Gender Barriers, Structural Transformation, and Economic Development*

Gaurav Chiplunkar
University of Virginia

Tatjana Kleineberg
World Bank

December 2023

Abstract

Using nationally representative data across many countries spanning five decades between 1970-2018, we document distinct *gender* patterns in the process of structural transformation – both across sectors and across occupations *within* sectors. While gender differences in occupational and sectoral choices have declined in all countries over time, differences remain salient even today, even in the most developed countries. Gender segregation in employment is largely driven by variation *across* sectors in low- and middle-income countries, but by variation across occupations *within* sectors in high-income countries. Gender wage gaps have declined more slowly and without a clear correlation with economic development. Interpreted through the lens of a Roy model and a development accounting framework, the reductions of gender barriers–defined as gender-specific wedges in occupational-and-sectoral preferences and remuneration–account on average for more than a third of the observed employment transitions towards non-agriculture, and for around 20% of real output growth in our sample that includes countries at different stages of economic development.

Keywords: Gender, Structural Transformation, Economic Development

*We would like to thank Cheryl Doss, Doug Gollin, Charles Gottlieb, Michael Peters, Barbara Petrongolo, Markus Poschke, Claudia Olivetti, and participants in conferences and seminars at the World Bank, Northwestern University, Cornell University, Dartmouth University, Boston University, University of Virginia, CEPR STEG Annual Workshop, NEUDC, ASSA, SED, and WADES for helpful suggestions and comments. Anushka Chawla, Sebastian Cubillos, and Facundo Ulivarri provided excellent research assistance.

Chiplunkar: Darden School of Business, University of Virginia, 100, Darden Blvd., Charlottesville, VA 22905. Email: ChiplunkarG@darden.virginia.edu. Kleineberg: Development Research Group, World Bank, 1818 H Street NW, Washington, DC 20433. Email: tkleineberg@worldbank.org.

1 Introduction

As countries grow richer, workers allocate from agriculture to non-agriculture sectors. The drivers of structural transformation and its implications on aggregate productivity and economic development have been extensively studied (cf. Kuznets (1973); Maddison (1980); Duernecker and Herrendorf (2022)). As main drivers of structural transformation, the literature focuses on sector-specific technological change, income effects in consumption, as well as differences in sectors' capital- or human-capital-intensity combined with the increase in the supply of capital and human capital (Herrendorf et al. (2013, 2014), Ngai and Pissarides (2007), Acemoglu and Guerrieri (2008), Caunedo and Keller (2023) Porzio et al. (2022)). In this paper, we show that changes in gender roles over the last decades across the globe have been an additional important driver of structural transformation. We first show that men and women make different occupational and sectoral choices and that employment transitions along countries' development path differ between men and women. Second, we show that gender wage gaps differ across sectors and occupations and are still large even today and even in the richest countries.

For example, despite improvements in better representation of women in professional and managerial roles, as well as within the service sector, there remains a prevalent over-representation of women in clerical positions (such as secretaries, receptionists, cashiers, etc.). Furthermore, despite the decline in gender wage gaps across countries in recent years, women continue to earn less than men, even in the most developed countries today. For example, women earn only around 85 cents for every dollar earned by a man in the developed G7 and OECD countries (Fabrizio et al., 2018; OECD, 2023). Second, similar trajectories of economic growth across countries have resulted in very different ways in which men and women have engaged in the labor market. For instance, while the expansion of the service sector in the United States contributed to a reduction in gender-based wage disparities and employment gaps (Ngai and Petrongolo, 2017), a similar service-led growth in India has resulted in falling female labor force participation over the past few decades. Lastly, gender differences in the labor market have important implications for aggregate productivity and economic development, as documented by Hsieh, Hurst, Jones and Klenow (2019) in the US.[1]

In this paper, we document stylized facts on the gender dimension of structural transformation, using comprehensive labor-force and census surveys across 91 countries and spanning almost five decades from 1970 to 2018 (272 country-years). We show that the canonical process of structural transformation, i.e., employment transitions from agriculture to non-agriculture with economic development, is primarily driven by men. We observe a very different trajectory for women: at low levels of economic development (measured by real GDP per capita), women move out of agriculture, but exit the labor force, only to re-enter the workforce at higher levels of development in the service sector. Unlike their male counterparts, the share of women working in the manufacturing sector remains small and constant across all levels of economic development. A novel contribution of this paper is to also document the gender aspect of occupation transitions as well, in addition to the sectoral ones. This focus is crucial because (as we will show later) a key component of understanding employment transitions between men and women as countries develop, is driven by occupational transitions *within* sectors, as opposed to those *across* sectors. For instance, while the likelihood of women working as managers or in clerical jobs (like secretaries, receptionists, etc.) is around

[1]Hsieh et al. (2019) show that a reduction in gender norms and wage discrimination in the US since 1960 has improved the talent allocation in the economy, potentially explaining around 20-40% of the aggregate growth over the last five decades.

half that of men in low-income countries, they are almost as likely to work as managers, and 2-3 times more likely to work in clerical jobs in rich countries.

To measure gender segregation across occupations and sectors more systematically, we compute the Theil entropy index of segregation. Similar to the Gini index, this metric quantifies the degree to which an observed distribution of men and women across occupation-sectors deviates from an egalitarian allocation. Our findings reveal an inverted-U shaped pattern in gender segregation and economic development. Gender segregation first increases between low- and middle-income countries, followed by a decline in high-income ones. This can be rationalized by the fact that most men and women work in agricultural jobs in low-income countries, resulting in low gender segregation. With economic development, women transition out of the labor force, while men move into manufacturing and service jobs, thus increasing gender segregation in middle-income countries. At advanced levels of economic development, segregation declines again as more men and women find employment in the service sector. A key advantage of the entropy index is that it can be additively decomposed into the share of gender segregation that can be attributed to variation across sectors as opposed to within a sector across occupations. We find that 40% of the overall gender segregation in low- and middle-income countries, and more than 60% in high-income countries, is driven by variation within sectors (across occupations). This underscores the need to incorporate the occupational dimension within sectors when examining employment choices of men and women (as advocated by Bandiera et al. (2022) as well). It is important, for example, to distinguish between a scenario where women in the service sector work on lower rungs of the job ladder, as opposed to another where they are also equally represented at the top as well.

After documenting the employment transitions for men and women, we narrow our focus to a select group of six countries (namely, India, Indonesia, Brazil, Mexico, Canada, and the U.S.) that we call our "core sample". While the core sample consists of countries that cover the spectrum of economic development, we restrict our attention to these countries primarily for three reasons: firstly, the differences in employment transitions documented earlier conflate inherent disparities between countries and the changes within countries over time. To be able to distinguish between these, our core sample provides us with at least one labor-force survey/census in each decade between 1970 to 2018. Secondly, and more importantly, these countries (unlike the larger sample) contain reliable measures of hourly wages for men and women, within each occupation and sector, spanning the five decades. We leverage these data to document novel patterns in how gender wage gaps across sectors and occupations have evolved within and across countries over time. Lastly, our quantification exercise (described below) requires us to bring in additional data (most notably sectoral real value added growth), which is also reliably available for these countries.

Turning to the results, we find that while the gender wage and employment gaps (the probability that a woman works in an occupation-sector as compared to a man) were large in the 1970s, they have significantly declined across almost all occupation-sectors and countries over time. Interestingly, while employment gaps decline more generally with economic development as well i.e., middle- and high-income countries have lower employment gaps as compared to low-income ones, wage gaps are strikingly similar in both poor and rich countries. Therefore, gender equity in employment can co-exist with large gender pay gaps, which is most salient for example, in professional occupations and (albeit to a lesser degree) within the service sector.

Several economic channels, standard in the literature, can help partially explain these gender differences in employment and wages. For instance, these differences could stem from technological change that is biased towards specific sectors or occupations where women have a comparative (dis)advantage, or from gender differences in the skill distribution that impacts the supply of effective human capital, or income effects that tilt consumption baskets towards the service sector, where women might have a comparative advantage. On the other hand, there could also be "non-economic" drivers that can make it more or less attractive for women (relative to men) to work in certain occupation-sectors, or in the labor force all together. These include (in a reduced-form way) "wage discrimination", where employers pay women less than men *conditional* on their marginal product i.e., level of human capital; and "gender norms", which are *excess* utility costs incurred by women for working in certain occupation-sectors, or in the labor force all together. These reflect differences in amenities, restrictive social norms, or other barriers faced by women that, *in addition* to the real wage differences across occupation-sectors, impacts their employment choices in them.

We therefore develop a theoretical framework that parsimoniously incorporates each of the above-mentioned channels, thus allowing us to study the role of gender in the economy and its aggregate implications. The aim of this theoretical exercise is twofold: first, to quantify the magnitudes of (and changes in) gender norms and wage discrimination across countries and over time, *after* taking into account differences in the economic fundamentals in them; and second, to simulate counterfactual growth trajectories for countries that can help assess the importance of gender barriers in impacting their process of structural transformation and economic growth. In a nutshell, individuals in our model differ by gender and ability (schooling). Conditional on the fundamentals of the economy, captured by the economic and non-economic channels described previously, they choose: (i) whether to participate in the labor force or not; and (ii) conditional on working, their occupation and sector of work. On the demand side, we incorporate non-homothetic preferences that allow for sectoral expenditure shares to vary by income (Alder et al., 2022; Fan et al., 2021). Wages and prices are then determined in equilibrium. The non-economic channels therefore distort workers' employment choices, thus misallocating talent in the economy, and potentially lowering productivity and growth. We then take the model to the data. Our calibration strategy imposes minimal restrictions on the fundamental economic and non-economic parameters i.e., we do not impose any dynamic restrictions on their evolution, their relationship with the level of economic development, or whether one gender systematically has an (dis-)advantage or faces more constraining barriers as compared to the other. In the spirit of an accounting exercise, gender barriers (norms and discrimination) within each occupation-sector (country-year) are then then calibrated as residual wedges that match the observed gender gaps in employment and wages, *after* accounting for economic factors. As we demonstrate in Section 7.2, despite being residuals, our estimates of gender barriers strongly correlate with empirically measured social norms and labor market constraints in the World Bank's "World, Business, and the Law" (WBL) database (World Bank, 2019, 2020; Hyland et al., 2020).

Turning to the results, we find that gender norms i.e, the excess utility costs faced by women for participating in occupation-sectors (relative to men), are on average lower in richer countries than poorer ones. While they have declined across all countries in our sample since the 1970s, the decline has been largest in middle-income (Brazil and Mexico) and high-income (Canada and USA) countries as opposed to low-income ones (India and Indonesia). Furthermore, gender norms in the service sector, and especially for clerical jobs, is lowest in middle- and high-income countries, while they are lowest in agricultural jobs in low-income countries. For wage discrimination, we find women in the 1970s earned around 75% of their marginal product (as compared to men) in agriculture and services, and around a two-third their marginal product in manufacturing. Wage discrimination was highest in trade and service occupations (38%) and professional/managerial jobs (20%), and lowest in clerical jobs (8%). While wage discrimination has decreased over time across most occupation-sectors and countries, it continues to persist even today, even in the most developed countries. Importantly, unlike gender norms, levels in female wage discrimination are strikingly similar across countries irrespective of their economic development i.e., today's developed countries do not perform better in fair remuneration for women than today's emerging countries.

Lastly, we examine the importance of these non-economic gender barriers in explaining the growth trajectories of countries over time. To do so, we hold wage discrimination and gender norms *fixed* in the first year (in the 1970s) for each country, while allowing for all other (economic) parameters to evolve according to their calibrated values. To assess the importance of gender barriers, we then compare the simulated counterfactual path of structural transformation and economic growth for each country, to their actual empirically observed trajectory. We find that changes in gender barriers have had large effects on impacting (female) labor force participation across countries (except India) as well as in driving structural transformation. They explain around half the observed increase in manufacturing employment, and around a third in services, driven a consequent reduction in agricultural employment as well as increased participation in the labor force. Across occupations as well, a reduction in gender barriers explain around a third of the increase in employment in professional and managerial jobs, and around 60% of the increase in trade/service occupations, driven substantially by a reduction of women in clerical jobs, where they have been historically over-represented. At the aggregate, a reduction in gender barriers account for 12%, 16%, and 22% of the growth in agricultural, manufacturing, services output (measured by real value-added) on average across countries in our core sample, which translates into 20% of the growth in aggregate output (defined as an expenditure-weighted average of sectoral output). This further masks important heterogeneity across countries. Gender barriers can explain only 2% of aggregate output growth in India, 16% in Indonesia, and around 25-30% in the other countries (see Table 3).

Put together, we draw the following conclusions from our analysis: first, there are distinct differences in how men and women transition across sectors as well as occupations through the process of economic development. Despite large reductions in gender barriers over time, they continue to persist even today, and even in the most advanced economies. However, a reduction in these gender barriers has had non-trivial effects on economic development. They have robustly enabled structural transformation in these countries, and can explain on average, around 20% of the economic growth in these countries over the last five decades. Lastly, important differences across countries (such as Brazil vis-a-vis India, for instance) continue to remain, which warrants further comparative micro-economic studies in them, to examine the sources of these similarities and differences. While this exercise is beyond the scope of this paper, it is a promising direction for future research. The paper is organized as follows: Section 2 discusses our contribution to the literature, Section 3 describes the data and Section 4 presents the stylized facts. Section 5 builds the theoretical model and Section 6 describes the identification and calibration exercise. Sections 7 and 8 discuss the results from the calibration and counterfactual simulations respectively, and Section 9 offers a short conclusion.

2 Literature

Our paper contributes to several strands of the literature that study structural transformation, occupational choices, talent allocation, and the importance of gender roles in driving economic choices and outcomes. A large literature studies workers' reallocation from agriculture to manufacturing and then to services. More recently, several papers have also focused on the role of gender in driving structural change (Cuberes and Teignier, 2014, 2016; Moro, Moslehi and Tanaka, 2017; Olivetti, 2014; Ngai and Petrongolo, 2017; Rendall, 2018), or a reversal of gender gaps in education as countries grow richer (Ying et al., 2023). We contribute to this literature by documenting gender differences in employment choices and wages across many countries and over a long time period. Our analysis also incorporates occupational transitions in addition to sectoral ones, uses a structural model to decompose these gender gaps between economic and non-economic channels, and quantify the importance of the latter in explaining growth in these countries.

Our paper relates to Lee (2022), who studies the effects of gender barriers on cross-country differences in agricultural productivity and finds that low-income countries have higher frictions for women in non-agricultural employment. Setting these frictions to US levels increases labor productivity by 21.3% and GDP per-capita by 3.6%. Our analysis on the other hand, studies the effects of gender barriers on occupational and sectoral choices, the (mis-)allocation of talent, and its macroeconomic implications. In that sense, it is similar to, and extends the analysis by Hsieh, Hurst, Jones and Klenow (2019), who show that a reduction in gender barriers explains 20-40% of overall growth in the US from 1960 to 2010. Our paper also relates to Gottlieb, Doss, Gollin and Poschke (2023), who undertake a similar cross-country exercise of examining gender roles and economic development. However, they focus on how the composition of a household impacts individual employment decisions and in particular, the gender time allocation within it between home and market work.

Lastly, a large literature measures and studies women's LFP which discusses the U-shaped pattern of female LFP over countries' development process. Goldin (1994) shows this pattern for the United States, and Heath and Jayachandran (2016); Fletcher et al. (2017); Mammen and Paxson (2000); Psacharopoulos and Tzannatos (1989) show it in other countries. We expand on these empirical facts by distinguishing between home and market sectors (agriculture, manufacturing, services), and several occupations within each sector, to relate these patterns directly to countries' process of structural transformation. Our model measures gender norms and wage discrimination for each occupation-sector, which allows us to simulate counterfactuals that quantify the importance of gender on countries' process of structural transformation.

3 Data

Data Sources and Sample of Countries

Our primary data source is the The Integrated Public Use Microdata Series (IPUMS International, 2020) that harmonizes individual-level data on education and employment variables from nationally representative censuses, household and labor force surveys for many countries and over time. We use a sample of 91 countries and 273 country-years ranging from 1960-2018 to document how occupational and sectoral employment changes for men and women along countries' development spectrum. On average, we have 3 rounds of data for each country. While there is better coverage across countries in recent decades (1990 onwards), coverage across countries in older decades is non-trivial as well (See Table A.2.1).

A key requirement for our quantitative exercise (described in Section XX onwards) is the availability of high-quality data on hourly wages in each occupation-sector-gender category over a long period of time (1970-2015). We therefore focus our attention on six countries (34 country-years) for which this is available, namely: India, Indonesia, Mexico, Brazil, Canada, and the United States. Put together, this sample covers a wide range of the income spectrum and accounts for 25-30% of the world population. We also complement the IPUMS data with data on sectoral value-added which we obtain from the Economic Transformation Database (ETD). For India and Indonesia, we complement the IPUMS data with data from labor force surveys (PLFS for India and SAKERNAS for Indonesia) to extend the time coverage to the most recent years (2018). Appendix E provides an elaborate discussion on the data construction.

Classification of Sectors and Occupations

We aggregate the harmonized sector classifications into (a) Agriculture; (b) Manufacturing and (c) (Market) Services as shown in Table A.2.3 and discussed in Herrendorf, Rogerson and Valentinyi (2013) and Herrendorf and Schoellman (2018). We create a category "Home Sector" to which we attribute unemployed and inactive individuals. This classification follows a recent literature that examines the role of gender and occupation choices (Moro, Moslehi and Tanaka, 2017; Ngai and Petrongolo, 2017), and is consistent across countries (Bridgman et al., 2018). Occupations are classified with the 1 digit ISCO-88 occupation codes as reported in Table A.2.4. We aggregate the top three occupation codes (managers, professionals, and technicians) due to small sample sizes for some country years.

