Gender Roles and the Misallocation of Labour Across Countries*

Nava Ashraf† Oriana Bandiera‡ Virginia Minni§
Víctor Quintas-Martínez¶

April 28, 2022

Click here for the latest version

Abstract

This paper asks whether the gendered division of work inside and outside the home leads to the misallocation of labor. Using personnel data of a multinational firm covering 100K employees in 101 countries over 5 years together with labor force participation data we show that women are more positively selected than men: the productivity of the average female worker is higher than that of the average male worker, and this gap is decreasing in women's participation in the labor force. Structural estimates indicate that equalising barriers to labor force participation would increase firm productivity by 32% keeping employment and the wage bill constant.

*We thank our partners at the MNE for sharing the data, Sam Bowles, Anton Heil, Erik Hurst, Rocco Macchiavello, Andrea Prat, Imran Rasul and seminar participants at CIVOE, Harvard, HBS, IIES, LSE, MIT Sloan, NBER Summer Institute, Northwestern, Princeton, Toulouse, UCL, UMichigan for useful feedback. We thank Rishabh Malhotra, Donata Schilling, Misha Seong for great research assistance. All errors are our own.

†STICERD and Department of Economics, LSE. n.ashraf1@lse.ac.uk
‡STICERD and Department of Economics, LSE. o.bandiera@lse.ac.uk
§STICERD and Department of Economics, LSE. v.m.minni@lse.ac.uk
¶Department of Economics, MIT. vquintas@mit.edu

1 Introduction

One of the surest ways to increase national income is to create (..) employment for women outside the home (Lewis, 1954)

The gender division of labour inside and outside the home varies across countries and time, but always within the confines of norms that assign the largest share of housework to women. Understanding the consequences of this division requires knowing whether potential productivity for work inside and outside the home is the same across genders. If men have an advantage for work outside the home, the observed allocation of labor along gender lines is efficient.[1] If not, gender roles lead to misallocation as some of the women who stay at home would be more productive in the workforce and some of the men who are currently in the workforce would be more productive at home.

This paper investigates the effect of gender roles on misallocation using microdata on the gender pay gap among the universe of employees of a large MNE combined with labor force participation (LFP) across cohorts in the 101 countries where the MNE operates.[2] To quantify the cost of misallocation, we structurally estimate the parameters of the firm's wage policy and, matching both the mean and the variance of wages, we back out the implied productivity by gender as well as the average pay and the utility of staying at home.

Following the intuition of Olivetti and Petrongolo (2008), we illustrate how the sign of the correlation between female labor force participation (FLFP) and the gender pay gap in the firm is a reflection of whether the underlying distribution of potential productivity differs by gender, through the prism of a Roy (1951) model of the choice between work inside or outside the home. If variation in FLFP is due to variation in productivity differences, and pay reflects productivity, high FLFP

[1]Women might be less productive than men in work outside the home for many reasons. There is a long literature on technologies that drove specialisation into home and outside work along gender lines because of their reliance on physical force (Boserup, 1970; Goldin, 1995; Foster and Rosenzweig, 1996; Alesina et al., 2013). Centuries of such specialisation could have solidified men's advantage in market versus home production, perhaps through intergenerational transmission of skills within gender. It could also be that current workplace organizations are set up optimally for men, and thus may make women less productive.

[2]In doing so, we complement single country studies (in the U.S., Mulligan and Rubinstein, 2008; Blau and Kahn, 2006; and in the U.K., Blundell et al., 2007).

will imply high female productivity and high female pay. On the contrary, if the underlying distribution is the same and variation in FLFP is due to variation in the barriers that women face, high FLFP will be symptomatic of low barriers, low female productivity and low pay. Thus, a negative correlation between the FLFP and the gender pay gap implies positive selection. In aggregate data, the relationship is confounded by differences in sector and industry choice (Goldin, 2014; Card et al., 2015; Wiswall and Zafar, 2017; Strittmatter and Wunsch, 2021). In our context, this is shut down by design as we observe women and men doing exactly the same job, with the same contract, in the same firm and in the same country.

To estimate the correlation between the gender pay gap and FLFP, we match the age cohorts in the firm personnel records with the average LFP in the country in the decade of labor market entry by gender and we compute the labor force participation ratio (female over male LFP). The analysis is then conducted on an unbalanced panel of 100,819 unique employees over 5 years in four age cohorts and 101 countries. The different countries and cohorts allow us to exploit variation in LFP both across and within countries.

Our main finding is that when the LFPR is at its lowest level, the gender pay gap is "inverted". That is, women are paid more than men with the same experience, same tenure and working in the same function. The gap shrinks as the LFPR increases and it converges to standard levels for the industry (negative 10%) when and where the LFPR is highest. The sign and size of the effect are similar whether we use within country or within cohort variation for identification, thus ruling out country level confounders.

The finding that women are more positively selected than men into the workforce is symptomatic of misallocation because the marginal woman who would move from working inside to working outside the home is more productive in the workforce than the marginal man who would move in the opposite direction.

The second part of the paper quantifies the productivity cost of this misallocation by means of a structural model of the firm pay policy. The value added of the model, relative to the reduced form estimates and to the existing literature, is that we can leverage our individual level data to separately identify gender differences in fixed pay, due to - for instance - discrimination, from differences in pay due to productivity. To mirror the fact that pay is set globally and is merit based, we assume that pay equals to an average wage plus a reward proportional to marginal productivity. This gives us a relationship that links salary in the MNE to the cost of working outside the home. Both variables have a component which is common to all people of the same gender in the same country (for instance, country-specific social norms about gender), and a component that is specific to the individual (their own productivity and preferences). Because we have individual-level pay data, we can separate the common component of salary from the reward for individual-specific performance. We can use this to identify productivity thresholds that split the population into workers and homemakers.

We find that, in line with the gender pay gap, there is a gender productivity gap which is positive and large (.9 SD) at low levels of the LFPR and decreases to .3 SD at high levels. The estimated productivity gap is larger in countries with weaker gender equity labor laws and with more conservative gender norms.

Alternatively, we can relax the assumption of equal preferences and use the model to back out the difference in preferences that would make the current LFP gap optimal. The implied preference gap is several orders of magnitude larger than any other gender gap in preferences estimated in the literature to date.

To probe the external validity of our estimates we extract balance sheet data from ORBIS to cover all manufacturing firms in the same countries where the MNE operates: this yields a sample of 2 million firms in 158 SIC3 sectors. We find that the estimate of misallocation we obtain from the MNE data is strongly correlated with the productivity of other firms in the economy, especially those that operate in the same sector as our MNE.

With these estimates in hand, we analyse three counterfactuals. First, given the productivity differences between men and women, we ask whether the firm could effectively "undo" the gender norm by changing the terms of the contract to attract more women. To do so, we derive the contract that maximises productivity while keeping employment and the wage bill constant in each country. We find that the optimal contract has a lower base pay and a steeper performance gradient than the observed contract. This brings the firm's gender ratio close to one and increases productivity by 22%. However, we note that such contract would massively increase inequality within and between genders; most notably the difference in pay between women and men would go up by 78%. This captures both differences in performance for the same job and differences in jobs as more able women climb the corporate ladder faster. Whilst it is, thus, theoretically possible for the firm to adopt policies that compensate for societal norms, such a steep performance gradient would create a high level of inequality among employees, where the stronger the gender norms the higher the inequality within the firm. It may quite possibly also be unsustainable for the firm: in order to hire more women without excessively increasing inequality, they would have to increase women's pay without decreasing men's pay.

The second counterfactual eliminates gender norms, that is we simulate what would happen if we could eliminate gender differences in the cost of working outside the home. Overall, eliminating gender differences would, by definition, bring the pay gap to zero and it would increase productivity by 32% while keeping the wage bill and employment constant. The productivity gains are due to both high productivity women joining and low productivity men leaving. We note that the mirror image of the gender tax that women have to pay to work outside the home is the equivalent tax that men have to pay to work inside the home independently of their skills and preferences. Thus, eliminating gender norms will also eliminate misallocation in work inside the home, by allowing the men who wish to do so to specialize in home production.

The third counterfactual simulates the effect of more stringent labor regulations that make it harder to link pay to performance. We show that this leads to a larger intake of lower ability workers who, by selection, are more likely to be male. These findings underscore the importance of taking into account the effects of labor regulation through selection (Besley and Burgess, 2004; Propper and Reenen, 2010).

The remainder of the paper proceeds as follows. Section 2 presents the institutional context of the MNE, and describes the data sources. Section 3 introduces the model and provides reduced form evidence on the link between the LFPR and the gender pay gap. In Section 4, we calibrate the parameters of the model from the firm's personnel data and country-cohort level LFPR and Section 5 uses our estimates to evaluate the effect of different counterfactuals. Section 6 concludes discussing welfare implications and other issues for further research.

2 Context and data

2.1 Context

We collaborate with a private consumer goods MNE with headquarters in Europe and offices in more than 100 countries worldwide as illustrated in Figure A.1. In 2019, the last year in our sample, the firm had a turnover of €20+ billion and employed over 120,000 workers, of which approximately 55% were white collars. We focus on white collar workers because blue collar workers are only observed in two thirds of countries where the MNE has production activities. Typical white collar jobs in the MNE involve sales, engineering, marketing, HR, R&D for product development and general managerial activities. The workers have homogeneous levels of human capital as applications require a college degree, and most employees have degrees in either business administration (50%) or engineering (20%).

2.2 Data

Pay data: Our sample covers the universe of employees between 2015 and 2019. We focus our analysis on local employees (non-expats), resulting in 100,819 distinct regular full-time workers over 2015-2019 in 101 countries (303,756 employee-year observations). The company is organized into a hierarchy of work levels that goes from work level 1 to 6 (C-Suite). Our main outcome variable is total compensation in logs (fixed plus variable pay), which is measured in euros for all countries. We look at four 10-year age cohorts within the company, 18-29, 30-39, 40-49 and 50-59[3]. Overall, average tenure is 8.5 years, 40% of workers are aged between 30-39 and the majority of workers are in WL1 (79%). Table 1 presents summary statistics

[3]We do not have more granular data on age cohorts because of data privacy clauses. Due to a limited sample size of workers in age cohorts above 50-59, we only consider workers up to the 50-59 cohort.

separately by gender at the gender-cohort-country-tenure level, which is the relevant unit of analysis used in the structural estimation[4]. On average, women earn less than men, have lower tenure and a lower probability of being in the top echelons of the firm. They are more likely to major in economics and human sciences while less likely to major in STEM.

Country-cohort level data: We combine the firm's administrative records with country-cohort data on labor force participation rates of males and females from the World Bank. In particular, we match the age cohorts in the firm with the average LFP rate in the country in the decade of labor market entry, separately by gender. For example, employees of age 18-29 are associated to the LFP rates of the 2010-2020 decade while employees of age 30-39 are associated to the LFP rates of the 2000-2010 decade and so on. We also make use of some additional publicly available cross-country data including: the World Values Survey, the Restrictive Labor Regulations Index from the World Bank, the Women, Business and the Law Index from the World Bank, and occupation-level wages (ISCO08) from the International Labor Organization (ILO).

Figure 1 illustrates the variation in the gender pay gap and the LFPR (female to male labor force participation, $\frac{FLFP}{MLFP}$), across countries and cohorts. Throughout, we define the gender pay gap as the difference between women's and men's total salary expressed as a percentage of men's salary.

3 Testing for misallocation

3.1 Framework

To formalize our intuition about differential selection into the labor force, consider a basic two-sector Roy (1951) model (as formalized by Borjas, 1987). Suppose that the utility from working outside the home is equal to earnings and that these are a

[4]As we explain in Section 4, we let the structural parameters vary by gender, cohort, country, tenure in the firm to control for confounders and allow for the wage policy to differ across countries as well as take into account factors such as on-the-job learning.

(log-)linear function of individual i's productivity, A_i:[5]

$$y_i^1 = \alpha^1 + \beta^1 A_i, \tag{1}$$

The term α^1 is the unconditional average wage and β^1 is the return to productivity. We interpret deviations from that average wage as arising from individual differences in productivity.

Similarly, we model worker i's value of housework as:

$$y_i^0 = \alpha^0 + \nu^0 N_i, \tag{2}$$

where N_i captures sources of individual heterogeneity in the value of staying out of the labor force that are, without loss of generality, independent of productivity. Here, α^0 captures the unconditional average value of staying out of the labor force (e.g. social norms that affect all women).

We make the distributional assumption:

$$\begin{bmatrix} A_i \\ N_i \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

equally distributed in all our sample. A difference with the canonical Roy-Borjas model is that, as a result, y_i^1 and y_i^0 are independent. This asssumption will allow us, in section 4, to identify the parameters of the model with individual-level data only on those who are working at the firm. It has also become a standard assumption in the literature on misallocation with structural models of sectoral choice (e.g. Hsieh et al., 2019).

Individual i self-selects to work outside the home if and only if:

$$y_i^1 \geq y_i^0$$

$$\Leftrightarrow \quad \eta_i \equiv \frac{\beta^1 A_i - \nu^0 N_i}{\sqrt{(\beta^1)^2 + (\nu^0)^2}} \geq \frac{\alpha^0 - \alpha^1}{\sqrt{(\beta^1)^2 + (\nu^0)^2}} \equiv \xi.$$

[5]This could be micro-founded, for example, by assuming that workers are paid their marginal product of labor, that all workers supply the same amount of hours (full-time) and that the individual production function is Cobb-Douglas, $F(K, l_i) = e^{z_i} K^\alpha l_i^{1-\alpha}$, where $z_i \sim \mathcal{N}(\mu, \sigma^2)$.

Because $\eta_i \sim \mathcal{N}(0,1)$, that happens with probability $1 - \Phi(\xi)$.

Moreover, we have that:[6]

$$\begin{aligned}\mathbb{E}[A_i \,|\, \eta_i \geq \xi] =& \frac{\text{cov}(A_i, \eta_i)}{\text{var}(\eta_i)} \mathbb{E}[\eta_i \,|\, \eta_i \geq \xi] \\ =& \frac{\beta^1}{\sqrt{(\beta^1)^2 + (\nu^0)^2}} \frac{\phi(\xi)}{1 - \Phi(\xi)}.\end{aligned}$$

3.2 Identification

The model above suggests a reduced-form test that allows us to test the null of no misallocation by exploiting cross-country and cross-cohort variation in LFP and wages. Under the null, *male advantage*, men and women have the same opportunities, but men have a higher productivity at work than women. Under the alternative hypothesis, *barriers to entry*, men and women have the same underlying distribution of productivity at work, but women face greater barriers to labor force participation. These two views are illustrated in Appendix Figure A.2.

Under the null, countries with higher LFP are those where the mean productivity of women is higher, so the pay gap would be positively correlated with the LFPR. Under the alternative, the opposite is true: although the underlying distribution of productivity is the same, the productivity of the average woman in the labor force is higher the higher the barriers to entry, so that the pay gap would be negatively correlated with the LFPR.

