Gender Gaps across the Spectrum of Development: Local Talent and Firm Productivity*

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

April 2025

Abstract

We ask whether the gendered division of work affects firm productivity across the spectrum of economic development. Personnel records of over 100,000 individuals hired by a global firm that operates in 100 countries reveal that the performance of female employees is higher where women are underrepresented in the candidate pool. This implies productivity gains from hiring more women, but realizing them would require increasing women's pay relative to men. The findings highlight how unequal gender norms in local labor markets create an equity-efficiency trade-off inside the firm, particularly in low-income countries with conservative gender norms.

JEL Codes: O12, O15, M5

*We thank our partners at the MNE for sharing the data, Sam Bowles, Stefano Caria, Anton Heil, Jonas Hjort, Erik Hurst, Rocco Macchiavello, Andrea Prat, Imran Rasul, and seminar participants at CIVOE, Harvard, HBS, IIES, IZA, LSE, MIT Sloan, NBER Summer Institute, Northwestern, Princeton, Toulouse, UCL, UMichigan for useful feedback. We thank Rishabh Malhotra, Donata Schilling, Misha Seong, Wenzhi Wang for excellent 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

§University of Chicago, Booth School of Business, CEPR and IZA. virginia.minni@chicagobooth.edu

¶Department of Economics, MIT. vquintas@mit.edu

1 Introduction

A firm seeking to hire employees in India or Turkey will choose from a pool of applicants where men outnumber women by a factor of four. A firm seeking to hire for the same job in Canada or Ethiopia will choose from a pool of applicants with a similar number of men and women. How do restrictions in the pool of talent by gender affect the productivity of firms?

The answer relies on measuring how men and women select differently into local labor markets and how this difference affects the talent pool from which firms hire. To the extent that gender norms create greater barriers to entry into the labor force for women, only those motivated and able to overcome those barriers will select into the labor force (Olivetti and Petrongolo, 2016; Mulligan and Rubinstein, 2008).[1] Testing for potential positive selection of women is important because it would imply that firms could increase productivity by changing the composition of the workforce.

Measuring potential differential selection by gender is difficult because applicants may look the same on observables at the stage of hiring (Sarsons, 2017). Once inside the firm, the same gender norms that created barriers to labor force participation create barriers to pay and career progression.

In this paper, we propose a method to measure potential positive selection by studying productivity by gender in a multinational firm where women face the same barriers inside the firm, but differential barriers in their local labor market by country and cohort. We combine individual-level data from the personnel records of a multinational company with macro-level, local labor market data on the labor force participation across 4 cohorts and the 101 countries where the multinational operates. The personnel records contain the individual earnings and career paths of approximately 100,000 employees.

We match the firm's administrative data with aggregate data at the gender-age cohort-country level from the World Bank on the ratio of women to men in the labor force, the labor force participation ratio (henceforth LFPR). We use the labor participation data to measure selection into the labor force across different countries, age cohorts, and genders. Importantly, the LFPR varies among countries with similar levels of income. Thus, differential labor force participation across genders does not solely capture the level of economic development.[2]

[1] This has a parallel in the ordeals literature on targeting, where by design ordeals (barriers) change the composition of applicants through self-targeting (Alatas, Purnamasari, Wai-Poi, Banerjee, Olken and Hanna, 2016). While ordeals can improve efficiency, in this case, barriers can distort the allocation of talent as they act differentially on a dimension - gender - that is orthogonal to productivity.

[2] For instance, Algeria and Ecuador had a similar level of GDP per capita in the 1980s, but the former's LFPR was less than 1/4 of the latter's LFPR.

Having data from the same firm operating in many countries allows us to disentangle firm policy from local conditions since the multinational adheres to global policies determined at its headquarters. On the other hand, local firms are embedded in local contexts, which can confound demand and supply factors. An additional advantage of our data is that we observe women and men with the same education, same tenure, and working in the same function in the same firm. Hence, the gender earnings gap is not influenced by differences in occupational choices that make comparisons between genders difficult.

We start by documenting stylized facts in the firm workforce. First, we show that the gender ratio in the firm mirrors the gender ratio in labor force participation across labor markets. In other words, the firm hires fewer women in countries and cohorts with low LFPR. This is in line with the firm using the same selection process in all countries; that is, they do not employ more pro-women policies in countries where norms keep women inside the home (or vice versa). The evidence is supported by the accounts of the firm's HR managers and is consistent with similarly centralized policies in the sample of 1,213 multinationals analyzed by Hjort, Li and Sarsons (2020). It also indicates that local economic conditions shape the firm workforce through the variation in labor supply by gender in each country.

Second, we analyze how differences in LFPR correlate with gender differences in pay and performance within the firm. Like Olivetti and Petrongolo (2008), we find that the gap between female and male earnings (henceforth, the gender pay gap) monotonically decreases as the LFPR increases. That is, when the LFPR is at its lowest, the gender pay gap is "inverted", and 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. Since we observe employees of different ages in the firm, we can exploit both cross-country and cross-cohort variation in the LFPR. The sign and size of the effect are similar whether we use within-country or within-cohort variation for identification, ruling out country-level confounders. An explanation of this counterintuitive negative correlation between women's pay in the firm (relative to men) and the LFPR is the presence of positive selection into the firm by women. Under the logic of selection, when a group faces a high cost to join the workforce, only the individuals whose returns are high enough will do so. This implies that the marginal, and hence the average, productivity of female employees is decreasing in the share of women working in the firm, as is standard in models of selection on gains (Lazear, 2021).

Our third fact corroborates the intuition of differential selection by gender using

other performance metrics. In low LFPR countries and cohorts, women are over-represented in the highest rungs of the hierarchy and are more likely to be promoted relative to their counterparts in high LFPR countries and cohorts. Relatedly, women are over-represented in the top decile of the wage distribution and under-represented in the bottom decile when LFPR is low. Moreover, there is a negative correlation between women's average performance and the LFPR, and it originates from the bottom percentiles of the wage distribution. The wages of the women at the bottom decile of the wage distribution decrease as the LFPR increases, while the wages of the women at the top decile remain constant. Under the rationale of positive selection, as the LFPR increases, the ability of the marginal female employees falls, in line with the existence of an ability threshold above which women work inside the firm.

Overall, the reduced form evidence is consistent with the interpretation that the women whom we observe in the firm in low LFPR countries had the ability to overcome higher barriers and are thus positively selected. This raises the question of why the firm does not use the differences in labor force participation rates by gender to its advantage. In particular, the firm could increase average productivity by "undoing" gender differences in local contexts and hiring more women. In the second part of the paper, we structurally estimate the parameters of the firm's compensation policies using a simple 2-sector Roy model of occupational choice. The model allows us to understand to what extent the firm's existing wage policy is optimal in terms of maximizing the firm's productivity and to simulate the effects of alternative firm wage policies and local economic conditions on firm productivity and pay inequality.

We first estimate the individual-level ability and the parameters of the firm's pay policy. We assume that log-pay consists of a component that is common to all people of the same gender in the same country-cohort group (for instance, country-specific regulations) and a component that is specific to the individual and proportional to his or her productivity. The structural estimates leverage the significant advantage of our data: we observe several employees in each country, gender, and cohort, allowing us to estimate both a fixed parameter common to all employees in the same cell (e.g., discrimination based on gender) and, using the variation in wages within cell, differences in individual productivity. Moreover, as the pay data comes from the firm's administrative records, we do not rely on noisily reported wages. In our case, there is no measurement error in the salary information.

Although we assume that the underlying distribution of productivity is the same for every individual, the productivity of the average worker may still differ across genders, countries, and age cohorts because of differential selection into the labor force. We use the variation in labor force participation (LFP) at the gender-cohort-country

level to measure gender differences in the selection into the labor force. We find that, on average, across countries and cohorts, female employees have 0.10 higher ability (in units of standard deviations of the underlying productivity variable) than their male counterparts. The difference is larger in countries with weaker gender equity labor laws and with more conservative gender norms; for instance, the ability of women in the firm is 0.18 higher than that of men in the firm in countries with below-median LFPR, whereas this gap decreases to 0.07 in countries with above-median LFPR. The estimated calibrated ability correlates with other measures of worker performance that are not used in the estimation. With these estimates in hand, we analyze the effect of a number of counterfactual scenarios on firm productivity and inequality.

The first counterfactual compares the observed wage policy to the optimal contract that maximizes average productivity while keeping employment and the wage bill as binding constraints. We show that, given the productivity differences between men and women, the firm could increase productivity if they were to change the terms of the wage contract to attract more women. 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 50% on average. However, we note that such a contract would significantly increase inequality within and between genders; most notably, the difference in pay between women and men would go up by 73 log points. This captures both differences in performance for the same job and differences in jobs as more able women climb the corporate ladder faster. Thus, whilst it is theoretically possible for the firm to benefit from policies that compensate for unequal labor market conditions, these policies would create a high level of inequality among employees of different genders.

In order to hire more women without excessively increasing inequality, the firm could increase women's pay without decreasing men's pay or have the same wage policy across genders. We estimate the effects of this in a second counterfactual. We let the firm optimize the wage policy with an additional constraint that imposes the wage policy parameters to be the same across genders. The increase in productivity is about half the size compared to the unconstrained wage policy. Moreover, there is still a large gender pay gap of 24 log points because women's ability is higher on average.

These counterfactuals underscore the significance of a trade-off between efficiency and equality among workers *within the same firm*, which is inverted when compared to the efficiency-equality trade-off observed outside the firm. In the labor market, efficiency can be achieved by restoring equity and closing the gap in labor force participation between genders (Hsieh, Hurst, Jones and Klenow, 2019).[3] Instead, within a

[3]Hsieh et al. (2019) finds that reducing misallocation by lowering barriers across gender and race

single organization that faces unequal participation in the labor market by gender, efficiency can only be achieved by tolerating high within-firm inequality across genders. When we mute differences in gender LFP by equating the outside option of both genders, the trade-off between productivity and inequality disappears, and the firm can increase productivity by 11% and eliminate pay inequality by increasing the female share among its employees.

The final counterfactual simulates the effect of more stringent labor regulations that make it harder to link pay to performance. We find that firm productivity decreases by 17%, and there is a reduction in the difference in pay between women and men. The results cast new light on pro-worker labor policies and the gender earnings gap. We show that more stringent labor regulations may hurt the minority group, especially when the barriers to labor market entry are higher. Intuitively, most pro-worker measures, such as restrictions on hiring and firing, make it harder to link pay to performance (Propper and van Reenen, 2010) and this leads to a larger intake of lower-productivity workers who, by selection, are more likely to be the majority group.

The main contribution of our paper is to show that gender disparities in local labor markets affect firm productivity by creating an equity-efficiency tradeoff within the firm. This phenomenon is likely to continue to grow in importance as firms continue to increase in size and expand into new regions (Hsieh and Rossi-Hansberg, 2023; Hazell, Patterson, Sarsons and Taska, 2022).

First, our analysis contributes to the literature on how firms set wages and organize their economic activity across space and how this is shaped by local economic conditions (Grossman and Helpman, 2008; Blinder and Krueger, 2013; Aghion, Bloom, Lucking, Sadun and Van Reenen, 2021). We show the consequences on firm productivity and inequality of across-country differences in the worker gender composition of local labor markets. In doing so, we elucidate how the tradeoff between equity and efficiency inside the same organization can be opposite to the usual economy-wide equity-efficiency tradeoff across organizations. In the macroeconomy, bringing convergence in labor force participation by gender or race contributes to higher growth via an improved allocation of talent (Hsieh et al., 2019). Conversely, a single firm facing unequal labor pools by demographic groups can potentially benefit from the resulting misallocation of talent by leveraging the positive selection of minority groups to its advantage. And yet, fully capitalizing on this would require tolerating significant within-firm wage disparities among different groups, which may prove unfeasible. A growing strand of research in economics - as well as in psychology, sociology, and or-

groups accounts for 41.5% of the increase in GDP per capita in the United States between 1960 and 2010.

ganizational behavior - has documented that individuals care about their pay relative to that of their co-workers (Charness and Kuhn, 2011; Fehr, Goette and Zehnder, 2009; Lemieux, MacLeod and Parent, 2012; Breza, Kaur and Shamdasani, 2018).

Second, our paper connects the literature on the barriers to female labor force participation (Goldin, 1995; Jayachandran, 2015; Olivetti and Petrongolo, 2016; Bursztyn, Cappelen, Tungodden, Voena and Yanagizawa-Drott, 2023) to the literature on the impact of diversity for firm productivity (Alesina and La Ferrara, 2005; Hamilton, Nickerson and Owan, 2012; Hjort, 2014; Bertrand and Duflo, 2017; Marx, Pons and Suri, 2021). Seen through the lens of selection, the link between diversity and productivity is underpinned by the traits of the minority due to the barriers they had to overcome rather than a direct "treatment" effect of diversity on productivity through, for instance, role model effects or changes in culture. In our data, we observe women and men with the same education, same tenure, and working in the same function in the same firm. This means that the gender gaps we observe are not affected by different occupational choices, which often complicate gender comparisons (e.g., Blau, 1977; Goldin, 2014; Card, Cardoso and Kline, 2016; Wiswall and Zafar, 2018; Andrew, Bandiera, Costa-Dias and Landais, 2021).

Across countries and time, the division of labor inside and outside the home has remained within the confines of norms that assign the largest share of housework to women. Because of this, the under-representation of women in spheres of influence and employment has led to significant efforts in both the private and public sectors to address the gap through extensive diversity initiatives (Bertrand, 2011; Olivetti and Petrongolo, 2016; Bertrand, 2020). Supporters of these initiatives argue that diversity per se could be beneficial for productivity and profits due, for example, to the nature of the production function or role model effects (Lazear, 1999; Athey, Avery and Zemsky, 2000; Hong and Page, 2001). We contribute to the debate by showing that to understand the consequences of gender parity on firms, it requires to first understand the selection into the workforce of women and men.

The paper is organized as follows. Section 2 presents the institutional context of the MNE and describes the data sources. Section 3 provides reduced form evidence on the link between the LFPR and differential proportion, pay, and performance in the firm between genders. Section 4 introduces the model that we then estimate using the firm's personnel data and country-cohort level LFPR, and Section 5 uses our estimates to evaluate the effects of different counterfactuals. Section 6 concludes by discussing policy implications and other issues for further research.

2 Context and data

2.1 Context

We collaborate with an MNE with headquarters in Europe and offices in more than 100 countries worldwide, as illustrated in Figure A.1. The MNE produces consumer goods; in 2019, it 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%).

To preserve the confidentiality of the firm, we generally refrain from reporting country labels for the entire sample of countries in the figures.