Our main analysis focuses on seven occupation categories: (1) Professionals, which include managers, technicians, senior officials, legislators, directors, etc.; (2) Clerks, which include secretaries, librarians, cashiers, etc.; (3) Service Workers, which include travel, housekeeping and personalcare workers, along with those in shop, sales and service jobs; (4) Skilled Agricultural Workers, which include those in subsistence and market-oriented agricultural production; (5) Crafts and Trades Workers such as builders, painters, blacksmiths, electricians, etc.; (6) Plant and Machine Operators such as those workers in mining, metal, glass, wood, etc.; and (7) Elementary Occupation Workers such as street vendors, domestic helpers, prters, manual laborers, etc. Some occupation-sector-gender categories are very sparsely populated, so that we limit the agricultural sector to two occupations: skilled agricultural workers and elementary occupations. The manufacturing and the service sector consist of six occupations as we attribute all "skilled agricultural workers" to the agricultural sectors. Home is modeled as a separate sector.

4 Empirical Facts

4.1 Sectoral and Occupational Transitions by Gender

We begin by documenting the employment transitions for men and women across sectors and occupations across countries at various stages of economic development.

We first discuss sectoral transitions i.e., how the employment shares of men and women in the four sectors (agriculture, manufacturing, services and home sector) change with a country's level of economic development (measured by log real GDP per-capita). In Figure 1(a), which shows the employment transitions for men, we see the standard pattern of structural change that has been extensively documented in the literature: (male) employment transitions away from agriculture, and towards manufacturing and services, as countries grow richer. Moreover, the share of men in the home sector is low and relatively constant across different stages of economic development. On the contrary, sectoral transition for women is very different (Figure 1(b)). At low levels of economic development, women first leave agriculture, but mostly leave the labor force all together (sorting into the home sector). At higher levels of economic development, women then re-enter the labor force to work (mostly) in the service sector. Hence, female labor force participation (FLFP) follows U-shape pattern across countries' GDP per capita, a feature that is well documented in the literature.

Next, we explore transitions in occupations, a dimension that has not received adequate attention in the context of structural transformation. Analyzing gender-specific transitions in occupations (as compared to sectors) is equally, if not more important for two key reasons. Firstly, as we will elaborate on later, a non-trivial portion of the variation in gender-specific employment transitions through the process of economic development can be attributed to changes in occupations rather than sectors. Moreover, it offers additional insights from a policy standpoint. For instance, let us examine Figure A.1.1, which plots how the ratio of female to male employment shares in managerial and clerical occupations (in manufacturing and services) change with economic development. A ratio of more (less) than 1 indicates that women are more (less) likely to work in that occupation relative to men. In poor countries therefore, we see that women are half as likely to work in both managerial and clerical occupations as compared to men. As countries develop, we see that for every 10 men working as managers, there are around 8-9 women. However, for every man working in a clerical occupation, there are around 2.5-3 women. This indicates therefore that as individuals transition out of agriculture, men and women do not advance up the job ladder (into managerial roles) equally–women primarily take on clerical positions, while men secure the managerial ones. In line with this examination, Figures 1(c) and 1(d) report the gender-specific employment shares across occupations at various stages of economic development. The employment share (for both men and women) in agricultural jobs declines with economic development (similar to the agriculture sector discussed above). Consequently, employment shares in other occupations increase. However, men tend to be over-represented in trade and service jobs (like builders, painters, etc.), and machine-operators (mostly in manufacturing), while women tend to be over-represented in clerical jobs as well as in the home sector.

Figure 2 underscores the importance of examining occupational transitions by demonstrating differential gender-specific transitions *within sectors* as well. Specifically, *conditional* on working in the manufacturing sector (Figures 2(a) and 2(b)), 75-80% of men and women work as trade and service workers in poor countries. While the share of workers in these occupations decline with economic development, around half of the male manufacturing employment in rich countries continues to be engaged in trade and service jobs, as compared to only around 20% for women. Consequently, employment share in machine operators, professional, and elementary occupations increase for men, whereas women are over-represented in clerical occupations, along with elementary and professional occupations as well. As compared to manufacturing sector, occupations are less concentrated in the service sector for poor countries (Figures 2(c) and 2(d)). Approximately 50% of men and women work in trade and service workers, 20% in professionals, 20% in elementary occupations, and 10% clerks. At higher income levels, the share of service and elementary workers decrease while clerks and professionals increase. Within the service sector, women are once again heavily over-represented in clerical occupations, while men are over-represented in machine-operator, elementary workers, and professional/managerial occupations.

4.2 Gender Segregation Across and Within Sectors

Segregation indices offer an additional way of presenting gender segregation across occupations and sectors. One example is the Theil Index. The Theil Index measures an entropic "distance" that the population is away from the "ideal" egalitarian state, which equals an index value of 0. Higher numbers for the index indicate more segregation in the population of interest (see Appendix C.1 for mathematical details). Figure 3(a) plots the global Theil Index across all 91 countries in our sample against GDP per-capita. We see an inverted U-shape, where gender segregation across occupation-sectors first increases and then decreases with economic development. This is perhaps unsurprising, given that most individuals (irrespective of gender) work in agriculture in poor countries. The transition of women out of the labor force and men into manufacturing and services worsens segregation for middle-income countries, which declines for high-income countries as women re-enter the labor force in services. A key advantage of the Theil index is that it further be decomposed into the share of segregation that is due to across-sector variation as compared to the part that is due to within-sector (across-occupation) variation. Figure 3(b) plots the share of gender segregation that is explained by segregation across occupations within each sector. The within-sector component explains around 30-40% of the segregation in low- and middle-income countries, which increases to around 60% for high-income ones.[2] This resonates strongly with Bandiera et al. (2022), who examine how the organization of labor changes across jobs and economic development in over a 100 countries. Put together, the above discussion underscores the importance of examining the role of gender in occupational choices, along with sectoral ones.

4.3 Gender Gaps in Employment and Wages

The gender patterns in employment transitions discussed in the previous section provide important insights into how the process of structural transformation differs for men and women. However, two key questions remain: can these patterns be explained by differences across countries? or have patterns *within* the same country changed over time (and economic development)? Moreover, similar to employment transitions, how do gender wage gaps (in sectors and occupations) vary with economic development? We now turn our attention to examining these questions. Specifically, a key challenge we encounter is that wage data at the sector-occupation-gender-country-year level is not available across most countries in our previous sample. Therefore, we restrict our sample to six countries that will form our core sample, namely: India, Indonesia, Brazil, Mexico, Canada and the United States.

[2] In this analysis, changes in labor force participation are attributed to across-sector variation since we attribute individuals who are not in the labor force to the home sector (which has no occupations). To abstract from the extensive margin, we re-compute the Theil index conditional on labor force participation. For this, we find the same inverted U-shape for the global Theil index and the within-sector variation now becomes even more important, explaining between 40-70% of the overall gender segregation.

For these countries, we compute an employment gap, which is the probability that a woman (relative to a man) works in an occupation-sector (within a country-year).Similarly, we compute a wage gap, which is the ratio of the average hourly wage in each occupation-sector (within a country-year) earned by a woman as compared to a man. Note that in both cases, a ratio of 1 implies gender parity (in employment or wages), while a ratio less (more) than 1 indicates that men (women) work/earn more in that occupation-sector than women (men). Figure 4 then plots the employment and wage ratios against GDP per capita for these six countries. For ease of interpretation, we only report (and connect) the ratios in the first and last year for each country. This allows us to analyze differences across countries and changes within-country over on average over four decades between 1970-75 until 2010-18 (see Section 3 for details).

Gender Employment Gaps

Figure A.3.1(a) in the Appendix provides two insights on how the distribution of these employment gaps has moved between these decades (1970s and 2010s), after pooling all sector-occupation-countries. First, women were less likely to work than men in over 90% of occupations in the 1970s i.e., employment ratio is less than 1. Second, despite significant improvements (a rightward shift) of the distribution over time (Kolmogorov-Smirnov p-value=0.00), the probability of a woman working in an industry-occupation is still less than a man in around 80% of occupations.[3] As is intuitive, this is because women are non-trivially more likely to work at home than men. Table 1 summarizes this information across sectors and occupations over time. Specifically, we see that across these six countries, women were almost 8 times more likely to stay at home as compared to men in the 1970s. Correspondingly, the employment gaps were largest in agriculture and manufacturing (around 3 women per 10 men), followed by services (6.3 women per 10 men). Women were also under-represented in managerial, trade and service occupations (around 4 women per 10 men), while being over-represented in clerical jobs (12 women per 10 men). Over the four decades, employment gaps have reduced substantially, particularly with more women entering the labor force. There are around 9 women per 10 women working as managers, trade and service workers, especially in the service sector. However, employment gaps in agriculture, and manufacturing continue to remain. There is some variation across countries as well, as reported in Figures 4(a) and 4(b).[4] In high-income countries like Canada and the US–where FLFP is high–the service sector employs more women than men (ratio > 1). Meanwhile, India remains a stark exception, with low female labor force participation rates, and large employment gaps in the manufacturing and service sectors (only 3 women per 10 men) that have not improved over time.

Figure 4(b) shows the gender employment ratio across occupations. In all countries (except India), the female-to-male employment ratios increase over time in almost all occupations. A main driver is the increase in FLFP that is particularly strong during our sample period for Indonesia, Mexico, and Brazil. However, large gender differences remain across occupations (in levels and trends). In the US, Canada, Mexico and Brazil, clerks and professionals have the highest female-to-male employment ratios which increase over time. In the US, Canada, Mexico, and Brazil more women than men work as clerks (in the most recent period).[5] For Canada and the US this also applies to clerks and professionals. Indonesia and India start

[3] The lower panel of Figure A.3.2 shows that these patterns are not specific to the choice of our sample period as results are similar when we use data for all countries for 1980 to 2010.

[4] The graph excludes the home sector, where the share of women is multiple times that of men.

[5] For readability, we omit the employment ratio for clerks for the US and Canada as these countries have up to three times more female than male clerks.

with low female-to-male employment ratios in most occupations due to very low FLFP. India shows almost no improvement over time, while Indonesia shows a sharp increase in female employment–particularly among professionals and clerks.

Gender Wage Gaps

Similar to employment gaps, Figure A.3.1(b) shows that women earned less than men in over 95% of sector-occupations in the 1970s, with some where they earned only around 50c to $1 earned by a man. While these have improved over time (Kolmogorov-Smirnov p-value=0.00), especially those sector-occupations with large wage gaps, women continue to earn less than men in over 80% of sector-occupations. From Table 1, women earned on average around 60-70c for every $1 earned by a man in the same sector or occupation in the 1970s. Over time, these gaps have closed significantly. Nevertheless, women continue to earn on average around 75-80c for every $1 earned by a man in the same sector or occupation even today. Looking across countries (Figures 4(c) and 4(d)), gender wage gaps have improved across almost all sectors and occupations across countries.[6] Most notably though, the levels of these wage gaps are strikingly similar across the poorest and richest countries and there is no clear correlation with employment gaps in these sectors or occupations. For example, consider the wage gap in professional/managerial occupations in the most recent sample year: for each $1 earned by a man, a woman earned roughly between 70-80c across all countries, irrespective of their level of economic development. Meanwhile in many of these countries, professional occupations have the lowest employment gaps (as discussed earlier).

4.4 Key Takeaways and Model Implications

Our empirical patterns provide three key takeaways.

First, employment transitions across sectors and occupations along countries' development process have salient gender patterns. For sectors, men follow the standard patterns of structural transformation. Women follow a different pattern. When they exit agriculture they first sort into the home sector and then enter into the service sector at higher income levels. Female employment in manufacturing remains small and relatively constant across income levels. Across occupations, men are more likely to work in trade and service jobs, or as machine operators, or in managerial and professional occupations. On the other hand, clerical occupations grow fast and are very predominantly female-dominated.

Second, the global and within-sector Theil Index of segregation underscores the importance of taking into account the occupational transitions in the process of structural transformation as well. In particular, the index shows that variation across sectors and across occupations within sectors are both important in explaining the overall gender segregation in the economy. In particular, around 30-40% of gender segregation in low- and middle-income countries, and 60% in rich countries can be explained by segregation across occupations within a sector. This is an important dimension that we incorporate in our model.

[6] Wage ratios for clerks and professionals in India and Indonesia are an exception as they start with relatively large female-to-male wage ratios at the beginning of the sample period and then remain constant. Other countries then catch-up to these higher levels and the wage ratios of clerks and professionals converge to similar levels across all countries.

Third, gender gaps in employment and wages vary substantially across sectors and occupations. Employment gaps tend to be smaller in rich countries but wage gaps are strikingly similar in poor and rich countries. Both gaps improve over time for most countries, but sizable gender inequality still persists in the most recent period in all countries. A good gender balance in employment can co-exist with large gender pay gaps, which is most salient perhaps in professional and managerial occupations. In theory, this could driven by wage discrimination, or by sorting of workers, or both. Disentangling these channels is key to understanding the drivers behind these patterns–a feature that we will build into our theoretical framework.

The above takeaways point to the fact that several economic and ”non-economic” mechanisms can affect employment choices and wage gaps between men and women during the process of economic development. Economic drivers include technological change, which for example, can be biased towards specific sectors or occupations where women (or men) have a comparative advantage. In addition, workers' effective human capital in sectors or occupations can change due to higher education, or due to skill-biased technological change. Alternately, as documented by as recent literature (Alder, Boppart and Müller, 2022; Herrendorf, Rogerson and Valentinyi, 2013; Comin, Lashkari and Mestieri, 2021), income effects (as countries grow richer) can change consumers' consumption baskets, by shifting expenditure away from agriculture and towards services (where women might have a comparative advantage). The occupational mix of workers in sectoral production might also change with economic development. On the other hand, there could also be ”non-economic” drivers such as gender-specific amenities/barriers/wedges that can make it more or less attractive for women (relative to men) to work in certain occupations and/or sectors. In particular, we consider two such barriers that vary across occupation-sector-pairs: "wage discrimination" that allows employers to pay women less than men, conditional on their level of human capital, and "gender norms", which are gender-specific preferences or gender-specific amenities/barriers to working in certain occupation-sectors (or in the labor force all together).

To decompose these empirical patterns and quantify the importance of ”non-economic” channels, we develop (and estimate) a Roy model of occupational and sectoral choices, which incorporates each of the above channels. In particular, we allow gender barriers, productivity/skills, and returns to human capital to differ across occupations, sectors, countries, and over time. Gender barriers distort workers' occupational and sectoral sorting based on their comparative advantage which generates a misallocation of talent and lowers productivity and growth. We flexibly estimate the model to fit several key data moments separately for each country-year. In the spirit of an accounting exercise, we identify gender barriers for each occupation-sector as residual wedges that match the observed gaps in employment and wages, *after* accounting for the all the economic factors. Lastly, we then evaluate counterfactuals that quantify the importance of changes in these gender barriers on explaining a country's employment transitions, aggregate output growth, and welfare.

5 Model

We now describe the model set up, solve for individuals' employment choices and firms' production decisions, and define the equilibrium.

5.1 Setup and Preferences

Model Setup: The economy consists of O occupations and J sectors, namely agriculture (A), manufacturing (M), market services (S^m), and home services (S^h). There is a mass N_g of individuals of gender g, either male (m) or female (f). Each individual of gender g has an ability z and chooses to work in an occupation-sector-pair oj, with the "home sector" being one possible choice.

Non-Homothetic Preferences: Individuals have non-homothetic preferences in the PIGL class over agriculture, manufacturing and services $j = \{A, M, S\}$, where services are a composite of home and market services. Non-homothetic preferences allow for consumers' sectoral expenditure shares change with income. This is in line with a recent literature that has documened this channel as an important driver of structural transformation and hence labor allocation within the economy.[7] We can represent PIGL preferences with an indirect utility function:

$$V(I_{ojgz}, p) = \frac{1}{\eta}\left[\frac{I_{ojgz}}{P}\right]^{\eta} - D(p) \tag{1}$$

where $I_{ojg}(z)$ is the income earned by an individual of gender g and ability z who works in an occupation-sector pair oj and p is a vector of sectoral prices. P and $D(p)$ are homogeneous of degree zero and one respectively and are parameterized as:

$$P = \prod_{j=\{A,M,S\}} p_j^{\omega_j} \quad \text{and} \quad D(p) = \sum_{j=\{A,M,S\}} \nu_j \ln p_j$$

$$\text{where:} \sum_j \omega_j = 1 \quad \text{and} \quad \sum_j \nu_j = 0$$

Roy's Identity implies that individuals' sectoral expenditure shares are given by:

$$\varphi_j(I_{ojg}(z), p) = \omega_j + \nu_j \left[\frac{I_{ojgz}}{P}\right]^{-\eta} \tag{2}$$

See Appendix C.2 for the proof. Equation (2) highlights how income and prices affect sectoral expenditure shares. ω_j denotes the sectoral expenditure shares under homothetic preferences i.e., when $\nu_j = 0$, or in the limit when real income approaches infinity. ν_j therefore governs the direction of the income effect: an expenditure share increases with income if $\nu_j < 0$ (in services, for example) and decreases with income if $\nu_j > 0$ (in agriculture).

