Integration Costs and Missing Women in Firms around the World

By Conrad Miller, Jennifer Peck and Mehmet Seflek*

The female share of the labor force varies dramatically across regions, ranging from 46% in sub-Saharan Africa to 20% and 24% in the Middle East and North Africa (MENA) and South Asia.[1] Prior research argues that this variation can be explained in part by social norms regarding gender roles and their effects on labor supply decisions (e.g., Fernandez and Fogli, 2009; Alesina, Giuliano and Nunn, 2013). Recent work shows that these norms also constrain *firm* behavior and can have persistent effects on aggregate female labor *demand* (Miller, Peck and Seflek, 2022; Jayachandran, 2021).

In particular, where there are social norms that favor gender segregation, firms face additional costs to *integration*, defined here as employing both women and men. Workers, customers, and regulators may expect firms to establish gender-segregated facilities, including restrooms and workspaces. Firms may also segregate tasks or teams to limit interactions between men and women. For male-dominated firms, hiring women may necessitate changes in their workplace culture. In World Bank Enterprise 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. Firms may also need to restructure tasks and occupations to enable female employment or to reorganize their production process to comply with restrictions on women's working hours.[2] These types of restrictions all constrain the production process for integrated firms. A distinguishing feature of the costs associated with accommodating social norms is that they often have a significant *fixed* component. We study the consequences of these integration costs for women's employment.

The distribution of female employment across firms provides prima facie evidence that integration is costly for firms. In most of the world, all-male firms are rare—in sub-Saharan Africa, East Asia and the Pacific, Eastern and Central Europe, and Latin America and the Caribbean, only 2%–11% of medium (20–99 employee) manufacturing firms are all-male, and 1%–2% of large (100+ employee) firms are all-male.[3] This is unsurprising given that women make up 27%–41% of employees in these firms overall. However, all-male firms are dramatically more common in MENA and South Asia: 48% and 50% of medium firms and 23% and 29% of large firms are all-male. While women make up a smaller share of employees in surveyed firms in these regions (17% and 15%), the share of firms that are all-male is substantially larger than what one would expect by chance.[4] We show that this cannot be explained by differences in the female share of workers across regions or occupations. Instead, the mass of firms with zero female employees in MENA and South Asia

* Miller: Haas School of Business, UC Berkeley, Berkeley, CA 94720-1900, ccmiller@berkeley.edu. Peck: Swarthmore College, 500 College Avenue, Swarthmore PA 19081, jpeck1@swarthmore.edu. Seflek: Haas School of Business, UC Berkeley, Berkeley, CA 94720-1900, mseflek@berkeley.edu.

[1] As measured by the World Bank and International Labour Organization in 2018. According to the World Bank, unpaid workers, family workers, and students are often omitted from these figures.

[2] The majority of countries in MENA and South Asia prohibit women from working at night, with such regulations typically justified by concerns over safety (International Finance Corporation, 2013; World Bank, 2018).

[3] See calculations in Miller, Peck and Seflek (2022). All-female firms are even more rare: less than 1% of medium and large firms are all-female.

[4] For example, if the gender of each employee were an independent draw from a binomial distribution where the probability an employee is female is 0.169 (the female share of manufacturing employment in MENA), the probability that a firm with 50 employees is all-male is $(1 - 0.169)^{50} = 0.0001$.

strongly suggests that firms face an extensive margin decision of whether to integrate their workforce. We deploy methodology developed in Miller, Peck and Seflek (2022) that uses the distribution of female employment across firms to estimate the share of firms with binding integration costs and counterfactual female employment at all-male firms. We find evidence of significant costs in MENA and South Asia, consistent with regional variation in social preferences for gender segregation. Integration costs are more binding in countries with stronger gender-segregation preferences, and our measures of these costs are negatively correlated with female labor force participation (LFP).

I. Data

We use cross-country data from the World Bank Enterprise Surveys, which survey private sector companies across 139 countries between 2005 and 2018. We limit our sample to manufacturing firms, where surveys record the gender composition of employees by occupation (production or nonproduction). We drop surveys where information on gender composition is missing for more than 20% of firms, firms with missing data on gender composition or with fewer than five employees, and then drop surveys with fewer than 100 remaining firms. This leaves 105 surveys in 65 countries.[5] Survey participants are sampled by stratified random sampling with firm size categories, sectors, and within-country geographic regions as the strata. We use the provided survey weights so that firms are representative of the manufacturing sector in a given country and year.