2.2 Data

Personnel records: 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). For each employee, we observe: (1) work level and job title; (2) tenure in the firm and job; (3) annual performance score decided by the manager, and (4) total compensation (fixed plus variable pay in euros). The information on pay comes directly from the firm payroll records and hence does not have measurement error, unlike self-reported wages from labor surveys. Pay differences capture differences in performance between employees, encompassing off-peak salary increases as well as promotions. We look at four 10-year age cohorts within the company: 18-29, 30-39, 40-49, and 50-59[4].

A recent strand of evidence shows that multinationals pay higher wages in the countries where they operate (Verhoogen, 2008; Javorcik, 2015; Hale and Xu, 2016; Alfaro-Urena, Manelici and Vasquez, 2022; Hjort et al., 2020). We confirm that this is the case for our firm in Appendix Figure A.2 which shows that the firm's average wages are usually well above the countries' average wages, using both the average

[4]The age band classification is decided by the firm. 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.

wages in the manufacturing sector from the ORBIS database and the ILO estimates for white-collar employees. Table 1 presents summary statistics separately by gender at the gender-cohort-country level, which is the relevant unit of analysis used in the structural estimation.[5] The first two columns are for the full sample of workers, and columns 3 and 4 are for the sample used in the structural estimation, where we only consider the cells with at least 30 workers of each gender. The female cells show lower average pay, age, tenure, probability of being in managerial positions, and probability of experiencing fast promotions. Overall in the company, 40% of workers are aged between 30-39 and the majority of workers are in WL1 (>70%).

Country-cohort level data: We combine the firm's administrative records with country-cohort data on labor force participation rates (LFP) of men and women from the World Bank. The World Bank labor force participation data are from the ILO modeled estimates database where the latest edition covers the years from 1990 onwards.[6] We match the worker-level data from the MNE with the country-level LFP data by pairing the workers' country and age cohort. Hence, the variation in LFP rates is at the gender-country-age cohort level. In particular, we match the age-country cohorts in the firm with the average LFP rate in the country in the decade of labor market entry, separately by gender (i.e. upon finishing their college education when they are in their 20s years of age).[7]. In this way, we can leverage variation in the LFPR both at the country level and at the cohort level, as LFPR varies in both dimensions. For example, employees aged 18-29 are associated with the LFP rates of the 2010-2020 decade, while employees of ages 30-39 are associated with the LFP rates of the 2000-2010 decade, and so on.[8]

In our main analysis, we use the LFP data for all individuals because the LFP data for individuals with advanced education is missing for almost 20% of the sample, particularly from countries with low female labor force participation (FLFP). However, we conduct a robustness exercise when using the LFP data only for individuals with advanced education.

We also use other aggregate data for some additional analysis: GDP per capita at constant 2010 USD from the World Bank; the World Value Survey (waves 1-6); the Women, Business and the Law Index from the World Bank; the Restrictive Labor Regulations Index from the World Economic Forum; the gender development index (GDI),

[5] As we explain in section 4, we let the structural parameters vary by gender, cohort, and country in the firm to allow for the wage policy to differ across countries.

[6] Accessed here on March 22, 2020.

[7] Since our sample consists of college-educated workers, the decade of labor market entry is assumed to be their 20s.

[8] Since data on the LFP for the 1980-1989 decade is not available, we use a linear interpolation of the LFP for that decade.

which is the ratio of female/male Human Development Index from the United Nations Development Programme; and the female to male ratio in years of schooling and the percentage in post-secondary education from the World Bank.

Figure 1 plots moments of the distribution of the labor force participation ratio of women's LFP against men's LFP (the LFPR)[9] at different deciles of GDP per capita. It shows that while the mean exhibits the well-known U-shape pattern (Goldin, 1995), the distributions largely overlap: there is variation in LFPR at every level of GDP per capita. For instance, the interquartile range of LFPR is broadly similar across the deciles of GDP per capita. Moreover, the 90th percentile of LFPR only ranges between 1 and 0.9. This indicates that there are country-cohort cells with high LFPR at every level of economic development, instead of being solely concentrated in high-income countries. This is essential for the analysis that follows because it allows us to partial out the differences in national income among the countries where the MNE operates.

3 Facts

In this section, we document stylized facts of the employment and wages of women and men at the firm and how these vary across different labor markets. In particular, we look at how the gender ratio of the firm workforce and its gender pay and promotion gaps correlate with the LFPR in the country. The main findings are that the firm gender ratio closely follows the LFPR in the country, particularly so in low LFPR countries, and that, when the LFPR is at its lowest, the gender pay gap is "inverted", with women being paid more than men with the same experience, tenure and working in the same function.

We interpret the findings through the lens of a Roy model where men and women have the same productivity but women face barriers to working outside the home. The model predicts when barriers are high, only the women who are most productive in the workplace will be observed there. As barriers fall, so does the average productivity of women who enter the labor force and are observed in the workplace. In line with the prediction of the Roy model, women are overrepresented in the highest rungs of the hierarchy and more likely to be promoted relative to their counterparts in high LFPR countries. Relatedly, the difference between men and women is driven by individuals in the lowest deciles of pay or promotion: in low LFPR countries, the firm hires many men from the left tail of the ability distribution.

[9] We rescale women's LFP by men's LFP in order to account for country-cohort level differences in the probability of individuals to work. However, we note that the variation in men's LFP is tiny compared to the variation in women's LFP. In other words, the variation in LFPR is mostly driven by the variation in the female labor force participation, the numerator in the ratio.

3.1 Hiring gap

First, we investigate whether the proportion of men and women in the firm correlates with the level of LFPR in the different labor markets. Figure 2 panel (a) documents a large variation in female shares of employment within the MNE across countries, which matches the variation in LFPR across countries: the female to male ratio in the firm closely follows the same ratio in the labor force, particularly at lower levels of LFPR (<0.6). When the LFPR is above 0.6, the firm hires more women relative to men compared to the LFPR, although a formal test of differential slopes by above/below median LFPR does not indicate that the relationship between the firm and country ratios significantly varies with the level of LFPR. We also note that wages at hiring are the same between men and women (the gender pay gap in starting salary is equal to 0.0005 with a p-value of 0.926), as also confirmed by HR managers at the firm. Overall, this evidence is in line with the firm using the same selection process in all countries, rather than, for example, employing more pro-women policies in countries with low female labor force participation (or vice versa). In other words, the firm adopts the same headquarters personnel policies worldwide rather than adapting to local economic conditions - as broadly found in the literature on multinationals' wage and price setting procedures (DellaVigna and Gentzkow, 2019; Hjort et al., 2020). As a result, countries' LFPRs "bind"; that is, barriers at the country level constrain the firm's talent pool.

3.2 Pay gap

We turn to look at how the gender pay gap changes with the LFPR. We define the gender pay gap as the difference between women's wages and men's wages (in logs). Pay in the firm captures differences in performance between employees, encompassing off-peak salary increases as well as promotions. Figure 2 panel (b) estimates a kernel-weighted local polynomial of the pay gap on LFPR. 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 (Commission, 2019). Overall, Figure 2 panel (b) documents a negative correlation between the gender pay gap and the LFPR.

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 looks at the gender pay gap using different sources of variation, thus also allowing us to check whether the sign of the correlation is consistent regardless of the source of

variation that we use to identify it. We estimate the following model:

$$w_{iact} = \alpha LFPR_{ac} + \beta Female_i + \gamma LFPR_{ac} * Female_i + \mathbf{X}'_{iact}\boldsymbol{\Lambda} + \boldsymbol{\psi}_t + \epsilon_{iact} \quad (1)$$

where w is log wage of employee i in country c and year t for age group a. $\boldsymbol{\psi}_t$ represents 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. We include different controls in $\mathbf{X}_{iact}$: column 2 controls for a quadratic function of tenure and function fixed effects, column 3 adds GDP per capita in logs so to show that the variation in the LFPR is not only a function of country income, column 4 adds cohort fixed effects so to only exploit the variation across countries; column 5 replaces the cohort fixed effects with country fixed effects hence only exploiting the variation within countries. The comparison between columns 4 and 5 is particularly informative as it uses one source of variation at a time.

The estimates of γ are negative and precisely estimated in all specifications. In Appendix Table A.1, we use the LFP data for individuals with advanced education only.[10] The results are nearly unchanged when we adopt this measure. However, we lose almost 20% of the sample, particularly from countries with low FLFP. Hence, we employ the overall LFP as our default measure.[11] We conduct additional robustness checks in the appendix and the γ estimate is stable throughout.[12]

These gender differences are not present at the hiring stage - the gender pay gap in starting salary is equal to 0.0005 (p-value=0.926) - backing up the interpretation of the adoption of a global personnel policy set at headquarters. However, we next look at the pay progression for new hires. We do this in columns 6 and 7 of Table 2 by estimating the same specification as in column 1 for new hires only. Women in low LFPR country-cohorts display faster pay growth and a higher probability of promotion, and this positive gender gap in realized performance decreases as the LFPR

[10] As defined by the World Bank, an individual with advanced education has completed a short-cycle tertiary education or a college degree and/or above.

[11] The correlation between overall LFPR and LFPR for individuals with advanced education is 68%.

[12] 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 5 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 fixed pay rather than in variable bonus, which constitutes a much lower proportion of overall salary (the median ratio of bonus to fixed pay is 13%) — see Appendix Table A.4. In the company, pay summarizes altogether most differences in performance between employees, encompassing off-peak salary increases as well as promotions.

increases. This evidence on the sample of new hires supports the interpretation that the firm is not taking into account the implications on productivity that different LFP rates by gender might entail when making decisions on hiring and wage offers. It is only after some time at the firm that women and men *ex-post* pay gaps start to diverge the lower the LFPR is, indicating positive selection of women into the firm which is not accounted for at the hiring stage.

This perhaps counterintuitive inverted pay gap can be understood through the logic of positive selection. In countries with low LFPRs, there are higher barriers to entry into the firms for women, and only the most talented women overcome them. As a result, the lower the LFPR, the more positively selected the women hired within the MNE are compared to men. Thus, the average productivity of women in the workplace will be negatively correlated with their share in the labor force, as is standard in models of selection on gains (Roy, 1951). In other words, high barriers mean that only the most productive women are likely to be found in the workplace. As these barriers diminish, the average productivity of women entering the workforce also decreases. Building on this reasoning, we would expect women to be disproportionately represented at the upper echelons of the organizational hierarchy and to receive promotions more frequently than their male counterparts in countries with a low LFPR. We examine this hypothesis further in the following subsection.

3.3 Promotion gap

Figure 3 displays plots of women's performance measures against the LFPR. Panels (a) and (b) of Figure 3 show that women are over-represented at the top decile of the wage distribution and under-represented at the bottom decile when overall LFPR is low and converge when LFPR is close to 1. Moreover, Panels (c) and (d) show that when LFPR is low, women are over-represented in managerial positions and among those who get promoted quickly, but the two converge as participation rates get more equal. In Appendix Figure A.3, we show the plots with the men's performance measures, which display the opposite patterns: men's performance is positively correlated with the LFPR.

A Roy selection framework predicts that as the LFPR increases, women with lower ability enter the labor force while high-ability women already in the labor force are unaffected. We test this result empirically in Figure 4 by looking at how women's wages at different points of the distribution change as the LFPR increases. Since overall wage levels change across countries, we control for the respective wage measure for men. The first panel from the left shows that women's average wages decrease

as the LFPR increases. The remaining two panels indicate whether this decrease is driven by the bottom or the top of women's wage distribution. It comes from the bottom of the wage distribution: the 10th percentile of women's wages decreases with LFPR (second panel), while there is no impact on the 90th percentile (third panel).[13] These differential patterns at different levels of the wage distribution are consistent with women facing a higher bar for entering the firm in countries with lower LFPR. As the LFPR increases, lower-ability women start to enter, hence decreasing the wages at the bottom of the distribution, while high-ability women are not affected, hence leaving the wages at the top of the distribution unchanged.

The fact that the negative correlation is driven by the left tail of the wage distribution rules out alternative explanations that would also generate a negative correlation between average female performance and LFPR. For instance, this pattern rules out that the high productivity of women in low LFPR countries, and hence their high wages, is because women bring different valuable inputs to the firm, and therefore, the marginal value of these inputs is high when the share of women is low. If this were the case, we would find that the productivity of the top percentiles of women would decrease as LFPR increased, which is the opposite of what we find in Figure 4. It also rules out the argument that women's performance increases relative to men's because men in low LFPR countries have better outside options.[14]

Taken together, the evidence is consistent with positive selection into the labor force. That is, the average productivity of women who work outside the home is higher the higher the share of women who work inside the home. This implies that the firm could increase productivity by increasing their employees' gender ratio above that observed in the labor force. Moreover, the potential productivity gains would be higher where the female share in the labor force is lower. Seen through the lens of selection, the benefits of gender equality inside the firm are higher when gender equality outside the firm is lower. Yet, panel (a) of Figure 2 shows that the firm does not hire more women relative to their population share. This suggests that the cost of hiring more women is also higher when the gender ratio in the labor force is lower and hence that the net benefit is negative. These costs cannot be identified in the data because we only observe the state of the world in which they are not incurred. In what follows we combine theory and individual-level data to estimate the parameters of the firm's pay policy. We will then use these to create the counterfactuals we need to evaluate the

[13] Results are robust to using other percentiles, for example, the 25th percentile.

[14] Besides being inconsistent with the distributional changes, the assumption that men's outside option is decreasing in LFPR is inconsistent with the fact that multinationals pay higher wages and offer more amenities across the board (e.g., Verhoogen, 2008; Javorcik, 2015; Hale and Xu, 2016; Alfaro-Urena et al., 2022).

effect of alternative gender policies.

4 Model and estimation

Our goal in this section is to use the logic of selection to retrieve the parameters of the firm's pay policy and the value of outside options that generate the documented gender differences in pay and performance and use them to calibrate the firm's productivity under different counterfactual scenarios.

4.1 Framework

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 pay and that log-pay y_i^1 is a linear function of individual i's productivity, A_i:[15]

$$y_i^1 = \alpha^1 + \beta^1 A_i, \tag{2}$$

The term α^1 is the unconditional average wage and β^1 is the return to productivity. Under the assumption that deviations from the average wage arise from individual differences in productivity, the fact that we observe several employees in each country-cohort-gender cell implies that we can use this equation together with a distributional assumption on A_i to identify both the level of fixed pay and of performance pay for each cell. This is an advantage compared to aggregate data on average wages by group identity and it is important for our purpose because it implies that we can separate the component of pay that is specific to an individual woman, due to her A_i, from the component of pay that is specific to all women in the same country, due to discrimination for example.