Home and Market Services: We assume a preference structure similar to Ngai and Petrongolo (2017), where services are a CES composite of home and market services $S = \{S^h, S^m\}$ given by:

$$C_S = \left[\sum_{s' \in \{S^h, S^m\}} \alpha_{s'} C_{s'}^{\frac{\eta_s - 1}{\eta_s}}\right]^{\frac{\eta_s}{\eta_s - 1}}$$

[7] See Herrendorf, Rogerson and Valentinyi (2013); Alder, Boppart and Müller (2022); Comin, Lashkari and Mestieri (2021); Fan, Peters and Zilibotti (2021).

where η_s is the elasticity of substitution between home and market services and α_s are the preference weights across home and market services, with $\sum_{s'} \alpha_{s'} = 1$. Expenditure shares for home and market services are therefore given by:

$$\varphi_{s'} = \alpha_{s'} \left[\frac{p_{s'}}{P_S} \right]^{(1-\eta_s)} \times \varphi_S, \tag{3}$$

where φ_S is the expenditure share on services and $P_S = \left[\sum_{s'} \alpha_{s'} p_{s'}^{1-\eta_s} \right]^{\frac{1}{1-\eta_s}}$ is the CES price index. The proof is provided in Appendix C.3.

5.2 Income and Occupational Choice

We now define and solve workers' occupational and sectoral choice problem. Utility of a worker i of gender g and ability z who works in an occupation-sector pair oj is given by:

$$U^i_{ojg}(z) = V(I^i_{ojg}(z), p) - A_{ojg} + \varepsilon^i_{oj}$$

where $V(I^i_{ojg}(z), p)$ is workers' indirect utility as defined in Equation (1). A_{ojg} are gender-specific utility costs of working in an occupation-sector pair which capture a wide range of factors including amenities, preferences, norms or entry costs that can vary across occupation-sector pairs oj and across genders g. What we will refer to as "gender norms" for the rest of the paper are then the excess utility costs faced by women, relative to men, when working in an occupation-sector oj i.e., $\triangle A_{oj} = A_{ojf} - A_{ojm}$. Note that in the calibration exercise, these differences will be estimated flexibly. In particular, we will not impose any restriction on the direction of $\triangle A_{oj}$ i.e., in principle, they could be positive (women face excess costs), negative (women face lower costs), or zero (men and women have equal costs). We will also estimate these separately for each occupation-sector-country-year and therefore impose no dynamic or country-specific constraints, based on their level of economic development. Lastly, ε^i_{oj} are idiosyncratic preference shocks for working in each occupation-sector pair.

Assumption 1: *Preference shocks ε^i_{oj} are Extreme Value Gumbel distributed across occupation-sector pairs with a dispersion parameter σ_ε.*

With this assumption, we can express the probability of a worker (and hence the share of workers) of gender g and ability z who chooses to work in occupation-sector pair oj as:

$$\Pr(oj|g,z) = \frac{\exp\left[\frac{1}{\sigma_\varepsilon} V(I_{ojg}(z), p) - \frac{1}{\sigma_\varepsilon} A_{ojg} \right]}{\sum_{j'} \sum_{o'} \exp\left[\frac{1}{\sigma_\varepsilon} V(I_{o'j'g}(z), p) - \frac{1}{\sigma_\varepsilon} A_{o'j'g} \right]} \tag{4}$$

From the above equation, sorting of workers into occupation-sectors is therefore based on the real income earned in those occupation-sector, and the gender-specific utility cost (A_{ojg}) incurred for working in them.

Occupation-sector pairs differ in their returns to ability (κ_{oj}), so that $z^{\kappa_{oj}}$ are the effective units of human capital that a worker of gender g and ability z can supply to occupation-sector pair oj. For each unit of human capital, an occupation-sector pays a wage rate w_{ojg}, which can vary for men and women. We model female wage discrimination as a "wedge" τ_{oj} between women's wage rate and their marginal product so that: $\{w_{ojm}, w_{ojf}\} = \{w_{oj}, (1-\tau_{oj})w_{oj}\}$. This assumes that men receive the effective wage rate w_{oj}, while women are paid a fraction $(1-\tau_{oj})$ of their marginal product.[8] Like with gender norms ($\triangle A_{oj}$) above, we do not assume any direction (positive, negative, or zero) for τ_{oj} in the calibration, or impose any restrictions either across countries, or within countries over time. The total income of an individual of gender g and ability z working in an occupation-sector oj is therefore given by: $I_{ojg} = w_{ojg} \times z^{\kappa_{oj}}$.

5.3 Production, Aggregation and Equilibrium

Aggregate Supply of Human Capital: The aggregate supply of human capital in each occupation-sector oj is then a summation across individuals' employment choices and can be given by:

$$H_{oj} = \sum_g N_g \int_z \Pr(oj|g,z) z^{\kappa_{oj}} dF(z), \tag{5}$$

Production: A representative firm in each sector produces output Y_j with a Cobb Douglas production function, using as input the human capital from each occupation, so that:

$$Y_j = T_j \prod_o H_{oj}^{\gamma_{oj}} \tag{6}$$

γ_{oj} are therefore the expenditure shares across occupations within each sector (with $\sum_o \gamma_{oj} = 1$). T_j is sector-specific productivity, and H_{oj} is the total human capital that is supplied to oj in equilibrium (Equation (5)). Firms' profits are therefore given by:

$$\pi_j = p_j Y_j - \sum_o w_{oj} H_{oj}$$

where p_j are sectoral prices.

Aggregate Expenditure Shares: PIGL preferences allow us to aggregate sectoral expenditure shares across individuals in a tractable, closed-form way, despite being non-homothetic in nature. Aggregate expenditure shares in sector j, denoted by Φ_j, are given by:

$$\Phi_j = \omega_j + \nu_j \times \frac{P^\eta}{I} \times \sum_o \sum_j \sum_g \int_z I_{ojg}^{1-\eta} \times \Pr(g) \times \Pr(oj|g,z) dz, \tag{7}$$

See proof in Appendix C.4.

Equilibrium: Exogenous parameters of the model characterize preferences $\{\alpha_s, \eta_s, \omega_j, \nu_j, \eta\}$, the dispersion of preference shocks across occupation-sectors (σ_ε), the ability distribution $z \sim F(z)$, firm production

[8] This wedge can be micro-founded for example, from Hsieh, Hurst, Jones and Klenow (2019), where entrepreneurs have a disutility δ_{oj} of hiring women, which is compensated with the profits that arise from paying women below their marginal product.

$\{\gamma_{oj}, T_j, \kappa_{oj}\}$, and gender barriers $\{\tau_{oj}, A_{ojg}\}$. Given these parameters, the equilibrium in each country-year is defined by a vector of sectoral prices and occupation-sector-specific wage rates $\left\{\{p_j\}_{\forall j}, \{w_{oj}\}_{\forall oj}\right\}$ so that:

1. Workers make optimal consumption and employment choices.
2. Firms in each sector hire human capital from each occupation to maximize profits.
3. Labor markets clear in each occupation-sector pair equalizing human capital supply and firms' demand for it.
4. Good markets clear in each sector.

6 Model Calibration

We calibrate a set of parameters to the literature or to data moments outside of our model. The remaining parameters are then calibrated by fitting our model's equilibrium conditions to the data in an iterative algorithm which is described in Appendix D.

6.1 Parameters Calibrated Outside of the Model

Preferences: Preference parameters do not vary across countries or over time. For the CES preferences over home and market services, we follow Ngai and Petrongolo (2017) and set the elasticity of substitution $\eta_s = 2.3$. For the preference shocks across occupation-sector-pairs, we follow Hsieh et al. (2019) and set the dispersion parameter to $\sigma_\varepsilon = 2$. For the PIGL preferences, we follow Fan, Peters and Zilibotti (2021) and set $\eta = 0.4$. The remaining PIGL parameters $\{\omega_j, \nu_j\}$ are estimated by indirect inference.

Ability Distribution and Returns to Human Capital: Ability distributions vary across countries and over time and follow a log-normal distribution, so that $z \sim \ln N(\mu, \sigma_z)$. We calibrate $\{\mu, \sigma_z\}$ for each country-year from the moments of the observed schooling distribution. Note that the schooling distribution is exogenous in our model. This implies that it does not change in response to policy experiments that we examine in counterfactual simulations (such as eliminating gender barriers, for example). Gender differences in the endogenous acquisition of schooling/skills in response to different policy environments is definitely an important exercise, but we leave it for future work. Turning to the returns to human capital, occupation-sector specific returns (κ_{oj}) determine the income of an individual of gender g and ability z who works in that occupation-sector in the following way: $I_{ojg}(z) = w_{ojg} \times z^{\kappa_{oj}}$. This provides us with a structural equation that we can take to the data to estimate κ_{oj}, similar to Fan, Peters and Zilibotti (2021). Specifically, we use individual-level data for each country-year to estimate the following Mincerian wage regression:

$$\ln I_i = \alpha_{oj} + \kappa_{oj} \ln(\text{YrsSchool}_i) + \varepsilon_i, \tag{8}$$

where I_i is income/earnings and YrsSchool_i are years of schooling of an individual i. α_{oj} are occupation-sector fixed effects that capture average wages and other unobserved factor that can affect individuals' wages in each occupation-sector. The coefficient $\hat{\kappa}_{oj}$ is estimated for each occupation-sector-pair and corresponds through the lens of our model to the occupation-sector-specific returns to human capital. We estimate the above equation separately for each country-year.[9] We set $\kappa_{home} = 0$.

Occupational Expenditure Shares: For every country-year and each sector, we fix the occupational expenditure shares γ_{oj} to the corresponding wage expenditure shares that we observe in the data.

6.2 Parameters Calibrated using Model's Equilibrium Conditions

We estimate the remaining parameters by fitting our model's equilibrium conditions to key data moments. These are the PIGL preferences $\mathcal{U} = \{\omega_j, \nu_j\}_{\forall j}$, sectoral productivity $\mathcal{T} = \{T_j\}_{\forall j}$,, and gender barriers $\mathcal{B} = \left\{ \{\tau_{oj}\}_{\forall oj}, \{A_{ojg}\}_{\forall ojg} \right\}$, which consist of wage discrimination and gender norms. While the preference parameters ($\mathcal{U}$) do not vary across country-years, sectoral technology and gender barriers $\{\mathcal{T}, \mathcal{B}\}$ are estimated separately for each country-year. We provide the intuition of the estimation strategy here, while we leaving the detailed numerical procedure and iterative algorithm in Appendix D.

Wage Discrimination: We estimate wage discrimination τ_{oj} in each occupation-sector (and country-year) by matching the ratio of gender-specific average wages in the data. Note from in our model, we can write the observed wage gap (ratio of female to male wages) in the data as a function of female wage discrimination (τ_{oj}) and the (model-implied) gender gap in average human capital:

$$\underbrace{\frac{\overline{\text{wage}}_{ojf}}{\overline{\text{wage}}_{ojm}}}_{\text{Obs. Wage Gap}} = \underbrace{(1-\tau_{oj})}_{\text{Wage Discr}} \times \underbrace{\frac{\overline{H}_{ojf}}{\overline{H}_{ojm}}}_{\text{Avg. HC Gap}} . \tag{9}$$

where $\overline{H}_{ojg}$ is the average human capital that men or women supply to an occupation-sector oj, which is defined as: $\overline{H}_{ojg} = \int_z \Pr(oj|g,z) z^{\kappa_{oj}} dF(z)$. This measure of average human capital is endogenously generated in the model, and depends on the extent to which workers sort into an occupation-sector (due to other factors like gender norms, returns to ability, etc.) in equilibrium. As noted before, we have to impose that τ_{oj} for men is zero. However, we impose no such conditions for women.

Gender Norms: Individuals choose an occupation-sector pair oj based on real income and based on a gender-specific utility cost, which captures amenities and gender norms. From Equation (4), note that the probability that an individual sorts into an occupation-sector depends on:

$$\Pr(oj|g) \propto \left[\underbrace{V(I,p)}_{\text{Real Income}} - \underbrace{A_{ojg}}_{\text{Utility Cost}} \right]$$

[9]For our baseline model, we take κ_{oj} to be the same for both genders since these are the *fundamental* returns to schooling for an individual. In an alternate exercise, we estimate κ_{oj} separately for men and women (within each country-year). As reported in Figure A.1.5, there is little variation across gender. Our results are therefore robust to using gender-specific estimates of κ_{oj}.

where $V(I,p)$ is workers' indirect utility that further depends on her real income (Equation 1), and A_{ojg} represents the gender-specific utility cost of working in oj. For each gender therefore, we can only identify the utility costs A_{ojg} relative to an occupation-sector. We therefore normalize $A_{home,g} = 0$, which implies that individuals do not have any disutility from staying at home. A_{ojg} for other occupation-sectors then capture the utility cost of working in them relative to staying at home. To calibrate A_{ojg}, we use our model to compute worker's indirect utility $V(I,p)$ in each occupation-sector, which is a model-implied measure determined in equilibrium. We then use this measure, along with data on gender-specific employment shares $\Pr(oj|g)$ to infer A_{ojg} for each occupation-sector according to Equation (4).

Sectoral Productivity, Prices and Wages: We infer sectoral productivity T_j for each country-year to target the real sectoral value added growth in each sector (which we obtain from the Economic Transformation Database). Since we observe the changes and not levels of sectoral real value added for a country, we can only identify changes in, and not levels of sectoral productivity as well. Therefore, we normalize sectoral productivity in each sector and country to 1 in the first year.[10] Lastly, sectoral prices p_j, and occupation-sector wage rates w_{oj} are solved in equilibrium to clear the goods and labor market in each sector.

PIGL Parameters: The PIGL parameters $\{\omega_j, \nu_j\}$ are identified by indirect inference from Equation (7). For each sector, we regress the observed expenditure shares from the data on model-implied real income across the country-years of our sample. The constant of this regression identifies the PIGL expenditure shares ω_j, which are $\{0.7\%, 16.14\%, 83.16\%\}$ in agriculture, manufacturing, and services respectively. The coefficient on the real income measure identifies ν_j, which we estimate to $\{0.086, 0.125, -0.211\}$ in these sectors respectively.

7 Gender Barriers and Model Validation

7.1 Gender Barriers

Our estimates of wage discrimination (τ_{oj}) and excess gender norms ($\triangle A_{oj}$) for each occupation-sector-country-year quantify the part of gender differences in wages and employment that cannot be explained by differences in human capital or skill sorting. Figures A.3.2(a) and A.3.2(b) in the Appendix pool all sector-occupation-countries and report how the distribution of gender norms and wage discrimination respectively have changed between the 1970s and 2010s. Figure 5 then reports the evolution of these gender barriers across each country separately.

[10]Unlike models with homothetic preferences, this normalization impacts our calibrated estimates of $\triangle A_{oj}$ and τ_{oj}, as well as the extent to which gender barriers can explain sectoral output growth within a country over time. In Figures A.1.2, A.1.3, and A.1.4 in the Appendix, we re-estimate the entire model for alternate normalizations of $T_j = \{0.50, 1.00, \ldots, 5.00\}$ and show that the differences in our estimates for these barriers, both across countries and over time (Appendix Figures A.1.2 and A.1.3), as well as how much these barriers matter in the aggregate (Appendix Figure A.1.4) are quantitatively small and do not impact our interpretations.

Gender Norms: From Appendix Figure A.3.2(a), we see that women faced excess utility costs of participating in all occupation-sectors (relative to men) in the 1970s i.e., $\triangle A_{oj} > 0$. While these barriers have reduced across all occupation-sectors over time (a leftward shift of the distribution, Kolmogorov-Smirnov p-value=0.00), women continue to face excess barriers in them. Figures 5(a) and 5(b) then examine changes in these norms across sectors and occupations respectively for each country. The decline has been lowest in the low-income countries of India and Indonesia, while being larger in the middle-income (Brazil and Mexico), and high-income (USA and Canada) countries in our sample. Mirroring the empirical patterns on employment shares discussed previously (Figure 4), gender norms in the service sector, and especially for clerical jobs, is lowest in middle- and high-income countries. On the contrary, these norms are lowest for agriculture in low-income countries. Lastly, India, with its low female labor participation rates exhibits little changes in gender norms over these four decades.

Wage Discrimination: Appendix Figure A.3.2(b) reports the distribution of τ_{oj} across occupation-sector-countries in 1970s and 2010s. Mirroring the gender wage gaps (Appendix Figure A.3.1(b)), we see that women face wage discrimination in almost all occupation-sectors in the 1970s, which have substantially improved over time (Kolmogorov-Smirnov p-value=0.00), especially in those ones where they faced highest discrimination.[11] Nevertheless, women continue to face wage discrimination in most occupation-sector-countries even today. Table 2 reports the average τ_{oj} across countries in the 1970s (Column 3) and 2010s (Column 4). We see that on average, the wage penalty faced by women (after accounting for human capital differences) was around 25% in agriculture and services, and around a third in manufacturing in the 1970s. Furthermore, it was highest in trade and service occupations (38%) and professional/managerial jobs (20%), and lowest in clerical jobs (8%). To put it another way, women were paid 60%, 80%, and 92% of their marginal product in trade and service, managerial, and clerical jobs respectively. These penalties have reduced significantly over time, particularly in agriculture and manufacturing sectors, and trade and service occupations. However, they appear to have worsened in professional and managerial jobs (25%). Lastly, Figures 5(c) and 5(d) report the results across each country separately over time. While we see declines in wage discrimination across all countries, sectors, and occupations, of special note is their worsening in service sector jobs (managerial and clerical jobs in particular) in India and Indonesia. However, unlike gender norms, levels in female wage discrimination is strikingly similar across all countries in our sample. We see that there is no discernable correlation in wage discrimination and economic development. Today's developed countries do not perform better in fair remuneration than today's emerging countries.

7.2 Model Fit and Validation

Model Fit

Figure B1 pools the data across all occupations, sectors, countries, and years and shows a strong negative correlation between the gender norms ($\triangle A_{ojct}$) and the observed gender employment gaps in the data. Figure B2 documents the same for the correlation between wage discrimination (τ_{ojct}) and gender wage gaps. Figure B3 shows that our gender barriers also closely track the gender wage and employment gaps

[11]The lower panel of Figure A.3.2 shows that these patterns are not specific to the choice of our sample period as results are similar when we use data for all countries for 1980 to 2010.

for each country over time. As an additional model validation, Figure B4 shows that our model replicates the share of value-added across sectors as well.