We link our measure of integration costs to social preferences using data from the Arab Barometer, a cultural, religious, and political opinion survey that is run periodically across several MENA countries. The survey is designed to cover a representative sample of adults within each country and records gender segregation preferences, with all four survey groups asking about agreement or disagreement with a statement regarding gender mixing in university classes.[6] We use these responses to measure the share of respondents in each country that approve of gender mixing and match this information to the World Bank data.

II. Empirical Strategy and Results

If the costs of integration are largely fixed, firms will integrate only if their expected number of female employees under integration exceeds some threshold. Firms that have not paid these fixed integration costs are *ex-ante segregated*, i.e. face some binding fixed cost of integration. Our analysis follows the methodology developed in Miller, Peck and Seflek (2022) to test for integration costs and calculate ex-ante integration rates in Saudi firms. The empirical strategy is based on a partial equilibrium model of firm hiring in which (1) firms face an extensive margin integration decision (i.e. decide ex-ante whether to incur the integration costs associated with hiring both men and women) and (2) integration costs are largely fixed, so firms integrate only if they anticipate employing enough women to justify the costs. The probability that the top candidate for position i is female is modeled by a function $\theta(\cdot)$ of observable job characteristics, X_i. To test the null that all firms are ex-ante integrated, we estimate $\theta(X_i)$ using data on employees at all firms, simulate the distribution of female 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 then re-estimate $\theta(X_i)$ and use this estimate to measure ex-ante integration rates as a function of expected female employment under integration.

[5] Appendix Table A2 lists the surveys we include.

[6] The statement is "It is acceptable in Islam for male and female university students to attend classes together" in waves I (2006-2009) and IV (2016-2017), and "Gender-mixed education should be allowed in universities" in II (2010–2011) and III (2012–2014).

A. Testing the Null of No Binding Integration Costs

We begin by testing the null that integration costs are nonbinding for all firms separately by region. Following the procedure developed in Miller, Peck and Seflek (2022), we calculate a separate $\theta(X_i)$ function for each survey (country-year pair), estimating a logistic regression model of the form

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

and include location and occupation type (production or nonproduction) in X_i. We then simulate the distribution of female employment across firms under the null hypothesis given the estimated $\hat{\theta}^0(X_i)$ and compare that to the observed distribution.

Figure 1 compares the distributions for Ethiopia and India.[7] In Ethiopia our simulation matches the actual distribution reasonably well. In particular, we do not substantively underpredict the number of all-male firms. By contrast, the pattern in Egypt is similar to the results from Miller, Peck and Seflek (2022) in Saudi Arabia: we substantially underpredict the number of firms with zero female employees. In Egypt, we predict that 22% of firms will be all male; in fact, 63% are. We also overpredict the number of firms with few female employees.

One concern with this comparison across countries is that we may see more bunching at zero in India and Egypt simply due to *censoring* at zero. To address this concern, we examine the same distributions for larger firms, where almost no firms would have zero female employees under the null. In Appendix Figure A2, we plot the same distributions for firms with at least 50 employees. Across all countries, virtually no firms have zero female employees in our simulations, and in Ethiopia, China, Russia, and Brazil, virtually none of these firms are all male in practice. Yet, about 30% of such firms in India and Egypt are all male.

If some firms are ex-ante segregated, then $\hat{\theta}^0(X_i)$ will underestimate $\theta(X_i)$, since some of these segregated firms are included in its estimation. To correctly estimate $\theta(X_i)$ we therefore repeat the exercise limiting the data to integrated firms. This yields two functions that approximate $\theta(X_i)$ for ex-ante integrated firms: one that uses data on *ex-post* integrated firms ($\hat{\theta}^{EP}(X_i)$) and another that uses a structural model to jointly estimate $\theta(X_i)$ and the probability that a firm is ex-ante integrated as a function of firm characteristics, ($\hat{\theta}^S(X_i)$).[8] We use these approximations of $\theta(X_i)$ in the remainder of the analysis.