Symmetrically, we model worker i's log-value of housework y_i^0 as:

$$y_i^0 = \alpha^0 + \nu^0 N_i, \tag{3}$$

where N_i captures sources of individual heterogeneity in the value of staying out of the labor force. The common parameter α^0 captures the unconditional average value of staying out of the labor force (e.g. social norms that affect all women equally).

[15] 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 $z_i \sim \mathcal{N}(\mu, \sigma^2)$.

We make the following distributional assumption of joint normality:

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

Combining the payoffs with the distributional assumptions allows us to study the choice between being inside or outside the workforce. Here, ρ is the correlation between an individual's wage in the workforce and payoff outside of the labor force.

In particular, individual i chooses to work outside the home if and only if:

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

Because $\eta_i \sim \mathcal{N}(0,1)$, that happens with probability $1 - \Phi(\xi)$. It is straightforward to show that a higher α_0 due to, for instance, stronger norms about gender roles increases ξ and therefore reduces the probability of working outside the home $1 - \Phi(\xi)$. This means that we can use the observed labor force participation in each gender-country-cohort cell to back out the value of staying at home for the average person in that cell. If women are less likely to work outside the home, this will result in a higher estimated alpha zero for women.

Moreover, we have that:

$$\begin{aligned} \mathbb{E}[A_i \mid \eta_i \geq \xi] =& \frac{\operatorname{cov}(A_i, \eta_i)}{\operatorname{var}(\eta_i)} \mathbb{E}[\eta_i \mid \eta_i \geq \xi] \\ =& \frac{\beta^1 - \rho\nu^0}{\sqrt{(\beta^1)^2 + (\nu^0)^2 - 2\rho\beta^1\nu^0}} \lambda(\xi), \end{aligned}$$

where $\lambda(\cdot) \equiv \phi(\cdot)/[1 - \Phi(\cdot)]$ is the inverse Mills ratio. This implies that as long as $\beta^1 - \rho\nu^0 > 0$, the average ability of those who choose to work outside the home is increasing in ξ. Intuitively, as long as the returns from productivity are higher in the workplace than they are at home, people who join the labor force have higher productivity than those who stay at home.

4.2 Estimation

We estimate the model by letting the parameters vary by gender (g), cohort (a), and country (c) cells: $\alpha^1_{gac}, \beta^1_{gac}, \alpha^0_{gac}, \nu^0_{gac}, \rho_{gac}$. 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. Moreover, individual-level data allow us to separately identify gender differences in fixed pay from differences in pay due to differences in individual-specific performance.

We assume that all gac cells have the same underlying distribution of (A_i, N_i) at birth, normalized to be jointly standard normal (as in subsection 4.1),[16] but we allow the parameters in the wage and value of housework to be cell-specific. The goodness of fit plots of the wage distribution against the distribution implied by the model (Figure A.5) show that the distributional assumption of normality fits the data reasonably well, except when the data presents large departures from the normal distribution (e.g., bimodality).

We proxy $\Pr(y^1_{igac} \geq y^0_{igac})$ by country-cohort-level LFP data. We are thus 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 Appendix Figure A.2 (discussed in section 2) and in Figure 2 (discussed in section 3).

Theorem 6 in Heckman and Honoré (1990) characterizes what parameters can be identified in a two-sector log-normal Roy-Borjas model when only wages in one sector and the share of workers in that sector are observed. Under our distributional assumption, labor force participation, and the observed wage satisfy the following moment conditions:

$$LFP_{gac} = 1 - \Phi(\xi_{gac}) \tag{4}$$

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

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

$$\begin{aligned}\mathbb{E}[(y^1_{igac} - \mathbb{E}[y^1_{igac} \,|\, \text{employed}])^3 \,|\, \text{employed}] \\ = \theta^3_{gac}\left[2\lambda(\xi_{gac})^3 - 3\xi_{gac}\lambda(\xi_{gac})^2 + \xi^2_{gac}\lambda(\xi_{gac}) - \lambda(\xi_{gac})\right]\end{aligned} \tag{7}$$

where $\xi_{gac} \equiv (\alpha^0_{gac} - \alpha^1_{gac}) / \sqrt{(\beta^1_{gac})^2 + (\nu^0_{gac})^2 - 2\rho_{gac}\beta^1_{gac}\nu^0_{gac}}$ are the participation thresholds and $\theta_{gac} \equiv ((\beta^1_{gac})^2 - \rho_{gac}\beta^1_{gac}\nu^0_{gac}) / \sqrt{(\beta^1_{gac})^2 + (\nu^0_{gac})^2 - 2\rho_{gac}\beta^1_{gac}\nu^0_{gac}}$ is a measure of the sign and strength of selection.

[16] A potential concern could be that even if we take as given that innate talent is equally distributed at birth across genders, women and men face different opportunities to acquire human capital, which results in different distributions of actual productivity. However, we note that gender differences in educational attainment have been drastically reduced over the last decades and are much smaller compared to the gender gap in labor force participation (see Appendix Figure A.4). An active area of research is to understand why the convergence in educational qualifications has not translated into greater gender equality in labor market outcomes (Heath and Jayachandran, 2017; Jayachandran, 2021).

As discussed in Heckman and Honoré (1990), the LFP identifies the participation thresholds ξ_{gac}. The skewness of the observed wage distribution pins down the sign and strength of selection θ_{gac}. Our key parameters of interest are those of the firm's wage policy, α^1_{gac}, and β^1_{gac}, which are identified by the mean and variance of the observed wage distribution, respectively. Since we do not observe the value of staying at home y^0_{igac}, we cannot separately identify α^0_{gac}, ν^0_{gac} and ρ_{gac}.

We calibrate $(\alpha^1_{gac}, \beta^1_{gac}, \xi_{gac}, \theta_{gac})$ by matching the moments above to their empirical counterparts. 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. Since we observe the wage for those working in the firm, we can use the sample average and variance as the empirical analogs for $\mathbb{E}[y^1_{igac} \mid \text{employed}]$ and $\text{Var}(y^1_{igac} \mid \text{employed})$. To eliminate the effect of observables, as our measure of y^1_{igac}, we use the residuals of a regression of log(base pay + bonus) on year and function dummies, tenure, and tenure squared. Moreover, we exclude from the structural analysis all country-cohort cells with less than 30 male employees or less than 30 female employees, leaving 260 cells representing 84 countries. To calibrate the parameters in the participation decision, our empirical analog for LFP_{gac} is the World Bank LFP data in each gender, cohort, and country cell. Finally, beyond the moments that we targeted for calibration, Figure A.5 shows that the distributional assumption of normality fits the data reasonably well.

For some of our counterfactual exercises below, we will additionally need to calibrate α^0_{gac}, ν^0_{gac} and ρ_{gac}. These are not separately identified because we do not observe the value of staying at home. For the main results below, we adopt one of the normalizations suggested by Heckman and Honoré (1990), $\nu^0_{gac} \equiv 1$, which allows us to back out α^0_{gac} and ρ_{gac} from the composite parameters ξ_{gac}, θ_{gac}. In Appendix A.2 we show robustness to an alternative normalization.

4.3 Structural estimates

In this subsection, we report the structural estimates and correlate them with other data not used in the estimation to facilitate their economic interpretation.

Summary statistics of the calibrated parameters are presented in Panel A of Table 4. On average across countries and age cohorts, we estimate fixed pay α^1_{gac} to be higher for men and returns to productivity β^1_{gac} to be higher for women, albeit with a lot of variation across cells. The labor force participation thresholds, ξ_{gac}, are higher for females, as a direct consequence of having lower LFP than their male counterparts. Finally, we estimate the selection with respect to the productivity parameter, θ_{gac}, to

be positive on average, although it is also very heterogeneous across cells, including cells where it is negative.

Panel B of the same table summarizes the parameters of the value of staying out of the labor force, which are identified under an additional normalization $\nu^0_{gac} \equiv 1$. We exclude country-cohort cells for which no solutions or multiple solutions to the calibration moment equations exist, and cells for which we cannot compute the main set of counterfactuals discussed in section 5. We find that women have stronger average values of staying out of the labor force, α^0_{gac}, again consistent with having a lower LFP than their male counterparts. The correlation between the wage in the firm and the value of staying out of the labor force is similar on average for males and females, although with a lot of variation across cells.

4.3.1 Calibrated ability

Having obtained the parameters in the firm's wage policy, we can recover productivity as:

$$A_i = \frac{y^1_{igac} - \alpha^1_{gac}}{\beta^1_{gac}},$$

where y^1_{igac} is the residualized log-wage described before. Because we observe wages in the firm, we can impute productivity even though α^0_{gac}, ν^0_{gac} and ρ_{gac} are not separately identifiable (since we do not observe the value of staying at home y^0_{igac}). Under the normalization of $A_i \sim N(0,1)$, differences in productivity will be measured in units of standard deviations of the underlying productivity variable, which enters the log-wage equation.

Figure 5 plots the average calibrated productivity by the LFPR. Average productivity is approximately constant for male workers in most of the sample, whereas for female workers, average productivity is very high when FLFP is much smaller than MLFP (about 0.4 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 6 further shows that this is due to a shift in the entire productivity distribution of women. As the LFPR increases, there is a downward shift of the entire productivity distribution, except the right tail. The observation that changes occur in the left tail of the distribution aligns with the findings presented in section 3, which show that the ability of the marginal female worker decreases as the LFPR increases.

To validate our estimates we show that they correlate with other data not used to calibrate the model. We use two separate external variables to corroborate our calibrated productivity estimate using individual performance data from the firm's

records. The results for this exercise are in Figure 7. Panel (a) correlates calibrated productivity against the firm's performance score that a manager gives every year, and panel (b) against pay growth in the first year (for new hires). The correlation of both indicators with our calibrated productivity is positive and strong.

4.3.2 Interpretation

To aid the interpretation of the structural estimates, we show that the parameters of the home payoffs and the firm wage policy are correlated with country-level social norms and country-level labor laws, respectively.

Figure 8 panel (a) 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 that relate to gender norms: (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 negative correlation between disagreement with the statement and our gap in the average value of staying at home (the magnitude of the slope coefficient is at least -0.4 across the four questions, p-values<0.05). This corroborates the interpretation of the gender differences in home payoffs as the cost of gender norms for the average woman. Panel (b) of Figure 8 shows the gap in our calibrated $\alpha^0_{Ftac} - \alpha^0_{Mtac}$ against the Women, Business and the Law Index from the World Bank.[17] The figure shows that the gap $\alpha^0_{Ftac} - \alpha^0_{Mtac}$ is strongly negatively correlated with laws allowing or facilitating women's labor. Therefore, part of the restrictions to FLFP due to gender norms may be embedded in the laws of certain countries.

Figure 8 panel (c) plots our calibrated β^1_{gac} (which represents returns to productivity, and, in our model, is what generates dispersion in pay within gender-country-cohort cells) against the Restrictive Labor Regulations Index from the World Economic Forum.[18] Consistent with stricter labor regulation limiting performance pay, we find

[17] 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, parenthood, entrepreneurship, assets, and pensions — for current laws and regulations (i.e. religious and customary laws are not considered unless they are coded). Hyland, Djankov and Goldberg (2021) provides an overview of the data documenting how gender discrimination by law affects women's economic opportunities.

[18] The WEF Restrictive Labor Regulations Index is available for the period 2008–2020, and it is based on an annual survey of 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.

that our calibrated β^1_{gac} is lower in countries with a higher value of the index.

5 Counterfactuals

We use the model estimates to evaluate the effect of different counterfactuals on pay inequality and firm productivity. 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 maximization as the guiding criterion. Rather, we take the observed level of employment and the wage bill in each country-cohort group as binding constraints. In practice, these are determined by a maximization problem that we do not observe. This is equivalent to assuming that the firm sets the optimal scale of operation and then decides who to hire to maximize productivity.

Our first counterfactual asks whether, under these constraints, the firm maximizes productivity. Second, we quantify the impacts on firm pay inequality and productivity of equating women's and men's selection in the firm. Finally, we simulate the effects of stricter labor laws. The results for all counterfactuals are summarized in Table 5.

5.1 Optimal firm wage policy

To compute the optimal wage policy, we let the firm choose α^1_{gac}, β^1_{gac} to maximize the productivity of its employees ($LFP \cdot \mathbb{E}[A_i \mid \text{empl'd}]$), subject to two constraints: (i) total employment cannot be smaller than at baseline, and (ii) the total wage bill cannot be greater than at baseline. 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 adjust α^1_{gac} and β^1_{gac} in a way that increases productivity without changing employment, at the cost of paying much higher wages. The precise expressions for the objective function and the constraints are detailed in Appendix A.3.

To compute the optimal wage policy we need to know the structural parameters α^0_{gac}, ν^0_{gac} and ρ_{gac}, which are pinned down only with an additional normalization, as discussed in subsection 4.2. Our main specification here imposes $\nu^0_{gac} \equiv 1$. We discuss an alternative calibration strategy in Appendix A.2. We are able to calibrate the parameters for both men and women and to compute the first set of counterfactuals in 114 cells representing 67 countries.[19]

[19]For some cells, there is no solution $\rho_{gac} \in [-1, 1]$ to the system of equations. We also drop cells with

Figure 9, 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 country-cohort cells: to maximize productivity the firm should attract more women (exploiting the fact that female workers are more positively selected than male workers). Row 2b of Table 5 presents some summary measures of the results for this counterfactual. Compared to the baseline (row 1b), the gender ratio would increase from 0.71 to 0.92 on average across cells. The firm achieves that by decreasing fixed pay for men (α^1) while increasing it for women, and generally increasing variable pay (except in cases where the LFP and wage bill constraints become binding). The optimal policy effectively undoes differences in LFP and leads to higher productivity in every country as shown in Figure 9 panel (e). On average, the firm could increase productivity by 0.16 standard deviations of the underlying productivity variable, or around 50% of the baseline value.

Why is the firm not setting the optimal wage policy (α^1_{gac}, β^1_{gac})? A possible answer is that 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 and cohorts, it would be productivity-maximizing to pay women more to attract more of them, 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 country-cohort cells, the gender pay gap (Female – Male) would have to increase by 73 log-points (Figure 9, panel (d) and Table 5 rows 1b and 2b) from -0.10 to 0.63. Since individual productivity is not directly observable, such an increase might not be socially acceptable. To add to the challenge, we note that the pay inequality would be higher, especially in the countries where female labor force participation is the lowest, that is where barriers against women working are the highest.

To further investigate this efficiency-equality tradeoff, we compute a second counterfactual, where we let the firm optimize the wage policy subject to an equality constraint in the wage parameters: $\alpha^1_{Fac} = \alpha^1_{Mac}$ and $\beta^1_{Fac} = \beta^1_{Mac}$ for all a, c. That is, we require the firm to set the same wage policy for males and females within a country-cohort cell so that workers with equal ability are paid the same. The results for this exercise are in Figure 10 and row 6c of Table 5. We are able to compute this optimal policy with equality in 88 cells representing 55 countries.[20] The changes go in

multiple solutions. That leaves 137 cells. Of those, we can compute all the counterfactual wage policies presented in Table 5, Panel B in 114 cells.