Validating the Model Estimates with Measures of Social Norms

While our estimates for gender norms and wage discrimination are accounting residuals in our model, we now examine whether they capture measurable changes in women's underlying social norms and labor market constraints. We take advantage of the World Bank's "World, Business, and the Law" (WBL) database (World Bank, 2019, 2020; Hyland, Simeon and Goldberg, 2020). The WBL data set evaluates 35 aspects of countries' legal code to create 8 indicators, which measure gender equality in the labor market, at the workplace, and in the legal code across 190 countries and over five decades (1970-2020). The indicators include answers to gender normative questions such as "Can a woman get a job in the same way as a man?" or "Can a woman work at night in the same way as a man"?. We regress our estimated values of gender barriers (both norms and discrimination) on the WBL indicators in these countries as follows:

$$y_{ojct} = \alpha_c + \alpha_t + \text{WBL}_{ct} + \ln \text{GDP p.c.}_{ct} + \varepsilon_{ct}, \tag{10}$$

where the dependent variable y_{ojct} is either female wage discrimination τ_{ojct} or gender norms $\triangle A_{ojct}$ and where WBL_{ct} is a specific indicator from the WBL data for a country c in year t. We control for real GDP per-capita and for country and year fixed effects to control for unobserved time-invariant country-specific social norms, or other unobserved (economic or social) changes over time. We cluster standard errors at the country-level.

Panel A of Table B1 shows a strong negative relationship between our estimated gender norms ($\triangle A_{ojct}$) and 5 WBL indicators which measure gender equality in mobility and at the workplace (e.g., equality in getting a job, working at night, working in industrial and 'dangerous' jobs). Panel B of Table B1 examines the correlation between our estimated gender norms and 5 measures of gender equality in the household (e.g., equality as household head, rights to remarry and ownership, and legislations against domestic violence). We find a negative correlation in 4 out of the 5 indicators which are statistically significant for 3 out of the 5 indicators. Panel A of Table B2 shows a negative correlation between estimated wage discrimination (τ_{ojct}) and the 5 WBL indicators on gender equality in mobility and at the workplace (as described above). In Panel B, we regress our estimated wage discrimination on WBL indicators that measure gender equality at the workplace including whether it offers paid maternity leave. We find a negative and significant correlation only for the provision of paid maternity leave.

Put together, the results show that the gender barriers (estimated as residuals in our model) contain important information about and correlate with changes in underlying social and gender norms that are measured in the data across countries and over time.

8 Importance of Gender Barriers in Explaining Economic Growth

Given our estimates of gender barriers, we can now answer the following question: for each country, how much of structural transformation and sectoral output growth over the last four decades be attributed to changes in gender barriers? In our model, the gendered employment transitions from agriculture and homework to manufacturing and services can be driven by multiple channels: changes in gender barriers (wage discrimination or gender norms: τ_{oj}, A_{ojg}), changes in production technologies (occupation-, sector-, or skill-biased technological change: γ_{oj}, T_j, κ_{oj}), changes in human capital (μ_{zg}, σ_{zg}), or by income effects which affect consumers' expenditure shares (ω_j, ν_j).

To assess how much changes in gender barriers contribute to each country's process of structural transformation and sectoral output growth, we simulate counterfactuals that hold these gender barriers constant at their values in the first year of each country, while allowing other parameters to evolve according to their calibrated values from the data. First, we fix only female wage discrimination (τ_{oj}), second only gender norms ($\triangle A_{oj}$), and third, gender norms and female wage discrimination simultaneously. For each counterfactual, we solve for workers' employment choices, wages, and sectoral prices which are consistent with the general equilibrium of the model. To quantify the importance of changes in gender barriers, we then compare the counterfactual path to the empirical path in the data. Specifically, let g is the annualized growth rate for some variable of interest (employment share, sectoral output, etc.) observed in the data and $\hat{g}$ be the growth rate in a counterfactual where gender barriers are fixed to their initial values. Then $1 - \hat{g}/g$ is the fraction of growth that can be attributed to changes in gender barriers over time.

Impact of Changing Gender Barriers In Explaining Employment Transitions Across Countries

We first examine the importance of changes in gender barriers on sectoral and occupational employment shares. In Figure 6, we take the average across the six countries and show how sectoral and occupational employment shares change in the baseline and in each counterfactual over four decades (1970-2015). For example, in the baseline data (black bars), labor force participation increased on average by 0.16 p.p. each year. For our median sample period of approximately 40 years, this corresponds to a total change of 6.6 p.p. A counterfactual simulation that fixes wage discrimination (τ_{ojct}) to the values in 1970 would have generated a 0.14 p.p. annualized increase in the LFPR which implies that the reduction in wage discrimination can explain around 12.5% (1-0.14/0.16) of the observed increase in LFP. While this magnitude is non-trivial, changes in gender norms had much larger effects. If gender norms had not changed since their initial values in the 1970s, LFP would have actually decreased by 0.20 p.p. each year or by 0.23 p.p. if neither wage discrimination nor gender norms had changed. Changes in gender norms also had large effects on sectoral employment. Employment share in agriculture decreased at 2.23 p.p. each year on average across countries in our sample. Without changes in wage discrimination and gender norms, the employment share in agriculture would have decreased even faster–around 2.4 p.p. a year. These changes in agriculture and LFPR would have consequently slowed employment growth in manufacturing and services. For example, the employment share in services would have grown by only 0.82 p.p. instead of 1.20 p.p. per year, implying that changes in gender norms explain around a third, or 31.5% (1-0.82/1.20) of the observed growth in service employment. Similarly, changes in gender norms explain around half of the observed growth in the employment share in manufacturing. Figure 6(b) shows the effects for employment by occupation. Changes in gender barriers explain around a third of the changes in professional and managerial occupations, and machine and elemenatary occupations, and around 60% in trade/service occupations. Employment in clerical occupations would have actually decreased by 0.13 p.p. per year in the absence of any changes to gender barriers over time. Most of these effects are again driven by changes in gender norms.

Impact of Changing Gender Barriers In Explaining Growth in Real Sectoral Output Across Countries

We now turn to calculating how much of the growth in real sectoral output (measured by real value added) between 1970s and 2010s can be explained by changes in gender barriers (both norms and discrimination). We report these results in Table 3. Columns (1)-(3) report the fraction of output growth in agriculture, manufacturing and services that are explained by gender barriers, while Column (4) report the the same for aggregate output. Aggregate output is calculated as an expenditure-share weighted average of the sectoral output. We report the results for each country, along with a sample average. Turning to the results, from Column (1), changes in gender barriers account for 12% of growth in agricultural output on average. This ranges from 5-8% in India and Indonesia, to over a quarter in Brazil, to 10-13% for the other countries. From Column (2), changes in gender barriers account for 16% of growth in manufacturing output on average. For low-income countries (India and Indoensia), this is only around 5%, while around 20-25% for the middle- and high-income countries. From Column (3), 22% of the growth in services output can be attributed to changes in gender barriers on average. This is lowest in India (2%), around a third in Brazil, and around 20-25% for the other countries. Overall from Column (4), changes in gender barriers from 1970-2015 explain around 20% of changes in aggregate output. However, there is a large variation across countries, ranging from 2% in India, 16% in Indonesia, and round 25-30% in the other countries. Put together, we conclude that they were an important driver of reallocation of workers across sectors as well as had non-trivial effects on output growth.

9 Conclusion

The paper documents stylized facts about the gender dimension of structural transformation across multiple countries over the last five decades. We find substantial gender gaps in employment and wages across occupations-sector pairs, which narrow over time, but still persist today even in the most developed countries. To quantify the effects of gender barriers on economic outcomes, we develop a general equilibrium Roy model that incorporates standard economic drivers of occupational and sectoral choices as well as gender barriers through the form of wage discrimination and gender norms. We estimate the model for six countries across five decades, and use our estimated model for a counterfactual analysis. We find that the reduction in gender barriers over the last five decades had large effects on sectoral employment changes and output growth. The importance of changes in gender barriers varies across sectors and countries with larger effects for the service sector and small effects for agriculture.

Our analysis (intentionally) does not propose specific policies that could bolster gender parity in the labor market, but we view our quantitative model as a useful framework that allows decomposing observable changes in empirical data patterns into a part that is due to standard economic channels and another part that is due to changes in gender barriers. In addition, our general equilibrium framework is useful to aggregate changes in individual choices to quantify the macroeconomic and sectoral effects of changing gender barriers. Future research should explore the underlying factors that led to larger declines in gender barriers in some countries (like Brazil) than in others (like India).

References

Acemoglu, Daron and Veronica Guerrieri, "Capital Deepening and Nonbalanced Economic Growth," *Journal of Political Economy*, 2008, *116* (3), 467–498.

Alder, Simon, Timo Boppart, and Andreas Müller, "A theory of structural change that can fit the data," *American Economic Journal: Macroeconomics*, 2022, *14* (2), 160–206.

Bandiera, Oriana, Ahmed Elsayed, Anton Heil, and Andrea Smurra, "Economic Development and the Organisation of Labour: Evidence from the Jobs of the World Project," *Journal of the European Economic Association*, 2022, *20* (6), 2226–2270.

Bridgman, Benjamin, Georg Duernecker, and Berthold Herrendorf, "Structural transformation, marketization, and household production around the world," *Journal of Development Economics*, 2018, *133*, 102–126.

Caunedo, Julieta and Elisa Keller, "Capital-Embodied Structural Change," Technical Report, Unpublished Manuscript 2023.

Comin, Diego, Danial Lashkari, and Martí Mestieri, "Structural change with long-run income and price effects," *Econometrica*, 2021, *89* (1), 311–374.

Cuberes, David and Marc Teignier, "Gender inequality and economic growth: A critical review," *Journal of International Development*, 2014, *26* (2), 260–276.

— **and** — , "Aggregate effects of gender gaps in the labor market: A quantitative estimate," *Journal of Human Capital*, 2016, *10*, 1–32.

Duernecker, Georg and Berthold Herrendorf, "Structural transformation of occupation employment," *Economica*, 2022, *89* (356).

Fabrizio, Stefania, Lisa Kolovich, and Monique Newiak, "Pursuing Women's Economic Empowerment," *International Monetary Fund*, 2018.

Fan, Tianyu, Michael Peters, and Fabrizio Zilibotti, "Service-Led or Service-Biased Growth? Equilibrium Development Accounting across Indian Districts," Technical Report, National Bureau of Economic Research 2021.

Fletcher, Erin, Rohini Pande, and Charity Maria Troyer Moore, "Women and work in India: Descriptive evidence and a review of potential policies," 2017.

Goldin, Claudia, "The u-shaped female labor force function in economic development and economic history," *Working Paper 4707, National Bureau of Economic Research*, 1994.

Gottlieb, Charles, Cheryl Doss, Douglas Gollin, and Markus Poschke, "Understanding the gender division of work across countries," Technical Report 2023.

Heath, Rachel and Seema Jayachandran, "The causes and consequences of increased female education and labor force participation in developing countries," Technical Report, National Bureau of Economic Research 2016.

Herrendorf, Berthold and Todd Schoellman, "Wages, human capital, and barriers to structural transformation," *American Economic Journal: Macroeconomics*, 2018, *10* (2), 1–23.

__ **, Richard Rogerson, and Akos Valentinyi**, "Two perspectives on preferences and structural transformation," *American Economic Review*, 2013, *103* (7), 2752–89.

__ **,** __ **, and** __ , "Growth and Structural Transformation," *Handbook of Economic Growth*, 2014, *2*, 855–941.

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

Hyland, Marie, Djankov Simeon, and Pinelopi Koujianou Goldberg, "Gendered laws and women in the workforce," *Unpublished Manuscript*, 2020.

IPUMS International, "Minnesota Population Center, Integrated Public Use Microdata Series, International," *Version 7.3 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D020.V7.3*, 2020.

Kuznets, Simon, "Modern economic growth: findings and reflections," *The American economic review*, 1973, *63* (3), 247–258.

Lee, Munseob, "Allocation of Female Talent and Cross-Country Productivity Differences," *Unpublished Manuscript*, 2022.

Maddison, Angus, "Economic growth and structural change in the advanced countries," *Western Economies in Transition, eds.: I. Leveson and W. Wheeler. London: Croom Helm*, 1980.

Mammen, Kristin and Christina Paxson, "Women's work and economic development," *Journal of economic perspectives*, 2000, *14* (4), 141–164.

Moro, Alessio, Solmaz Moslehi, and Satoshi Tanaka, "Does home production drive structural transformation?," *American Economic Journal Macroeconomics*, 2017, pp. 116–146.

Ngai, L. Rachel and Barbara Petrongolo, "Gender Gaps and the Rise of the Service Economy," *American Economic Journal: Macroeconomics*, 2017, *9* (4), 1–44.

__ **and Christopher A. Pissarides**, "Structural Change in a Multisector Model of Growth," *American Economic Review*, March 2007, *97* (1), 429–443.

OECD, "Gender Wage Gap Indicator," *doi: 10.1787/7cee77aa-en (Accessed on 04 November 2023)*, 2023.

Olivetti, Claudia, "The female labor force and long-run development: the american experience in comparative perspective," *Human Capital in History: The American Record, Ed. L. Platt Boustan, C. Frydman R.A. Margo University of Chicago Press*, 2014.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo, "The Human Side of Structural Transformation," *American Economic Review*, August 2022, *112* (8), 2774–2814.

Psacharopoulos, George and Zafiris Tzannatos, "Female labor force participation: An international perspective," *The World Bank Research Observer*, 1989, *4* (2), 187–201.

Rendall, Michelle, "Female market work, tax regimes, and the rise of the service sector," *Review of Economic Dynamics*, 2018, *28*, 269–289.

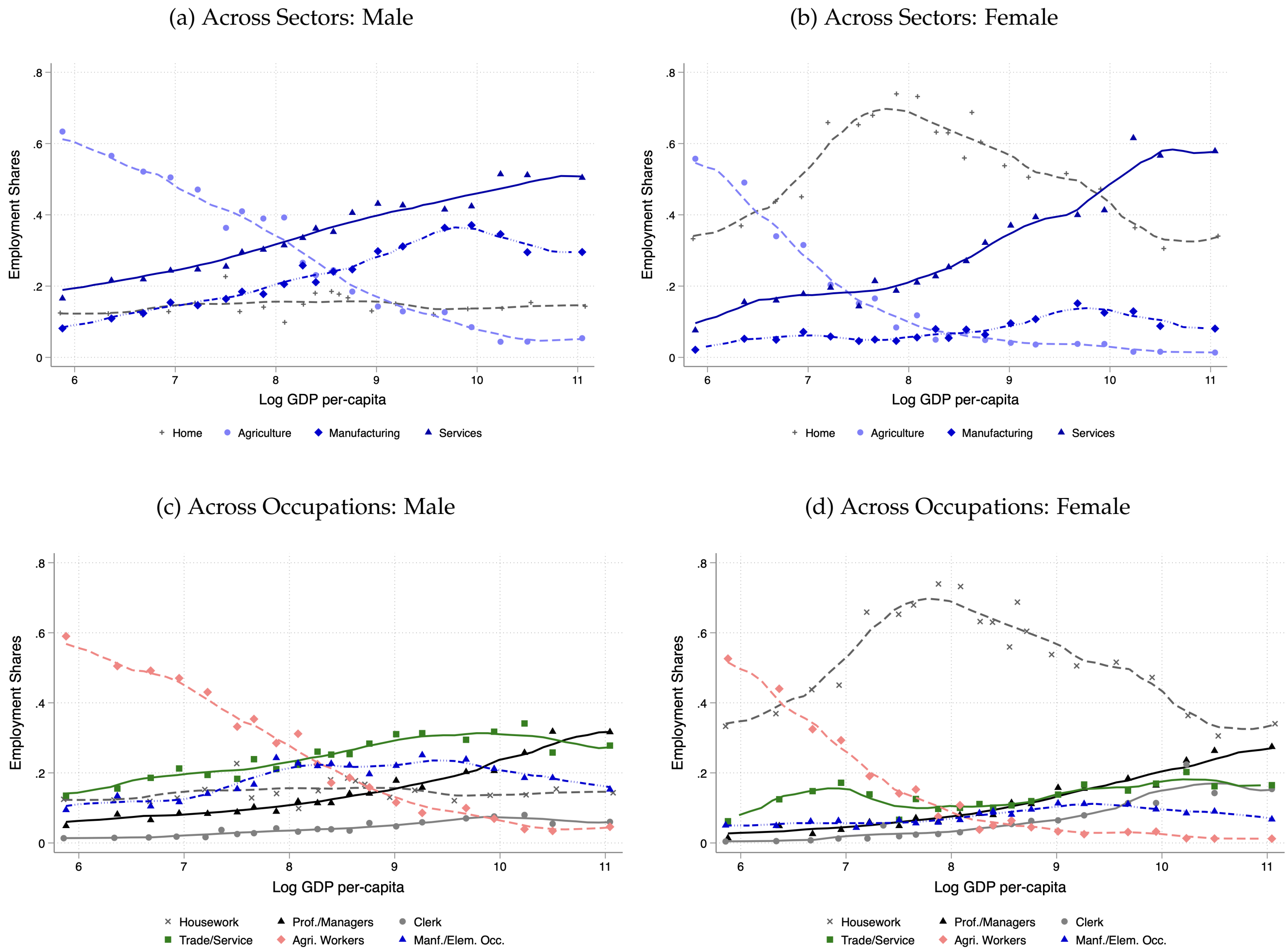
World Bank, "Women, Business and the Law 2019: A Decade of Reform," *Technical Report*, 2019.