B. Ex-Ante Integration Rates

We next calculate ex-ante integration rates as a function of $\bar{\theta}_j n_j$, a firm's expected number of female employees if ex-ante integrated. We use our estimates of $\theta(X_i)$ to estimate counterfactual female employment at non-integrated firms:

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

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. We estimate $\theta(X_i)$ and $\bar{\theta}_j$ using ex-post integrated firms. In Figure 2, we plot ex-ante integration rates for our six most populous countries in the World Bank data. In Ethiopia, China, Russia, and Brazil, ex-ante integration rates are close to 100% for all values of $\bar{\theta}_j^{EP} n_j$. Ex-ante integration rates are uniformly lower in Egypt and India. As shown below, this pattern corresponds with more general regional differences in integration.

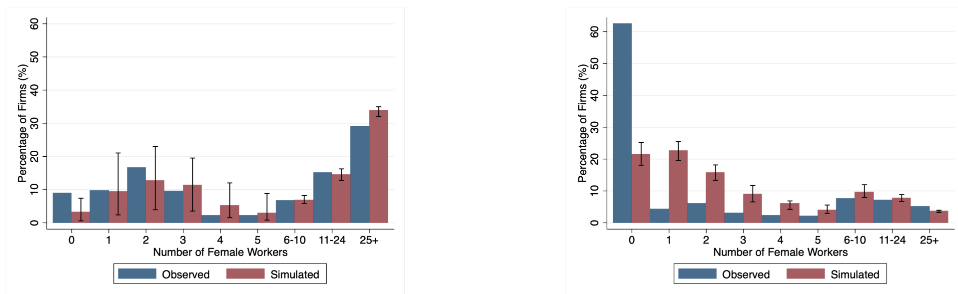
We construct three measures to summarize integration costs by country. We calculate average ex-ante integration rates in two ways: first using $\hat{\theta}^{EP}(X_i)$ based on data from ex-post integrated firms and second using $\hat{\theta}^S(X_i)$ from overall ex-ante integration rates estimated using the structural approach described in Miller, Peck and Seflek

[7] Appendix Figure A1 compares the distributions for the largest countries in each region represented in our data: Ethiopia, China, Russia, Brazil, India, and Egypt.

[8] See Miller, Peck and Seflek (2022) for details on how this structural model is specified and estimated.

Figure 1. : Distribution of Female Employment across Firms: Ethiopia and Egypt

(a) Ethiopia (b) Egypt



Note: These figures compare observed and simulated distributions of female employment across firms in Ethiopia and India. The simulated distributions are simulated under the null that no firms face binding integration costs.

(2022). For our third measure, we calculate the implied ex-ante integration rate for a "representative" firm with $\bar{\theta}_j^{EP} n_j = 10$.

Social preferences are correlated with our measures of ex-ante integration rates within the MENA region. Figure 3 plots ex-ante integration rates by country (derived using ex-post integrated firms) against support for gender mixing from the Arab Barometer surveys. Countries with less support for gender mixing have lower ex-ante integration rates. This is consistent with our motivating hypothesis that integration is costly for firms where social norms for gender segregation are strong.

Ex-ante integration costs are also correlated with women's labor market outcomes across countries. Figure 4 reports OLS estimates of a regression of LFP measures on all three measures of integration costs and plots country-level female LFP rates against overall ex-ante integration rates (derived using ex-post integrated firms). There are two main points to note. First, ex-ante integration rates are lowest in MENA and South Asia. Ten of the fourteen MENA and South Asian countries in our data have average ex-ante integration rates below 60%, and no other countries fall in this range. Second, ex-ante integration rates are positively correlated with female LFP rate, both in absolute terms (columns (1) and (2)) and relative to male LFP (columns (3) and (4)).[9] The positive relationship between ex-ante integration rates and women's economic engagement also holds within regions (see columns (2) and (4), specifications that include region fixed effects).