Assume, for simplicity, that the only parameters that vary across gender, countries and cohorts are the unconditional mean productivity (that we denote by μ_F and μ_M for females and males, respectively) and the mean value of staying at home (α_F^0, α_M^0), which in turn determines the participation thresholds ξ_F and ξ_M. Using the results from the previous subsection, we can write the pay gap as:

$$\mathbb{E}[y_{iF}^1 \,|\, \text{empl'd}] - \mathbb{E}[y_{iM}^1 \,|\, \text{empl'd}] =$$

$$= \underbrace{\mu_F - \mu_M}_{\text{Male advantage } (\uparrow \text{ in LFPR})} + \frac{(\beta^1)^2}{\sigma_\eta} \underbrace{[\lambda(\xi_F) - \lambda(\xi_M)]}_{\text{Barriers to entry } (\downarrow \text{ in LFPR})}$$

where $\lambda(\cdot) = \phi(\cdot)/(1 - \Phi(\cdot))$ is the inverse Mills ratio.

[6] To see why, realize that we can write $A_i = \text{cov}(A_i, \eta_i)/\text{var}(\eta_i) \times \eta_i + u_i$, where u_i is the population OLS projection error. By the normality assumption, u_i and η_i are not just uncorrelated, but also stochastically independent.

This decomposition allows us to test the null hypothesis by using cross-country and cross-cohort variation in LFP and looking at the gender pay gap in our firm. Under the null, the first term induces a positive correlation between the pay gap and the LFPR, whereas the second term is zero. If there are barriers to entry, the second term induces a negative correlation between the pay gap and the LFPR, while the first term is zero. Therefore, the sign of the correlation between the pay gap and the LFPR is a test for misallocation.

This test is valid if the following three conditions hold. The first is that selection into the labor force is positive, i.e. those who work in the firm are, on average, more productive than those who stay at home. This is the case in our model since we assume that productivity at home and in the firm are independent. In practice, this will be true if returns to productivity are greater at work than at home, where productivity is defined as all those idiosyncratic factors that increase the observed wage. Under this condition, there is a negative correlation between gender-specific LFP and average wages in the firm.

The second is that variation in LFPR does not shift the firm's wage policy. For example, we could erroneously conclude that the male advantage effect dominates if the differences in fixed pay in the firm between women and men $\alpha_F^1 - \alpha_M^1$ are decreasing in LFPR (i.e. women are paid more on average compared to men in countries where the LFPR is relatively low). This illustrates the advantage of using data from one firm relative to aggregate data on the pay gap. The latter might be decreasing in LFPR because the marginal woman might choose sectors and jobs with lower pay. In contrast, identification here only requires that the firm does not discriminates more in favor of women (or against men) in countries where the LFPR is lower, which is a much weaker assumption and one that is supported by the evidence on gender discrimination and labor market outcomes across countries (Hyland et al., 2021).

The third is that we need selection into our firm to reflect selection into the labor force at all levels of LFPR. This requires that workers of both genders weakly prefer to work at our firm than at any other firm and that the firm does not observe workers' productivity differently across genders and levels of LFPR when hiring. A recent strand of evidence shows that MNEs pay higher wages in developing countries (Hjort et al., 2020; Alfaro-Urena et al., 2019), and Appendix Figure A.3 shows that the firm's average wages are usually well above the countries' average wages using both the average wages in the manufacturing sector from the ORBIS database and the ILO estimates for white-collar employees. This assumption rules out that the best available local jobs vary by gender and that the best job for men is preferred to our firm. This implies that differences in selection come entirely from the left hand side of the productivity distribution, that is the threshold for working is lower for men than it is for women, while all top men and women prefer working for the MNE. Figure 2 documents a large variation in female shares of employment within the MNE across countries, which matches the variation in LFPR across countries. The relationship between the firm and country ratios does not vary with the level of LFPR, which is confirmed by a formal test of differential slopes by above/below median LFPR. This suggests that the selection of men and women in the MNE follows closely their LFP decisions and hence that the countries' LFPRs "bind". Under the assumption that the firm adopts the same personnel policies worldwide, which is supported by the evidence from the literature on MNEs' practices[7], this relation suggests that women face country-specific barriers to entry in the labor market stemming from different gender attitudes across countries.

3.3 Empirical test

Figure 3 implements the misallocation test derived above by estimating a kernel-weighted local polynomial of the pay gap on LFPR. Pay captures differences in performance between employees, encompassing off-peak salary increases as well as promotions. The correlation is negative: at low levels of LFPR we observe an "inverted pay gap", that is women earn between 22% and 45% more than men; the gap falls as the LFPR increases and it plateaus at around -10% when the LFPR reaches 0.8. This is close to the average gender pay gap of -16% in Europe (European Commission, 2019). Overall, Figure 3 documents a negative correlation between the gender pay gap and the LFPR. In light of the selection model in Section 3, we infer that barriers to entry are the dominating force in generating gender differences in LFP. In countries with low LFPR the productivity threshold for women to work is so high that only the most talented women exceed it. As a result, the lower the LFPR, the more positively selected the women hired within the MNE are, compared to men. To corroborate the gender gap finding, we show in Appendix Figure A.4 that the gender gap in fast-track promotions, defined as major promotions achieved at a relatively

[7]Several papers suggest that multinationals transpose domestic wages and organisational practices to their foreign subsidiaries (Bloom and Van Reenen, 2007; Alfaro-Urena et al., 2019; Hjort et al., 2020).

younger age, exhibits a very similar pattern.

We can check whether the sign of the correlation between the pay gap and the LFPR is consistent regardless of the source of variation that we use to identify it.

Table 2 identifies the correlation using different sources of variation. We estimate the following model:

$$w_{iact} = \alpha LFPR_{ac} + \beta Female_i + \gamma LFPR_{ac} * Female_i + \mathbf{X'}_{iact}\mathbf{\Lambda} + \boldsymbol{\psi}_t + \epsilon_{iact} \quad (3)$$

where w is log wage of employee i in country c and year t for age group a. $\boldsymbol{\psi}_t$ represent year fixed effects to take out year-level macro shocks and $\mathbf{X}_{iact}$ is a vector of controls. We cluster standard errors at the same level as the RHS variable, that is country-cohort. The coefficient of interest is γ which measures the change in the pay gap as LFPR increases. Under the identifying assumptions above, $\gamma < 0$ implies that the barriers to entry effect dominates and $\gamma > 0$ implies that the comparative advantage effect dominates.

Table 2 estimates the model in Equation 3 with different controls in $\mathbf{X}_{iact}$: column 2 controls for a quadratic function of tenure and function (department) fixed effects, column 3 adds cohort fixed effects so to only exploit the variation across countries; column 4 replaces the cohort fixed effects with country fixed effects hence only exploiting the variation within countries. The comparison between columns 3 and 4 is particularly informative as it uses one source of variation at the time. The fact that the γ coefficient is stable does reassure us that it is, in fact, measuring selection. Finally, column 5 estimates the same specification as in column 3 (with controls and cohort fixed effects) for new hires only. In so doing, we check whether γ is capturing gender differences in learning about employee's productivity and dismissals decisions. The estimates of γ are negative and precisely estimated in all specifications.[8]

[8] We conduct a number of additional robustness checks. Table A.2 in the Appendix shows that the patterns in Table 2 hold if we add fixed effects for the geographical region, and if we split the sample by lower income and higher income countries (as defined by the World Bank). In Appendix Table A.3 we report the results when converting wages from euros into PPP 2017 \$, using the PPP conversion rates of the ICP at the World Bank. The gender gap is unaffected and the only change is the magnitude of the coefficient on LFPR which shrinks to the level found when controlling for country fixed effects (column 4 in Table 2). This is what we would expect as differences in PPP exchange rates would not affect cross-country comparisons of the gender gap. Finally, we note that the results are driven by differences in pay rather than bonus, which constitutes a much lower proportion of overall salary (the median ratio of bonus to fixed pay is 13%) — see Appendix Table A.1. In the company, pay summarizes altogether most differences in performance between employees, encompassing off-peak salary increases as well as promotions.

4 Quantifying misallocation

In this section, we take the model of worker selection into the labor force that we introduced in section 3 to the data. The value added of the model, relative to the reduced form estimates and to the existing literature, is that we can leverage our individual level data to separately identify gender differences in fixed pay (due to -for instance - discrimination) from differences due to productivity. Normally this cannot be estimated with aggregate data as it only contain the average wage while individual data can be used to compute the variance which maps directly to the variation in pay due to differences in productivity. We first discuss how to calibrate the parameters of the model from the firm's personnel data and country-cohort level LFP. Next, we present the results of our calibration and validate them against different variables and independent datasets.

4.1 Model calibration

For our structural exercise, we will use the two-sector Roy-Borjas model presented in section 3. We let the parameters vary by country (c), cohort (a), tenure in the firm (t)[9] and gender (g) cells: $\alpha^1_{gtac}, \beta^1_{gtac}, \alpha^0_{gtac}, \nu^0_{gtac}$. This flexibility allows us to control for a variety of confounders: for example, we do not make assumptions as to whether the firm has the same wage policy for men and women or across countries, and the fact that we estimate β^1 by tenure group allows for on-the-job learning in a way that is possibly correlated with productivity. We keep the normalization that (A_i, N_i) have the same jointly standard normal underlying distribution across $gtac$ cells, since differences in that distribution would be absorbed in our calibration by $\alpha^1_{gtac}, \beta^1_{gtac}, \alpha^0_{gtac}$ and ν^0_{gtac}.

In what follows, we make three assumptions. First, as with any other binary choice models, the scale parameter of the selection Probit of our Roy model is not identified. Hence, we normalize it to 1 for calibration purposes, so that $\text{Var}(\eta_{igtac}) = \text{Var}(\beta^1_{gtac}A_i - \nu^0_{gtac}N_i) = (\beta^1_{gtac})^2 + (\nu^0_{gtac})^2 = 1$.

Second, we do not have individual-level data for those not working at the firm, so we are not able to identify the correlation between A_i and η_{igtac} (as one would do with a standard Heckman selection model). Hence, we are going to assume that A_i and N_i are independent.

Finally, again because we have data only on those in our firm, we need to obtain a

[9]We aggregate tenure in groups of two years each.

measure of the probability that a worker will want to work for our firm, i.e. $\Pr(y^1_{igtac} \geq y^0_{igtac})$. We proxy that by country-cohort-level LFP data. In other words, we are identifying the probability that an individual chooses to work for our firm with the probability that he or she chooses to work for *any* other firm. For that, we need to assume that working at our firm is weakly preferred to working at other firms for all individuals. We provide supporting evidence of this in section 3.

Under our distributional assumptions, labor force participation and the observed wage satisfy the following moment conditions:

$$\Pr(\text{employed}) = 1 - \Phi(\xi_{gac}) \tag{4}$$

$$\mathbb{E}[y^1_{igtac} \,|\, \text{employed}] = \alpha^1_{gtac} + (\beta^1_{gtac})^2 \lambda(\xi_{gac}) \tag{5}$$

$$\text{Var}(y^1_{igtac} \,|\, \text{employed}) = (\beta^1_{gtac})^2 + (\beta^1_{gtac})^4 \left[\xi_{gac}\lambda(\xi_{gac}) - \lambda(\xi_{gac})^2\right] \tag{6}$$

where $\lambda(\cdot) \equiv \phi(\cdot)/(1 - \Phi(\cdot))$ is the inverse Mills ratio and $\xi_{gac} \equiv \alpha^0_{gtac} - \alpha^1_{gtac}$ are the participation thresholds. Together with the restriction $(\beta^1_{gtac})^2 + (\nu^0_{gtac})^2 = 1$, we have four equations in four unknown parameters $(\alpha^1_{gtac}, \beta^1_{gtac}, \alpha^0_{gtac}, \nu^0_{gtac})$ for each gender, tenure level, age cohort and country cell.

In order to calibrate those parameters, we match the moments above to their empirical counterparts. Since we observe the wage for those working in the firm, we can use the sample average and variance as the empirical analogs for $\mathbb{E}[y^1_{igtac} \,|\, \text{employed}]$ and $\text{Var}(y^1_{igtac} \,|\, \text{employed})$. To eliminate the effect of observables, we use the residuals of a regression of log(base pay + bonus) on year and function dummies as our measure of y^1_{igtac}. To calibrate the parameters in the participation decision, we use World Bank LFP data in each gender, cohort and country cell, LFP_{gac}, as our empirical analog for Pr(employed). Table 3 provides a summary of the parameters of the model and the empirical target that each parameter tries to match in our calibration strategy.

4.2 Calibrated parameters

Once we have obtained the parameters in the firm's wage policy, we can recover productivity as

$$\widehat{A}_i = \frac{y^1_{igtac} - \alpha^1_{gtac}}{\beta^1_{gtac}},$$

where y^1_{igtac} is the residualized log-wage described in section 3. Figure 4 plots the average calibrated productivity for our sample of firm workers by LFPR. Average productivity is approximately constant for male workers in our countries, whereas for female workers, average productivity is very high when FLFP is much smaller than MLFP (about 0.8 standard deviations higher than for men), and decreases as the LFPR approaches 1. This result is consistent with selection with respect to productivity, so the lower a group's LFP, the more positively selected they are.

Figure 5 further shows that this is due to a shift of the whole productivity distribution. For men, the distribution of calibrated productivity is very concentrated and somewhat left-skewed for the lowest levels of LFPR, becoming more disperse as the LFPR increases. For women, as the LFPR increases, there is a downward shift of the entire distribution but the right tail.

If we believe that the underlying distributions of A_i and N_i in the population are the same for men and women, differences in LFP across genders must be due to the cost of social norms. In other words, if men and women have the same distribution of productivity[10] and preferences for staying at home, the only reason for their LFP to differ is that they must face different payoffs from working outside the home.

4.3 Productivity or preferences?

Alternatively, we can relax the assumption of equal preferences for housework and use the model to back out the difference in preferences that would make the current LFP gap optimal. If the observed LFPs are optimal, so that the marginal man and woman have the same productivity, we should have $\mu_N^F - \mu_N^M = \Phi^{-1}(MLFP) - \Phi^{-1}(FLFP)$, where μ_N^g is the average N_i of gender g (assuming that the standard deviation is the same).

Figure 6 plots these average preference gaps by country. Two points are of note: first, differences between genders are large - well over $1S.D.$ for the top decile of countries, and at least $0.5S.D.$ for most of the sample. Second, differences across countries are large, for instance the interquartile and the interdecile ranges are $0.45S.D.$ and $1.18S.D.$, respectively. These patterns are in sharp contrast with the existing evidence on gender differences in economic preferences across countries where they are much smaller. Figure A.5 in the appendix shows the distribution of gender differences in risk aversion, altruism,

[10] Gender differences in educational attainment have been drastically reduced over the last decades and are much smaller compared to the LFP gap (see Appendix Figure A.6).

trust, patience, positive and negative reciprocity from Falk et al. (2018) for the same set of countries. Overall, considering the interquantile range, for differences in preferences to rationalise our results we would need these to be at least three times larger than any other economic preference.

4.4 Validation

To validate our estimates we show that they correlate with other data not used to calibrate the model. We use three sets of external variables to validate our three key estimates: (i) individual performance data from the firm's records to validate calibrated productivity; (ii) country level labor laws that constrain the firm personnel policy to validate the parameters of the wage policy and (iii) country level social norms to validate our estimate of the difference in home payoffs.

The results for the first exercise are in Figure 7. Panel (a) validates calibrated productivity against the firm's performance appraisals that a manager gives every year, panel (b) against pay growth in the first year (for new hires), and panel (c) against an objective productivity measure for the sales department (based on reaching specific targets). The correlation of all three indicators with our calibrated productivity is positive and strong.