__ , "Women, Business and the Law 2020," *Technical Report*, 2020.

Ying, Feng, Jia Ren, and Michelle Rendall, "The reversal of the gender education gap with economic development," *Unpublished Manuscript*, 2023.

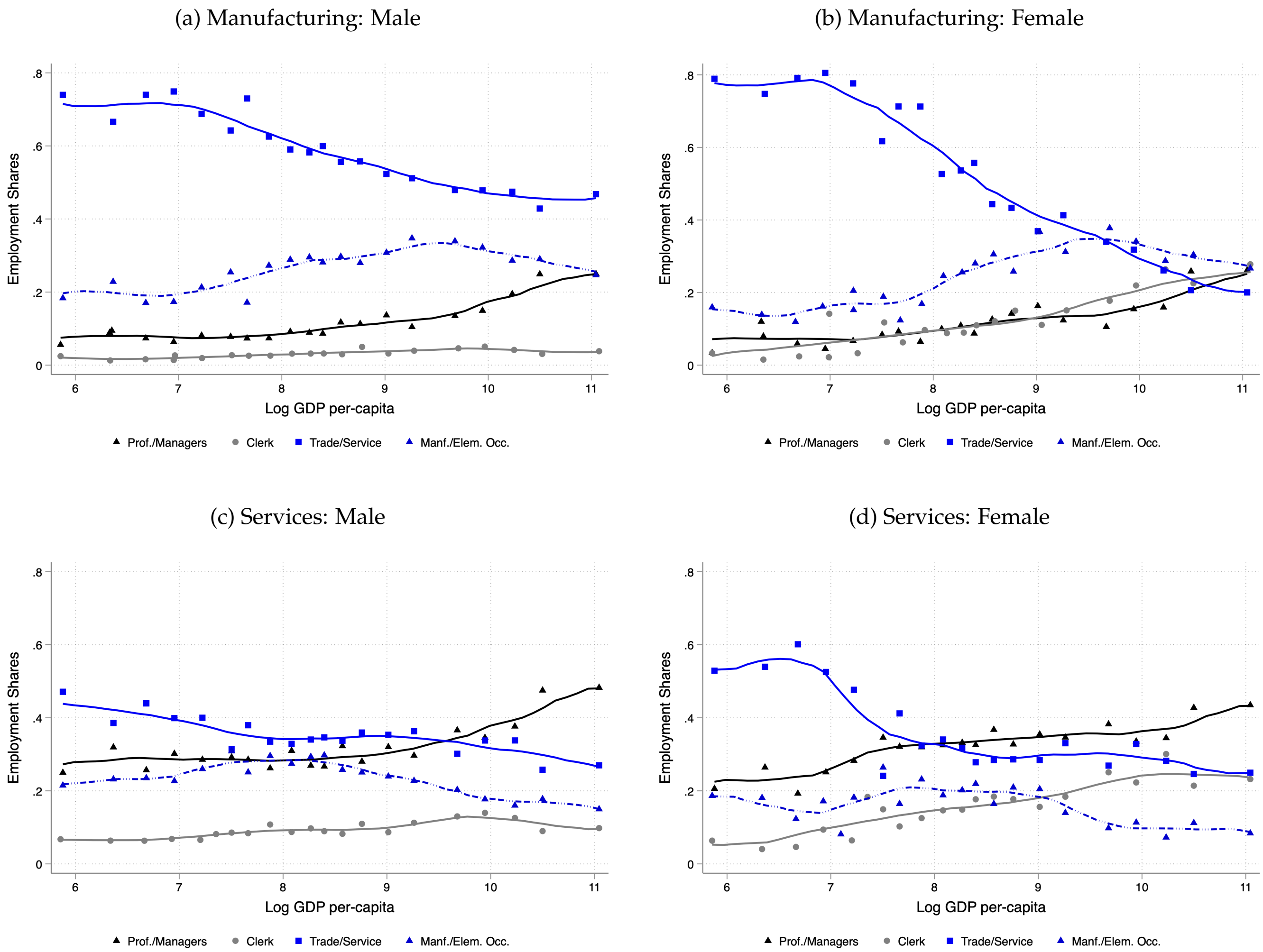
Figures

Figure 1: Sectoral and Occupational Employment by Gender



Notes: This figure is a non-parametric plot of the employment shares of men and women in each sector and occupation against the log of real GDP per-capita in 2010 US dollars. The sample pools all available country years from the IPUMS data. Figures (a) and (c) report the transitions for men, while (b) and (d) report these transitions for women.

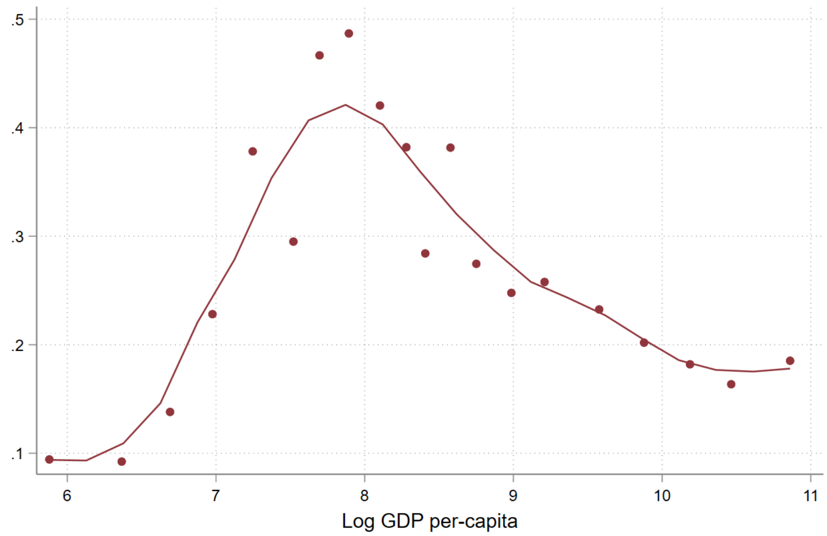
Figure 2: Occupation Structure within Manufacturing and Services by Gender



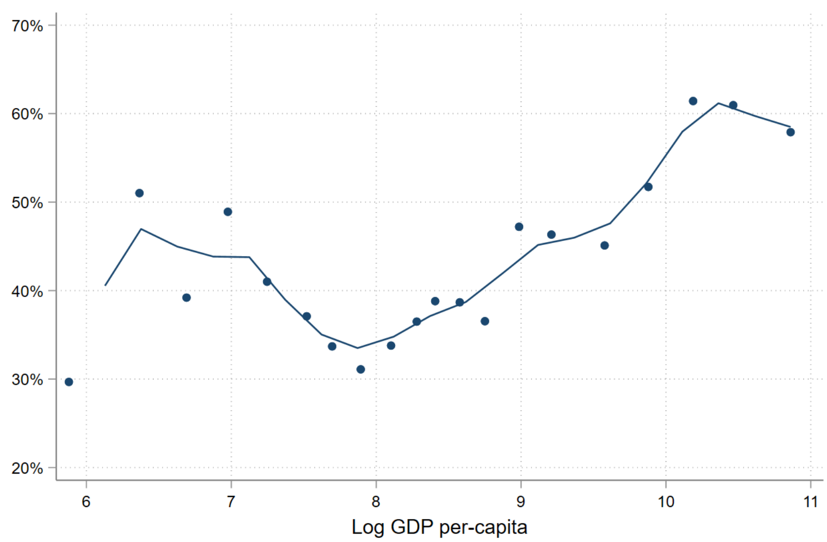
Notes: This figures is a non-parametric plot of the share of men and women in each occupation against the log of real GDP per-capita. Figures (a) and (b) show the occupational distribution among workers in the manufacturing sector. Figures (c) and (d) show the distribution for workers in the service sector. The sample pools all available country years from the IPUMS data.

Figure 3: Gender Segregation across Sectors and Occupations

(a) Theil Index Level



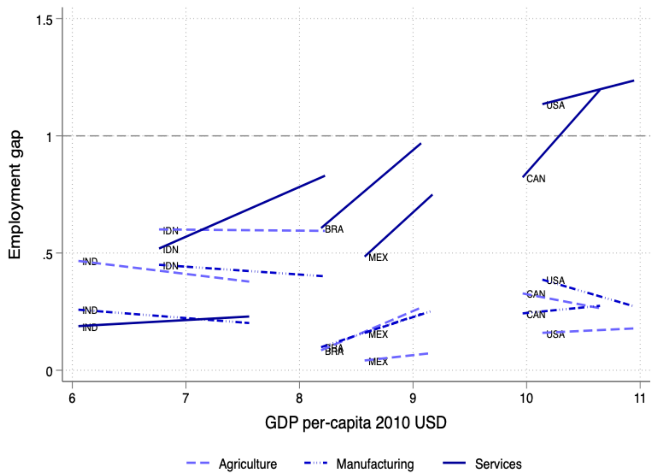
(b) Within Sector



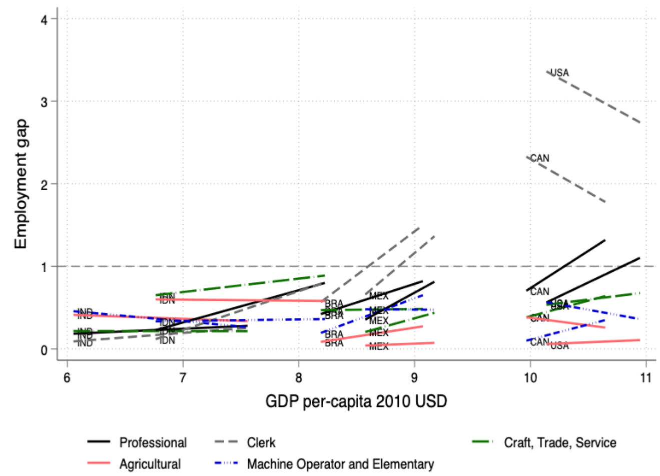
Notes: This figure plots the Theil Index that measures gender segregation across occupation-sector pairs against the lof og real GDP per capita. Figure (a) plots the level of the segregation Index and figure (b) plots the share of segregation that is explained by segregation across-occupations within-sectors.

Figure 4: Employment and Wage Ratios over Time

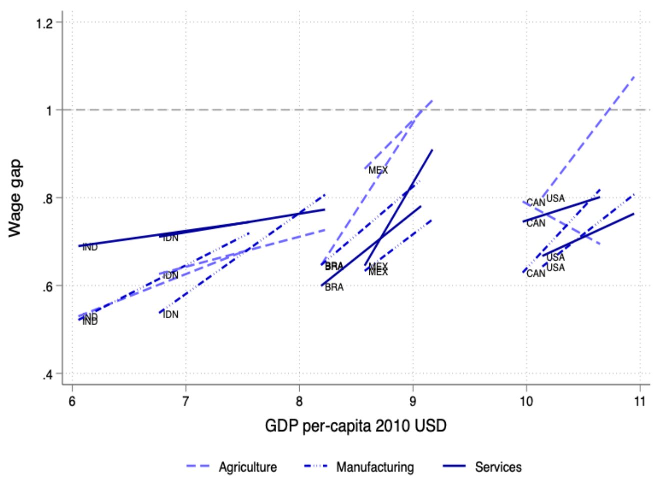
(a) Employment Ratio by Sector



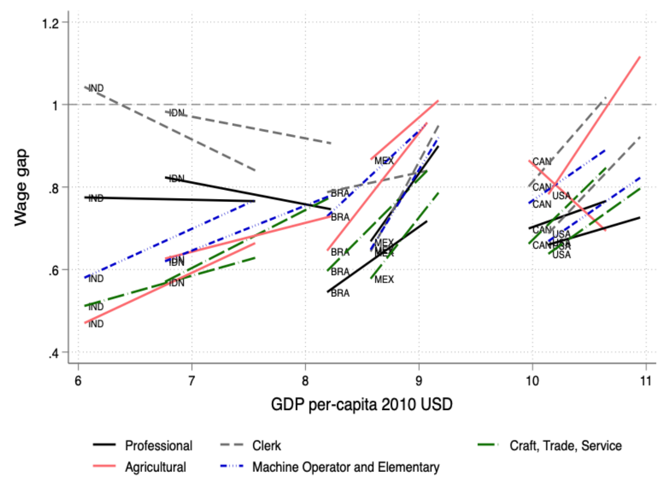
(b) Employment Ratio by Occupation



(c) Wage Ratio by Sector



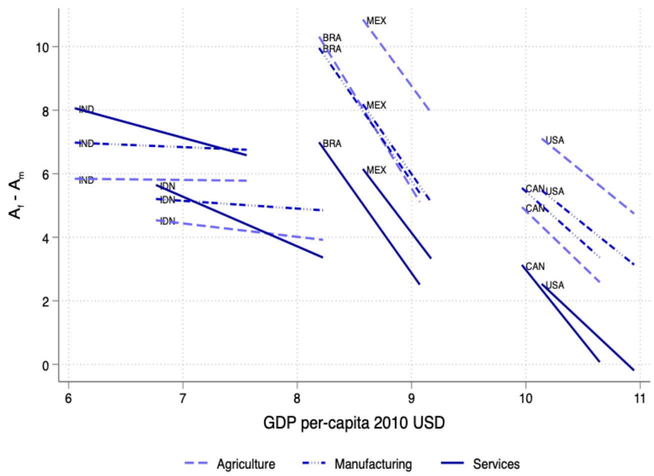
(d) Wage Ratio by Occupation



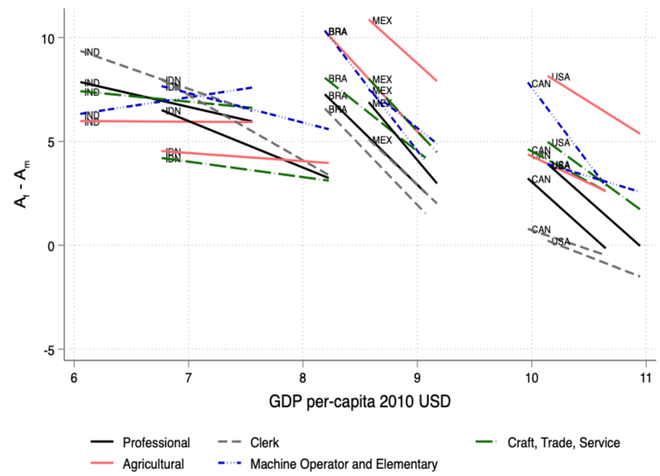
Notes: This figure plots the employment and wage ratios for selected countries over time against the log of real GDP per-capita in constant 2010 US dollars. The time period covers between 1970 and 2018 depending on data availability and the horizontal dimension of the graph shows how fast countries grew during the sample period. Employment ratios divide the share of women working in an occupation-sector by the share of men. Wage ratios divide the average wage of women in an occupation-sector by the average wage of men. Figures (a) and (c) show these ratios by sector while figures (b) and (d) show them by occupation. Figure (b) excludes the clerk occupation for the US and Canada as their employment ratios exceed 2 which makes the graph hard to read.

Figure 5: Gender Norms and Wage Discrimination Across Countries Over Time

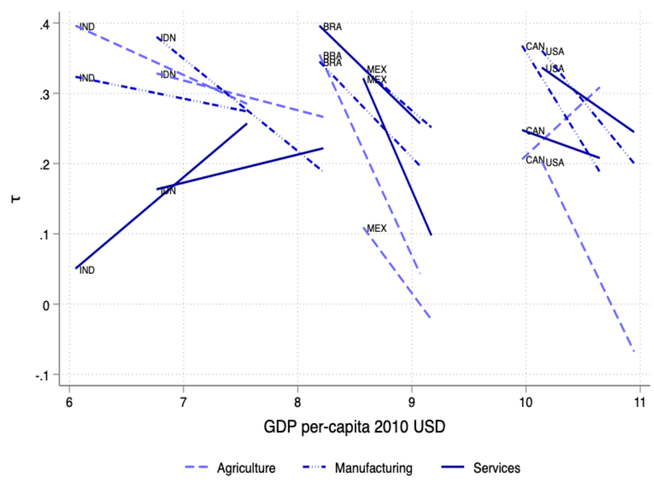
(a) Gender Norms ($\triangle A_{oj}$) by Sector



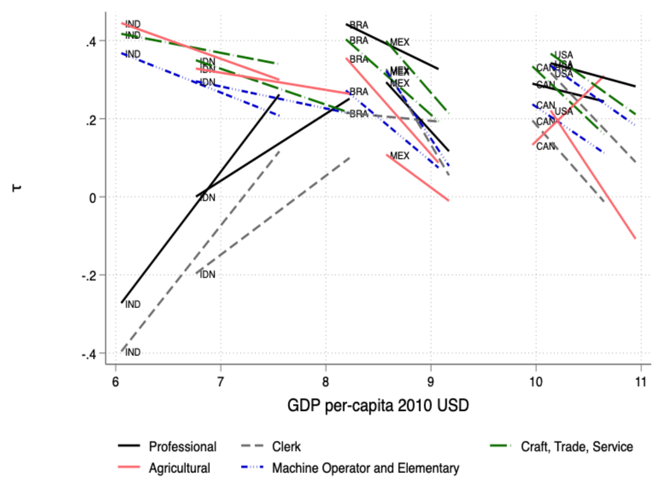
(b) Gender Norms ($\triangle A_{oj}$) by Occupation



