributions at the 1%

significance level.

By contrast, the simulated distribution for baseline integrated firms matches the observed distribution relatively well. Across all simulations, a Kolmogorov-Smirnov test fails to reject equality of the observed and simulated distributions at the 1% significance level. The average p-value is 0.10.

Consistent with our interpretation of bunching as evidence for the presence of ex-ante segregated firms, there is little evidence of bunching at firms that are likely ex-ante integrated.

H   Appendix: Aggregate Effects of Nitaqat

It is unclear what would happen in the aggregate if integration costs were eliminated across the labor market. Would aggregate demand for female labor increase or the female share of the workforce and gender differences in wages change? As in Becker (1957), integrated firms may be sufficiently numerous or large to absorb female labor so that the existence of constrained male-only firms has no bearing on female wages and employment. On the other hand, in the presence of search frictions or insufficient entry or growth of integrated firms, integration costs will reduce aggregate demand for female labor.

   To assess the aggregate consequences of integration costs, exogenous variation in integration costs across labor markets is needed. Lacking such variation, we take a different approach. We examine the labor market response to a policy that reduces the share of firms that face binding integration costs and assess the effects of the policy on female employment and the gender wage gap. If the policy increases female employment or relative wages, this would suggest that the presence of integration costs depresses those outcomes. The logic of our approach is to essentially use the policy as an instrument for (binding) integration costs. The exclusion restriction implicit in our argument is that the policy only affects our outcomes of interest by reducing the set of firms with binding integration costs. We discuss this exclusion restriction below.

   In particular, we investigate the aggregate effects of Nitaqat. As discussed in Section 4.4, Nitaqat induced many firms to integrate and hire women. We show this in Figure H.1, which plots over time the share of firms with at least five Saudi employees that employ both Saudi men and women. There is a clear trend break that begins just as Nitaqat is implemented, followed by a flattening which occurs soon after a doubling of the effective minimum wage for Saudis in the private sector. We discuss the effects of this minimum wage increase in more detail below.

FIGURE H.1
INTEGRATION RATES AND FEMALE SHARE OF WORKFORCE OVER TIME
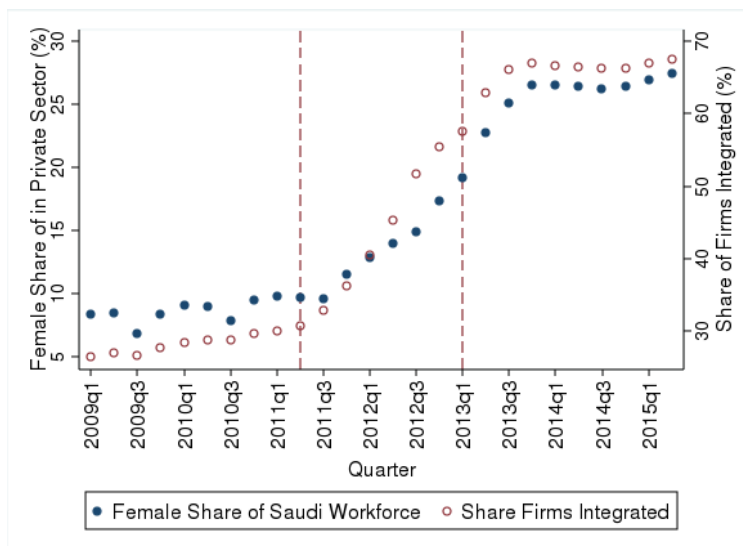
Note: This figure plots the female share of full-time Saudi workers and the share of firms that employ both men and women, both on a quarterly basis from the GOSI matched employee-employer data. For the latter outcome, firms are restricted to those with at least five Saudi employees. The vertical lines correspond to the implementation of Nitaqat (in Q2 of 2011) and the increase in the de facto minimum wage (in Q1 of 2013). This figure does not include employees in the security and military sectors. Source: Ministry of Civil Service.

Next, we explore how the female share of the workforce evolves in response to Nitaqat. In Section 4.4 we document that among firms that are all male in January 2009, Below Quota firms increase their female share of hires relative to Above Quota firms. Figure H.1 plots the female share of Saudis in the private sector over time, pooling employment at all firms. The overall pattern matches that of integration rates. Nitaqat led to a dramatic increase in the female share of Saudis in the private sector, from 10% in 2011 to 27% in 2015. This increase occurs primarily within sectors, as measured by industry and occupation. This increase in the female share of Saudis in the private sector is not offset by a decrease in the public sector; in fact, the female share of Saudis working in the public sector also increases over this period, from 33% in 2011 to 40% in 2015.

While this increase is striking, it does not necessarily indicate an important role for integration costs. Nitaqat may also increase the female share of employment through a *price effect*. Suppose aggregate male labor supply is inelastic relative to aggregate female labor supply, perhaps due to the relative scarcity of available and qualified male workers. Then Nitaqat may bid up men's relative wages, increasing relative demand for female labor. This would be a violation of our exclusion restriction: a path through which Nitaqat increases the female share of employment that is unrelated to integration costs per se.

However, the evidence suggests that changing prices are not the driving force behind the dramatic increase in the female share of the workforce. In fact, the gender wage gap *decreases* following Nitaqat. Moreover, after the effective minimum wage reduces the wage gap even further, the female share of the workforce remains elevated. This is illustrated in Figure H.2, which plots the female-male wage gap over time. The figure includes two measures of the gender wage gap: (1) the raw difference in average log wages for women and men and (2) the gap within labor market entry cohorts.