[20]The main obstacle is that, once we impose equality of wage parameters, in some cells, it is not possible to satisfy the LFP and wage bill constraints, so the feasible region is empty.

the same direction as the previous counterfactual, but they are smaller in size. First, it is still the case that the firm should increase fixed pay for women and decrease it for men (Figure 10 panel a), while generally increasing variable pay (Figure 10 panel b), compared to the baseline wage policy. Second, the gender ratio (Female/Male) increases (Figure 10 panel c), but not as much as in Counterfactual I (by 0.07 on average, across cells, comparing rows 1c and 6c of Table 5). Third, the increase in productivity (Figure 10 panel e) is only about half the size, 0.07 standard deviations on average or 18% of the baseline value. Moreover, even though the wage policy is the same for men and women, there is still a sizeable increase in the pay gap (24 log points), because female workers have higher productivity on average. Hence, even under an equal wage policy constraint, gender pay inequality in the firm would be large due to gender differences in ability that result from the differential selection by gender into the labor force.

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 with all men working under them. Note that this would be a very different scenario than that of many policies that 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.

5.2 Equalizing gender barriers

Our next set of counterfactuals assesses the effect of women's differential selection into the labor force on the firm's productivity. We set the value of the staying-at-home parameters of women (α^0, ν^0, ρ) equal to those of men (within the same country-cohort cells). We discuss effects both in the short run, that is, keeping the pay policy of the firm fixed (baseline wage policy), and in the long run, when the firm can optimally adjust its policy to the new environment.

Figure 11 plots the short-run effects of eliminating gender norms on the LFPR, the pay gap, and 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 in most country-cohort cells, so the ratio increases as well, by an average of 0.21 (comparing rows 1b and 3b of Table 5). These changes in LFP reduce the pay gap (by 2 log-points on average across countries and cohorts) because the women who enter the labor force

are generally less able on average than those who were already working at baseline, so women's average wages decrease. The effects on productivity, which is given by the product $LFP \cdot \mathbb{E}[A_i \mid \text{empl'd}]$, are theoretically ambiguous. On the one hand, the firm is averaging over a larger pool of workers, since the LFP increased (i.e. the first factor is higher in the product $LFP \cdot \mathbb{E}[A_i \mid \text{empl'd}]$). On the other hand, conditional on being in the labor force, female workers are less positively selected (i.e. the second factor $\mathbb{E}[A_i \mid \text{empl'd}]$ goes down). On average across cells, we find that average productivity decreases by 0.03 standard deviations. 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 equalized, the productivity cost rises.

The trade-off between productivity and inequality when the outside option is equalized, however, disappears if the firm is allowed to adjust the wage policy optimally, without (Figure 12) or with (Figure 13) equality constraints. First, notice that when the value of the outside option is equalized, the equal wage policy constraint is not binding.[21] As such, the Female/Male employment ratio becomes 1 and the gender pay gap disappears (panels c and d of Figure 12 and Figure 13). Equating the outside option of women to that of men makes it less costly for the firm to attract female workers, which in turn allows to decrease fixed pay (panel a) and increase variable pay (panel b) to attract higher ability workers without violating the employment and wage bill constraints. The increase in productivity is larger when the firm is constrained to use equal wage policies to begin with, 0.09 standard deviations, comparing 7c vs. 6c in Table 5 (vs. 0.04 without this constraint, comparing 4c vs. 2c in Table 5).

5.3 Stringency of labor regulations

Another reason why the firm might not be setting the optimal wage policy may be labor regulations, which can limit variable pay and pay inequality even within gender by imposing a cap on how high β^1_{gac} can be. We investigate this further in this subsection. We ask what the effect of stricter labor laws, that limit performance pay, would be on 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: $\beta^1_{gac} \leq \max\{\beta^1_{Fa,FRA}, \beta^1_{Ma,FRA}\}$, where $\beta^1_{Fa,FRA}, \beta^1_{Ma,FRA}$ 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 $(\alpha^0_{gac}, \nu^0_{gac}, \rho_{gac})$ unchanged at baseline, but we allow for the LFP of each group to respond optimally to the change in incentives.

[21] For that reason, rows 4c and 7c of Table 5 are exactly the same.

The results are plotted in Figure 14 and summarized in row 5b of Table 5. Productivity decreases in every cell (on average, by 0.08 standard deviations compared to row 2b in the same table). Strict labor regulations hurt high-ability female workers the most since the firm cannot rely on variable pay incentives to attract them anymore. As a result, the pay gap decreases by 6 log points. This result underscores a common drawback of several labor market policies that aim to bring equality between genders by treating both equally, despite them being different due to differential selection. At the same time, to still satisfy the LFP constraint, the firm is forced to increase fixed pay (in most cells), more so for women than for men, so that the gender ratio increases by 0.07 on average.

6 Conclusion

This paper shows that the traditional division of labor, where men work in the market and women at home, affects firm productivity through the differential participation of women and men in the workforce. That is, when female labor force participation is low compared to men's, the few women who work outside the home are those with the highest productivity. We find that this inequity in the labor market could potentially positively impact firm productivity by bringing in highly talented women. However, the firm does not use it to its full advantage, as doing so would imply a large increase in within-firm inequality between genders; this is particularly so where gender disparities in society are the largest. Precisely where the productivity gains are highest are where the costs in terms of intra-organization inequality are largest. This tension between efficiency and equality limits the extent to which a single firm can gain by going against discriminatory norms (Becker, 1971).

Understanding differential selection by gender, or indeed by any under-represented group, is key to informing personnel policy as well as broader labor market policies aimed at addressing the gender pay gap. Our counterfactual estimates highlight the complementarity between public policies that attenuate gender barriers and firm policies to maximize productivity.

The inability of firms to compensate for labor market gender disparities influences how they design their policies at the hiring, pay, and promotion stages. Aiming for gender equity — in pay, promotions, and dismissals — can turn out to be inequitable because selection generates different distributions of productivity between genders. Perhaps counter-intuitively, gender equity policies might end up hurting women as they limit the firm's ability to reward performance.

This paper focuses on firm wage and promotion policies and does not consider

firm hiring strategies. Nonetheless, applying the principle of self-selection could address a frequent issue faced by firms: that potential productivity is not fully revealed at the point of hiring. In particular, by positive selection, under-represented groups should have higher productivity, other things equal. This implies that between two potential hires with the exact same observable qualities, the minority candidate has, on average, better unobservables. Awareness of the evidence of positive selection of under-represented groups could change how quality can be inferred, particularly if quality is not perfectly observable or objectively measured. If made aware of the "distance traveled" — that the very presence of a member of an under-represented group has information on the talent of that member — managers might change hiring decisions and screening criteria. The implications of increased awareness of the positive selection effect of minority candidates by leaders for the recruitment, promotion, and productivity of under-represented groups are left to future research.

7 References

Aghion, Philippe, Nicholas Bloom, Brian Lucking, Raffaella Sadun, and John Van Reenen, "Turbulence, firm decentralization, and growth in bad times," *American Economic Journal: Applied Economics*, 2021, *13* (1), 133–169.

Alatas, Vivi, Ririn Purnamasari, Matthew Wai-Poi, Abhijit Banerjee, Benjamin A Olken, and Rema Hanna, "Self-targeting: Evidence from a field experiment in Indonesia," *Journal of Political Economy*, 2016, *124* (2), 371–427.

Alesina, Alberto and Eliana La Ferrara, "Ethnic diversity and economic performance," *Journal of economic literature*, 2005, *43* (3), 762–800.

Alfaro-Urena, Alonso, Isabela Manelici, and Jose P Vasquez, "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-firm Linkages," *The Quarterly Journal of Economics*, 2022, *137* (3), 1495–1552.

Andrew, Alison, Oriana Bandiera, Monica Costa-Dias, and Camille Landais, "Women and men at work," *IFS Deaton Review of Inequalities*, 2021.

Athey, Susan, Christopher Avery, and Peter Zemsky, "Mentoring and diversity," *American Economic Review*, 2000, *90* (4), 765–786.

Becker, Gary S, *The economics of discrimination*, University of Chicago press, 1971.

Bertrand, Marianne, "New perspectives on gender," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1543–1590.

_ , "Gender in the twenty-first century," in "AEA Papers and proceedings," Vol. 110 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2020, pp. 1–24.

_ **and Esther Duflo**, "Field experiments on discrimination," *Handbook of economic field experiments*, 2017, *1*, 309–393.

Blau, Francine D, *Equal pay in the office*, Lexington Books, 1977.

Blinder, Alan S and Alan B Krueger, "Alternative measures of offshorability: a survey approach," *Journal of Labor Economics*, 2013, *31* (S1), S97–S128.

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

Breza, Emily, Supreet Kaur, and Yogita Shamdasani, "The morale effects of pay inequality," *The Quarterly Journal of Economics*, 2018, *133* (2), 611–663.

Bursztyn, Leonardo, Alexander W Cappelen, Bertil Tungodden, Alessandra Voena, and David H Yanagizawa-Drott, "How Are Gender Norms Perceived?," *National Bureau of Economic Research*, 2023.

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 2016, *131* (2), 633–686.

Charness, Gary and Peter Kuhn, "Lab labor: What can labor economists learn from the lab?," *Handbook of labor economics*, 2011, *4*, 229–330.

Commission, European, "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).

DellaVigna, Stefano and Matthew Gentzkow, "Uniform pricing in us retail chains," *The Quarterly Journal of Economics*, 2019, *134* (4), 2011–2084.

Fehr, Ernst, Lorenz Goette, and Christian Zehnder, "A behavioral account of the labor market: The role of fairness concerns," *Annu. Rev. Econ.*, 2009, *1* (1), 355–384.

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*, 2014, *104* (4), 1091–1119.

Grossman, Gene M and Elhanan Helpman, "Fair wages and foreign sourcing," *Harvard University Press*, 2008.

Hale, Galina and Mingzhi Xu, "FDI effects on the labor market of host countries," *2016-25*, 2016.

Hamilton, Barton H, Jack A Nickerson, and Hideo Owan, "Diversity and productivity in production teams," in "Advances in the Economic Analysis of participatory and Labor-managed Firms," Vol. 13, Emerald Group Publishing Limited, 2012, pp. 99–138.

Hazell, Jonathon, Christina Patterson, Heather Sarsons, and Bledi Taska, "National wage setting," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2022, (2022-150).

Heath, Rachel and Seema Jayachandran, "The causes and consequences of increased female education and labor force participation in developing countries," in "The Oxford Handbook of Women and the Economy," Oxford University Press, 2017, pp. 345–367.

Heckman, James and Bo Honoré, "The Empirical Content of the Roy Model," *Econometrica*, 1990, pp. 1121–1149.

Hjort, Jonas, "Ethnic divisions and production in firms," *The Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

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

Hong, Lu and Scott E Page, "Problem solving by heterogeneous agents," *Journal of economic theory*, 2001, *97* (1), 123–163.

Hsieh, Chang-Tai and Esteban Rossi-Hansberg, "The industrial revolution in services," *Journal of Political Economy Macroeconomics*, 2023, *1* (1), 3–42.

_ , 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.

Javorcik, Beata S, "Does FDI bring good jobs to host countries?," *The World Bank Research Observer*, 2015, *30* (1), 74–94.

Jayachandran, Seema, "The roots of gender inequality in developing countries," *economics*, 2015, *7* (1), 63–88.

_ , "Social Norms as a Barrier to Womens Employment in Developing Countries," *IMF Economic Review*, 2021, *69* (3), 576–595.

Lazear, Edward P, "Globalisation and the market for team-mates," *The Economic Journal*, 1999, *109* (454), 15–40.

_ , "Why are some immigrant groups more successful than others?," *Journal of Labor Economics*, 2021, *39* (1), 115–133.

Lemieux, Thomas, W Bentley MacLeod, and Daniel Parent, "Contract form, wage flexibility, and employment," *American Economic Review*, 2012, *102* (3), 526–531.

Marx, Benjamin, Vincent Pons, and Tavneet Suri, "Diversity and team performance in a Kenyan organization," *Journal of Public Economics*, 2021, *197*, 104332.

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.

_ **and** _ , "The evolution of gender gaps in industrialized countries," *Annual review of Economics*, 2016, *8*, 405–434.

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.

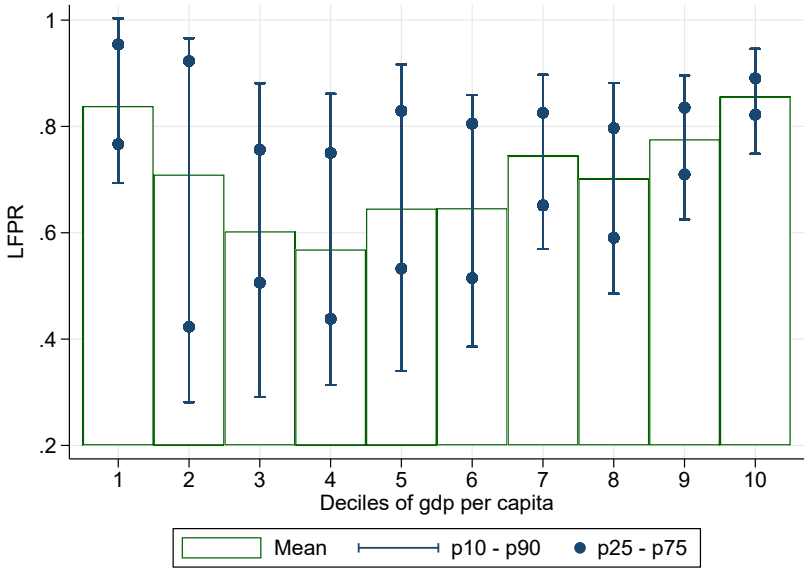
Sarsons, Heather, "Interpreting signals in the labor market: evidence from medical referrals," *Working paper*, 2017.

Verhoogen, Eric A, "Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector," *The Quarterly Journal of Economics*, 2008, *123* (2), 489–530.

Wiswall, Matthew and Basit Zafar, "Preference for the Workplace, Investment in Human Capital, and Gender*," *The Quarterly Journal of Economics*, 08 2018, *133* (1), 457–507.