III. Conclusion

In countries where there are social norms for gender segregation, firms face costs to employing both men and women that are largely fixed. We use a joint test for whether integration costs bind for any firm and a methodology for evaluating the firm-level consequences of those costs using cross-sectional World Bank survey data covering 65 countries. We find that a large fraction of firms in MENA and South Asia employ only men due to integration costs, but integration costs do not constrain firms in other regions. This is consistent with regional variation in social preferences for gender segregation. Within MENA, we find that variation in integration costs across countries is consistent with local preferences. Our results suggest that integration costs prevent some firms from hiring superior female candidates, with significant consequences for women's labor market outcomes. Overcoming these barriers may be key to increasing women's employment and labor force participation in areas where these constraints bind.

[9] The relationship is similar when we examine employment rates rather than LFP rates (see Appendix Table A3).

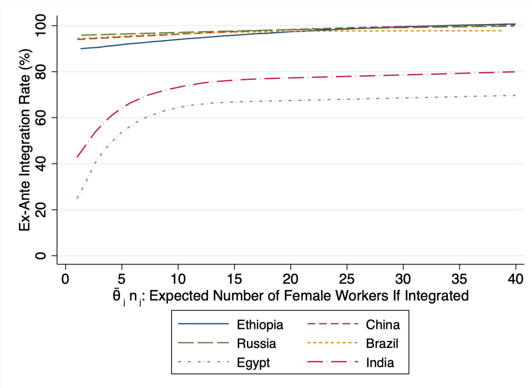


Figure 2. : Integration Rates by $\bar{\theta}_j n_j$

Note: This figure depicts the relationship between ex-ante integration rates and $\bar{\theta}_j n_j$, a firm's expected number of female employees if ex-ante integrated.

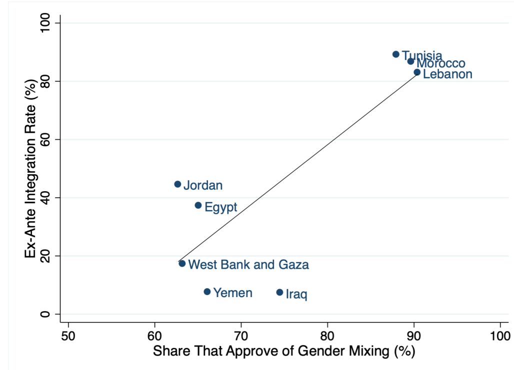
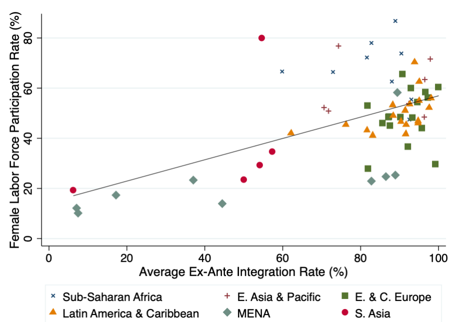


Figure 3. : Integration Rates and Support for Gender Mixing in MENA

Note: This figure plots ex-ante integration rates $(\hat{\theta}^{EP}(X))$ against local support for mixed-gender university classes across MENA countries.



	LFP_F		$LFP_F - LFP_M$	
Ex-ante integration rate:	(1)	(2)	(3)	(4)
Overall $(\hat{\theta}^{EP})$	0.425**	0.278**	0.453**	0.298**
	(0.079)	(0.089)	(0.067)	(0.075)
Overall $(\hat{\theta}^{S})$	0.615**	0.287**	0.638**	0.292*
	(0.133)	(0.142)	(0.117)	(0.123)
Rep. firm	0.489**	0.337**	0.496**	0.355**
$(\bar{\theta}_j^{EP} \cdot n_j = 10)$	(0.081)	(0.093)	(0.070)	(0.077)
Region FEs		✓		✓
Observations	65	65	65	65

Figure 4. : Female Labor Force Participation and Integration Rates Across Countries

Note: This figure plots female LFP rates against ex-ante integration rates for 65 countries. Ex-ante integration rates are derived using country-specific $\hat{\theta}^{EP}(X_i)$, with estimates of $\theta(X_i)$ derived using ex-post integrated firms. Table reports OLS estimates of regressions of female LFP measures on ex-ante integration rates. LFP_F is the average percentage of women age 15+ in the labor force, and $LFP_F - LFP_M$ is the difference between female and male LFP. Robust standard errors are in parentheses. ˜ p<0.1; * p<0.05; ** p <0.01.