Figure 8 panel (a) shows the results for the second exercise. It plots our calibrated β^1_{gtac} (which represents returns to productivity, and, in our model, is what generates dispersion in pay within gender-country-cohort-tenure cells) against the Restrictive Labor Regulations Index from the World Economic Forum.[11] Consistent with stricter labor regulation limiting performance pay, we find that our calibrated β^1_{gtac} is lower in countries with a higher value of the index.

Figure 8, panel (b) plots the gap in our calibrated $\alpha^0_{Ftac} - \alpha^0_{Mtac}$ (that we interpret as the cost of gender norms for the average woman) against the Women, Business and the Law Index from the World Bank.[12] The figure shows that the gap $\alpha^0_{Ftac} - \alpha^0_{Mtac}$ is

[11]The WEF Restrictive Labor Regulations Index is available for the period 2008–2020 and it is based on an annual survey on the most problematic factors for doing business (e.g. corruption, taxes, inflation, etc.). The survey is administered to a representative sample of around 15,000 business executives in 150 countries. The Restrictive Labor Regulations Index includes measures related to labor-employer relations, wage flexibility, hiring and firing practices, performance pay, labor taxes, attraction and retention of talent.

[12]The WB Women, Business and the Law Index covers 190 countries through the period 1971–2020 and is structured around the life cycle of a working woman. It consists of eight indicators constructed around women's interactions with the law — mobility, pay, workplace, marriage,

strongly negatively correlated with laws allowing or facilitating women's labor. Therefore, part of the restrictions to FLFP due to gender norms may actually be embedded in the laws of certain countries.

Finally, Figure 9 shows the gap in our calibrated average value of staying at home, $\alpha^0_{Ftac} - \alpha^0_{Mtac}$, against the responses to four questions in the World Value Survey: (1) "Men make better business executives than women do," (2) "Pre-school child suffers with working mother," (3) "Being a housewife is just as fulfilling as working," (4) "When jobs are scarce, men should have more right to a job than women." For all four questions, we see a strong positive correlation between agreement with the statement and our gap in the average value of staying at home.

4.5 Beyond the firm

4.5.1 Adjusted pay gaps

One implication is that gender pay gaps are much larger once we account for differential selection. Our model implies the following relationship between the moments of the observed and population pay distributions:

$$\mathbb{E}[y^1_{igtac} \,|\, \text{empl'd}\,] = \mathbb{E}[y^1_{igtac}] + (\beta^1_{gtac})^2 \lambda(\xi_{gac}).$$

If women are more positively selected into the labor force compared to men, the observed pay gap will underestimate the true average pay gap (i.e. the pay gap that the average woman in the population would face vis-à-vis the average man). Given external estimates of the observed pay gap, we could use our calibrated parameters to adjust them for selection as follows:

$$\begin{aligned}\underbrace{\mathbb{E}[y^1_{iFtac}] - \mathbb{E}[y^1_{iMtac}]}_{\text{Adjusted Gender Gap}} = & \underbrace{\mathbb{E}[y^1_{iFtac} \,|\, \text{empl'd}\,] - \mathbb{E}[y^1_{iMtac} \,|\, \text{empl'd}\,]}_{\text{Unadjusted Gender Gap}} \\ & - \underbrace{[(\beta^1_{Ftac})^2 \lambda(\xi_{Fac}) - (\beta^1_{Mtac})^2 \lambda(\xi_{Mac})]}_{\text{Adjustment term}}\end{aligned}$$

parenthood, entrepreneurship, assets and pensions — for current laws and regulations (i.e. religious and customary laws are not consider unless they are coded). Hyland et al. (2021) provide an overview of the data documenting how gender discrimination by law affects women's economic opportunities.

We use the above procedure to adjust the ILO estimates of the pay gap for ISCO-08 categories 1–5 (white collars).[13] The results are in Figure 10. Once we adjust gender gaps for selection, we get much larger magnitudes (in absolute value), up to four times the size of the un-adjusted gap. As expected, the difference is larger for countries with low LFPR, since those are the ones where the adjustment term is larger.

4.5.2 Other firms

To gauge the external validity of our estimates we extract balance sheet data for all manufacturing firms from the ORBIS database for our sample countries and years, and we test whether our estimate of the gender productivity gap (the average productivity of women minus the average productivity of men at the country level) correlates with the productivity of other firms in the economy. We obtain a cross-section of firms based on the latest year of balance sheet reporting between 2012 and 2019. The sample contains 2 million firms in 158 SIC3 sectors (Figure A.7 illustrates the coverage of manufacturing employment in ORBIS compared to World Bank manufacturing employment shares). Intuitively, our estimate of the gender productivity gap captures differences in productivity due to differences in barriers at the societal level as well as idiosyncratic differences due to the MNE pay and recruitment policies. If the latter dominates, the estimated productivity gap will not be correlated with the productivity of other firms. Table 4 reports the estimates of the following model:

$$y_{isc} = \alpha l_{isc} + \beta k_{isc} + \gamma(A_F - A_M)_c + \delta X_c + \epsilon_{isc} \tag{7}$$

where y_{isc}, l_{isc}, k_{isc}, are, respectively, log operating revenue, log employment and log capital of firm i in sector s in country c. Since our estimated productivity gap varies at the country level, we control for GDP and the LFPR in X_c and cluster standard errors by country. The coefficient of our estimated productivity gap is negative, precisely estimated, and quite large, likely due to correlated unobservables at the country level. To estimate correlations within country, we split sectors into two groups: those where the MNE operates and those where it does not. In column 3 we estimate a fixed effect model that identifies the correlation between our productivity gap and the productivity of firms in the same sector minus the correlation with productivity of firms in different sectors.

[13]For 68 countries not in our firm sample, we impute the selection adjustment term based on a cubic regression on the LFPR.

The estimated elasticity in column 3 is about half of that to capital.

Table 5 estimates the model in equation 7 at the sector level, on both the mean and the variance of the productivity gap. In line with the findings at the firm level, average productivity is negatively correlated with the calibrated productivity gap. Moreover, the dispersion of productivity between firms in the same sector and country (a rough measure of misallocation) is higher in countries where the productivity gap is higher.

5 Counterfactuals

We use the model estimates to evaluate the effect of different counterfactuals on the LFPR, the pay gap and welfare. To do this we need to take a stance on how the firm responds to changes in the environment. Since we do not observe the production function of the firm nor the elasticity of demand they face we cannot use profit maximisation as the guiding criterion. Rather, we take the observed level of employment and the wage bill in each country as binding constraints.[14] Our first counterfactual asks whether, under these constraints, the firm maximises productivity. Secondly, we quantify the effect of misallocation on the firm's productivity. In the third counterfactual, we simulate the effect of stricter labor laws.

5.1 Does the firm maximise productivity?

To compute the optimal wage policy, we let the firm choose α^1_{gtac}, β^1_{gtac} to maximize the productivity of its employees, subject to two constraints: (i) keeping total employment (or the labor force) fixed, (ii) keeping the total wage bill constant.[15] The exact details of the maximization problems are discussed in Appendix A.3.1.

Figure 11, panels (a) and (b) compare the calibrated wage policy parameters to the solution of the optimization problem described above. We can see that these do not coincide, and for some countries they are quite far apart. The difference between the optimal and the observed parameters follows the same pattern in most countries: to

[14]In practice these are determined by a maximisation problem that we do not observe. This is equivalent to assume that the firm sets the optimal scale of operation in each country and then decides who to hire to maximise productivity.

[15]Both constraints are necessary to make sure that we obtain a sensible solution. Without the employment constraint, the firm can increase average productivity just by hiring fewer people (because of positive selection). Without the wage bill constraint, the firm can raise both α^1_{gtac} and β^1_{gtac} in a way that increases productivity without changing employment, at the cost of paying much higher wages.

maximise productivity the firm should increase the fixed pay of women α_F^1, decrease the fixed pay of men α_M^1 and increase variable pay β^1 for both genders.

Under the optimal policy the firm equates the participation thresholds of men and women, so that the LFPR gets close to 1 as shown in Figure 11 panel (e). While equating marginal productivity across groups is obviously the solution to the unconstrained problem, this exercise tells us that it is possible to do so while simultaneously keeping employment and the wage bill constant. The optimal policy effectively undoes differences in LFP and leads to higher productivity in every country as shown in Figure 11 panel (c). On average, the firm could increase productivity by 22%.

Why is the firm not setting the optimal α_{gtac}^1, β_{gtac}^1? A possible answer is that while the average pay is constant by assumption, the optimal policy generates a stark increase in inequality between genders. Indeed, because women in the labor force are more positively selected in most countries, it would be productivity-maximizing to pay women more, both in terms of fixed and variable pay, so that the pay gap between females and males would be even larger than what we observe. Without any change in norms, on average over all countries, the gender pay gap (Female – Male) would have to increase by 78% (Figure 11, panel (d)). Since individual productivity is not directly observable, such increase might not be acceptable.

Another important reason may be labor regulations, which limit variable pay and pay inequality even within gender. We address this in subsection 5.3.

Finally, the fact that the firm would be better off hiring more women (in most cases, hiring as many women as men) suggests that quotas would not bind. However, meeting them would require a large increase in inequality between genders, with steep rewards for talent, in order to sufficiently attract women into the labor force. This could be as stark as, for example, most leadership positions being held by women while all men would work as subordinates. Note that this would be a very different scenario than that of many policies which prescribe equality in pay and rewards between genders, and would imply an equal number of men and women in top-level positions, as well as in lower-level positions. [16]

[16] If the policy were to combine gender quotas with the imposition of equal pay, rather than rewarding talent regardless of gender, the firm will reduce mean productivity to minimise labour costs. Results available on request.

5.2 How much does misallocation cost?

Our next counterfactual quantifies the effect of misallocation on the firm's productivity and on workers' welfare. In our framework the true value of staying at home for a woman equals that of a man with same observable characteristics and (A_i, N_i) type, $y^{0,*}_{iFtac} = y^{0}_{iMtac}$. However, when entering the labor force, women must pay a "gender norms tax" τ_{tac} (as in Hsieh et al. (2019)) proportional to their true value of staying at home, so that when making the decision of whether to enter the labor force or not they take into account $y^{0}_{iFtac} = (1 + \tau_{tac})y^{0,*}_{iFtac}$. In this counterfactual we eliminate the gender norms tax by setting the value of the staying at home parameters of women (α^0, ν^0) equal to those of men (within the same country, cohort and tenure cells).

We discuss effects both in the short run, that is, keeping the pay policy of the firm fixed, and in the long run, when the firm can optimally adjust its policy to the new environment. To do this we must take a stance on the firm's optimisation procedure. In light of the results in subsection 5.1 we let the firm choose α^1_{gtac}, β^1_{gtac} to maximize the productivity of its employees, subject to three constraints: (i) keeping total employment (or labor force) fixed, (ii) keeping the total wage bill constant, (iii) a bound on the pay gap in the firm. This bound is chosen to maximize the goodness of fit between the constrained optimal and calibrated wage policy parameters at baseline. The exact details of the maximization problems are discussed in Appendix A.3.1.

Figure 12 plots the short-run effects of eliminating gender norms on the LFPR, the pay gap and average productivity. In the short run, MLFP does not change (because men's value of staying at home and wages stay the same), and FLFP increases, so LFP increases overall, which means that the ratio increases as well. These changes in LFP reduce the pay gap because the women that enter the labor force are less able on average than those who were already working at baseline, so women's average wages decrease. For the same reason, average productivity decreases, since MLFP does not change, but FLFP increases and the entrants are, on average, lower productivity than those who were already in the labor force. These findings mirror those in subsection 5.1: the baseline policy sacrifices productivity to bound pay inequality when entry barriers vary by gender, and hence, once barriers are equalised, the productivity cost rises.

The trade-off between productivity and inequality when the gender tax is set to zero however disappears in the long run (Figure 13). If the firm can adjust its policy, it will equate the participation thresholds so that the LFPR becomes 1. The fixed total employment constraint means that MLFP decreases by the same amount as FLFP increases. In the long run, because we equate the parameters of staying at home, it is optimal for the firm to have the same wage policy for men and women, which eliminates the pay gap. Again, the constant wage bill constraint means that this is just a redistribution from men to women. Finally, in a world without norms tax, the firm replaces less able men with more able women, and average productivity increases.

Figures 14 shows, by baseline LFPR, the average productivity of LFP entrants and leavers. Average productivity of LFP entrants is, because of positive selection, lower in those countries where the LFPR changes are largest. However, those entrants are on average much more productive than the leavers they replace. Figure 15 shows the productivity gains of eliminating gender norms by countries. The average across countries is 32%, although there is substantial heterogeneity, with some countries, such as Pakistan or Sri Lanka, obtaining potentially an increase of up to 90% in productivity.

5.3 Stricter labor laws

In a third counterfactual, we ask what the effect of stricter labor laws, that limit performance pay, would be for average productivity in the firm. To answer this question, we consider the constrained optimal wage policy and add an additional cap on returns to productivity: $\beta^1_{gtac} \leq \max\{\beta^1_{Fta,FRA}, \beta^1_{Mta,FRA}\}$, where $\beta^1_{Fta,FRA}, \beta^1_{Mta,FRA}$ denote the corresponding parameters for France (since France is the 95th percentile of the WEF Restrictive Labor Regulations Index). We leave the value of staying at home parameters $(\alpha^0_{gtac}, \nu^0_{gtac})$ unchanged at baseline, but we allow for the LFP of each group to respond optimally to the change in incentives.

The results are plotted in Figure 16. Average productivity of women does not change, whereas average productivity of men decreases substantially. Intuitively, limiting performance pay hurts the firm because they cannot screen out lower-productivity men, whereas it does not affect women because lower-productivity women are out of the labor force even at baseline.

6 Conclusion

We have shown that women are more positively selected into the labor force than men, leaving significant productivity gains unrealized. Understanding differential selection by gender, or indeed by any underrepresented group, is key to informing personnel policy as well broader labor market policies.

While our data is well suited to quantify the cumulative productivity loss due to the gendered division of labor, it does not shed light on the individual components that make up gendered labor division, such as barriers to education, how work affects marriage prospects, child penalties and so on. We calibrate gains from eliminating gender norms at every stage; if we eliminated barriers to LFP without eliminating barriers to education the gains would of course be smaller because the pool of qualified women applicants would be smaller. We thus do not know ex ante what the size of the pool of qualified women would be if certain gender barriers were eliminated but others were not. This precludes us from analysing the trade-off between diversity and quality that is at the core of popular policy tools such as quotas and affirmative action.

More generally our data is not well-suited to evaluate welfare impacts. Eliminating the norms that lead to higher barriers for women has two consequences for welfare. First, the reallocation of labor inside and outside the home creates winners and losers. The former are the women who are strongly suited for work outside the home who transition from inside to outside as well as the men who are strongly suited for work inside the home who move in the opposite direction. However, men who get crowded out by women in the workplace might lose. The extent to which this happens depends on the second consequence of eliminating misallocation, which is the increase in efficiency and productivity overall, which generates a virtuous circle through increases in demand. We cannot assess these effects because we do not know the production function of the firm, so we cannot say how the firm would adjust employment in response to a convergence of gender norms. In summary, we are able to quantify the effect of misallocation on the share of the pie that everybody gets but we cannot quantify the effect of misallocation on the size of the pie.