8 Figures

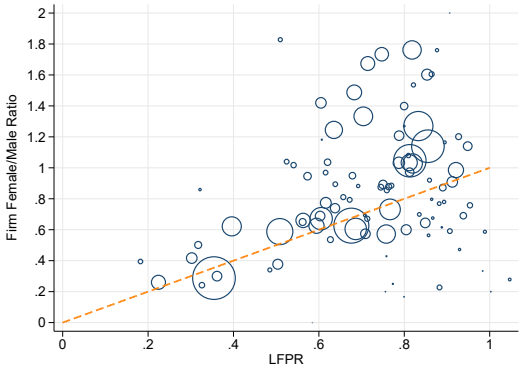
Figure 1: LFPR and GDP per capita



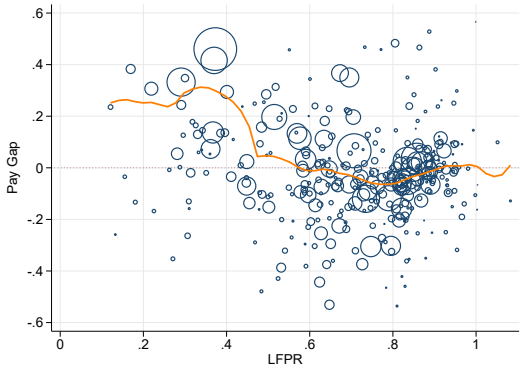
Notes. The figure plots the average, the interdecile range, and the interquartile range of the LFPR across deciles of GDP per capita across countries and cohorts. GDP per capita is at constant 2010 USD and is taken from the World Bank. The unit of observation is a country-cohort.

Figure 2: Gender ratio/pay gap and LFPR

(a) Gender ratio gap

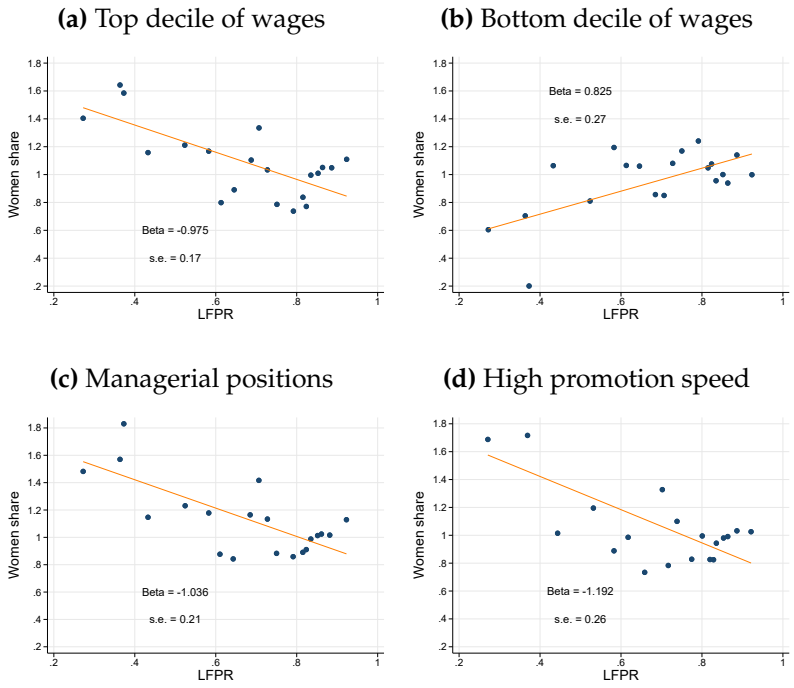


(b) Gender pay gap



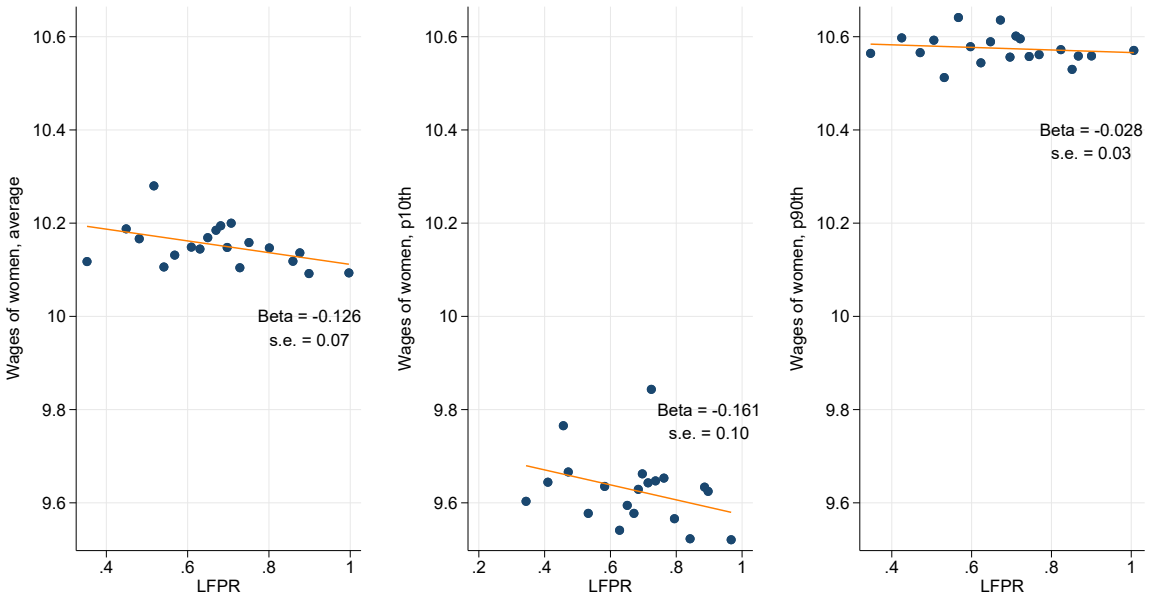
Notes. In panel (a), y-axis corresponds to the female/male employment ratio in the MNE while the x-axis corresponds to the LFPR in the countries. Each circle is a country and the size depends on the number of firm employees in the country. The orange line represents the 45-degree line. The unit of observation is a country. Panel (b) plots the gender pay gap (the difference between women's and men's salaries expressed as a % of men's salary) against the LFPR. Each circle represents a country-cohort pair, which size depends on the number of firm employees in the country-cohort cell. 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 3: Gender promotion gap and LFPR



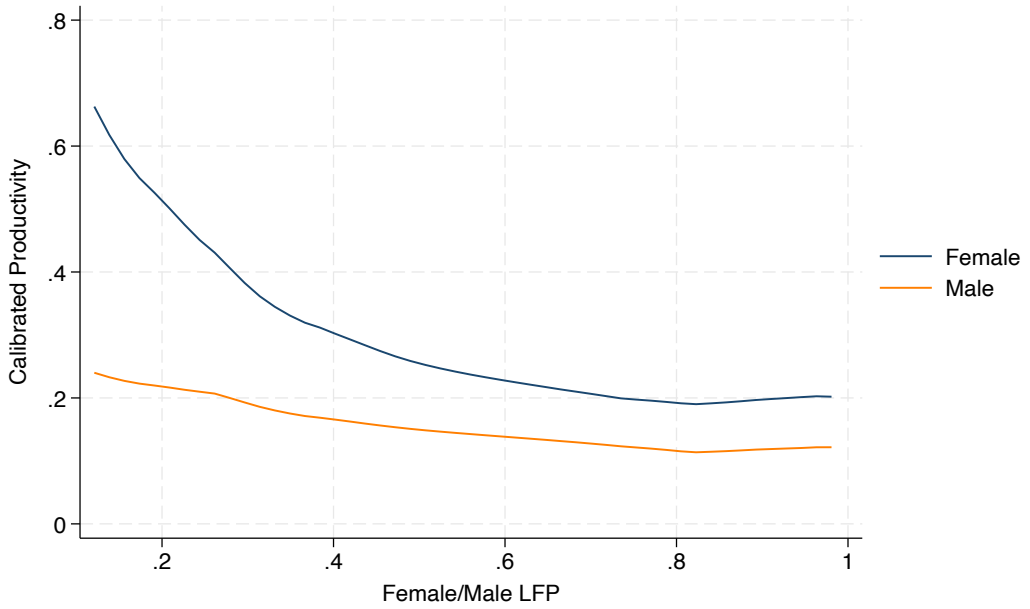
Notes. The figures are binned scatterplots and a linear fit of the share of women with different performance metrics (as a proportion of the female share in the firm) against the LFPR, in each country-cohort cell. Panel (a) looks at the top decile of wages; Panel (b) at the bottom decile; Panel (c) looks at the share of women in managerial positions; and Panel (d) at promotion speed to managerial positions, based on average promotion rates in the firm by labor market experience. In the regressions, we use analytical weights by employee size of each cohort-country cell. The unit of observation is a country-cohort.

Figure 4: Women's wages and LFPR



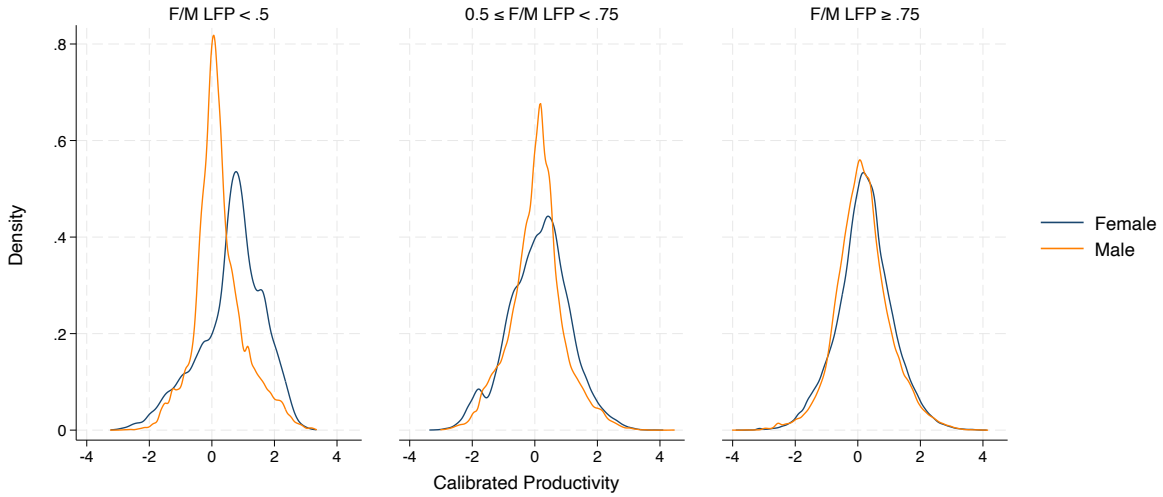
Notes. The figures are binned scatterplots and a linear fit of women's wages against the LFPR. The first figure from the left plots women's average wages, the middle one plots the 10th percentile, and the last one plots the 90th percentile. In the regressions, we control for the respective measure of men's wages, and we use analytical weights by employee size of each cohort-country cell and robust standard errors. The unit of observation is a country-cohort.

Figure 5: Calibrated productivity by gender and LFPR



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

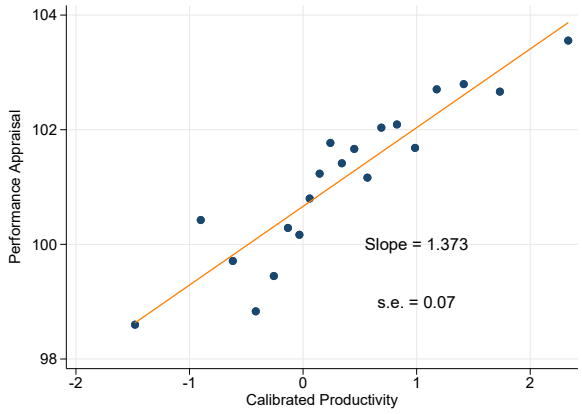
Figure 6: Calibrated productivity distribution by gender and LFPR



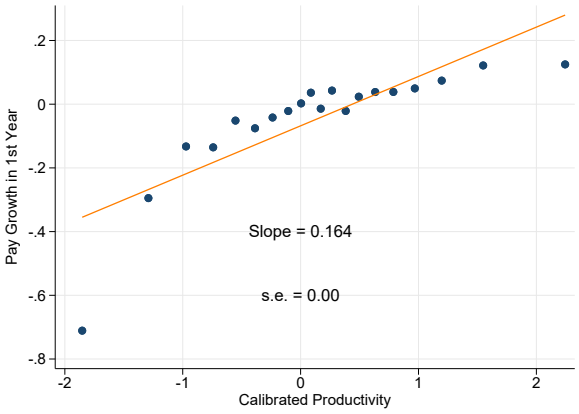
Notes. The figure plots a kernel density estimate of calibrated productivity for our sample of firm workers by three LFPR groups: $[0, .5)$, $[.5, .75)$ and $[.75, 1]$. The unit of observation is a worker.

Figure 7: Calibrated productivity and non-targeted performance measures

(a) Performance score given by manager



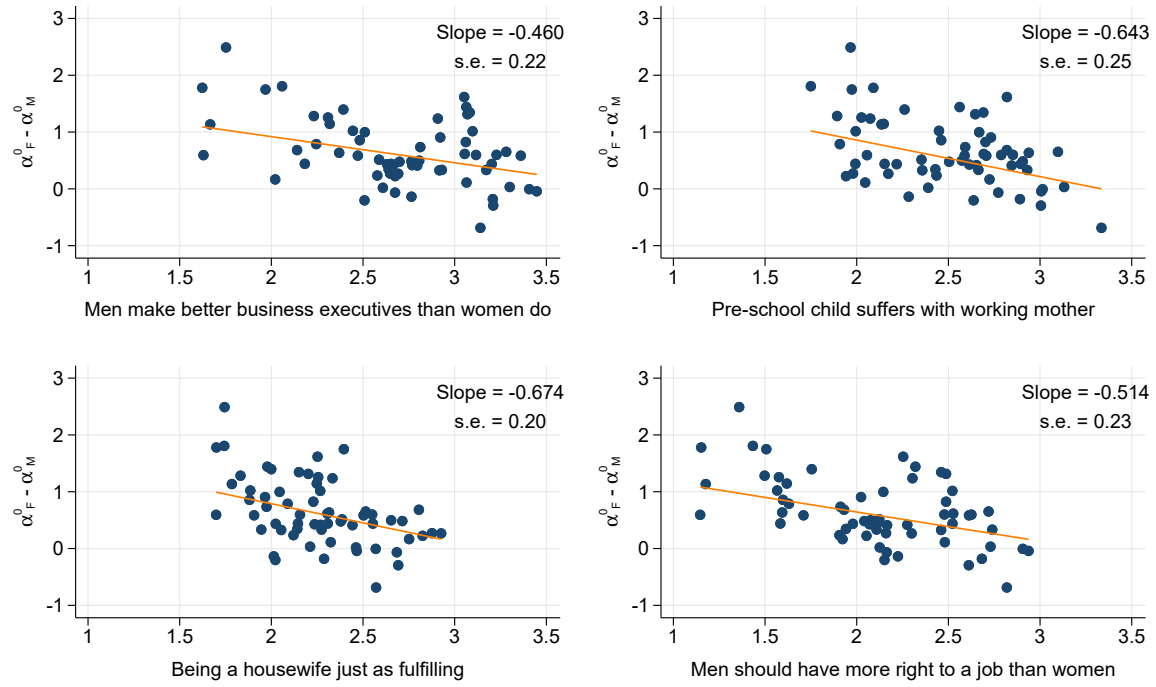
(b) Pay growth in the first year (new hires)



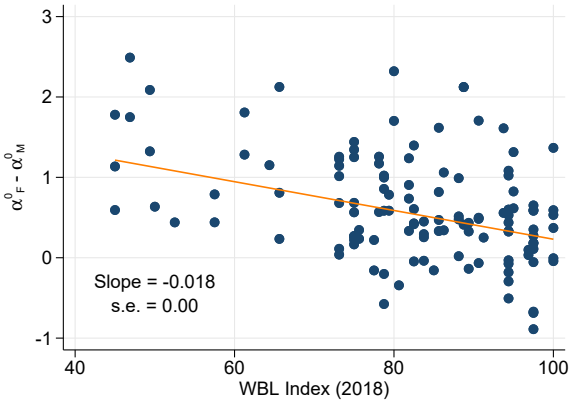
Notes. The figures are binned scatterplots and a linear fit of other performance measures (performance score, pay growth for new hires) against our calibrated productivity. The unit of observation is a worker-year. The performance score is an annual performance rating given by the line manager to each subordinate (continuous variable with values ranging between 0 and 150).