(c) Wage Discrimination (τ_{oj}) by Sector



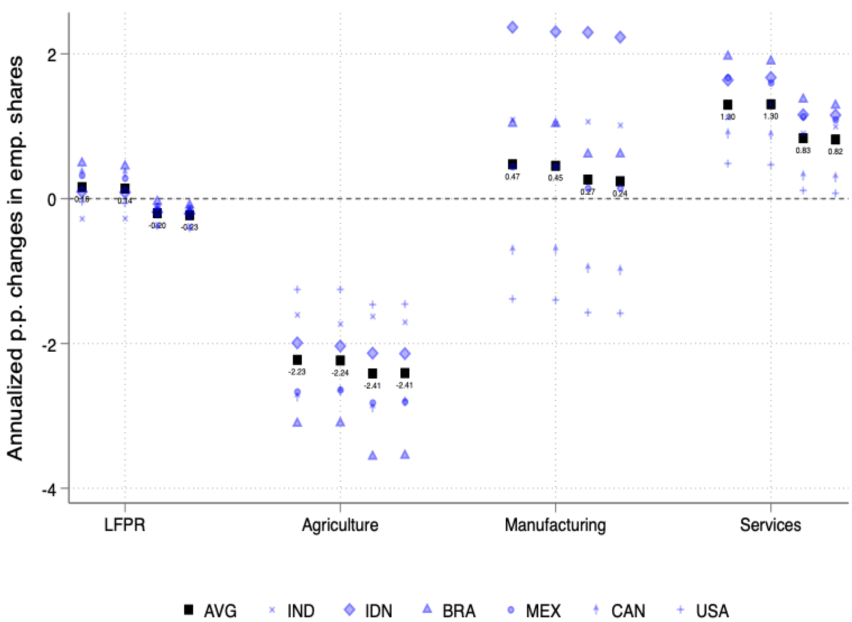
(d) Wage Discrimination (τ_{oj}) by Occupation



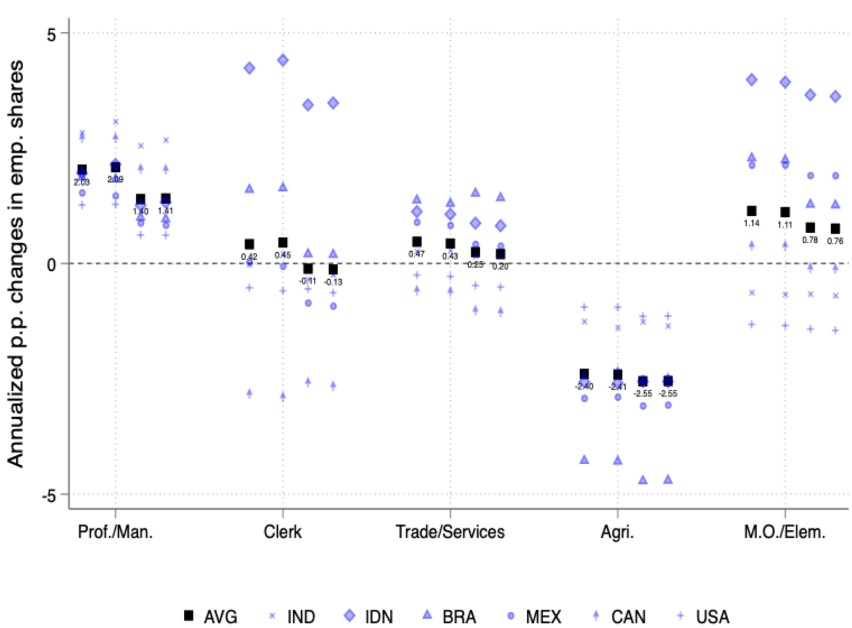
Notes: This figure plots the estimated gender norms ($\triangle A_{oj}$) and wage penalties (τ) for countries in our core sample for the first and last year in our sample, against the log of real GDP per-capita in constant 2010 US dollars. The horizontal dimension of the graph shows how fast countries grew during the sample period. Figures (a) and (c) show gender norms and wage discrimination by sector while figures (b) and (d) show them by occupation.

Figure 6: Gender Barriers and Change in Employment Shares Across Sectors and Occupations

(a) Employment Shares Across Sectors



(b) Employment Shares Across Occupations



Notes: This figure reports the average annualized percentage point changes (between the first and last year) in sectoral and occupational employment shares across countries. Figure (a) shows changes across sectors, while Figure (b) shows changes across occupations. The labor force participation rate (LFPR) is defined as 1- share of individuals in the home sector. The four columns for each sector/occupation is the baseline empirical change, along the three counterfactual simulations as follows: (i) baseline specification; (ii) fixing τ_{ojct}; (iii) fixing $\triangle A_{ojct}$; (iv) fixing both τ and $\triangle A$.

Tables

Table 1: Gender Employment and Wage Ratios

	Employment Ratio		Wage Ratio	
	1970-75	2010-2018	1970-75	2010-2018
	(1)	(2)	(3)	(4)
Home	8.33	3.74		
Agriculture	0.28	0.29	0.71	0.87
Manufacturing	0.27	0.27	0.60	0.79
Services	0.63	0.87	0.68	0.80
Professional	0.41	0.86	0.70	0.77
Clerk	1.19	1.41	0.82	0.91
Craft, Trade, Service	0.41	0.56	0.59	0.78
Agricultural	0.26	0.27	0.71	0.86
Machine Op. & Elementary	0.36	0.41	0.67	0.86

Notes: Columns (1)-(2) and (3)-(4) report the employment and wage ratios in the first and last year, averaged across countries. The employment ratio divides the share of women working in an occupation-sector by the share of men. The wage ratio divides average wage of women in an occupation-sector by the average wage of men. A ratio below 1 implies lower employment (or lower wages) for women relative to men.

Table 2: Gender Norms and Female Wage Discrimination

	Gender Norms ΔA		Wage Discrimination τ	
	1970-75	2010-2018	1970-75	2010-2018
	(1)	(2)	(3)	(4)
Agriculture	7.27	5.01	0.27	0.14
Manufacturing	6.89	4.77	0.35	0.22
Services	5.42	2.61	0.25	0.21
Professional	5.94	2.42	0.18	0.25
Clerk	5.01	1.90	0.08	0.09
Craft, Trade, Service	6.22	3.79	0.38	0.22
Agricultural	7.37	5.13	0.27	0.14
Machine Op. & Elem.	7.26	4.59	0.31	0.14

Notes: Columns (1)-(2) and (3)-(4) report the gender norms ($\triangle A_{ojct}$) and wage discrimination (τ_{ojct}) faced by women in the first and last year, averaged across countries.

Table 3: Change in Sectoral Output Explained by Gender Barriers

	Sectoral Output			Aggregate
	Agri.	Manf.	Services	Output
	(1)	(2)	(3)	(4)
IND	0.05	0.04	0.02	0.02
IDN	0.08	0.06	0.19	0.16
BRA	0.27	0.23	0.34	0.29
MEX	0.11	0.18	0.28	0.23
CAN	0.10	0.24	0.27	0.28
USA	0.13	0.19	0.24	0.24
AVG	0.12	0.16	0.22	0.20

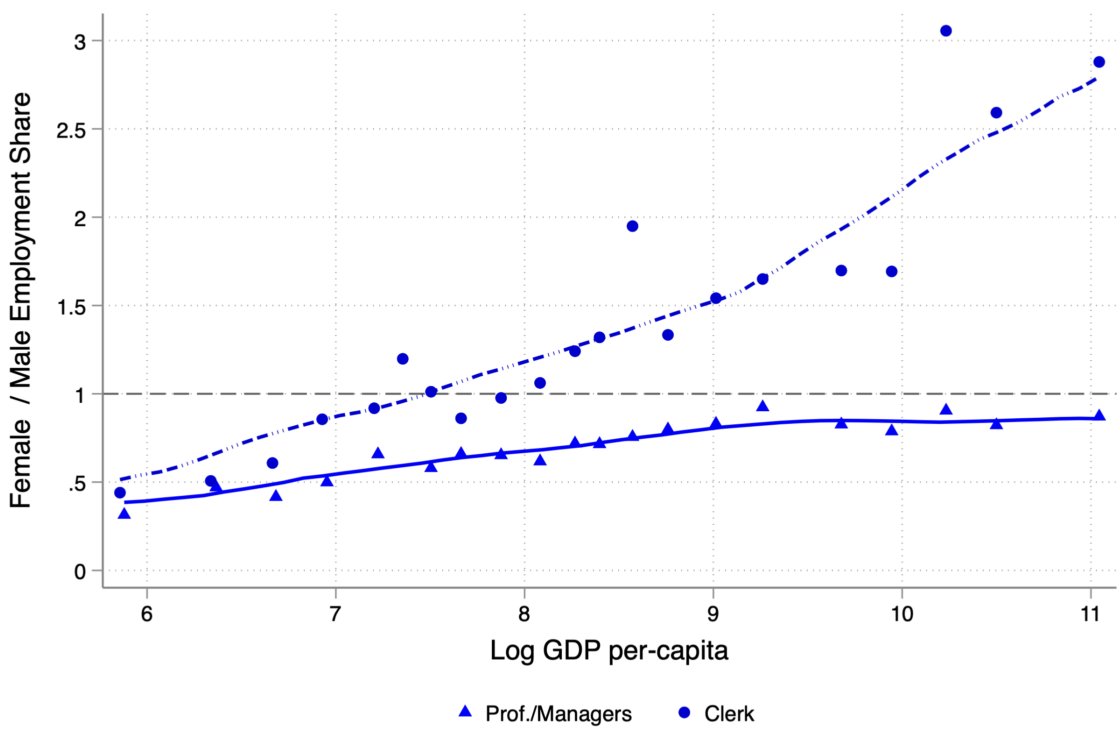
Notes: This table reports the share of sectoral output growth that is explained by changes in gender barriers. To calculate this share, we compute $(1 - \hat{g}/g)$ where g is the output growth observed in the data and $\hat{g}$ is the counterfactual output growth when gender barriers are fixed at their initial values.

ONLINE APPENDIX: NOT FOR PUBLICATION

A Tables and Figures

A.1 Appendix Figures

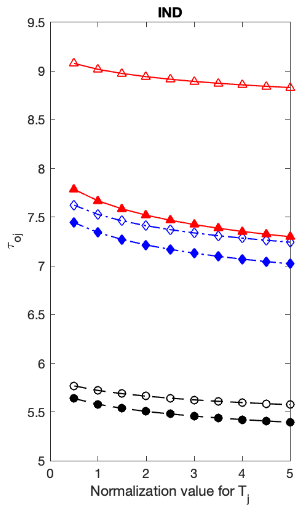
Figure A.1.1: Employment Gap in Professional and Clerical Occupations



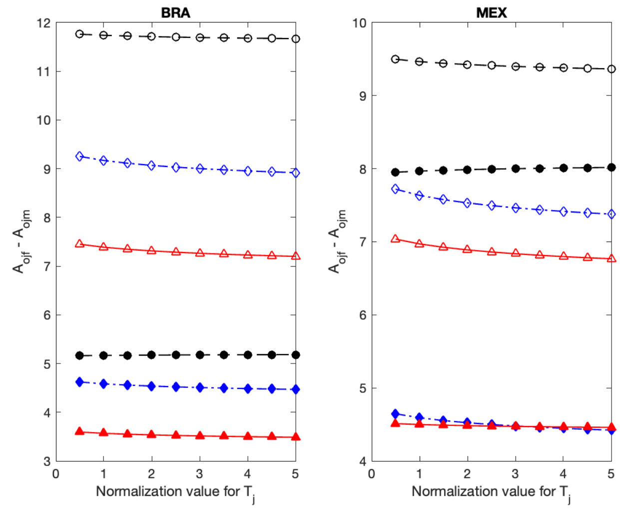
Notes: The above figures reports a non-paramemtric fit between the ratio of the female to male employment share in professional and managerial occupations (blue triangles) and clerks (blue rounds) on the vertical axis, and log of real GDP per-capita (in 2010 USD) is on the vertical axis.

Figure A.1.2: Estimates of Gender Norms ($\triangle A_{oj}$) For Alternate Normalizations of T_j

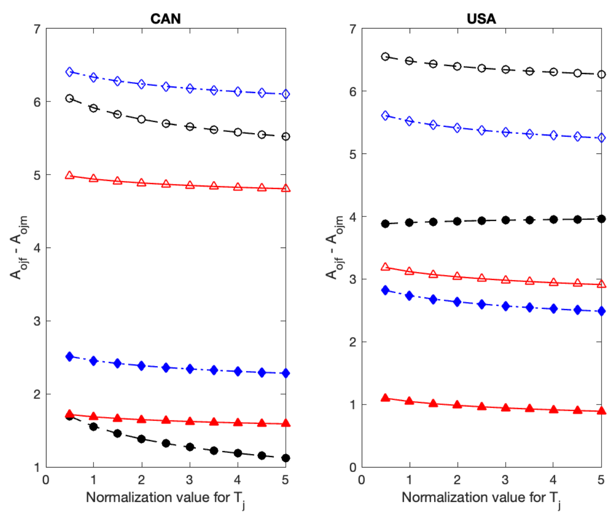
(a) India and Indonesia



(b) Brazil and Mexico



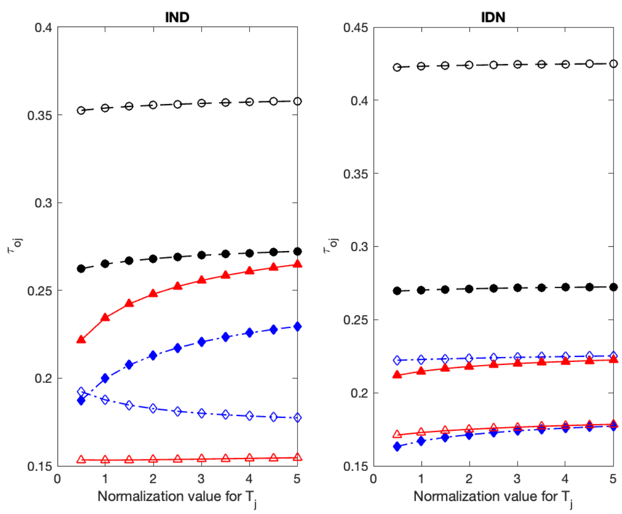
(c) Canada and USA



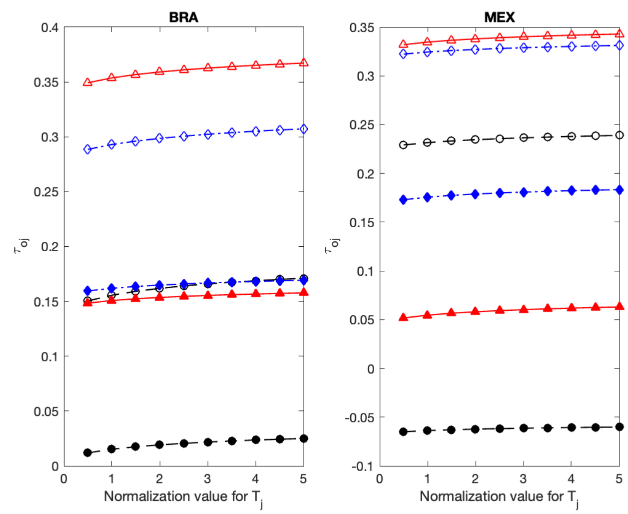
<u>Notes:</u> This figure plots the difference in gender norms ($\triangle A_{oj}$) for each sector (averaged across occupations) for different normalization of T_j, ranging from 0.5 to 5 (on the horizontal axis). Black dots represent the difference in gender norms in agriculture, blue diamonds in manufacturing, and red triangles in services respectively. Unfilled markers represent the first survey year for the country (usually in the 1970s), while the filled markers represent the last survey year (in the 2010s). Figure (a) plots low-income countries of India and Indonesia, Figure (b) plots middle-income countries of Brazil and Mexico, while Figure (c) plots high-income countries of Canada and USA.