Our findings have three implications for the design of firms' diversity policies at the hiring and compensation stages. The first is that, other things equal, the average female applicant is more productive than the average male applicant. This implies that even if the firm does not observe potential productivity at the point of hiring, the logic of selection indicates that underrepresented groups will have higher productivity. This leads to the second implication, namely that paying women hires more is justifiable on the grounds of productivity, regardless of whether diversity has an effect of its own. The third implication is that aiming for gender equity — in pay, promotions, dismissals — can turn out to be inequitable because selection generates different distributions of productivity. Perhaps counter intuitively, gender equity policies might end up hurting women as they limit the firm's ability to reward performance. As things stand, greater rewards for objective performance could induce more women to enter the labor market (and stay there). Rewards need not be monetary; indeed in the presence of social norms where women may bear disproportionate burden of childcare and there are transaction costs, resources that would enable women to better manage childcare could have equally large productivity gains for the firm through the selection margin.

Awareness of the evidence of positive selection of under-represented groups changes how quality can be inferred, particularly if quality is not perfectly observable or objectively measured: the very presence of a member of an under-represented group should change one's prior on the talent of that member, simply through the logic of positive selection. The implications of increased awareness of this by leaders for the recruitment, promotion and productivity of under-represented groups is left to future research.

References

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the origins of gender roles: Women and the plough," *The quarterly journal of economics*, 2013, *128* (2), 469–530.

Alfaro-Urena, Alonso, Isabela Manelici, and Jose P Vasquez, "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-firm Linkages," *Working paper*, 2019.

Besley, Timothy and Robin Burgess, "Can Labor Regulation Hinder Economic Performance? Evidence from India," *The Quarterly Journal of Economics*, 2004, *119* (1), 91–134.

Blau, Francine D. and Lawrence M. Kahn, "The U.S. Gender Pay Gap in the 1990s: Slowing Convergence," *Industrial and Labor Relations Review*, 2006, *60* (1), 45–66.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *The quarterly journal of Economics*, 2007, *122* (4), 1351–1408.

Blundell, Richard, Amanda Gosling, Hidehiko Ichimura, and Costas Meghir, "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds," *Econometrica*, 2007, *75* (2), 323–363.

Borjas, George J., "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 1987, *77* (4), 531–553.

Boserup, Ester, *Woman's role in economic development*, Routledge, 1970.

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*," *The Quarterly Journal of Economics*, 10 2015, *131* (2), 633–686.

European Commission, "Report on gender pay gap in EU countries," 2019. `https://ec.europa.eu/info/policies/justice-and-fundamental-rights/gender-equality/equal-pay/gender-pay-gap-situation-eu_en#relatedlinks`, (accessed October 10, 2020).

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global evidence on economic preferences," *The Quarterly Journal of Economics*, 2018, *133* (4), 1645–1692.

— , — , **Thomas J Dohmen, David Huffman, and Uwe Sunde**, "The preference survey module: A validated instrument for measuring risk, time, and social preferences," 2016.

Foster, Andrew D and Mark R Rosenzweig, "Comparative advantage, information and the allocation of workers to tasks: Evidence from an agricultural labour market," *The Review of Economic Studies*, 1996, *63* (3), 347–374.

Goldin, Claudia, "The U-shaped female labor force function in economic development and economic history," 1995.

— , "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, April 2014, *104* (4), 1091–1119.

Hjort, Jonas, Xuan Li, and Heather Sarsons, "Across-Country Wage Compression in Multinationals," Technical Report, National Bureau of Economic Research 2020.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow, "The Allocation of Talent and US Economic Growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Hyland, Marie, Simeon Djankov, and Pinelopi Koujianou Goldberg, "Do Gendered Laws Matter?," 2021.

Lewis, W. Arthur, "Economic Development with Unlimited Supplies of Labour," *The Manchester School*, 1954, *22* (2), 139–191.

Mulligan, Casey B. and Yona Rubinstein, "Selection, Investment, and Women's Relative Wages over Time," *The Quarterly Journal of Economics*, 2008, *123* (3), 1061–1110.

Olivetti, Claudia and Barbara Petrongolo, "Unequal Pay or Unequal Employment? A CrossCountry Analysis of Gender Gaps," *Journal of Labor Economics*, 2008, *26* (4), 621–654.

Propper, Carol and John Van Reenen, "Can Pay Regulation Kill? Panel Data Evidence on the Effect of Labor Markets on Hospital Performance," *Journal of Political Economy*, 2010, *118* (2), 222–273.

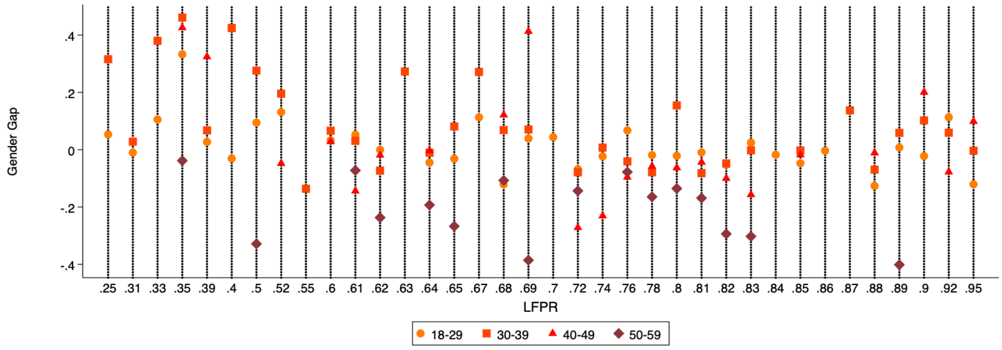
Roy, Andrew Donald, "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Strittmatter, Anthony and Conny Wunsch, "The Gender Pay Gap Revisited with Big Data: Do Methodological Choices Matter?," *arXiv preprint arXiv:2102.09207*, 2021.

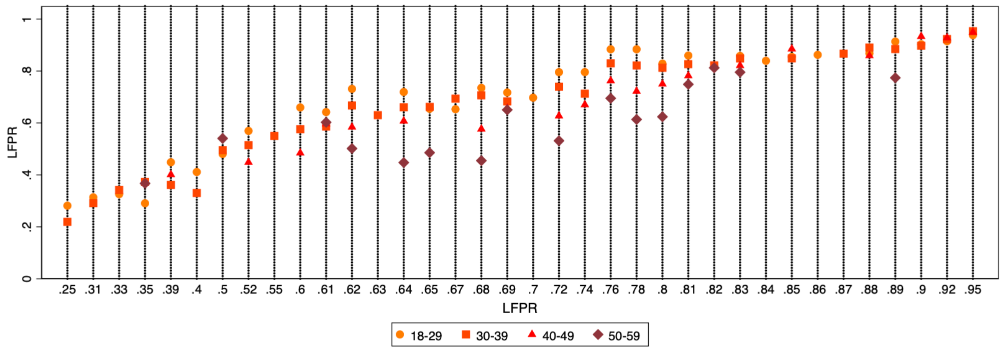
Wiswall, Matthew and Basit Zafar, "Preference for the Workplace, Investment in Human Capital, and Gender*," *The Quarterly Journal of Economics*, 08 2017, *133* (1), 457–507.

7 Figures

Figure 1: Gender gap and LFPR, across countries and cohorts



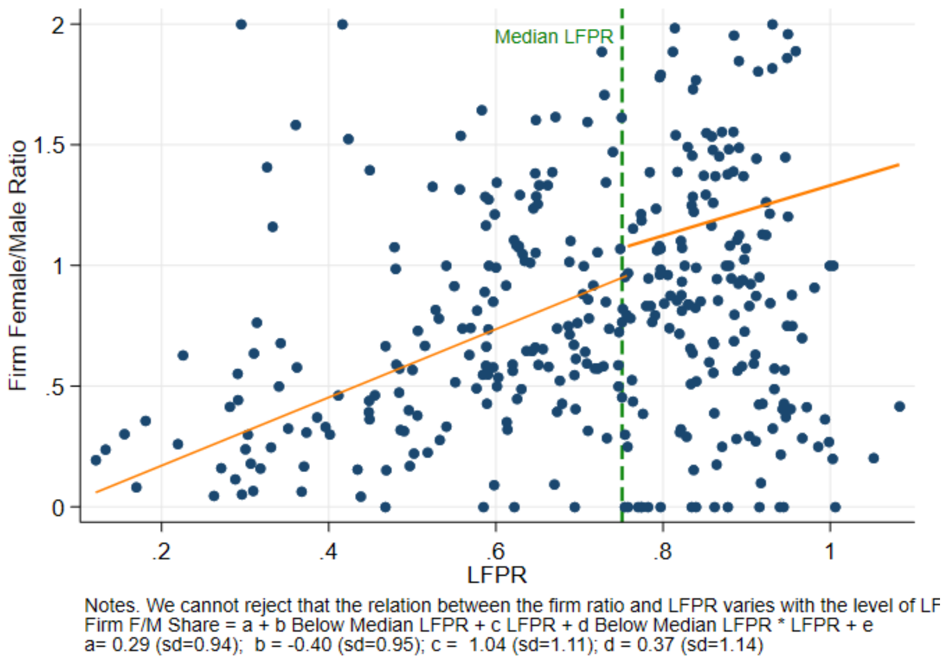
(a) A. Gender pay gap



(b) B. Female/Male LFP (LFPR)

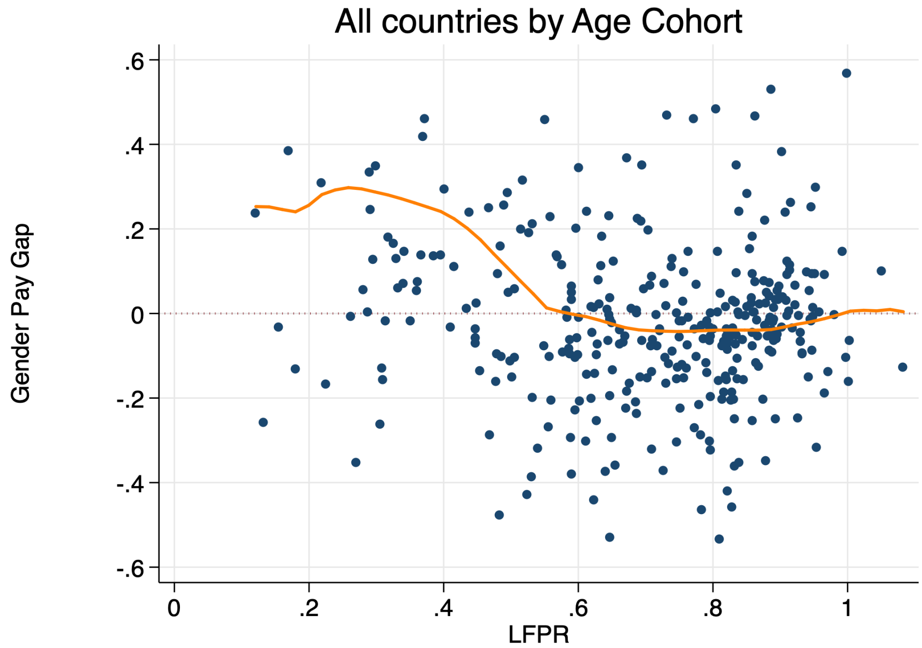
Notes. Panel A plots the gender pay gap (the difference between women and men's salary expressed as a % of men's salary) across countries and cohorts. Panel B plots the LFPR ratio (female to male LFP) across countries and cohorts. Both panels only include the largest 58 countries in terms of employee size. The x-axis is the LFPR in each country.

Figure 2: Female/Male in MNE vs LFPR



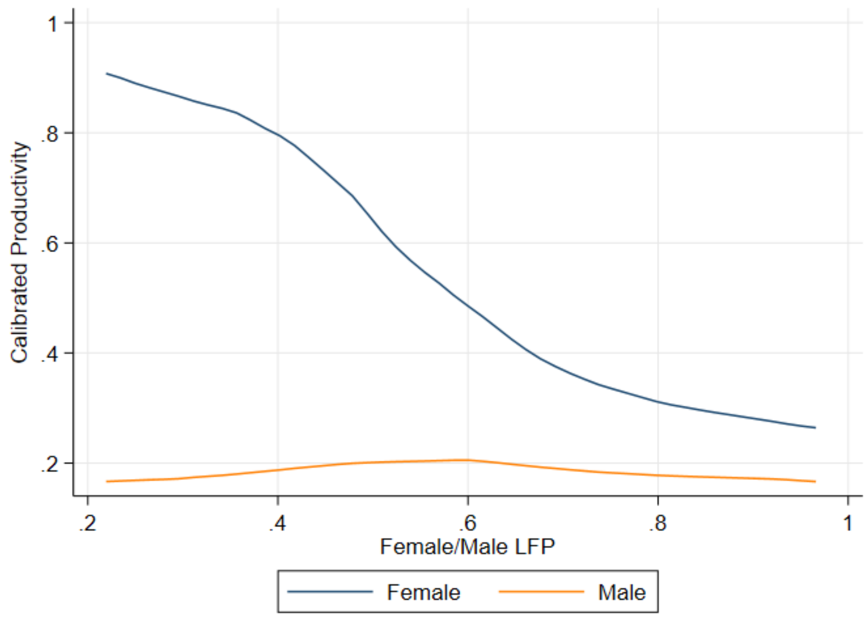
Notes. Each dot represents a country-cohort pair; cohort 1 is the 18-29 age group, cohort 2 is 30-39 age group, cohort 3 is 40-49 age group and cohort 4 is 50-59. The y-axis corresponds to the female/male employment ratio in the MNE while the x-axis corresponds to the LFPR in the countries. The orange lines represent the lines of best linear fit separately by whether the LFPR of each country-cohort is above or below the median LFPR (indicated by the green dashed line). In the regression, we use analytical weights by employee size of each cohort-country cell and robust standard errors.

Figure 3: Gender pay gap and LFPR



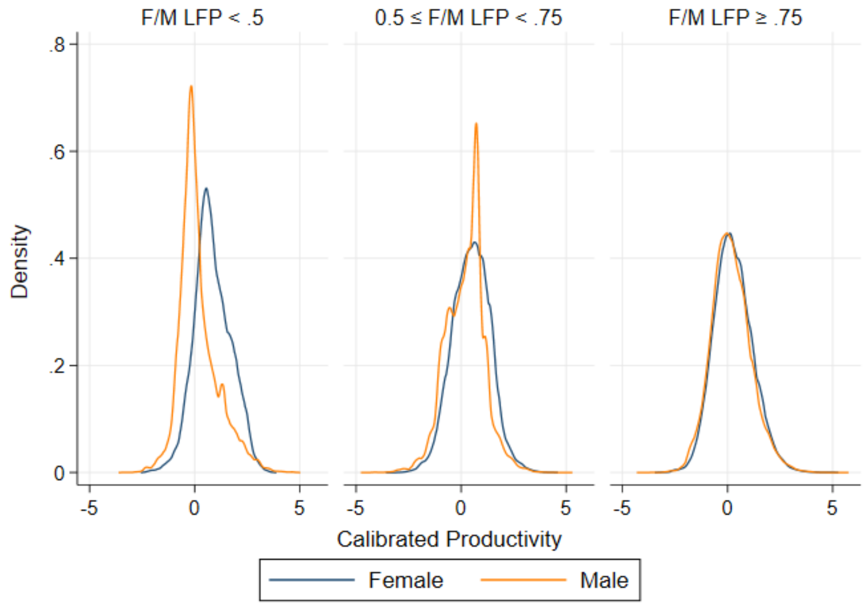
Notes. This figure plots the gender pay gap (the difference between women and men's salary expressed as a % of men's salary) against the LFPR. Each dot represents a country-cohort pair. The orange line represents the smoothed values of a kernel-weighted local polynomial regression using analytical weights by employee size of each cohort-country cell.