Figure 8: Calibrated parameters, norms and labor regulations

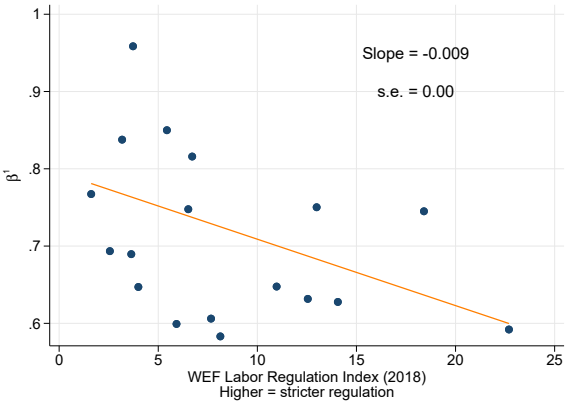
(a) World Value Survey



(b) Women, Business and the Law

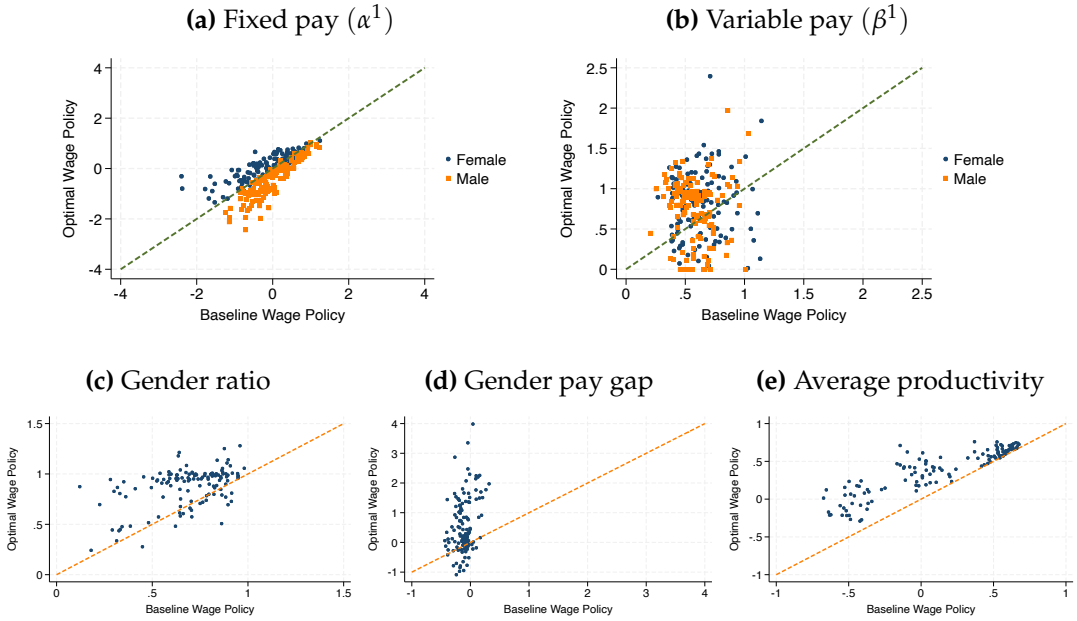


(c) Restrictive Labor Regulations Index



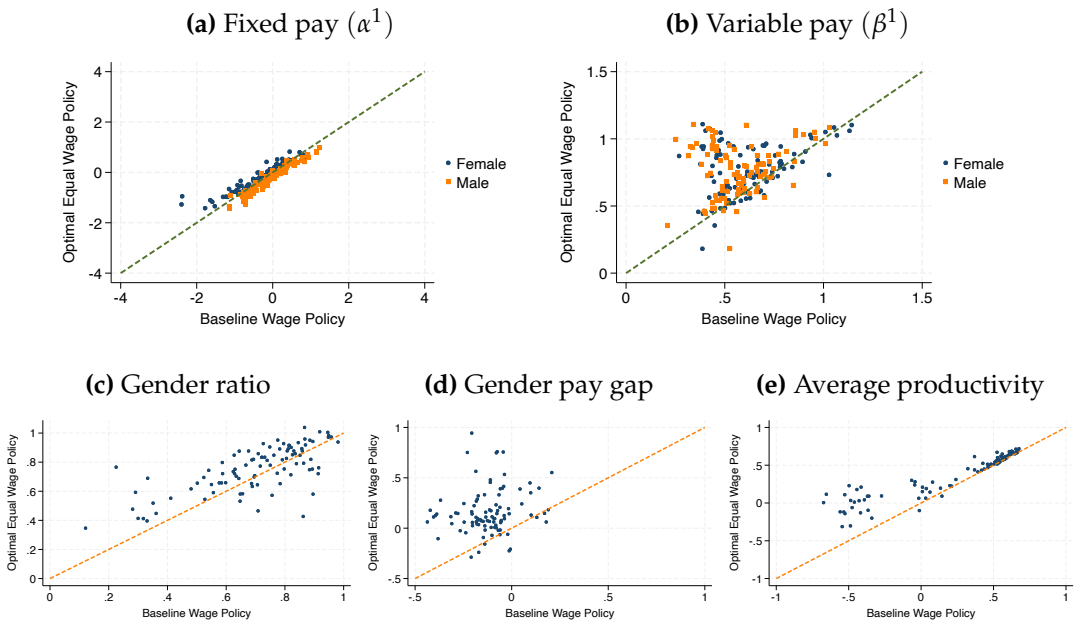
Notes. Panel (a) shows scatterplots and fitted linear regressions of the gap in calibrated $\alpha^0_{Ftac} - \alpha^0_{Mtac}$ against four questions in the World Value Survey: (1) "Men make better business executives than women do," (2) "Pre-school child suffers with working mother," (3) "Being a housewife 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 cell. Panel (b) shows the scatterplot and fitted linear regression of the gap in calibrated $\alpha^0_{Ftac} - \alpha^0_{Mtac}$ against the WB Women, Business, and the Law index in 2018. A higher value means fewer legal constraints on women's work. Each dot is a country cell. In all cases, the regression line uses cell size (at the firm) as analytic weights. Panel (c) shows the binned scatterplot and a linear fit of our calibrated parameters β^1_{gac} against the Restrictive Labor Regulations Index in 2018, using gender-country-cohort cell size (at the firm) as analytic weights.

Figure 9: Counterfactual I - optimal gender wage policy



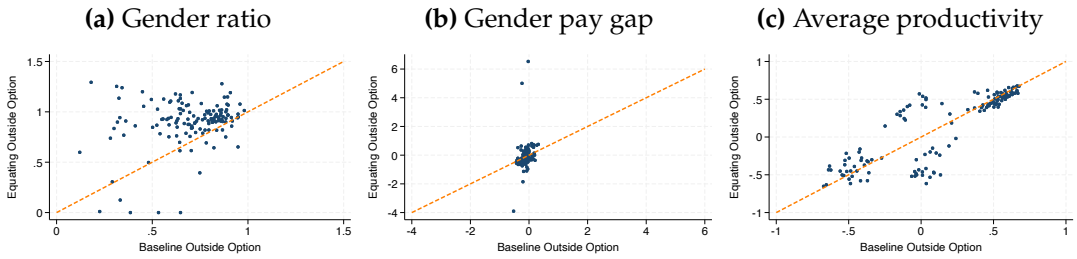
Notes. The figures compare different outcomes (female to male employment ratio, pay gap, and average productivity) and the wage policy parameters (α^1, β^1) at baseline vs. the optimal wage policy (see main text for details). Each dot represents a country-cohort cell and the 45-degree line is included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

Figure 10: Counterfactual II - optimal gender wage policy with equality constraint



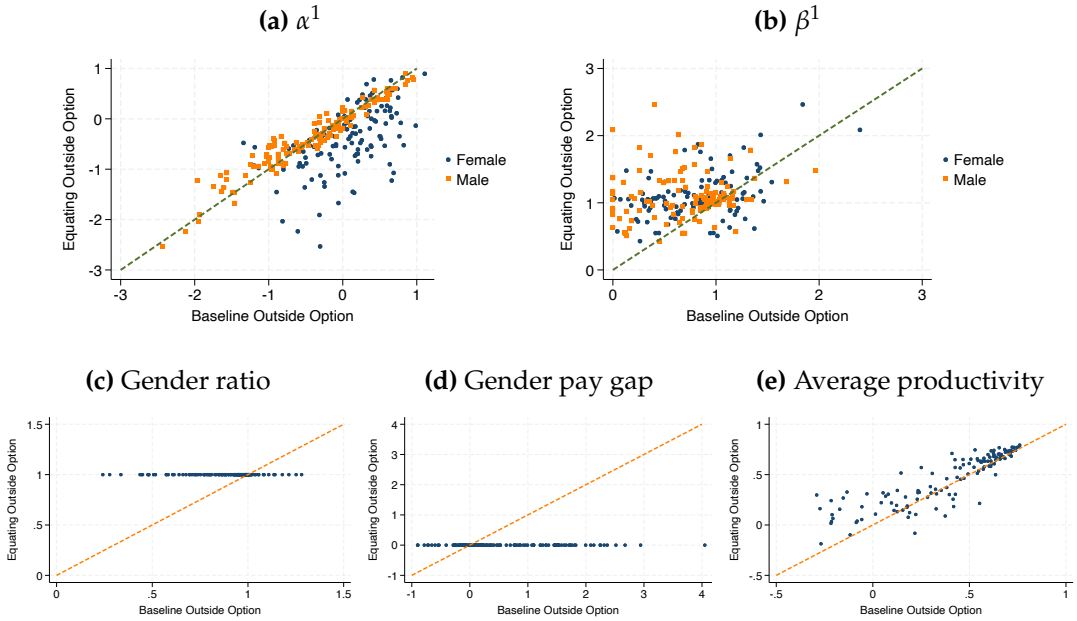
Notes. The figures compare different outcomes (female to male employment ratio, pay gap, and average productivity) and the wage policy parameters (α^1, β^1) at baseline vs. the optimal wage policy with equality constraint (see main text for details). Each dot represents a country-cohort cell and the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

Figure 11: Counterfactual III - gender equality in the labor force (baseline wage policy)



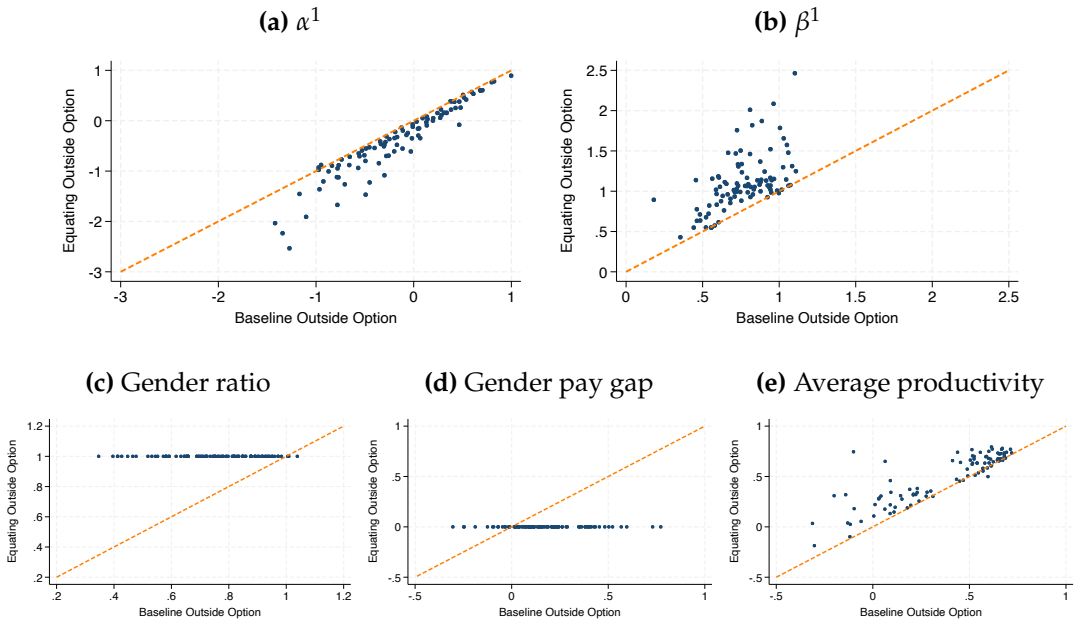
Notes. The figures compare different outcomes (female to male employment ratio, pay gap, and average productivity) under the baseline outside option parameters (α^0, ν^0, ρ) to the counterfactual where these are equalized at the male levels. We keep the wage policy of the firm fixed at the calibrated baseline parameters. Each dot represents a country-cohort cell and the 45-degree line is included.