REFERENCES

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

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

International Finance Corporation. 2013. *Investing in Women's Employment: Good for Business, Good for Development.* World Bank.

Jayachandran, Seema. 2021. "Social Norms as a Barrier to Women's Employment in Developing Countries." *IMF Economic Review.*

Miller, Conrad, Jennifer Peck, and Mehmet Seflek. 2022. "Missing Women, Integration Costs, and Big Push Policies in the Saudi Labor Market." *American Economic Journal: Applied Economics*, 14(2).

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

Online Appendix

Table A1—: 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

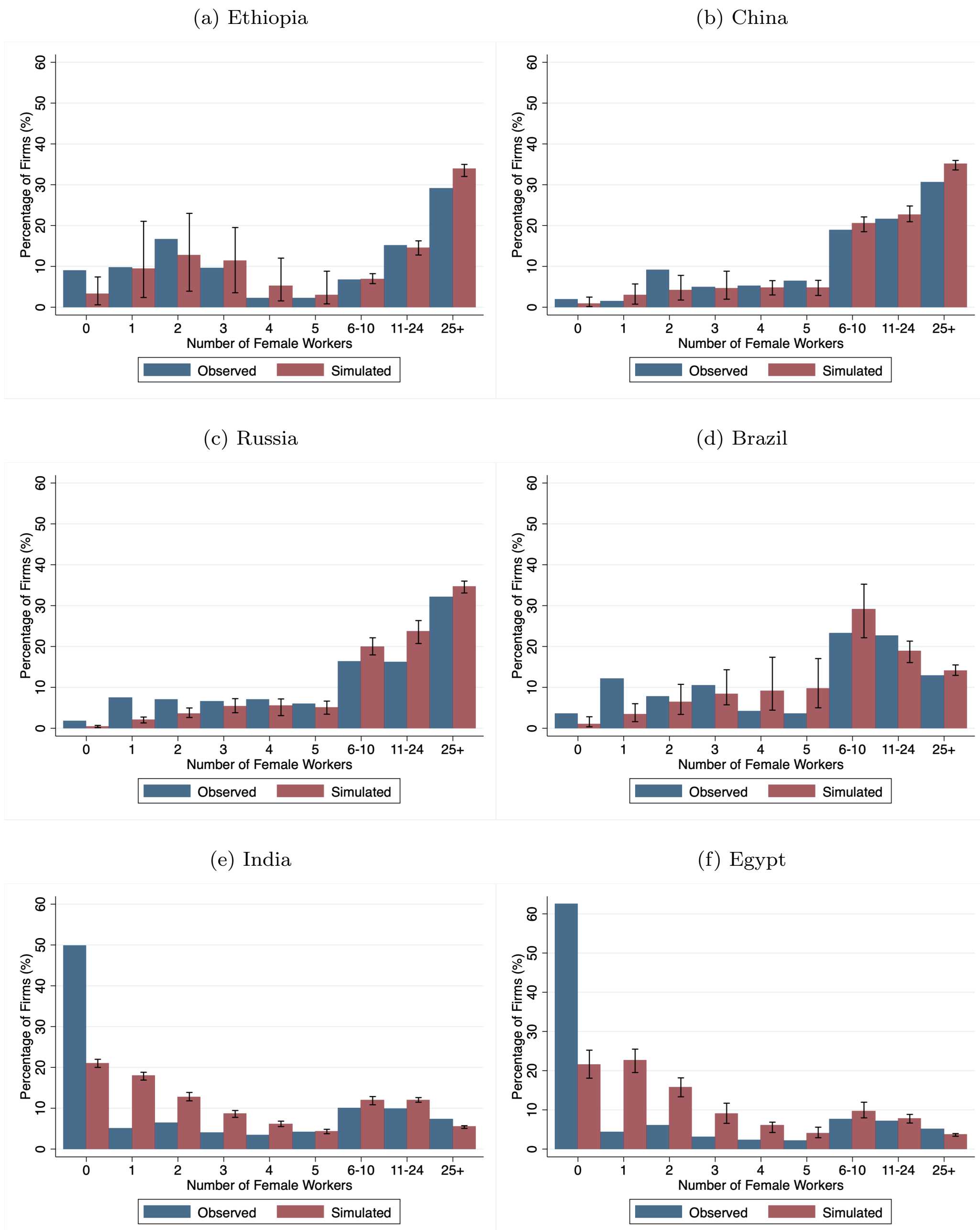
Note: Table reproduced from Miller, Peck and Seflek (2022). All-male share of firms calculated from World Bank Enterprise Survey, 2006 – 2018. Female share of labor force is derived from 2018 World Bank Development Indicators for the same countries and is not restricted to manufacturing.