Figure 4: Calibrated productivity: average by LFPR



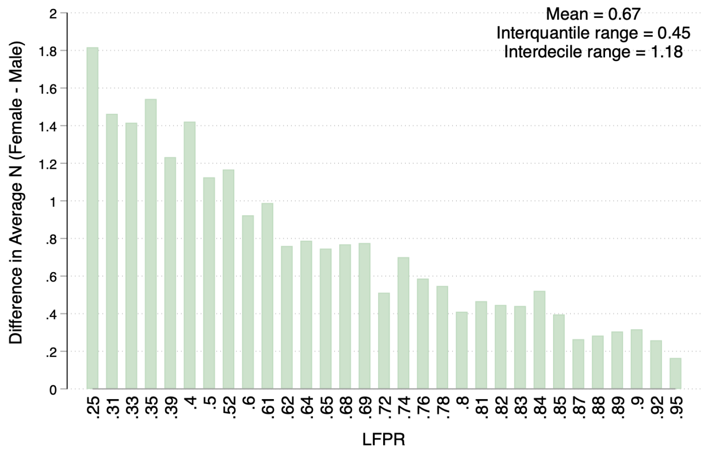
Notes. The figure plots average calibrated productivity for our sample of firm workers by Female/Male LFP, smoothed through a local linear regression.

Figure 5: Calibrated productivity: density by LFPR group



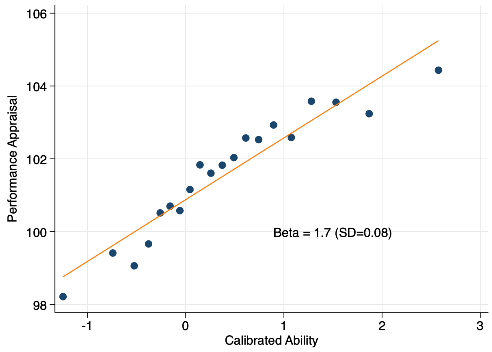
Notes. The figure plots a kernel density estimate of calibrated productivity for our sample of firm workers by three LFPR groups: $[0, .5)$, $[.5, .75)$ and $[.75, 1]$.

Figure 6: Counterfactual preference gap

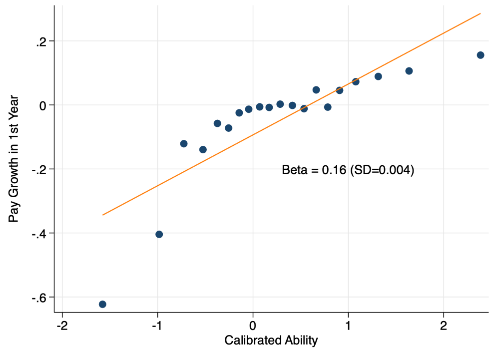


Notes. The figure plots the difference in means of N between genders that make the observed LFP gap optimal, computed as explained in the main text. The x-axis is the LFPR in each country.

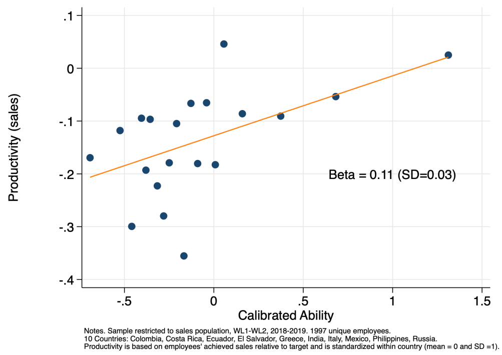
Figure 7: Validation of our calibrated productivity against other performance measures



(a) Performance appraisals



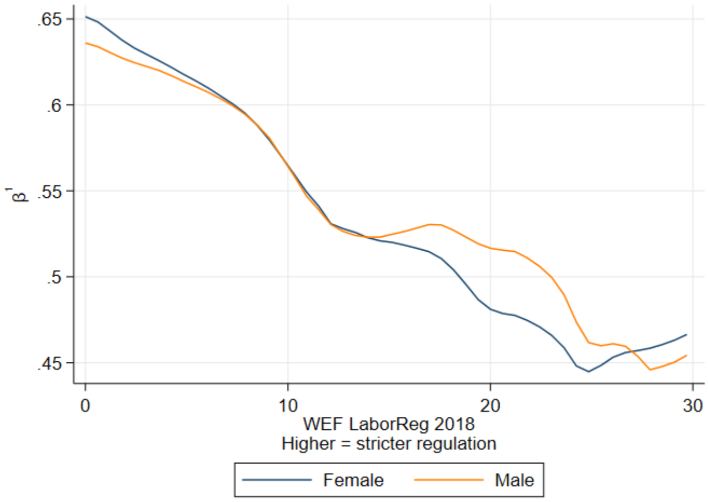
(b) Pay growth in the first year (new hires)



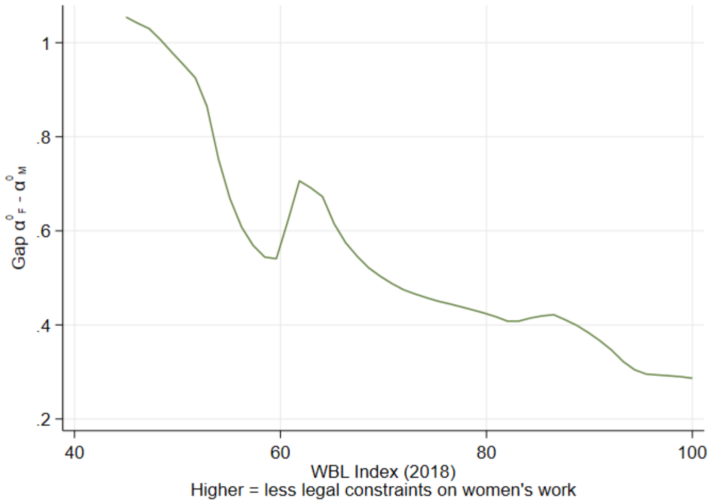
(c) Productivity (sales)

Notes. The figures are binned scatterplots and a linear fit of other performance measures (performance appraisals, pay growth for new hires and objective productivity) against our calibrated productivity. The objective productivity measure is available only for the sales function in 10 countries, and is based on reaching set targets.

Figure 8: Validation of our calibrated parameters against labor regulations



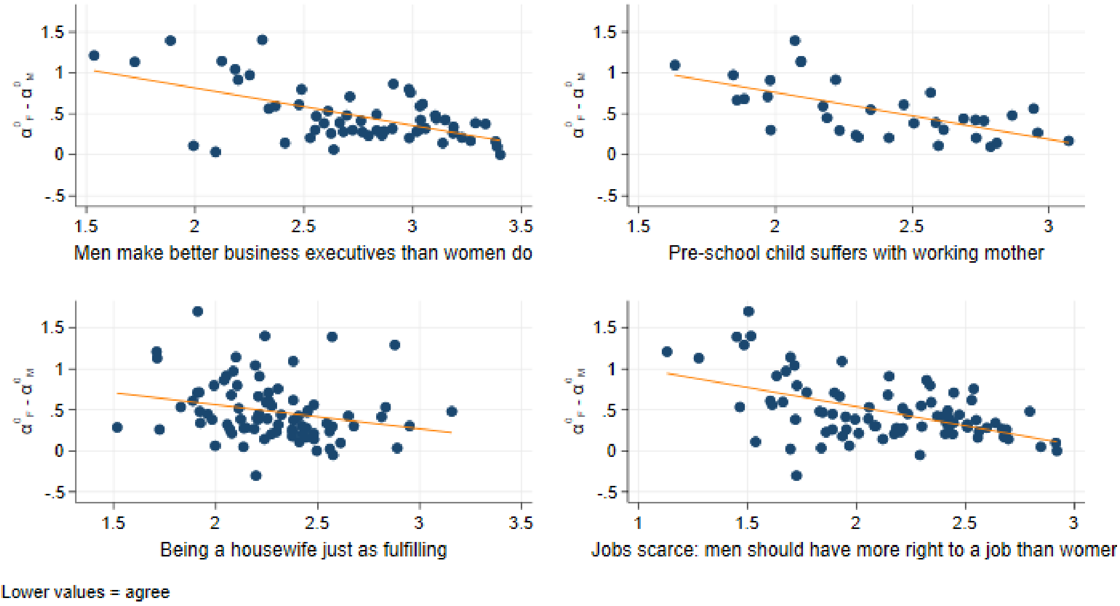
(a) Performance rewards and general labor regulation: β^1 against the Restrictive Labor Regulations Index (WEF)



(b) Value of staying at home and laws restricting or facilitating women's labor: $\alpha_F^0 - \alpha_M^0$ against the Women, Business and the Law Index (WB)

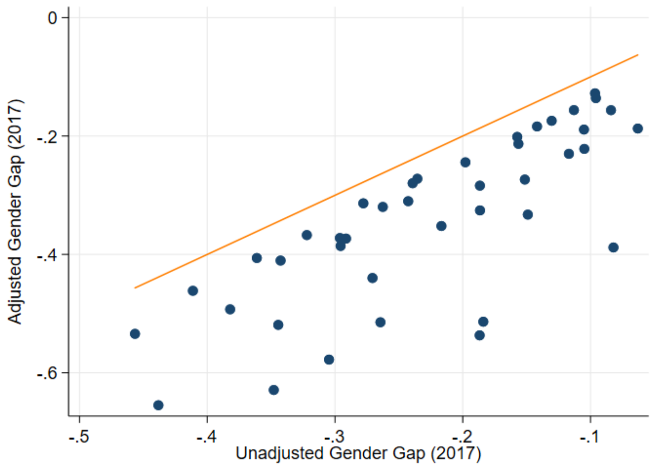
Notes. Both panels show local polynomials of our calibrated parameters against two indices related to labor regulations. Panel (a) plots the calibrated β^1_{gtac}, for men and women separately, against the WEF Restrictive Labor Regulations Index. Panel (b) plots the gap in calibrated $\alpha_F^0 - \alpha_M^0$ against the WB Women, Business and the Law index. Details about these indices are in the main text.

Figure 9: Validation of our calibrated parameters against values



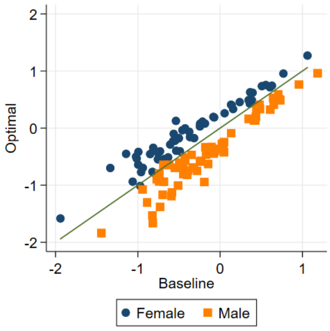
Notes. The figure shows scatterplots and fitted linear regressions of the gap in calibrated $\alpha^0_{Ftac} - \alpha^0_{Mtac}$ against four questions in the World Value Survey: (1) "Men make better business executives than women do," (2) "Pre-school child suffers with working mother," (3) "Being a housewife is just as fulfilling as working," (4) "When jobs are scarce, men should have more right to a job than women." For all questions, lower values of the index denote more agreement with the statement. Each dot is a country-cohort pair.

Figure 10: ILO gender pay gaps adjusted for selection

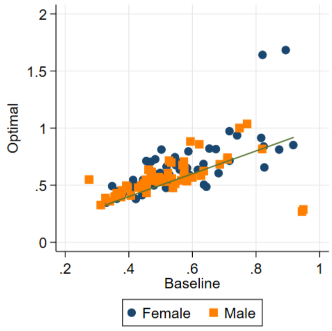


Notes. The figure is a scatterplot of the adjusted vs. unadjusted gender pay gaps. Each dot is a country. The original data (unadjusted) is from ILO, for ISCO-08 categories 1–5 (white collars). We adjust those for selection using our calibrated parameters by the procedure described in the main text.

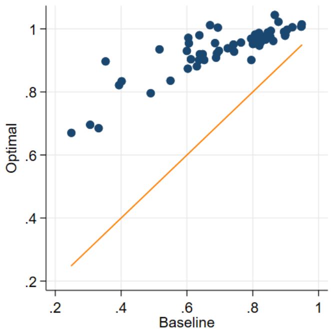
Figure 11: Baseline vs. optimal wage policy



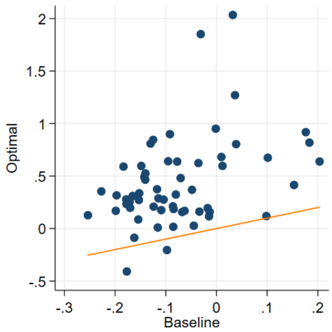
(a) α^1



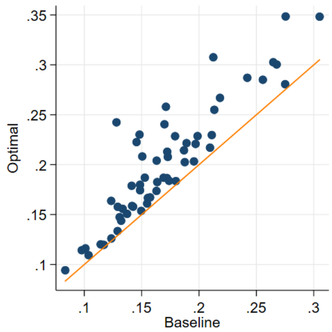
(b) β^1



(c) F/M LFPR



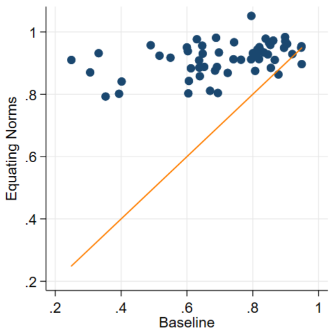
(d) Pay gap



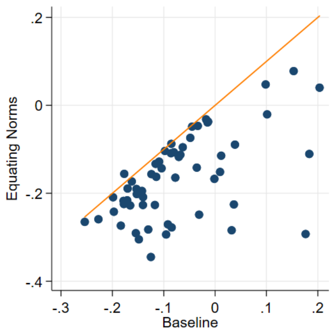
(e) Average productivity

Notes. The figures compare different outcomes (female to male LFP ratio, pay gap and average productivity) and the wage policy parameters (α^1, β^1) at baseline vs. the optimal wage policy (see main text for details). Each dot represents a country and the 45 degree line is included.