Figure A.1.3: Estimates of Wage Discrimination (τ_{oj}) For Alternate Normalizations of T_j

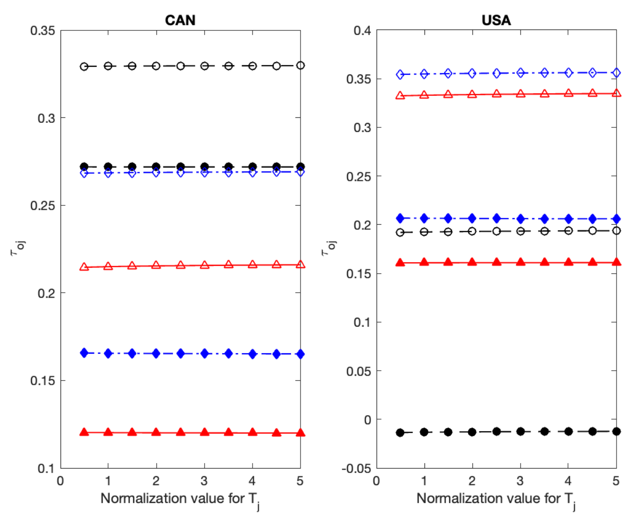
(a) India and Indonesia



(b) Brazil and Mexico

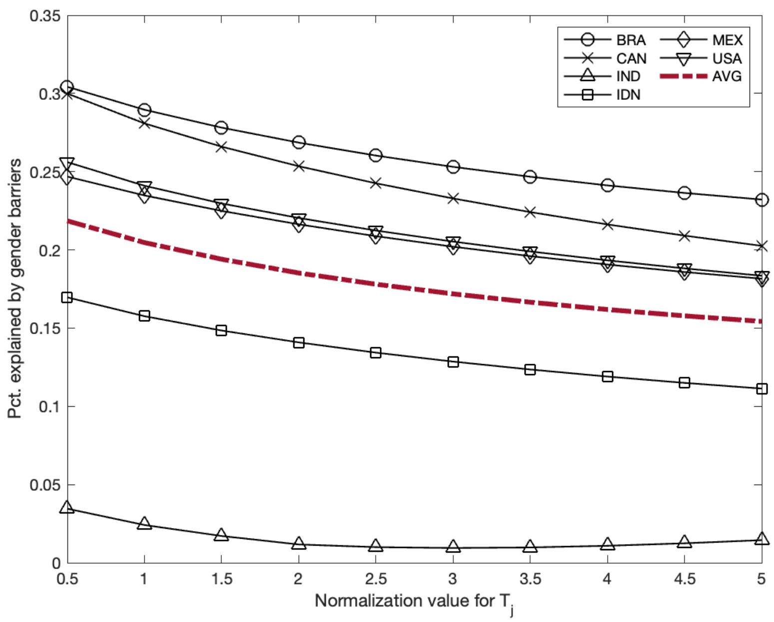


(c) Canada and USA



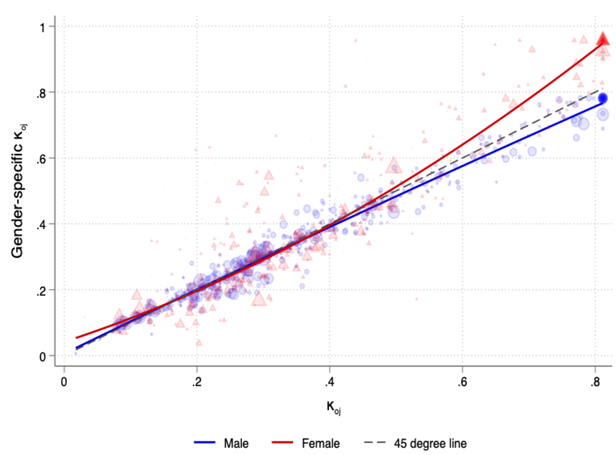
Notes: This figure plots the difference in wage discrimination (τ_{oj}) for each sector (averaged across occupations) for different normalization of T_j, ranging from 0.5 to 5 (on the horizontal axis). Black dots represent the difference in wage discrimination in agriculture, blue diamonds in manufacturing, and red triangles in services respectively. Unfilled markers represent the first survey year for the country (usually in the 1970s), while the filled markers represent the last survey year (in the 2010s). Figure (a) plots low-income countries of India and Indonesia, Figure (b) plots middle-income countries of Brazil and Mexico, while Figure (c) plots high-income countries of Canada and USA.

Figure A.1.4: Fraction of Aggregate Real Output Growth Explained by Gender Barriers For Alternate Normalizations of T_j



Notes: The above figure reports the fraction of aggregate real output growth in each country that is explained by changes in gender barriers (similar to Table 3) for different normalization of T_j ranging from 0.5 to 5.

Figure A.1.5: Comparison of κ_{oj} and Gender-Specific κ_{ojg}



Notes: The above figure reports the correlation between the returns to ability that is the same across gender (κ_{oj}) on the vertical axis with gender-specific returns (κ_{ojg}) on the vertical axis for men (blue circles) and women (red triangles). Each point is an estimate of κ obtained from Mincerian wage regressions described in Section 6. The dash black line is the 45 degree line.

A.2 Sample Details, Sector and Occupation Classification

Table A.2.1: Coverage of All Countries Across Decades

Decade	Country-Years	Percentage
1960-69	13	4.76
1970-79	30	10.99
1980-89	44	16.12
1990-99	57	20.88
2000-10	81	29.67
2010-18	48	17.58
Total	273	100

Notes: The above table reports the coverage of country-years in our data across decades.

Table A.2.2: Coverage of Six Countries

Country	Years	GDP p.c. in 2010
India	1983 to 2018	$1,357
Indonesia	1976 to 2018	$866
Mexico	1960 to 2015	$9,271
Brazil	1970 to 2010	$11,286
Canada	1971 to 2011	$48,464
USA	1960 to 2015	$48,467

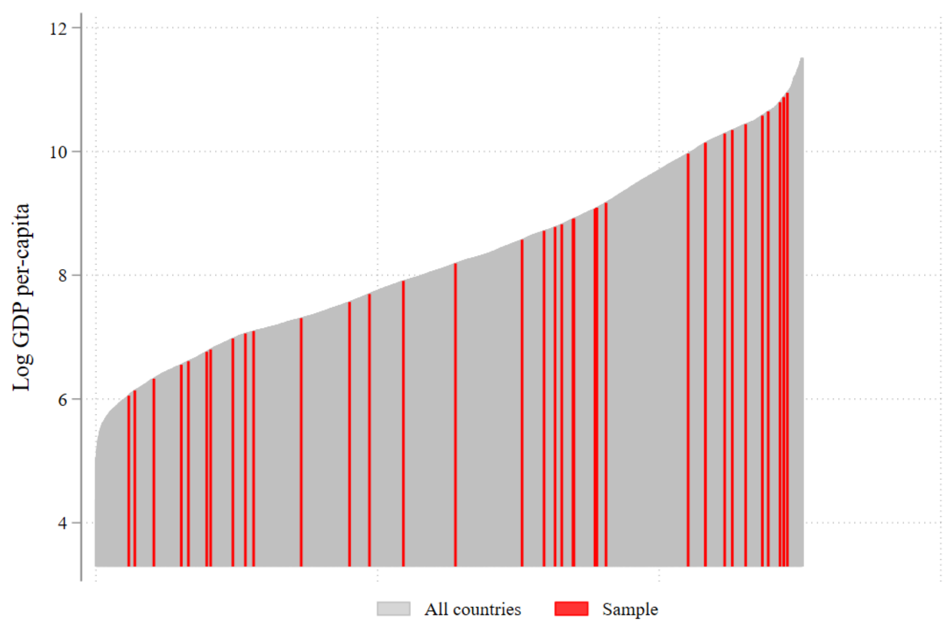
Notes: The above table reports the coverage of our data in the final sample of countries. Column 2 reports the years while Column 3 reports the GDP per-capita in 2010 US dollars from the World Bank data.

Table A.2.3: Classification of Sectors

Sector	IPUMS classification
Agriculture	Agriculture, Fishing and Forestry
Manufacturing	Mining, Construction, Electricity, Gas, Water
Market Services	Retail, Wholesale, Transport, Hotels, Education, Health
"Home Work"	Unemployed, Inactive or in Household Services

Notes: The above table shows the classification of sectors reported in the IPUMS data into Agriculture, Manufacturing, Market and Home services.

Figure A.2.6: Coverage of Countries



Notes: The above figure sorts all country-years by their GDP per-capita (USD 2010) and shows the coverage of countries in our final sample in red.

Table A.2.4: Classification of Occupations

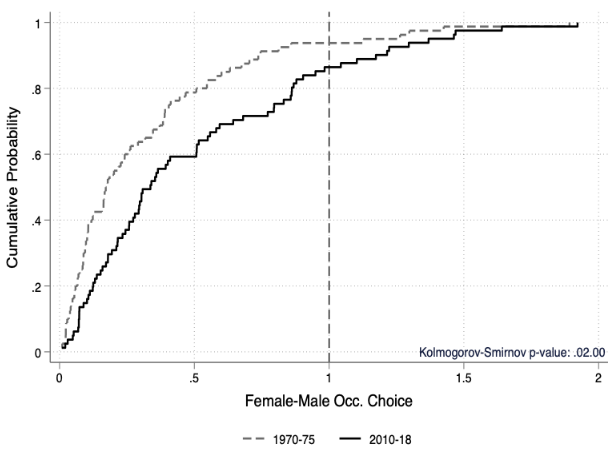
Code	Occupation	Classification	Sector	Details
1	Legislators, Senior Officials and Managers	Professional	M, S	Legislators and Senior Officials, General and Technical
2	Professionals	Professional	M, S	Managers, Professionals and
3	Technicians and Associate Professionals	Professional	M, S	Technicians
4	Clerks	Clerks	M, S	Secretaries, Librarians, Cashiers, Clerks
5	Service Workers and Shop and Market Sales	Services Workers	M, S	Travel, Housekeeping, Personalcare Workers, Shop and Market Sales and Service Workers
6	Skilled Agricultural and Fishery Workers	Skilled Agri.	A	Subsistence and Market-oriented Workers, Crop Growers and Animal Producers, Forestry and Fishery Workers
7	Crafts and Related Trades Workers	Craft/Trade Wrkrs	M, S	Builders, Painters, Blacksmiths, Electricians, Potters, Printers, Textile, Leather Workers
8	Plant and Machine Operators	Plant & Machine	M, S	Plant and Machine Operators in Mining, Metal, Glass, Wood, Chemical, Rubber, Transportation
9	Elementary Occupations	Elementary	M,S	Street Vendors, Domestic Helpers, Porters, Doorkeepers, Garbage Collectors, Manual and Transportation Laborers
10	Armed forces	Drop		
11	Other occupations, unspecified or n.e.c.	Drop		
97	Response suppressed	Drop		
98	Unknown	Drop		
99	NIU (not in universe)	Drop		

Notes: The above table shows the classification of occupations as reported in the IPUMS data. For our analysis in the paper, we aggregate them based on the ISCO 88 classification (Column 1) as well as report the sectors covered for each occupation (Column 4). Details on the classification and occupations can be found here.

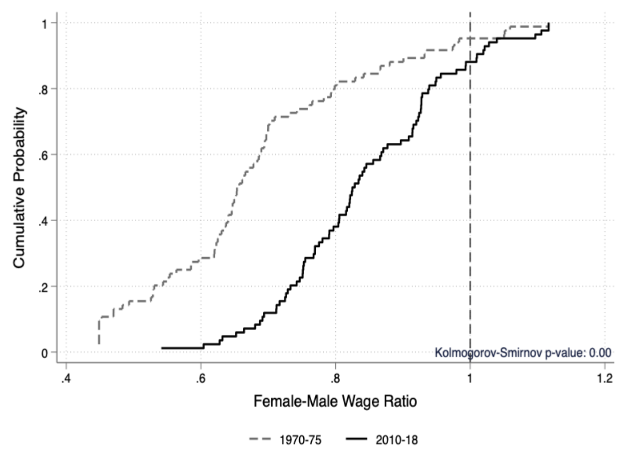
A.3 Distribution of Occupation and Wage Gaps, Gender Norms, and Wage Discrimination

Figure A.3.1: Distribution of Gender Employment and Wage Ratios

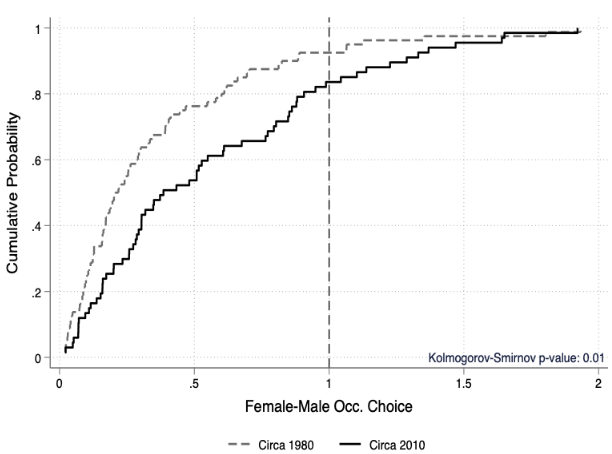
(a) Employment Ratio



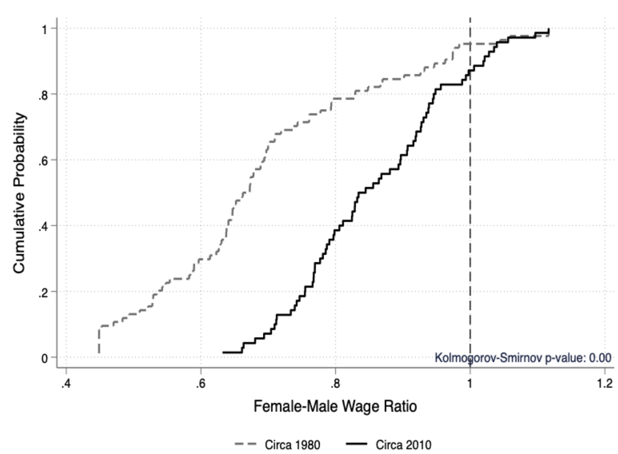
(b) Wage Ratio



(c) Employment Ratio (Circa 1980-2010)



(d) Wage Ratio (Circa 1980-2000)

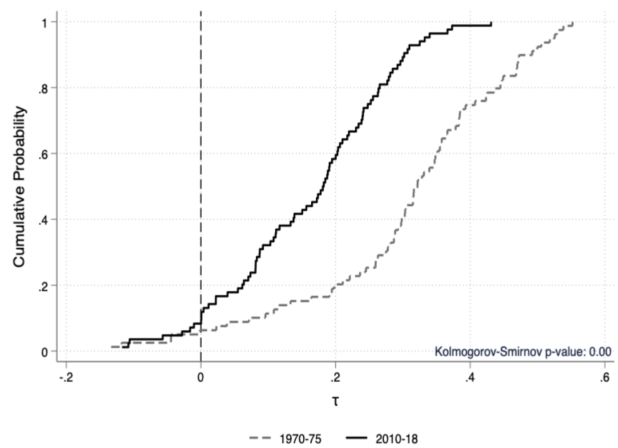


Notes: The above figure plots the CDF of the employment and wage gaps across all sectors and occupations. Figures (a) and (b) plot the employment and wage gaps using the first (dotted line) and last year (solid line) for each country. Figures (c) and (d) show the same distributions, but now using survey years across countries closest to 1980 and 2010. The employment ratio divides the share of women working in an occupation-sector by the share of men. The wage ratio divides the average wage of women in an occupation-sector by the average wage of men.

Figure A.3.2: Distribution of Gender Norms and Wage Discrimination

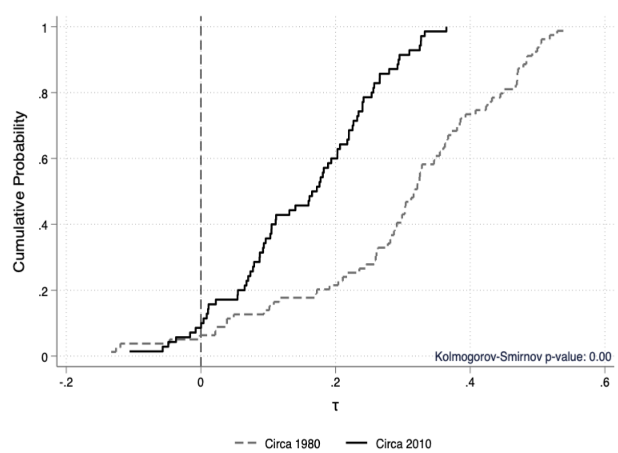
(a) Gender Norms ($\triangle A_{ojct}$)

(b) Wage Discrimination (τ_{ojct})



(c) Gender Norms (Circa 1980 and 2010)

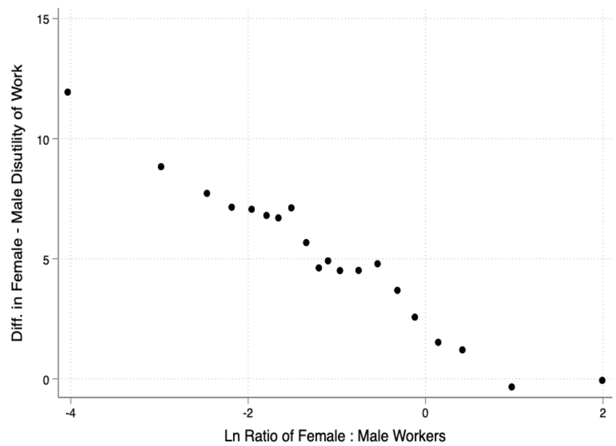
(d) Wage Discrimination (Circa 1980 and 2010)



Notes: The above figure plots the CDF of gender norms ($\triangle A_{ojct}$) and female wage discrimination (τ_{ojct}) across sectors and occupations. Figures (a) and (b) plot gender norms and wage discrimination using the first (dotted line) and last year (solid line) for each country. Figures (c) and (d) show the same distributions, but using survey years closest to 1980 and 2010.