Table A2—: List of World Bank Enterprise Surveys

Country	Region	Year	# Firms	% Female	Country	Region	Year	# Firms	% Female
Argentina	LAC	2006	559	23.8	Lao	EAP	2018	119	34.5
Argentina	LAC	2010	703	18.9	Lebanon	MNA	2013	174	21.0
Argentina	LAC	2017	571	20.8	Madagascar	AFR	2009	185	50.5
Armenia	ECA	2009	108	35.3	Mexico	LAC	2006	1060	35.9
Azerbaijan	ECA	2009	118	39.2	Mexico	LAC	2010	1065	31.5
Azerbaijan	ECA	2013	107	31.8	Mongolia	EAP	2009	126	56.7
Bangladesh	SAR	2007	1160	46.1	Mongolia	EAP	2013	106	51.7
Bangladesh	SAR	2013	1073	46.1	Morocco	MNA	2013	120	45.8
Belarus	ECA	2013	110	44.0	Myanmar	EAP	2014	314	58.2
Bosnia-Herzegovina	ECA	2009	112	36.3	Nepal	SAR	2009	122	12.7
Bosnia-Herzegovina	ECA	2013	103	37.3	Nepal	SAR	2013	231	14.9
Bolivia	LAC	2006	340	28.0	Nicaragua	LAC	2006	292	19.0
Bolivia	LAC	2010	106	20.2	Pakistan	SAR	2007	640	1.5
Botswana	AFR	2006	101	45.3	Panama	LAC	2006	223	27.3
Brazil	LAC	2009	1205	34.3	Panama	LAC	2010	105	33.0
Bulgaria	ECA	2007	501	52.2	Paraguay	LAC	2006	351	28.3
Chile	LAC	2006	602	25.7	Paraguay	LAC	2010	107	33.8
Chile	LAC	2010	755	19.0	Peru	LAC	2006	337	45.0
China	EAP	2012	1597	39.8	Peru	LAC	2010	715	25.4
Colombia	LAC	2006	588	53.0	Peru	LAC	2017	508	29.9
Colombia	LAC	2010	665	50.4	Philippines	EAP	2009	846	44.3
Colombia	LAC	2017	481	43.3	Poland	ECA	2009	108	43.7
Costa Rica	LAC	2010	285	22.5	Poland	ECA	2013	123	24.7
Croatia	ECA	2007	303	40.5	Romania	ECA	2009	135	44.7
Croatia	ECA	2013	109	42.0	Romania	ECA	2013	157	24.0
Dominican Republic	LAC	2010	109	27.2	Russia	ECA	2009	540	45.3
DRC	AFR	2006	128	10.1	Russia	ECA	2012	1106	42.6
DRC	AFR	2010	100	11.0	Serbia	ECA	2009	119	33.7
Ecuador	LAC	2006	336	24.2	Serbia	ECA	2013	105	40.8
Ecuador	LAC	2010	114	25.8	South Africa	AFR	2007	619	30.1
Egypt	MNA	2013	1535	11.4	Sri Lanka	SAR	2011	345	42.5
Egypt	MNA	2016	1063	13.0	Sweden	ECA	2014	277	22.8
El Salvador	LAC	2006	384	48.2	Tajikistan	ECA	2008	102	39.3
El Salvador	LAC	2010	121	44.0	Trinidad & Tobago	LAC	2010	110	29.1
El Salvador	LAC	2016	336	39.1	Tunisia	MNA	2013	280	43.1
Ethiopia	AFR	2011	218	44.6	Turkey	ECA	2008	699	27.7
Ethiopia	AFR	2015	340	37.0	Turkey	ECA	2013	872	23.0
Georgia	ECA	2008	104	38.9	Uganda	AFR	2006	254	20.7
Guatemala	LAC	2006	266	32.1	Uganda	AFR	2013	267	24.1
Guatemala	LAC	2010	326	30.3	Ukraine	ECA	2008	381	47.0
Guatemala	LAC	2017	118	33.3	Ukraine	ECA	2013	537	43.4
Honduras	LAC	2006	221	29.0	Uruguay	LAC	2006	315	38.9
Honduras	LAC	2010	111	28.4	Uruguay	LAC	2010	303	31.4
India	SAR	2014	6282	11.6	Uzbekistan	ECA	2008	116	37.8
Indonesia	EAP	2015	978	39.4	Uzbekistan	ECA	2013	116	32.8
Iraq	MNA	2011	377	1.6	Vietnam	EAP	2009	716	47.5
Israel	MNA	2013	170	29.0	West Bank & Gaza	MNA	2013	123	4.1
Jordan	MNA	2013	238	13.2	Yemen	MNA	2010	191	3.4
Kazakhstan	ECA	2009	147	40.8	Zambia	AFR	2007	276	12.2
Kazakhstan	ECA	2013	153	28.5	Zambia	AFR	2013	283	13.9
Kenya	AFR	2007	373	15.5	Zimbabwe	AFR	2011	332	16.2
Kenya	AFR	2013	338	19.0	Zimbabwe	AFR	2016	262	21.1
Kenya	AFR	2018	269	15.5					