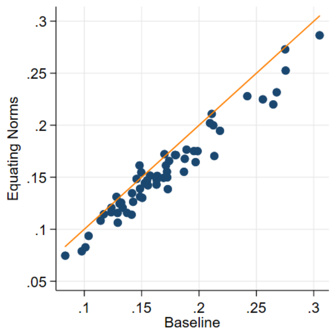
Figure 12: Baseline vs. equating norms (short run)



(a) F/M LFPR



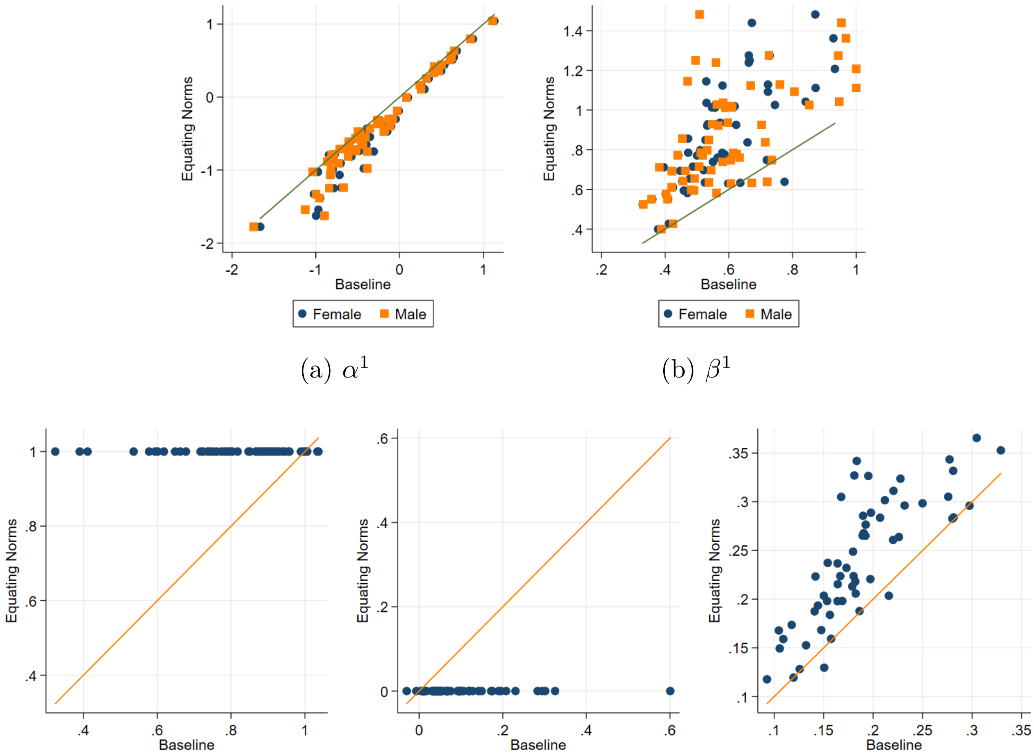
(b) Pay gap



(c) Average productivity

Notes. The figures compare different outcomes (female to male LFP ratio, pay gap and average productivity) under the baseline norm parameters (α^0, ν^0) to the counterfactual where these are equalized at the male levels. The "short-run" wage policy scenario keeps the wage policy of the firm fixed at the calibrated baseline parameters. Each dot represents a country and the 45 degree line is included.

Figure 13: Baseline vs. equating norms (long run)

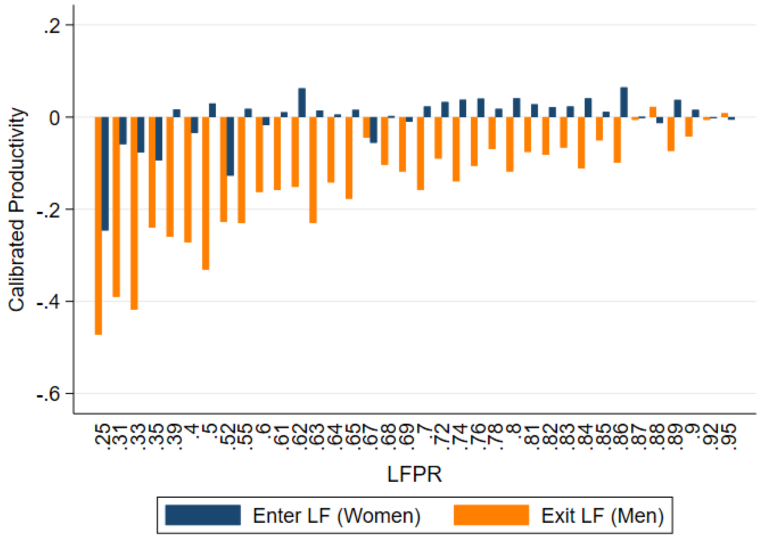


(a) α^1 (b) β^1

(c) F/M LFPR (d) Pay gap (e) Average productivity

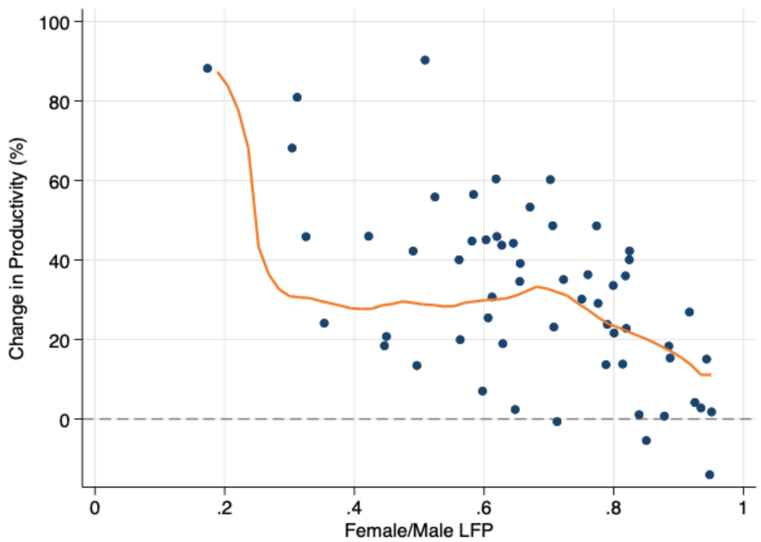
Notes. The figures compare different outcomes (female to male LFP ratio, pay gap and average productivity) and the wage policy parameters (α^1, β^1) under the baseline norm parameters (α^0, ν^0) to the counterfactual where these are equalized at the male levels. The "long-run" wage policy scenario lets the firm optimize the wage policy to maximize productivity under certain constraints (see the main text for details). Each dot represents a country and the 45 degree line is included.

Figure 14: Average productivity of entrants and leavers by country (long run)



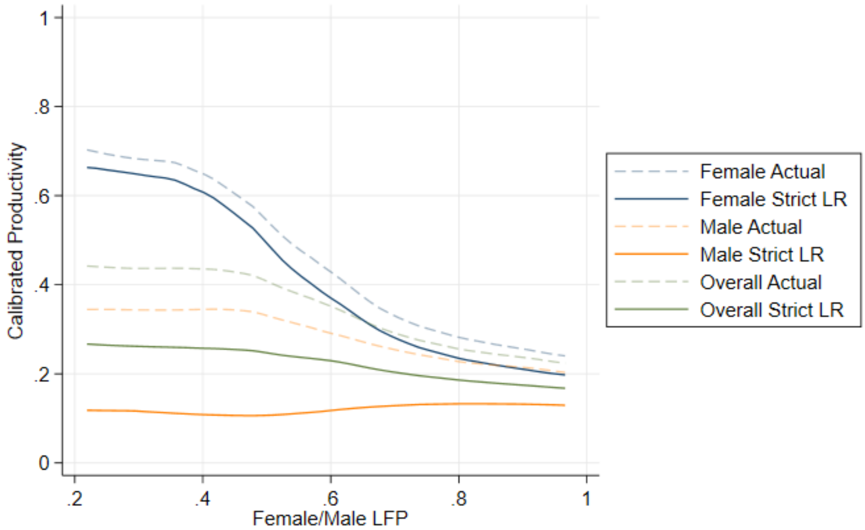
Notes. The figure plots the average productivity of LFP entrants against leavers in the long run wage policy response. The x-axis is the LFPR in each country.

Figure 15: Increase in average productivity by LFPR (long run)



Notes. The figure plots the percentage increase in average productivity when eliminating social norms. Each dot is a country. The average across countries is 32%.

Figure 16: Counterfactual average productivity with strict labor regulation



Notes. The figure plots a local polynomial regression of average productivity at baseline and under the strict labor regulation counterfactual, both estimated under the constrained optimal policy.

8 Tables

Table 1: Summary statistics

	(1) Male	(2) Female
Panel A: Pay, tenure and hierarchy		
Pay + Bonus (logs)	10.467	10.426
	(0.695)	(0.648)
Tenure	12.027	11.869
	(8.605)	(8.609)
Share in Work-level 1	0.757	0.786
	(0.286)	(0.272)
Share in Work-level 2	0.172	0.168
	(0.230)	(0.233)
Share in Work-level 3+	0.071	0.046
	(0.169)	(0.129)
Panel B: Age groups		
Share in Cohort 18-29	0.119	0.130
	(0.324)	(0.337)
Share in Cohort 30-39	0.250	0.256
	(0.433)	(0.436)
Share in Cohort 40-49	0.345	0.345
	(0.475)	(0.475)
Share in Cohort 50-59	0.286	0.269
	(0.452)	(0.444)
Panel C: College Major		
Econ, Business, and Admin	0.511	0.543
	(0.387)	(0.381)
Sci, Engin, Math, and Stat	0.300	0.211
	(0.347)	(0.304)
Social Sciences and Humanities	0.137	0.166
	(0.262)	(0.280)
Observations	3,338	3,103

Notes. This table reports summary statistics for the relevant sample of workers used in the analysis. An observation is a gender-cohort-country-tenure cell (tenure is binned in groups of 2 years each). This is the relevant unit in the structural estimation. Tenure is measured in years. Cohort refers to the age group and work level denotes the hierarchical tier (from level 1 at the bottom to level 6).

Table 2: Gender pay gap and LFPR

	Full Sample				New Hires
	(1)	(2)	(3)	(4)	(5)
	Pay + Bonus (logs)				
Female	0.377	0.256	0.253	0.197	0.210
	(0.142)	(0.0962)	(0.0941)	(0.0770)	(0.105)
LFPR	1.640	1.641	1.662	0.138	1.651
	(0.282)	(0.212)	(0.204)	(0.235)	(0.228)
Female × LFPR	-0.564	-0.471	-0.451	-0.376	-0.282
	(0.194)	(0.135)	(0.132)	(0.103)	(0.142)
Controls	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	Yes	No	Yes
Country FE	No	No	No	Yes	No
N	303756	303756	303756	303756	63887
R-squared	0.116	0.285	0.307	0.540	0.334

Notes. An observation is a worker-year. Year FE are included in every specification. Controls include: tenure, tenure squared and function FE. The last column reports estimates when restricting the sample to employees with no more than 1 year of tenure. Standard errors clustered at the country-cohort level.

Table 3: Summary of model parameters and empirical targets

Param.	Interpretation	Empirical Target
α^0_{gtac}	Unconditional average value of staying at home	LFP
α^1_{gtac}	Unconditional average log-wage	Average observed log-wage (controlling for selection)
β^1_{gtac}	Returns to productivity in the firm	Variance of the observed log-wage (controlling for selection)
ν^0_{gtac}	Dispersion of the idiosyncratic taste for staying at home	Not identified separately (Normalize $(\beta^1_{gtac})^2 + (\nu^0_{gtac})^2 = 1$)

Table 4: $\overline{A}_F - \overline{A}_M$ and log(operating revenue), ORBIS

	Log(OpRev)		
	(1)	(2)	(3)
Productivity gap	-9.151	-12.98	
	(1.281)	(1.879)	
Same SIC 3=1 × Productivity gap			-0.165
			(0.0385)
Log(employment)	0.607	0.616	0.610
	(0.212)	(0.215)	(0.00101)
Same SIC 3=1 × Log(employment)			0.148
			(0.00231)
Log(capital)	0.425	0.420	0.372
	(0.0608)	(0.0591)	(0.000664)
Same SIC 3=1 × Log(capital)			0.0157
			(0.00112)
Log(GDP)	0.532	0.517	
	(0.156)	(0.136)	
LFPR		-6.530	
		(3.357)	
Same SIC 3=1			-0.363
			(0.0140)
Country FE	No	No	Yes
R-squared	0.665	0.672	0.714
N	2239881	2239881	2239881

Notes. An observation is a firm in the ORBIS database. Cross-section based on latest year up to 2019, sample restricted to firms whose latest year is after 2011. Standard errors clustered at the country level in cols. 1 and 2 and robust in col. 3.

Table 5: $\overline{A}_F - \overline{A}_M$ and productivity dispersion, ORBIS

	Log(OpRev/emp.), Mean			**Log(OpRev/emp.), CV**		
	(1)	(2)	(3)	(4)	(5)	(6)
Productivity gap	-17.55	-11.39	-11.31	0.577	0.556	0.529
	(3.757)	(2.422)	(2.492)	(0.168)	(0.144)	(0.145)
Log(GDP)	0.868	0.483	0.483	0.00652	0.0111	0.0111
	(0.248)	(0.173)	(0.173)	(0.00782)	(0.00948)	(0.00948)
LFPR	-7.750	-6.892	-6.907	0.365	0.388	0.386
	(5.438)	(3.493)	(3.501)	(0.278)	(0.242)	(0.242)
Log(capital), Mean		0.458	0.449			
		(0.0651)	(0.0654)			
Same SIC 3=1 × Productivity gap			-0.379			0.143
			(0.613)			(0.0393)
Same SIC 3=1 × Log(capital), Mean			0.0622			
			(0.0192)			
Log(capital), CV					0.148	0.143
					(0.0851)	(0.0863)
Same SIC 3=1 × Log(capital), CV						0.0367
						(0.0321)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.751	0.890	0.891	0.485	0.523	0.529
N	2418	2418	2418	2418	2418	2418
Outcome mean	10.82	10.82	10.82	0.145	0.145	0.145
Outcome sd	2.006	2.006	2.006	0.0590	0.0590	0.0590

Notes. An observation is an industry (US SIC 3) -country cell in the ORBIS database. Analytics weights used. Measures based on cross-section of firms based on latest year up to 2019, sample restricted to firms whose latest year is after 2011. Standard errors clustered at the country level.

A Appendix

A.1 Framework and descriptive analysis

Figure A.1: The countries where the MNE operates and female to male LFP

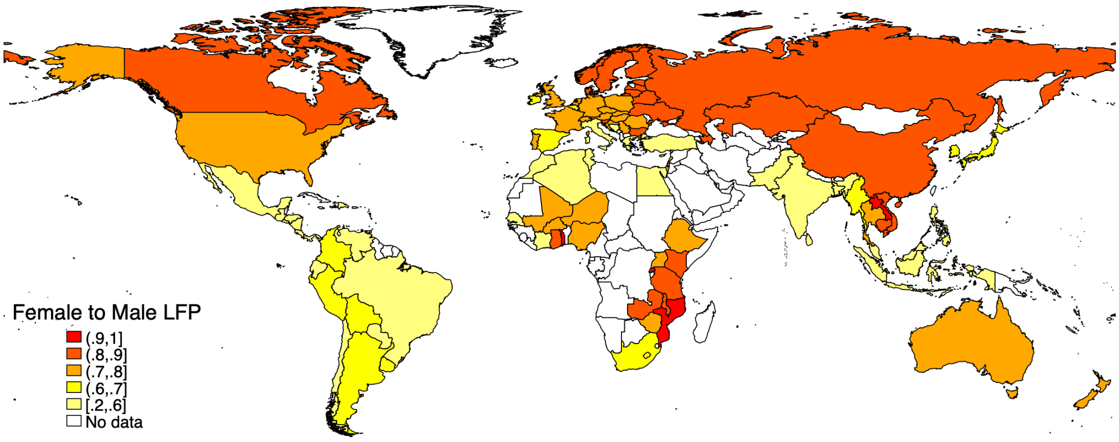
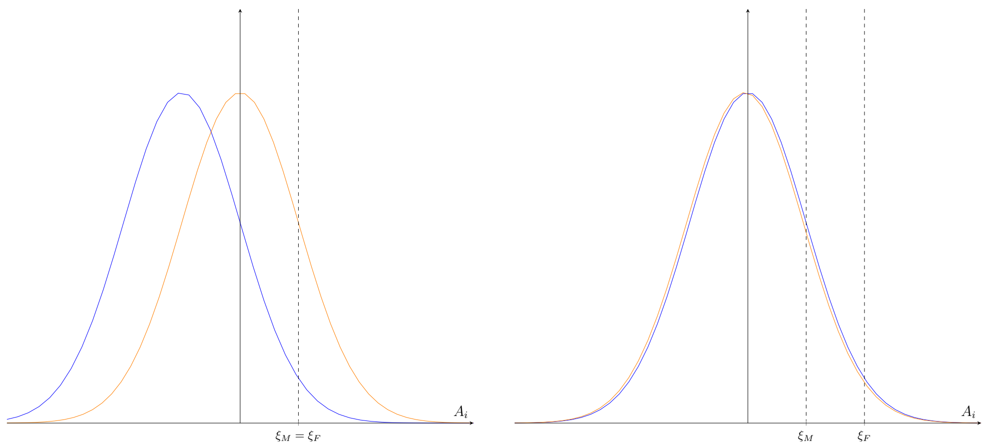


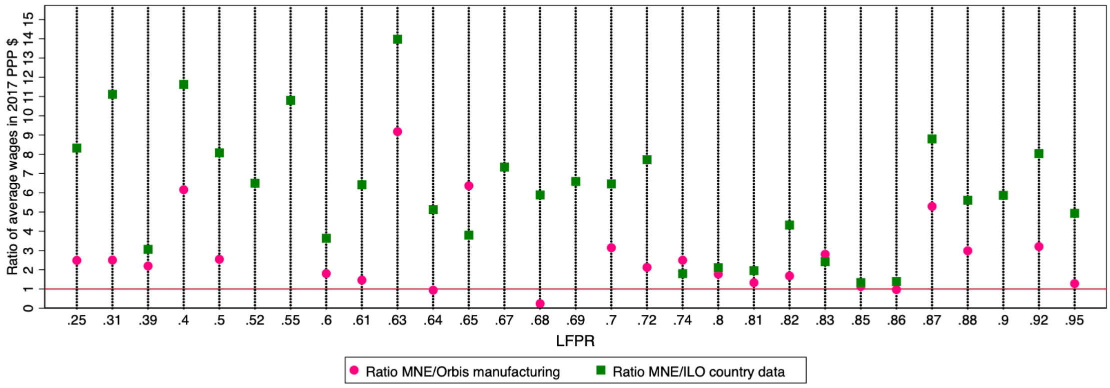
Figure A.2: Misallocation test



(a) H0: No misallocation, *male advantage*. The LFPR and the productivity gap at work are positively correlated. Women's distribution is the blue line while the men's distribution is the orange line.