Figure 12: Counterfactual IV - gender equality in labor force (optimized wage policy)



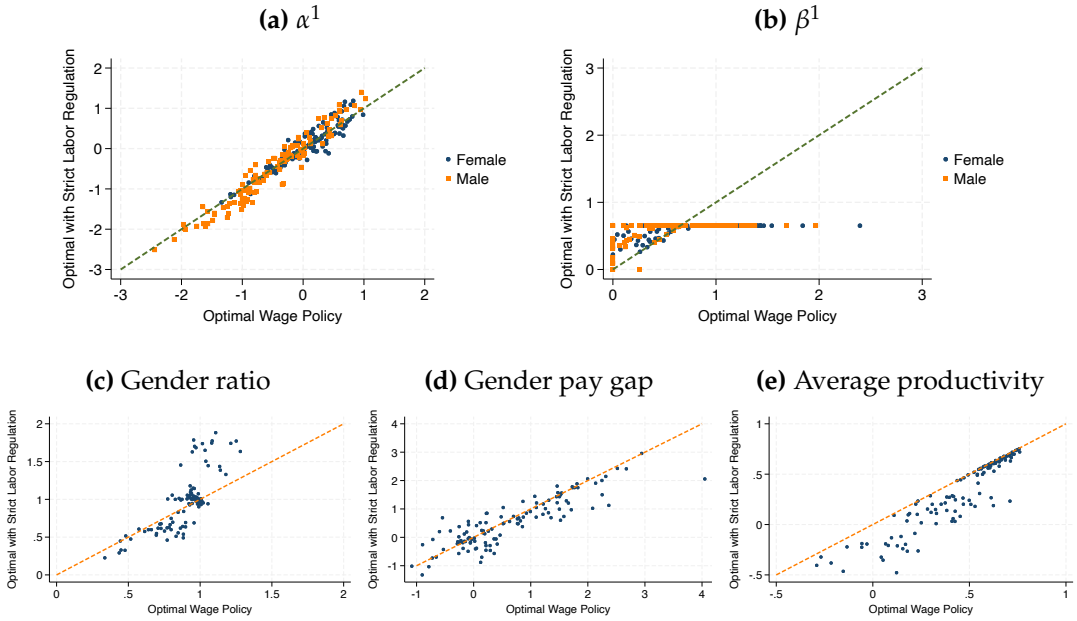
Notes. The figures compare different outcomes (female to male employment ratio, pay gap, and average productivity) and the wage policy parameters (α^1, β^1) under the baseline outside option parameters (α^0, ν^0, ρ) to the counterfactual where these are equalized at the male levels. In either case, we let the firm choose the optimal wage policy, as described in the main text. Each dot represents a country-cohort cell and the 45-degree line is included.

Figure 13: Counterfactual V - gender equality in labor force (optimized wage policy with equality constraint)



Notes. The figures compare different outcomes (female to male employment ratio, pay gap, and average productivity) and the wage policy parameters (α^1, β^1) under the baseline norm parameters (α^0, ν^0, ρ) to the counterfactual where these are equalized at the male levels. In either case, we let the firm choose the optimal wage policy, subject to an equality constraint, as described in the main text. Each dot represents a country-cohort cell and the 45-degree line is included.

Figure 14: Counterfactual VI - strict labor regulation (optimized wage policy)



Notes. The figures compare different outcomes (female to male employment ratio, pay gap, and average productivity) and the wage policy parameters (α^1, β^1), estimated under the optimal wage policy, to a scenario where we impose an additional cap on returns to ability β_1 (see main text for details). Each dot represents a country-cohort cell and the 45-degree line is included.

9 Tables

Table 1: Summary statistics

	Full Sample		Structural Sample	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
Pay + Bonus (logs)	10.331	10.274	10.383	10.352
	(0.636)	(0.581)	(0.597)	(0.549)
Age	39.581	39.057	38.485	38.485
	(11.397)	(11.320)	(10.742)	(10.742)
Tenure	9.132	8.877	9.164	8.995
	(7.337)	(7.263)	(7.245)	(6.957)
Share in Work-level 2+	0.192	0.156	0.193	0.177
	(0.218)	(0.178)	(0.173)	(0.157)
Share with Fast Promotions	0.280	0.251	0.301	0.268
	(0.342)	(0.324)	(0.335)	(0.322)
Share Top Performers	0.150	0.128	0.150	0.131
	(0.145)	(0.116)	(0.101)	(0.073)
Econ, Business, and Admin	0.515	0.567	0.517	0.559
	(0.324)	(0.305)	(0.305)	(0.286)
Share in Sales Function	0.514	0.380	0.486	0.336
	(0.269)	(0.255)	(0.213)	(0.186)
Observations	377	366	260	260
Median no. workers in cell	115	95	234	191

Notes. This table reports summary statistics for the relevant sample of workers used in the analysis. An observation is a gender-cohort-country cell. This is the relevant unit in the structural estimation. Columns 1 and 2 present results for the full sample. In Columns 3 and 4, sample is restricted to those cells used in structural estimation, where we excluded cells with fewer than 30 male or 30 female employees. Work level denotes the hierarchical tier (from level 1 at the bottom to level 6). The share of fast promotions only considers workers that achieve at least work-level 2 or higher. The sales function is the most common function (39%). Top performers are identified from the firm performance appraisal system based on the supervisor annual rating of worker performance.

Table 2: Gender pay gap and LFPR

	Full Sample					New Hires	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pay + Bonus (logs)					Pay Growth	Major promotion
Female	0.377	0.256	0.028	0.253	0.197	0.159	0.107
	(0.142)	(0.096)	(0.173)	(0.094)	(0.077)	(0.067)	(0.047)
LFPR	1.640	1.641	0.596	1.662	0.138	0.514	-0.028
	(0.282)	(0.212)	(0.153)	(0.204)	(0.235)	(0.151)	(0.080)
Female × LFPR	-0.564	-0.471	-0.274	-0.451	-0.376	-0.142	-0.112
	(0.194)	(0.135)	(0.116)	(0.132)	(0.103)	(0.087)	(0.062)
GDP per capita (logs)			0.275				
			(0.021)				
Female × GDP per capita (logs)			0.008				
			(0.016)				
Controls	No	Yes	Yes	Yes	Yes	No	No
Cohort FE	No	No	No	Yes	No	No	No
Country FE	No	No	No	No	Yes	No	No
N	303756	303756	302567	303756	303756	8274	8274
R-squared	0.116	0.285	0.435	0.307	0.540	0.103	0.056

Notes. An observation is a worker-year. Controls include: tenure, tenure squared, year FE and function FE. The last two columns report estimates when restricting the sample to new hires at the entry level observed for at least four years. *Pay growth* is computed as the difference in log pay between the last year a worker is observed and the first year a worker is observed. *Probability of promotion* equals 1 if the worker was promoted to work-level 2 during the sample period. Controlling for starting salary in columns 6 and 7. Standard errors clustered at the country-cohort level.

Table 3: Summary of model parameters and empirical targets

Param.	Interpretation	Empirical Target
A. Structural Parameters:		
α^1_{gac}	Unconditional average log-wage	Average observed log-wage (controlling for selection)
β^1_{gac}	Returns to productivity in the firm	Variance of the observed log-wage (controlling for selection)
α^0_{gac}	Unconditional average value of staying at home	Not separately identified. With an additional normalization, can be obtained from composite parameters ξ_{gac} and θ_{gac}
ν^0_{gac}	Dispersion of value of staying at home	
ρ_{gac}	Correlation between productivity and the value of staying at home	
B. Composite Parameters:		
ξ_{gac}	Participation threshold	LFP
θ_{gac}	Sign and strength of selection	Skewness of the observed log-wage (controlling for selection)

Notes. This table provides a summary of the parameters of the model and the empirical target that each parameter aims to match in the calibration strategy that we describe in subsection 4.2.

Table 4: Summary of Counterfactual Results

	Male				Female			
	mean	sd	min	max	mean	sd	min	max
A. Full Structural Sample (260 country-cohort cells)								
α^1_{gac}	−0.04	0.58	−1.25	1.93	−0.31	0.79	−2.40	1.41
β^1_{gac}	0.70	0.18	0.21	1.57	0.78	0.25	0.25	1.90
ξ_{gac}	−0.91	0.17	−1.44	−0.27	−0.12	0.41	−1.14	1.25
θ_{gac}	0.29	0.65	−1.48	1.36	0.34	0.64	−1.81	1.41
B. Counterfactual Sample I (114 country-cohort cells)								
α^0_{gac}	−1.13	0.67	−2.51	0.62	−0.48	0.60	−1.90	1.08
ν^0_{gac}	1.00	0.00	1.00	1.00	1.00	0.00	1.00	1.00
ρ_{gac}	−0.26	0.68	−0.98	1.00	−0.20	0.62	−1.00	1.00

Notes. This table summarizes the parameters by gender-country-cohort cells, using cell sizes as analytic weights. Panel A reports the parameters that are point identified without additional normalizations, as discussed in subsection 4.2, in the full structural sample (260 cells). Panel B reports the parameters relating to the value of staying out of the labor force, under the normalization $\nu^0_{gac} \equiv 1$, excluding cells in which (a) we cannot calibrate α^0, ν^0 and ρ, or (b) we cannot compute one of the optimal wage policies for the main set of counterfactuals.

Table 5: Summary of Counterfactual Results

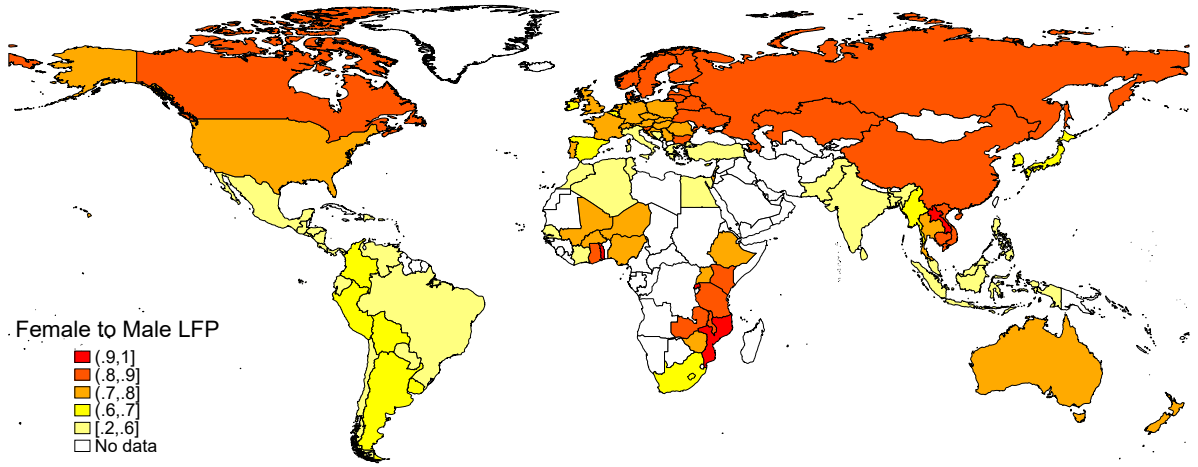
	Parameters		Results		
	α^1, β^1	α^0, β^0, ρ	Productivity	F/M Ratio	Pay Gap (F−M)
	A. Full Structural Sample (260 country-cohort cells)				
(1a)	Baseline	Baseline	0.25	0.67	−0.08
	B. Counterfactual Sample I (114 country-cohort cells)				
(1b)	Baseline	Baseline	0.31	0.71	−0.10
(2b)	Optimal	Baseline	0.47	0.92	0.63
(3b)	Baseline	Men's	0.28	0.92	−0.12
(4b)	Optimal	Men's	0.52	1.00	0.00
(5b)	Strict Labor Reg.	Baseline	0.39	0.99	0.57
	C. Counterfactual Sample II (88 country-cohort cells)				
(1c)	Baseline	Baseline	0.31	0.71	−0.10
(2c)	Optimal	Baseline	0.47	0.92	0.63
(3c)	Baseline	Men's	0.28	0.92	−0.12
(4c)	Optimal	Men's	0.52	1.00	0.00
(5c)	Strict Labor Reg.	Baseline	0.39	0.99	0.57
(6c)	Opt. with Equality	Baseline	0.46	0.78	0.14
(7c)	Opt. with Equality	Men's	0.52	1.00	0.00

Notes. This table summarizes the counterfactual results (average productivity, female-to-male employee ratio, and pay gap) across country-cohort cells, weighted by cell size. Productivity ($LFP \cdot \mathbb{E}[A_i \mid \text{empl'd}]$) is measured in units of standard deviations of the underlying latent variable A_i. The pay gap is the difference in log-wage (female − male). In the "optimal" wage policy, the firm maximizes ability subject to an employment and a wage bill constraint, see subsection A.3. In the "strict labor regulations" counterfactual, the firm solves the same problem as in the "optimal" wage policy, with an additional cap on returns to ability. In the "optimal with equality" wage policy, the firm maximizes ability subject to an employment, a wage bill constraint, and an equality between men and women restriction. Panel A reports the baseline results in the full structural sample (260 cells). Panel B reports a set of counterfactual results for 114 cells, which exclude those in which (a) we cannot calibrate α^0, ν^0, and ρ, or (b) we cannot compute one of the optimal wage policies for the counterfactuals reported. Panel C reports the same results for 88 cells, in which we can additionally compute the "optimal with equality" wage policy, primarily because the feasible region is empty (i.e., the firm cannot choose a wage policy that satisfies the LFP, the wage bill and the equality constraint at the same time).

A Appendix

A.1 Appendix figures and tables

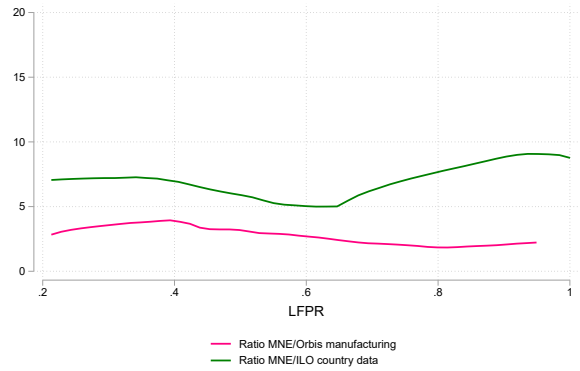
Figure A.1: The countries where the MNE operates and female to male LFP



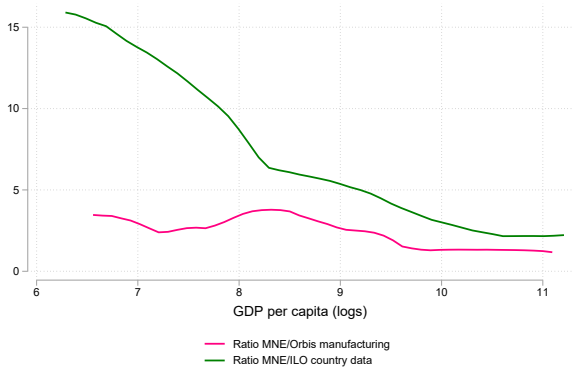
Notes. This figure plots the average labor force participation ratio (LFPR) in each country (averaged across cohorts).