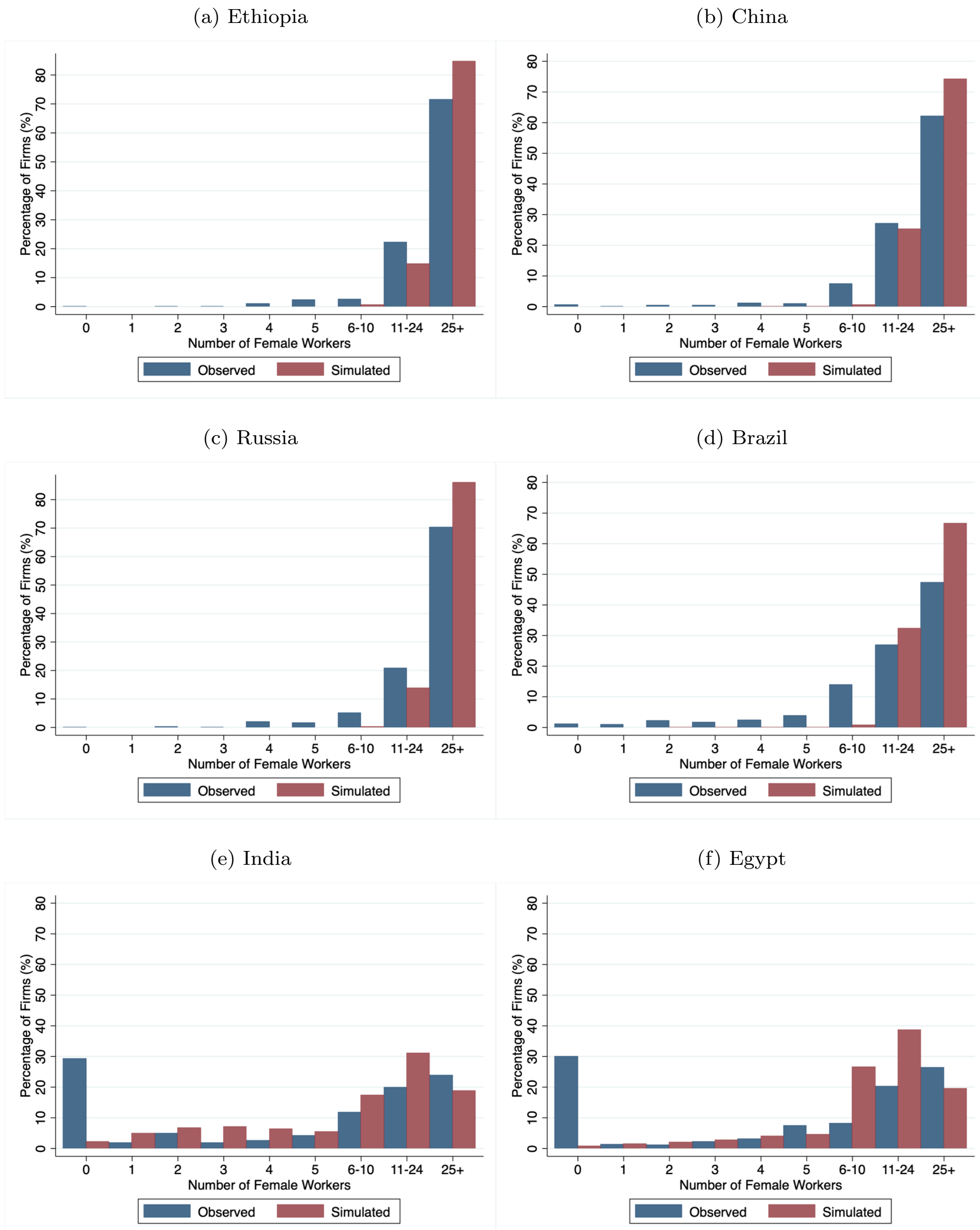
Note: This table lists the World Bank Enterprise Surveys that we include in our analysis. We limit our samples to manufacturing firms, where surveys include questions on the gender composition of employees by occupation. Next, we drop surveys where information on gender composition is missing for more than 20% of firms. In remaining surveys, we drop firms with missing data on gender composition or fewer than 5 employees. We then drop surveys with fewer than 100 remaining firms. This leaves us with 105 surveys in 65 countries. The six regions are: sub-Saharan Africa (AFR), East Asia and Pacific (EAP), Eastern and Central Europe (ECA), Latin America and Caribbean, Middle East and North Africa (MENA), and South Asia (SAR). '# of Firms' refers to the number of firms remaining in the survey following our sample restrictions. '% Female' is the female share of workers in these firms, weighted by firm sample weights.