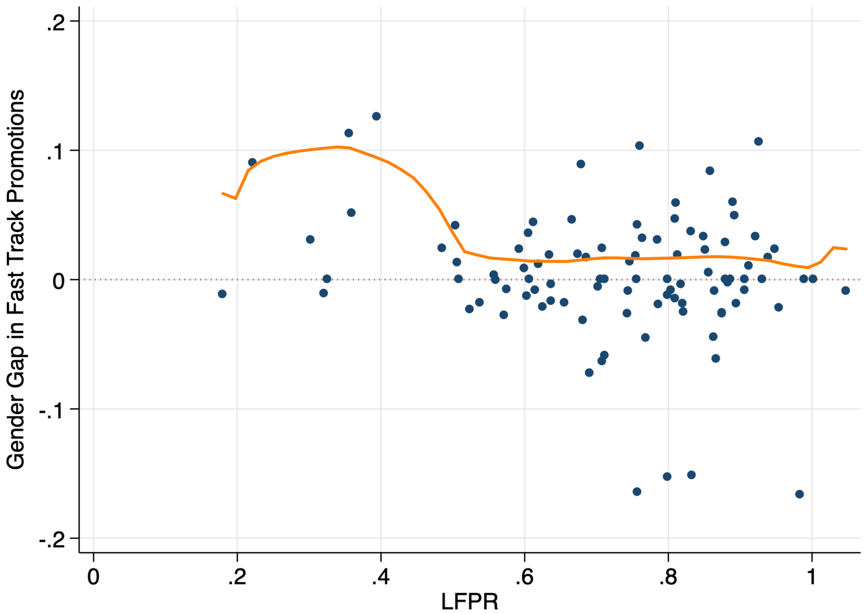
(b) H1: Misallocation, *barriers to entry*. The LFPR and the productivity gap at work are negatively correlated

Figure A.3: Average wages in firm and in country overall: a) ILO, white collar occupations only; and b) ORBIS, manufacturing sector only



Notes. This figure plots the ratio of average wage in the MNE and in the country overall: a) from the ORBIS database, considering the manufacturing sector only and b) from the International Labor Organization, considering white collar occupations only. Wages are measured in 2017 PPP $. The x-axis is the LFPR in each country.

Figure A.4: Gender gap in fast-track promotions and LFPR



Notes. This figure plots the gender gap in fast-track promotions against the LFPR. Each dot represents a country. The orange line represents the smoothed values of a kernel-weighted local polynomial regression using analytical weights by employee size of each cohort-country cell.

Table A.1: Gender pay gap and LFPR — fixed pay only

	Full Sample				New Hires
	(1)	(2)	(3)	(4)	(5)
	Pay (logs)				
Female	0.358	0.232	0.230	0.167	0.162
	(0.135)	(0.0923)	(0.0910)	(0.0741)	(0.0956)
LFPR	1.582	1.578	1.599	0.144	1.598
	(0.282)	(0.212)	(0.204)	(0.234)	(0.228)
Female × LFPR	-0.533	-0.436	-0.415	-0.328	-0.212
	(0.185)	(0.131)	(0.129)	(0.0997)	(0.130)
Controls	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	Yes	No	Yes
Country FE	No	No	No	Yes	No
N	303756	303756	303756	303756	63887

Notes. An observation is a worker-year. Year FE are included in every specification. Controls include: tenure, tenure squared and Function FE. The last column reports estimates when restricting the sample to employees with no more than 1 year of tenure. Standard errors clustered at the country-cohort level.

Table A.2: Gender pay gap and LFPR — by region and income group

	All	Region FE	Lower inc.	Higher inc.
	(1)	(2)	(3)	(4)
	Pay + Bonus (logs)			
Female	0.256	0.208	0.265	-0.0320
	(0.0962)	(0.0696)	(0.0503)	(0.111)
LFPR	1.641	0.698	0.547	1.384
	(0.212)	(0.186)	(0.134)	(0.338)
Female × LFPR	-0.471	-0.424	-0.408	-0.129
	(0.135)	(0.0996)	(0.0682)	(0.156)
Controls	Yes	Yes	Yes	Yes
Cohort FE	No	No	No	No
Country FE	No	No	No	No
N	303756	303756	71658	232098
R-squared	0.285	0.387	0.218	0.239

Notes. An observation is a worker-year. Year FE are included in every specification. Controls include: tenure, tenure squared and function FE. Standard errors clustered at the country-cohort level. Income group and geographical region are obtained from the World Bank.

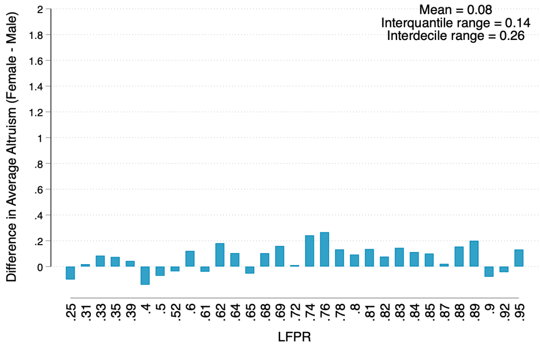
Table A.3: Gender pay gap and LFPR — PPP conversion

	Full Sample				New Hires
	(1)	(2)	(3)	(4)	(5)
	Pay + Bonus (logs), PPP 2017 USD				
Female	0.318	0.195	0.197	0.198	0.285
	(0.127)	(0.0795)	(0.0803)	(0.0762)	(0.0873)
LFPR	0.266	0.249	0.259	0.129	0.405
	(0.200)	(0.136)	(0.130)	(0.239)	(0.159)
Female × LFPR	-0.463	-0.372	-0.364	-0.381	-0.402
	(0.165)	(0.104)	(0.105)	(0.102)	(0.114)
Controls	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	Yes	No	Yes
Country FE	No	No	No	Yes	No
N	302789	302789	302789	302789	63598
R-squared	0.0143	0.164	0.173	0.339	0.169

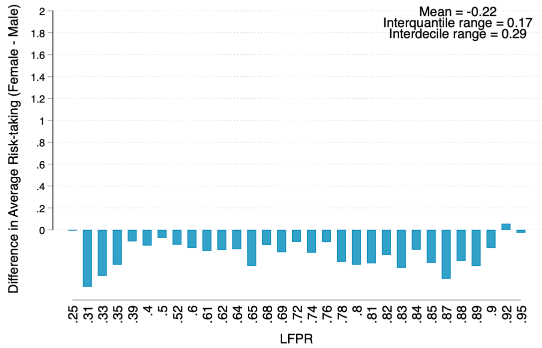
Notes. An observation is a worker-year. Year FE are included in every specification. Controls include: tenure, tenure squared and function FE. The last column reports estimates when restricting the sample to employees with no more than 1 year of tenure. Standard errors clustered at the country-cohort level. Wages are measured in PPP 2017 USD. Purchasing power parity (PPP) exchange rates are taken from the ICP (World Bank).

A.2 Structural model

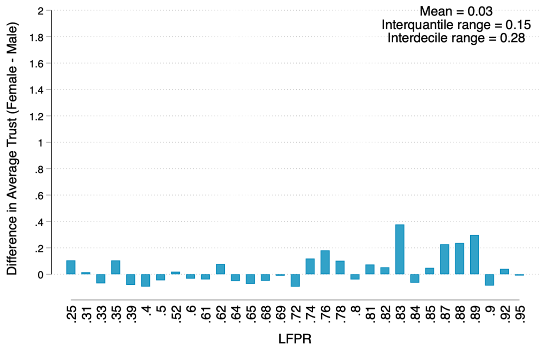
Figure A.5: Gender differences in preferences



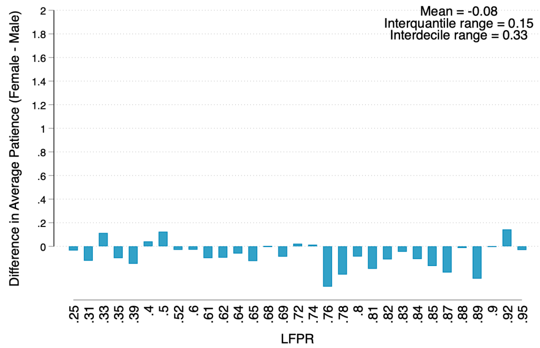
(a) Altruism



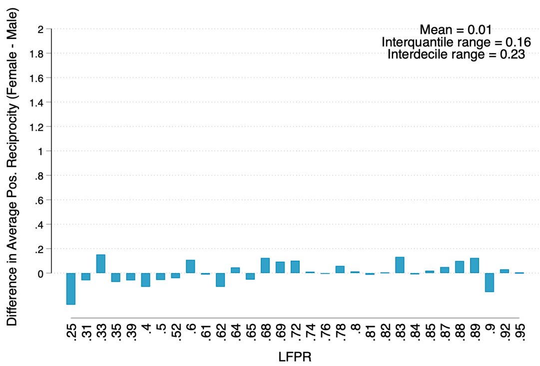
(b) Risk-taking



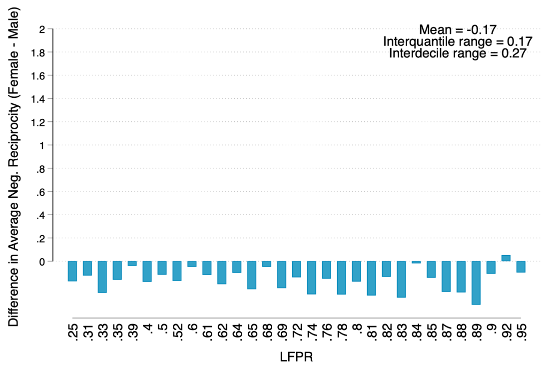
(c) Trust



(d) Patience



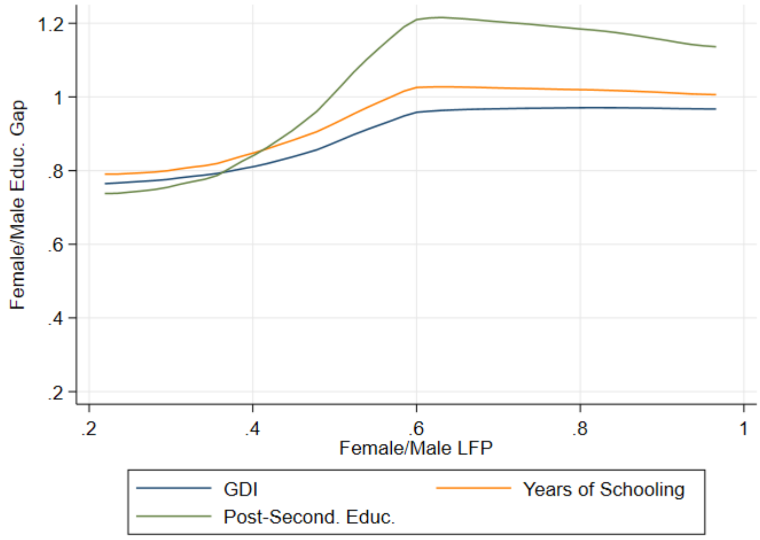
(e) Positive reciprocity



(f) Negative reciprocity

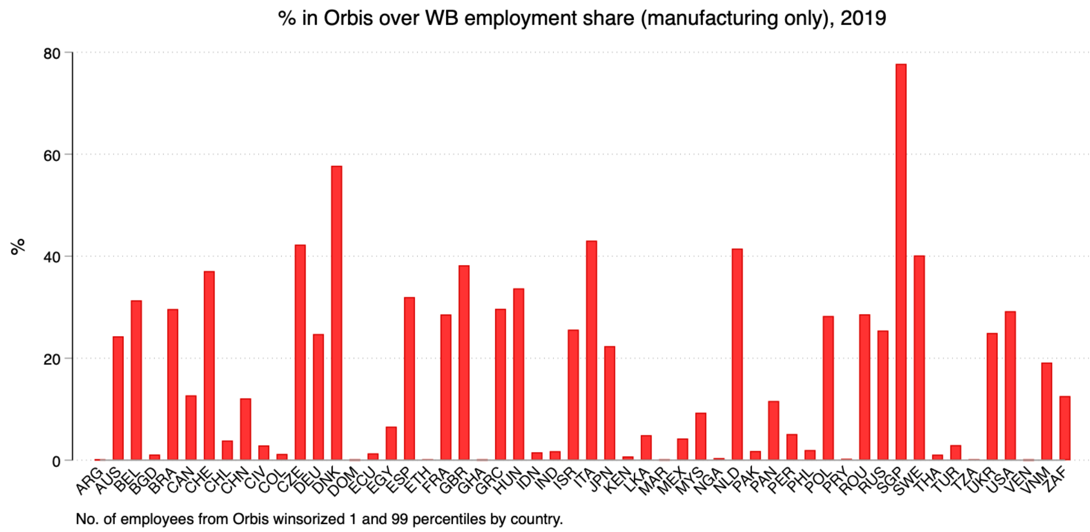
Notes. The figure plots the difference in means of economic preferences between genders. Data taken from the Global Preferences Survey (Falk et al. (2016), Falk et al. (2018)). The x-axis is the LFPR in each country.