FIGURE H.2
GENDER WAGE GAP OVER TIME
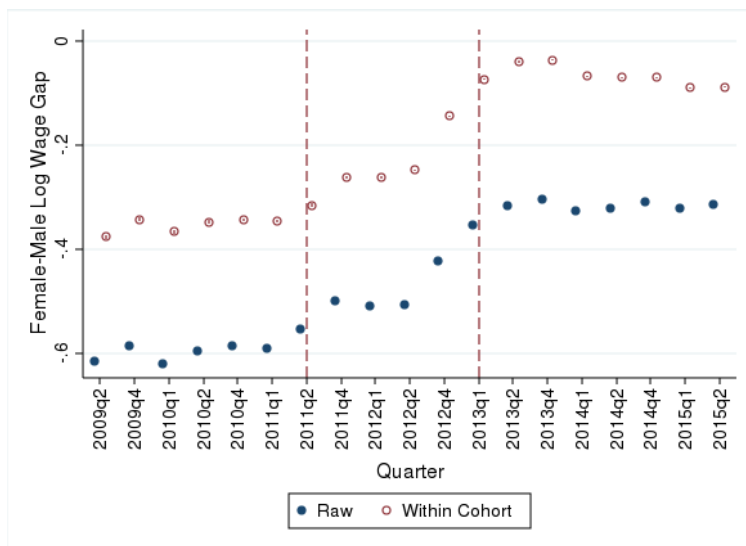
Note: This figure plots the female-male log wage gap on a quarterly basis. It includes both the raw log wage gap and the log wage gap controlling for cohort fixed effects, where cohorts refer to the year of the earliest start date for a worker as recorded in the GOSI data. The veritcal lines correspond to the implementation of Nitaqat (in Q2 of 2011) and the increase in the de facto minimum wage (in Q1 of 2013).

57

Prior to Nitaqat, the wage gap is relatively flat. The raw wage gap is 60 log points; within cohorts, the gap is about 35 log points. Following Nitaqat, but prior to the minimum wage increase, the wage gap decreases by about 10 log points. The 2013 minimum wage increase leads to a substantial reduction in the gender wage gap. Following its introduction, about 65% of women and 40% of men earn the new minimum wage. The raw wage gap drops to about 30 log points. Within cohorts, the wage gap drops to 4–9 log points. Yet, from Figure H.1, we can see that the female share of the private sector workforce is *increasing* over this period. This share stagnates beginning in 2013 but remains elevated thereafter. The fact that both female relative wages and employment increase is difficult to reconcile with a price-based explanation. Instead, the evidence is consistent with Nitaqat increasing relative demand for female labor by increasing the set of firms that integrate.

Finally, we explore the possibility that Nitaqat led to a shift in women's labor *supply*. Female labor force participation in Saudi Arabia is among the lowest in the world, at 17.8% in 2011 (GaStat, 2011). While this low rate is likely driven by multiple factors, one may be that households perceive that few firms are willing to hire women in the first place. Nitaqat may cause an outward shift in women's labor supply by increasing the set of firms that are ex-ante integrated. In fact, integration costs as a barrier to women's employment may generate feedback effects: women may only enter the labor market if enough firms have integrated, while firms only integrate if they can anticipate employing enough women to justify the costs of integration.

Unfortunately, we do not have data on labor supply decisions; in particular, we do not have data on anyone that is not employed in the private sector. Instead, we look at the response to Nitaqat for firms that had integrated prior to the policy's implementation. While Figure H.1 shows that the female share of the workforce is increasing, we expect this increase to be concentrated at firms induced to integrate by the policy. Firms that integrated prior to Nitaqat are already employing a mix of men and women and face an increase in the relative price of women. In the absence of a supply response, we would expect to see the female share of employment at these firms weakly decreasing.

Figure H.3 plots the female share of employment over time in firms that employed Saudis in January 2009, split by the firm's ex-post integration status in that month. For *both* sets of firms, there is a marked increase in the female share of employment beginning with Nitaqat's integration. As expected, the increase is larger for baseline-segregated firms. But for baseline integrated firms, the increase is also substantial: from 14.5% in Q1 of 2011 to 20.6% of Q1 2015.

FEMALE SHARE OF WORKFORCE BY BASELINE INTEGRATION STATUS
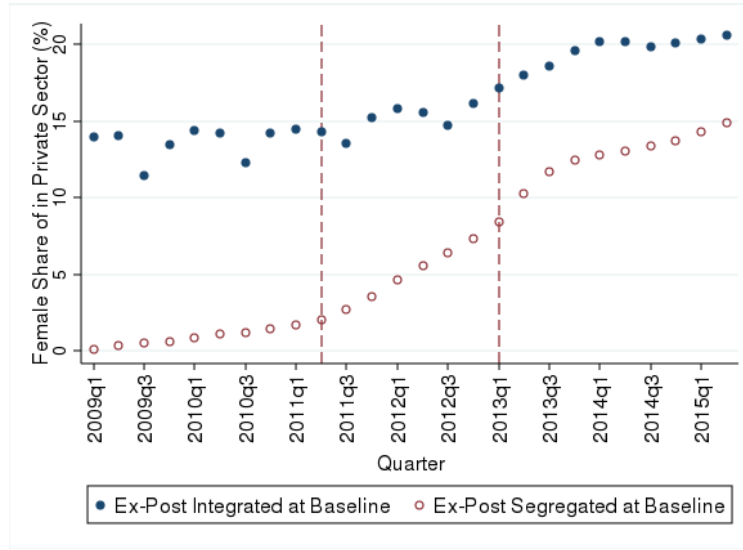
Note: This figure plots the female share of Saudi employment in the GOSI matched employee-employer data over time in firms that employed Saudis in January 2009, split by the firm's ex-post integration status in that month. These shares are measured on a quarterly basis. The vertical lines correspond to the implementation of Nitaqat (in Q2 of 2011) and the increase in the de facto minimum wage (in Q1 of 2013).