Figure A1. : Distribution of Female Employment across Firms, by Country



Note: This set of figures compares observed and simulated distributions of female employment across firms for six countries: Ethiopia, China, Russia, Brazil, Egypt, and India. The simulated distributions are simulated under the null hypothesis that no firm in that country faces binding integration costs. Sample selection and simulation details are described in Sections I and II.A.

Figure A2. : Distribution of Female Employment Across *Large* Firms, by Country



Note: This set of figures compares observed and simulated distributions of female employment across firms for six countries: Ethiopia, China, Russia, Brazil, Egypt, and India. We limit to firms with at least 50 employees. The simulated distributions are simulated under the null hypothesis that no firm in that country faces binding integration costs. Sample selection and simulation details are described in Sections I and II.A.

Table A3—: Female Employment and Integration Rates Across Countries

	EMP_F		$EMP_F - EMP_M$	
	(1)	(2)	(3)	(4)
Ex-ante integration rate:				
Overall ($\hat{\theta}^{EP}$)	0.388**	0.243**	0.420**	0.264**
	(0.083)	(0.097)	(0.065)	(0.073)
Overall ($\hat{\theta}^{S}$)	0.572**	0.262~	0.582**	0.237~
	(0.137)	(0.152)	(0.113)	(0.119)
Representative firm ($\theta_j^{EP} \cdot n_j = 10$)	0.461**	0.321**	0.459**	0.324**
	(0.085)	(0.101)	(0.069)	(0.075)
Region FEs		✓		✓
Observations	65	65	65	65

Note: Table reports OLS estimates of regressions of female employment measures on ex-ante integration rates. EMP_F is the average percentage of women age 15+ that are employed in the years the manufacturing surveys were conducted, averaged across years, and $EMP_F - EMP_M$ is the difference between female and male employment rates. Robust standard errors are in parentheses. ~ p<0.1; * p<0.05; ** p <0.01.