Figure A.6: The gender education gap versus the LFP gap



Notes. The figure plots the gender gap in education against the gender gap in LFP. We use a number of education measures: the gender development index (GDI, the ratio of female/male Human Development Index); years of schooling (female to male ratio) and the percentage in post-secondary education (female to male ratio). The GDI data is from the UNDP and the educational attainment data is from the World Bank.

Figure A.7: Manufacturing employment in ORBIS versus World Bank data (2019)



Notes. The figure plots the number of employees in manufacturing from the ORBIS firm-level data against the World Bank manufacturing employment numbers at the country level.

A.3 Counterfactuals

A.3.1 Optimization problems for the firm's wage policy under the counterfactuals

The optimal policy we consider in subsection 5.1 solves, for each country-cohort-tenure cell, the following program:

$$
\begin{aligned}
\max_{(\alpha^1_{gtac},\beta^1_{gtac})_{g\in\{F,M\}}} \quad & \left[1-\Phi\left(\widetilde{\xi}_{Ftac}\right)\right]\frac{\beta^1_{Ftac}}{\sigma_{Ftac}}\lambda\left(\widetilde{\xi}_{Ftac}\right)+\left[1-\Phi\left(\widetilde{\xi}_{Mtac}\right)\right]\frac{\beta^1_{Mtac}}{\sigma_{Mtac}}\lambda\left(\widetilde{\xi}_{Mtac}\right) \\
\text{subj. to:} \quad & 1-\Phi\left(\widetilde{\xi}_{Ftac}\right)+1-\Phi\left(\widetilde{\xi}_{Mtac}\right)=FLFP_{tac}+MLFP_{tac} \qquad \text{(i)} \\
& \left[1-\Phi\left(\widetilde{\xi}_{Ftac}\right)\right]\left[\alpha^1_{Ftac}+\frac{(\beta^1_{Ftac})^2}{\sigma_{Ftac}}\lambda\left(\widetilde{\xi}_{Ftac}\right)\right]+ \\
& \quad +\left[1-\Phi\left(\widetilde{\xi}_{Mtac}\right)\right]\left[\alpha^1_{Mtac}+\frac{(\beta^1_{Ftac})^2}{\sigma_{Mtac}}\lambda\left(\widetilde{\xi}_{Mtac}\right)\right]=\overline{y}^1_{Ftac}+\overline{y}^1_{Mtac} \\
& \qquad \text{(ii)}
\end{aligned}
$$

where $\widetilde{\xi}_{gtac}=(\alpha^0_{gtac}-\alpha^1_{gtac})/\sigma_{gtac}$, $\sigma_{gtac}=\sqrt{(\beta^1_{gtac})^2+(\nu^0_{gtac})^2}$. The objective is average productivity in the firm ($LFP\cdot\mathbb{E}[A\,|\,\text{empl'd}]$). Constraint (i) states that total employment (or LFP) should stay constant. Constraint (ii) states that the total wage bill should be unchanged.

The long-run constrained optimal policy considered in the counterfactuals in subsection 5.2 solves the program above, with the additional constraint:

$$
\left|\left[\alpha^1_{Ftac}+\frac{(\beta^1_{Ftac})^2}{\sigma_{Ftac}}\lambda\left(\widetilde{\xi}_{Ftac}\right)\right]-\left[\alpha^1_{Mtac}+\frac{(\beta^1_{Mtac})^2}{\sigma_{Mtac}}\lambda\left(\widetilde{\xi}_{Mtac}\right)\right]\right|\leq B \qquad \text{(iii)}
$$

Constraint (iii) places a limit B on how unequal the wage of the average men and women in the firm can be. The bound B is chosen to maximize the goodness of fit to the data, measured by the distance $\|\theta^*_B-\widehat{\theta}\|$ (where $\widehat{\theta}$ is the vector of calibrated $\alpha^1_{gtac}, \beta^1_{gtac}$ and θ^*_B is the solution of the constrained optimal policy problem with bound B).

A.3.2 Effects on workers' welfare

The effect of eliminating gender norms on wages (and hence on workers' utility) depends on the assumptions about the wage policy of the firm. If the wage parameters stay fixed at baseline levels, the only welfare effects are on the women who enter the labor force. By a revealed preference argument, those women gain because their wage in the firm is higher than their underlying value of staying at home (which is assumed to be the same as men's); the only reason why they were not in the labor force before at the current wages is the gender norm tax. There is no welfare effect on male workers nor on the women who are already in the labor force, since their wages are unchanged. Hence, the net welfare effect on workers is positive in the short run.

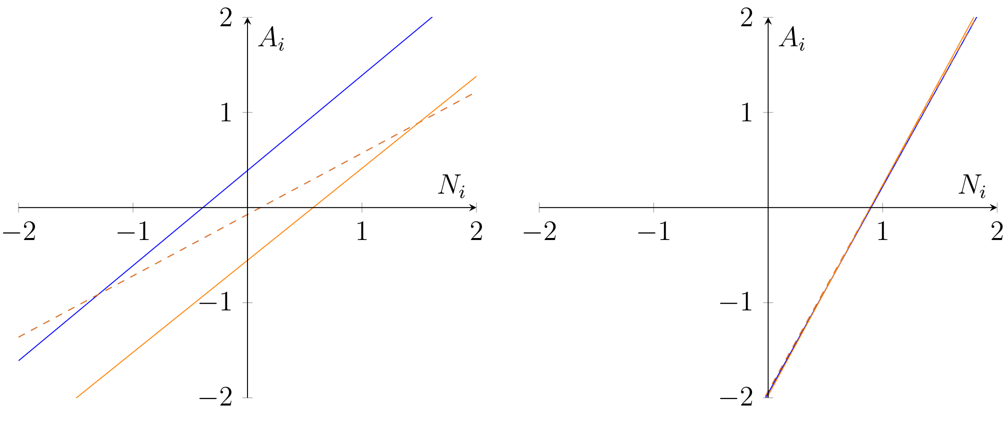
When the firm adjusts, in contrast, because we are imposing the constraint that the wage bill stays fixed, there are winners and losers of the policy, and men can be affected by the elimination of the gender norm tax too. In particular, by the same revealed preference argument as above, those who enter the labor force will gain, and those who exit the labor force will lose. The sign of the welfare change for those who are always in the labor force depends on how the firm optimally reacts to the elimination of gender norms: in particular, if β_1 increases, as we find in Figure 13 panel (b), higher-productivity workers will benefit from a wage increase and lower-productivity workers will suffer from a wage decrease.

To illustrate, we focus on the two countries on opposite ends of the LFPR distribution: India and Sweden. The solid lines in Figure A.8 represent the participation frontier, $\beta^1 A_i - \nu^0 N_i \geq \alpha^0 - \alpha^1$, for India and Sweden in the long-run horizon (averaging across cohort and tenure groups, weighted by size). The dashed lines in the same figure show the change after the gender norm tax is eliminated. LFP and the participation frontier are very equal in Sweden to begin with, hence the effect of equating α^0, ν^0 for men and women will be very small, and effectively there will be no welfare change. In contrast, in India the participation frontiers are very different to begin with, and equating the value of staying at home leaves room to attract high-productivity females into the labor force. Under the long-run wage policy we consider, once the gender norm tax is eliminated, the firm can maximize productivity by equating the wage parameters for men and women (which results, on average, in a wage increase for women and a wage decrease for men). The magnitude of this change depends on how constrained the firm was originally, as shown in Figure A.9.

The results for all countries are in Figure A.10, which plots the change in value of chosen option y^* (which can be interpreted as a percentage change, because it is in units of log-wages). For men, $y^* = \max\{y^0, y^1\}$. For women, $y^* = y^1_F \mathbf{1}\{y^1_F \geq y^0_F\} + y^0_M \mathbf{1}\{y^1_F < y^0_F\}$ (i.e. they make a decision based on y^0_F but the true value of staying at home is y^0_M, since we interpret the difference $y^0_F - y^0_M$ as the gender norm tax).

Whether it is men or women who lose depends strongly on how much flexibility the firm had to react to differences in gender norms in the first place, and it is best understood by looking at the parameters under the counterfactual policies in Figure 13. For women, the increase in LFP and in variable pay is big enough to offset the decrease in fixed pay, so that they are net winners. Men's LFP decreases and fixed pay falls enough that their overall welfare change is negative. Because we impose the constraint that the wage bill is kept constant, welfare changes are essentially a redistribution between female and male workers. It is important to note that this result is driven by the assumption that the firm cannot adjust employment levels. This is convenient for quantifying the pure effect of misallocation but it does not capture all the possible benefits of expanding women's access to the labor force. In a more general model where the firm can adjust employment in response to the change in gender norms the effect is unlikely to be zero sum.

Figure A.8: Examples: India and Sweden before and after equating norms

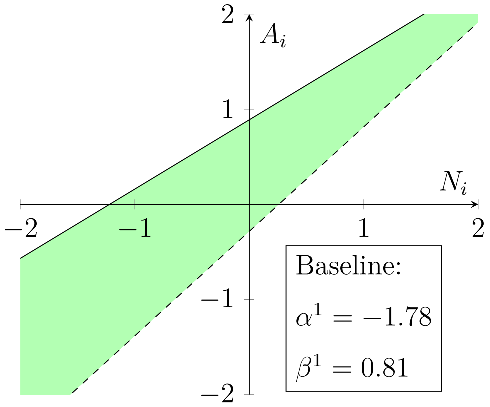


(a) India, constrained opt. policy

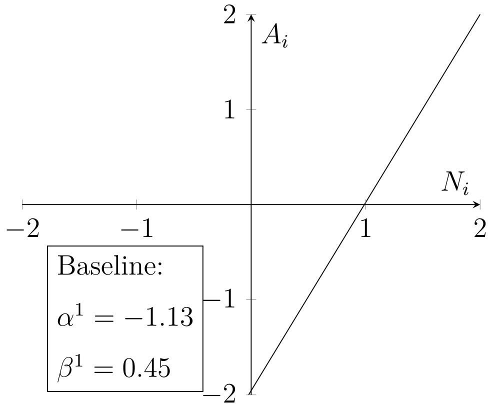
(b) Sweden, constrained opt. policy

Notes. The figures show the participation frontier, $\beta^1 A_i - \nu^0 N_i \geq \alpha^0 - \alpha^1$, for India and Sweden in the long-run horizon (averaging across cohort and tenure groups, weighted by size). Blue represents females, orange represents males. The solid line is the participation frontier with baseline α^0, ν^0; the dashed line is the counterfactual participation frontier after equating gender norms (and letting the wage policy respond). Those with (A_i, N_i) above and to the left of the participation frontier are in the labor force.

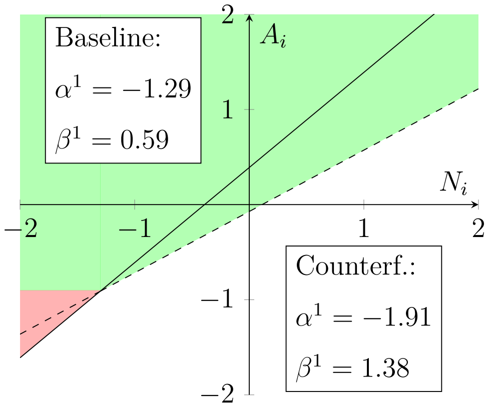
Figure A.9: Example: India, winners and losers under different wage policies



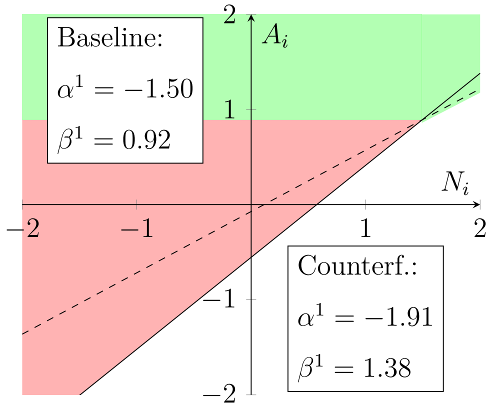
(a) Short run, females



(b) Short run, males



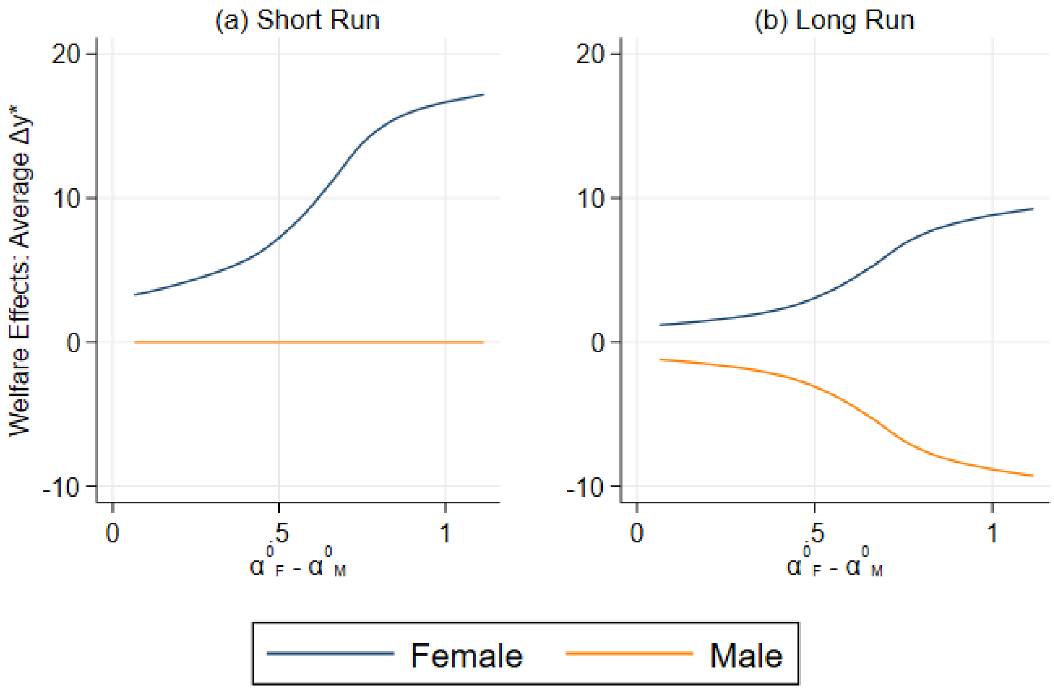
(c) Long run, females



(d) Long run, males

Notes. The figures show the participation frontier, $\beta^1 A_i - \nu^0 N_i \geq \alpha^0 - \alpha^1$, as well as the welfare effects for India under the two wage policy horizons we consider. The solid line is the participation frontier with baseline α^0, ν^0; the dashed line is the counterfactual participation frontier after equating gender norms (and letting the wage policy respond under the optimal policies). The area shaded green represents the population with (A_i, N_i) that benefits from equating gender norm, the area shaded red is the population that is harmed.

Figure A.10: Welfare effects of eliminating the gender norms tax



Notes. The figure shows the change in value of chosen option y^* (which can be interpreted as a percentage change, because it is in units of log-wages) by the difference $\alpha^0_F - \alpha^0_M$. For men, $y^* = \max\{y^0, y^1\}$. For women, $y^* = y^1_F \mathbf{1}\{y^1_F \geq y^0_F\} + y^0_M \mathbf{1}\{y^1_F < y^0_F\}$ (i.e. they make a decision based on y^0_F but the true value of staying at home is y^0_M, since we interpret the difference $y^0_F - y^0_M$ as the gender norm tax).