Figure A.2: Average wages in the firm and in the country overall: a) ILO, white collar occupations only; and b) ORBIS, manufacturing sector only

(a) Wage ratios against LFPR



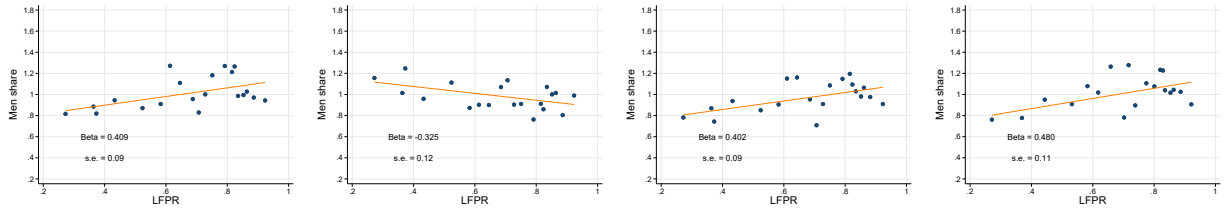
(b) Wage ratios against GDP per capita in logs



Notes. This figure plots the ratio of the 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 (ILO), considering white-collar occupations only. Wages are measured in 2017 PPP $. The x-axis is the LFPR (panel a) and the GDP per capita in logs (panel b) in each country.

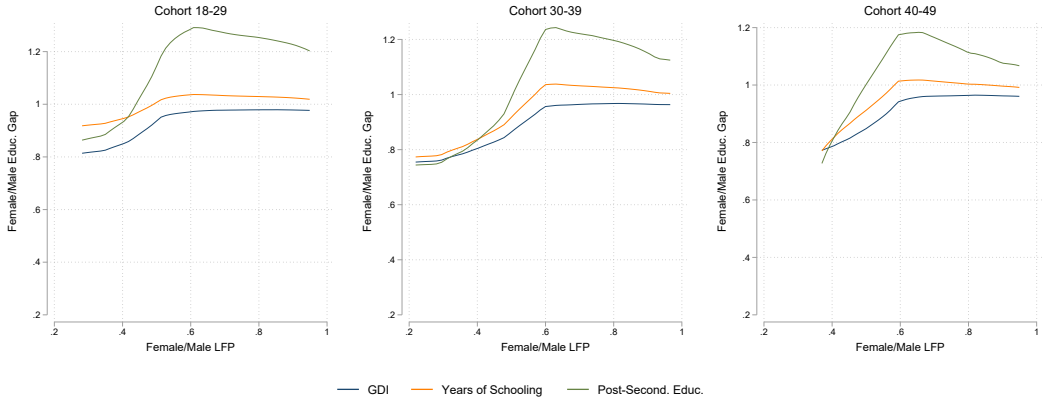
Figure A.3: Gender promotion gap and LFPR

(a) Share men, top decile of wages **(b)** Share men, bottom decile of wages **(c)** Share men, managerial positions **(d)** Share men, high promotion speed



Notes. The figures are binned scatterplots and a linear fit of the share of men with different performance metrics (as a proportion of the male share in the firm) against the LFPR, in each country-cohort cell. Panel (a) looks at the top decile of wages; Panel (b) at the bottom decile; Panel (c) looks at the share of men in managerial positions; and Panel (d) at promotion speed to managerial positions, based on average promotion rates in the firm by labor market experience. In the regressions, we use analytical weights by employee size of each cohort-country cell. The unit of observation is a country-cohort.

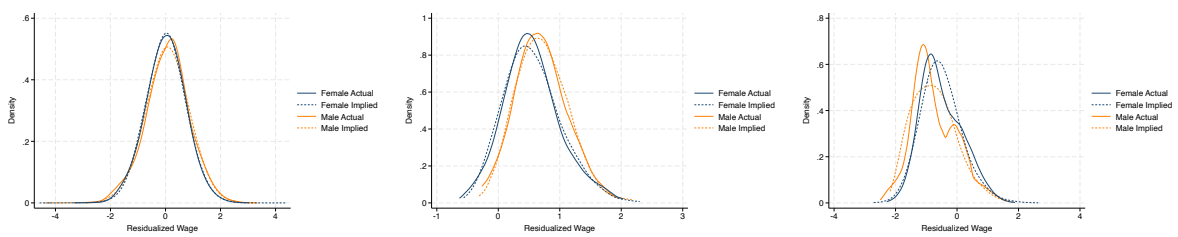
Figure A.4: Gender education gap and LFPR



Notes. The figure plots the gender gap in education against the gender gap in LFP for each country separately by age cohort. We use three distinct 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. The unit of observation in each plot is a country.

Figure A.5: Goodness of fit: Wage distribution in India and Sweden

(a) Overall **(b)** Sweden **(c)** India



Notes. The figures compare the actual distribution of residualized log wages to the distributions implied by our calibrated model (a) across all countries, (b) in Sweden, and (c) in India. Wages (in logs) are residualized using year and function dummies, tenure, and tenure squared, as detailed in sub-section 4.2.

Table A.1: Gender pay gap and LFPR — LFP for population with advanced education

	Pay + Bonus (logs)				
	(1)	(2)	(3)	(4)	(5)
Female	0.427	0.269	0.161	0.303	0.381
	(0.173)	(0.082)	(0.150)	(0.079)	(0.073)
LFPR, advanced education	1.776	1.692	-0.138	1.684	0.073
	(0.188)	(0.113)	(0.222)	(0.116)	(0.273)
Female × LFPR, advanced education	-0.496	-0.375	-0.484	-0.401	-0.522
	(0.198)	(0.100)	(0.140)	(0.096)	(0.088)
GDP per capita (logs)			0.358		
			(0.033)		
Female × GDP per capita (logs)			0.020		
			(0.016)		
Controls	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	No	Yes	Yes
Country FE	No	No	No	No	No
N	251336	251336	250152	251336	251336
R-squared	0.155	0.315	0.469	0.332	0.570

Notes. An observation is a worker-year. The LFPR is computed using the LFP for individuals with advanced education (short-cycle tertiary education or college degree and/or above). Controls include: tenure, tenure squared, year FE and function FE. 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)
Female	0.256	0.208	0.265	-0.032
	(0.096)	(0.070)	(0.050)	(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)	(0.100)	(0.068)	(0.156)
Controls	Yes	Yes	Yes	Yes
N	303756	303756	71658	232098
R-squared	0.285	0.387	0.218	0.239

Notes. An observation is a worker-year. Controls include: tenure, tenure squared, year FE 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

	Pay + Bonus (logs), PPP 2017 USD				
	(1)	(2)	(3)	(4)	(5)
Female	0.318	0.195	0.312	0.197	0.198
	(0.127)	(0.080)	(0.141)	(0.080)	(0.076)
LFPR	0.266	0.249	0.085	0.259	0.129
	(0.200)	(0.136)	(0.155)	(0.130)	(0.239)
Female × LFPR	-0.463	-0.372	-0.275	-0.364	-0.381
	(0.165)	(0.104)	(0.083)	(0.105)	(0.102)
GDP per capita (logs)			0.044		
			(0.018)		
Female × GDP per capita (logs)			-0.021		
			(0.013)		
Controls	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	No	Yes	Yes
Country FE	No	No	No	No	No
N	302789	302789	301600	302789	302789
R-squared	0.014	0.164	0.169	0.173	0.339

Notes. An observation is a worker-year. Controls include: tenure, tenure squared, year FE and function FE. 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).

Table A.4: Gender pay gap and LFPR — fixed pay only

	Pay (logs)				
	(1)	(2)	(3)	(4)	(5)
Female	0.358	0.232	-0.023	0.230	0.167
	(0.135)	(0.092)	(0.162)	(0.091)	(0.074)
LFPR	1.582	1.578	0.501	1.599	0.144
	(0.282)	(0.212)	(0.145)	(0.204)	(0.234)
Female × LFPR	-0.533	-0.436	-0.241	-0.415	-0.328
	(0.185)	(0.131)	(0.114)	(0.129)	(0.100)
GDP per capita (logs)			0.283		
			(0.020)		
Female × GDP per capita (logs)			0.011		
			(0.015)		
Controls	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	No	Yes	Yes
Country FE	No	No	No	No	No
N	303756	303756	302567	303756	303756
R-squared	0.110	0.285	0.448	0.309	0.552

Notes. An observation is a worker-year. Controls include: tenure, tenure squared, year FE and function FE. Standard errors clustered at the country-cohort level.

A.2 Model calibration

As discussed in 4.2, α^1_{gac} and β^1_{gac} are identified from the moments of the observed wage distribution, but we cannot identify the parameters α^0_{gac}, ν^0_{gac} and ρ_{gac} separately, because we do not observe the value of staying at home. For our main specification, following Heckman and Honoré (1990), we normalize $\nu^0_{gac} \equiv 1$, which allows us to calibrate α^0_{gac}, and ρ_{gac}. An alternative normalization suggested by Heckman and Honoré (1990) is $\rho_{gac} = 0$, implying that productivity in the firm is independent of the value of staying at home. Figure A.6 shows that, regardless of whether we normalize ν^0_{gac} or ρ_{gac}, we obtain very similar values of α^0_{gac}. The correlation between α^0_{gac} in the main calibration and in the alternative calibration is 0.95 for women and 0.96 for men. Additionally, the results of our main counterfactual exercises are robust to this alternative calibration, as shown in Table A.5, to be compared with Panel C of Table 5.

Figure A.6: α^0_{gac} in the main and alternative calibration

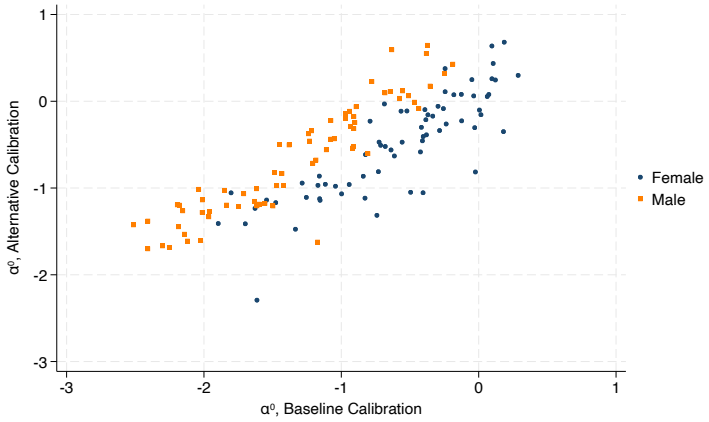


Table A.5: Summary of Counterfactual Results, Robustness

	Parameters		Results		
	α^1, β^1	α^0, β^0, ρ	Productivity	F/M Ratio	Pay Gap (F−M)
	Robustness Sample (60 country-cohort cells)				
(1)	Baseline	Baseline	0.50	0.69	−0.10
(2)	Optimal	Baseline	0.57	0.94	0.30
(3)	Baseline	Men's	0.51	0.81	−0.16
(4)	Optimal	Men's	0.61	1.00	0.00
(5)	Strict Labor Reg.	Baseline	0.55	0.98	0.24
(6)	Opt. with Equality	Baseline	0.55	0.86	0.06
(7)	Opt. with Equality	Men's	0.61	1.00	0.00

Notes. This table summarizes the counterfactual results (average productivity, female-to-male employee ratio, and pay gap) across country-cohort cells, weighted by cell size. Productivity ($LFP \cdot \mathbb{E}[A_i \mid \text{empl'd}]$) is measured in units of standard deviations of the underlying latent variable A_i. The pay gap is the difference in log-wage (female − male). In the "optimal" wage policy, the firm maximizes ability subject to an employment and a wage bill constraint, see subsection A.3. In the "strict labor regulations" counterfactual, the firm solves the same problem as in the "optimal" wage policy, with an additional cap on returns to ability. In the "optimal with equality" wage policy, the firm maximizes ability subject to an employment, a wage bill constraint, and an equality between men and women restriction. We report the same results as in Table 5 for 60 country-cohort cells, in which we can calibrate the parameters of the value of staying out of the labor force and compute all the counterfactual wage policies.

A.3 Optimization problems for the firm's wage policy

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

$$\max_{(\tilde{\alpha}^1_{gac},\tilde{\beta}^1_{gac})_{g\in\{F,M\}}} \quad \widetilde{LFP}_{Fac}\widetilde{A}_{Fac} + \widetilde{LFP}_{Fac}\widetilde{A}_{Mac}$$

subj. to:

$$\widetilde{LFP}_{Fac} + \widetilde{LFP}_{Mac} \geq LFP_{Fac} + LFP_{Mac} \tag{i}$$

$$\widetilde{LFP}_{Fac}\widetilde{W}_{Fac} + \widetilde{LFP}_{Mac}\widetilde{W}_{Mac} \leq LFP_{Fac}W_{Fac} + LFP_{Mac}W_{Mac} \tag{ii}$$

where

$$\widetilde{LFP}_{gac} = 1 - \Phi\left(\frac{\alpha^0_{gac} - \tilde{\alpha}^1_{gac}}{\sqrt{(\tilde{\beta}^1_{gac})^2 + (\nu^0_{gac})^2 - 2\rho_{gac}\tilde{\beta}^1_{gac}\nu^0_{gac}}}\right)$$

$$\widetilde{A}_{gac} = \frac{\tilde{\beta}^1_{gac} - \rho_{gac}\nu^0_{gac}}{\sqrt{(\tilde{\beta}^1_{gac})^2 + (\nu^0_{gac})^2 - 2\rho_{gac}\tilde{\beta}^1_{gac}\nu^0_{gac}}}\lambda\left(\frac{\alpha^0_{gac} - \tilde{\alpha}^1_{gac}}{\sqrt{(\tilde{\beta}^1_{gac})^2 + (\nu^0_{gac})^2 - 2\rho_{gac}\tilde{\beta}^1_{gac}\nu^0_{gac}}}\right)$$

$$\widetilde{W}_{gac} = \tilde{\alpha}^1_{gac} + \tilde{\beta}^1_{gac}\widetilde{A}_{gac}$$

are the probability of being employed, average ability conditional on being employed, and average wage conditional on being employed, and LFP_{gac}, W_{gac} are the observed LFP and average wages. 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 be equal to or greater than the total LFP at baseline. Constraint (ii) states that the total wage bill should be equal to or smaller than the total wage bill at baseline.

The optimal wage policy with equality constraint solves the program above with the additional restriction that $\tilde{\alpha}^1_{Fac} = \tilde{\alpha}^1_{Mac}$ and $\tilde{\beta}^1_{Fac} = \tilde{\beta}^1_{Mac}$ for all a, c. That is, the firm needs to have the same wage policy for males and females within a country-cohort cell so that workers with equal ability are paid the same.