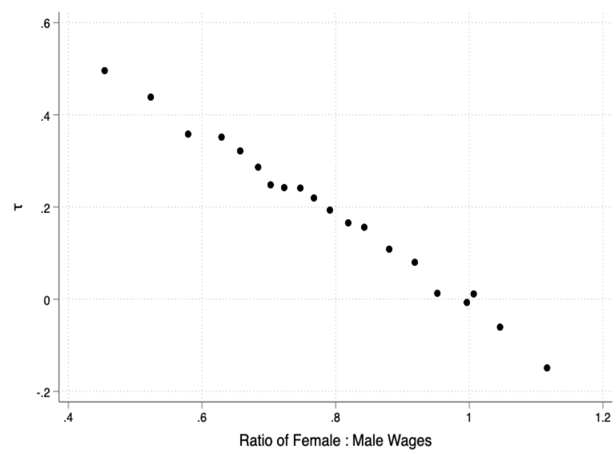
B Model Fit and Correlations with Social Norms

Figure B1: Correlations of Excess Gender Norms ($\triangle A$) and Gender Employment Gaps



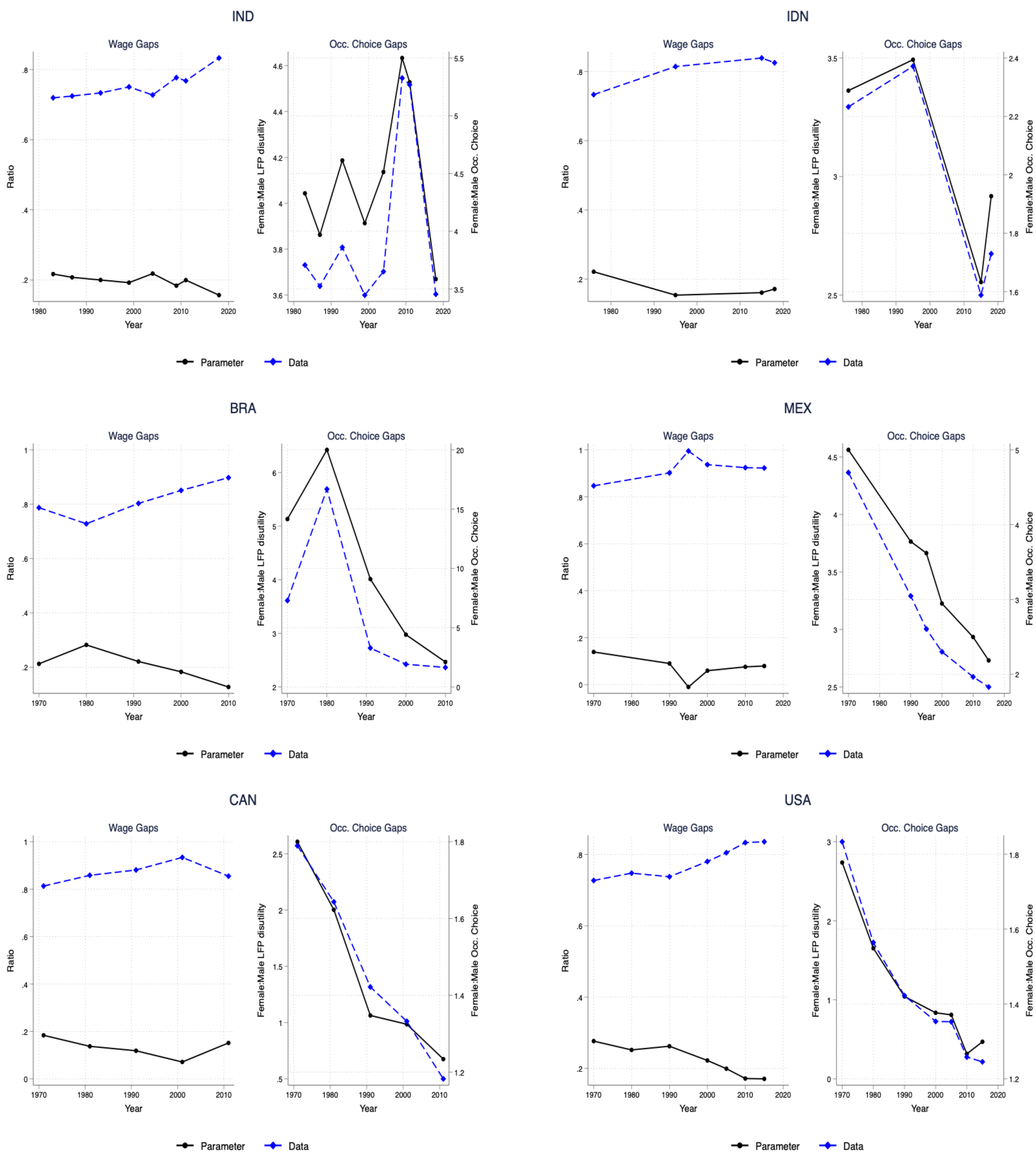
Notes: The above figure shows a binned scatter plot of the correlation between our estimated gender norms between men and women $\Delta A_{oj} = A_{ojf,ct} - A_{ojm,ct}$ and the observed log male to female workers in an occupation-sector across all country-years.

Figure B2: Correlations Wage Penalties (τ) and Gender Wage Gaps



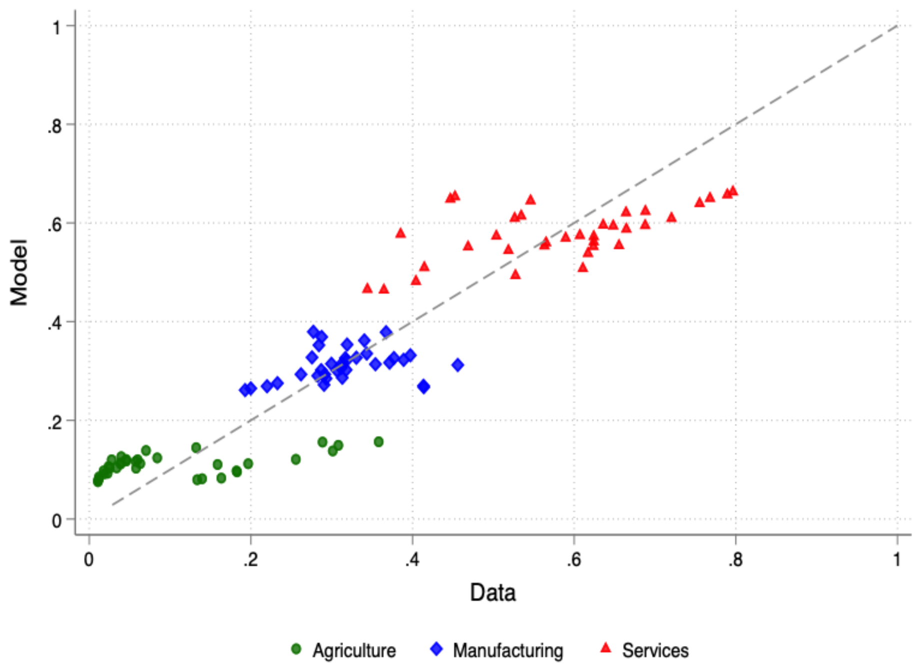
Notes: The above figure shows a binned scatter plot of the correlation between our estimated female wage penalty τ and observed wage gaps in all occupation-sectors and across all country-years.

Figure B3: Correlation of Calibrated Parameters and Targeted Data Moments for Specific Countries Over Time



Notes: The above figures plot the average wage and employment ratios (dash blue line) and the estimated values of wage discrimination (τ_{oj}) and gender norms ($\Delta A_{oj} = A_{ojf} - A_{ojm}$) (solid black lines) over time for each country.

Figure B4: Correlations in Value-Added Shares in the Model and Data



Notes: The above figure shows the correlation between the share of value-added in agriculture (green dot), manufacturing (blue diamond) and services (red triangle) in the data (horizontal axis) and the model (vertical axis) across all country-years in our sample.

Table B1: Correlation between $\triangle A$ and "World, Business, and the Law" Indicators

	Coefficient	S.E.	p-value
	(1)	(2)	(3)
Panel A. Gender Equality in Mobility and LFP			
Index of Mobility/LFP	-0.76	(0.05)	0.00***
Can a woman get a job in the same way as a man?	-0.48	(0.15)	0.09*
Can a woman work at night in the same way as a man?	-0.76	(0.05)	0.00***
Can a woman work in a job deemed dangerous in the same way as a man?	-0.64	(0.04)	0.00***
Can a woman work in an industrial job in the same way as a man?	-0.64	(0.05)	0.01***
Panel B. Household Norms			
Index of Household Norms	-0.40	(0.16)	0.13
Can a woman be head of household in the same way as a man?	-0.63	(0.09)	0.02**
Is there legislation specifically addressing domestic violence?	0.26	(0.06)	0.06*
Does a woman have the same rights to remarry as a man?	-0.12	(0.09)	0.33
Do men and women have equal ownership rights to immovable property?	-0.63	(0.09)	0.02**
Ln(1+Paid Maternity Days Leave)	-0.11	(0.04)	0.10*
Observations	525		

Notes: This table shows the OLS correlation between gender norms $\triangle A_{ojct}$ (which we standardize to mean 0 and std dev 1) and indicators of the "World, Business, and the Law" database as described in Equation (10). Standard errors are clustered at the country level. *** is p¡0.01, ** is p¡0.05 and * is p¡0.1.

Table B2: Correlation between τ and "World, Business, and the Law" Indicators

	Coefficient	S.E.	p-value
	(1)	(2)	(3)
Panel A. Gender Equality in Mobility and LFP			
Index of Mobility/LFP	-0.04	(0.01)	0.05**
Can a woman get a job in the same way as a man?	-0.05	(0.02)	0.12
Can a woman work at night in the same way as a man?	-0.01	(0.02)	0.67
Can a woman work in a job deemed dangerous in the same way as a man?	-0.02	(0.01)	0.06*
Can a woman work in an industrial job in the same way as a man?	-0.04	(0.02)	0.13
Panel B. Gender Equality at the Workplace			
Index of Workplace Equality	-0.05	(0.03)	0.27
Does the law prohibit discrimination in employment based on gender?	-0.02	(0.02)	0.27
Ln(1+Paid Maternity Days Leave)	-0.02	(0.00)	0.00***
Observations	525		

Notes: This table shows the OLS correlation between female wage discrimination (τ_{oj}) and indicators of the "World, Business, and the Law" database as described in Equation (10). Standard errors are clustered at the country level. *** is p¡0.01, ** is p¡0.05 and * is p¡0.1.

C Mathematical Appendix

C.1 Theil Index and Decomposition

The Theil Index of segregation is defined by:

$$T_{oj} = \sum_j \sum_o \frac{N_{oj}}{N} \log\left(\frac{N^f/N}{N^f_{oj}/N_{oj}}\right),$$

where N_{oj} is the number of workers in occupation o and sector j, N is the total population, and N^f is the total number of women in the population. A larger number implies more gender segregation across occupations and sectors. In the case of complete gender equality in employment choices, the ratio in the bracket is equal to 1 so that the whole index becomes equal to 0. The Theil index is additively decomposable into segregation across-sector and within-sector (across-occupation) in the following way:

$$T_{oj} = T_j + \sum_j \frac{N_j}{N} T^j_o,$$

where T_j is the Theil index for gender segregation across sectors and T^j_o is the Theil index for gender segregation across occupations in each sector j, which are defined as:

$$T_j = \sum_j \frac{N_j}{N} \log\left(\frac{N^f/N}{N^f_j/N_j}\right) \quad \text{and} \quad T^j_o = \sum_o \frac{N_{oj}}{N_j} \log\left(\frac{N^f_j/N_j}{N^f_{oj}/N_{oj}}\right).$$

C.2 Deriving Aggregate Sectoral Expenditure Shares

Individual Sectoral Expenditure Shares

From Roy's Identity, the Marshallian demand is:

$$\begin{aligned} x_j &= -\frac{\partial V/\partial p_j}{\partial V/\partial I} \\ \Rightarrow \varphi_j(I,p) &= -\frac{\partial V/\partial p_j}{\partial V/\partial I} \times \frac{p_j}{I} \end{aligned}$$

From Equation (1), we get:

$$\begin{aligned} \frac{\partial V}{\partial I} &= \frac{1}{I} \times \left(\frac{I}{P}\right)^\eta \\ \frac{\partial V}{\partial p_j} &= -\frac{\omega_j}{p_j}\left(\frac{I}{P}\right)^\eta + \frac{\nu^h_j}{p_j} \end{aligned}$$

Defining $P = \prod_j p_j^{\omega_j}$, and substituting in the Roy's identity above, we get that individuals' expenditure share for a sector j is given by:

$$\varphi_j(I_{ojg}(z), p) = \omega_j + \nu_j^h \left(\frac{I_{ojgz}}{P}\right)^{-\eta}$$

C.3 CES Preferences over Home and Market Services

For a sector k where $m \in hs, ms$ i.e., home and market services, the individual's optimization problem can be given by:

$$\min \sum_m p_k C_k$$
$$\text{s.t. } C_s = \left[\sum_k \alpha_k^{\frac{1}{\eta_s}} C_k^{\frac{\eta_s - 1}{\eta_s}}\right]^{\frac{\eta_s}{\eta_s - 1}}$$

Let $P_S = \left[\sum_k \alpha_k p_k^{1-\eta_s}\right]^{\frac{1}{1-\eta_s}}$ and λ be the Lagrange multiplier. Taking the first-order condition and solving we have:

$$\lambda p_k = \alpha_k^{\frac{1}{\eta_s}} \times \left(\frac{C_k}{C_S}\right)^{-\frac{1}{\eta_s}}$$
$$\Rightarrow C_k = \alpha_k (\lambda p_k)^{-\eta_s} C_S$$
$$\therefore \frac{C_{hs}}{C_{ms}} = \frac{\alpha_{hs}}{\alpha_{ms}} \times \left(\frac{p_{hs}}{p_{ms}}\right)^{-\eta_s}$$
$$\Rightarrow \frac{\varphi_{hs}}{\varphi_{ms}} \equiv \frac{P_{hs} C_{hs}/I}{P_{ms} C_{ms}/I} = \frac{\alpha_{hs}}{\alpha_{ms}} \times \left(\frac{p_{hs}}{p_{ms}}\right)^{1-\eta_s}$$

Lastly, substituting back in the constraint, we have:

$$C_S^{\frac{\eta_s - 1}{\eta_s}} = \lambda^{1-\eta_s} \left\{\sum_k \alpha_k p_k^{1-\eta_s}\right\} C_S^{\frac{\eta_s - 1}{\eta_s}}$$
$$\lambda = \left[\sum_k \alpha_k p_k^{1-\eta_s}\right]^{\frac{-1}{1-\eta_s}} = 1/P_S$$
$$\Rightarrow C_k = \alpha_k \left(\frac{p_k}{P_s}\right)^{-\eta_s} C_s$$
$$\Rightarrow \varphi_k \equiv \frac{p_k C_k}{I} = \alpha_k \left(\frac{p_k}{P_s}\right)^{1-\eta_s} \varphi_S$$

C.4 Aggregation of Sectoral Expenditure Shares

Given PIGL preferences, the expenditure share of an individual of gender g, ability z, working in occupation-sector oj is given by (see Equation (2)):

$$\varphi_j(I_{ojgz}, p) = \omega_j + \nu_j^h \left(\frac{I_{ojgz}}{P}\right)^{-\eta}$$

Therefore, the total expenditure on a sector j and the total income in the economy (across all occupations and gender) can be given by:

$$
\begin{aligned}
E &= \sum_j \sum_o \sum_g N_g \int_z I_{ojg}(z) \Pr(oj|z,g) dF(z) \\
E_j &= \sum_o \sum_g N_g \int_z \varphi_j \times I_{ojg}(z) \Pr(oj|z,g) dF(z) \\
&= \sum_g N_g \left\{ \omega_j \sum_o \int_z I_{ojg}(z) \Pr(oj|z,g) dF(z) + \nu_j P^\eta \sum_o \int_z I_{ojg}(z)^{1-\eta} \Pr(oj|z,g) dF(z) \right\}
\end{aligned}
$$

Define:
Aggregating over returns to ability: $E(z^x) = \int_z z^x \Pr(oj|z,g) dF(z)$
Total supply of human capital: $H_{ojg} = E(z^{\kappa_{ojg}}$ (from Equation (**??**)
Total income in a sector j: $I_j^{TOT} = \sum_o \sum_g w_{ojg} H_{ojg}$
Total income in the economy: $I^{TOT} = \sum_j I_j^{TOT}$
Income share of a sector j: $\iota_j = I_j^{TOT} / I^{TOT}$

Therefore:

$$
\begin{aligned}
\Phi_{jg} = \frac{E_j}{E} &= \omega_j \times \frac{\sum_o \int_z I_{ojg}(z) \Pr(oj|z,g) dF(z)}{\sum_j \sum_o \int_z I_{ojg}(z) \Pr(oj|z,g) dF(z)} + \nu_j P^\eta \times \frac{\sum_o \int_z I_{ojg}(z)^{1-\eta} \Pr(oj|z,g) dF(z)}{\sum_j \sum_o \int_z I_{ojg}(z) \Pr(oj|z,g) dF(z)} \\
&= \omega_j \times \frac{\sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}})}{\sum_j \sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}})} + \nu_j P^\eta \times \frac{\sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}(1-\eta)})}{\sum_j \sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}})} \\
&= \omega_j \times \frac{\sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}})}{\sum_j \sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}})} \\
&+ \nu_j \times \underbrace{\frac{\sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}(1-\eta)})}{\left[\sum_j \sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}}) \right]^{1-\eta}}}_{=g(z^{\kappa_{ojg}})} \times \left[\frac{\sum_j \sum_o \sum_g N_g w_{ojg} E(z^{\kappa_{ojg}})}{P} \right]^{-\eta} \\
&= \omega_j \iota_j + \nu_j \times g(z^{\kappa_{ojg}}) \times \left[\frac{I^{TOT}}{P} \right]^{-\eta}
\end{aligned}
$$

D Estimation Algorithm

Here we describe the numerical procedure that we use to estimate key model parameters by fitting our model's equilibrium conditions to data moments. With this procedure, we estimate the following parameters: PIGL preferences $\mathcal{U} = \{\omega_j, \nu_j\}_{\forall j}$, sectoral productivity $\{B_j\}_{\forall j}$, and gender barriers $\mathcal{B} = \left\{ \{\tau_{oj}\}_{\forall oj}, \{A_{ojg}\}_{\forall ojg} \right\}$. To calibrate these parameters, we exactly match data on men's and women's occupation-sector choices $\Pr(oj|g)$, gender gaps in average hourly wages ($\overline{\text{wage}}_{ojf}/\overline{\text{wage}}_{ojm}$) and growth in sectoral real value added ΔY_j. We use data on sectoral value added shares to calibrate the PIGL preference parameters.

Outer loop: Guess sectoral productivity B_j and PIGL parameters (ω_j, ν_j).

Inner loop: Guess occupation-sector wage rates w_{oj}, sectoral prices p_j, amenities A_{ojg}, and female wage discrimination τ_{oj}.

Step 1: Compute income $I_{ojgz} = w_{ojg} z^{\kappa_{oj}}$ and indirect utility $V(I_{ojgz}, p)$ for each gender-ability type gz using:

$$V(I_{ojgz}, p) = \frac{1}{\eta} \left[\frac{I_{ojgz}}{\prod_j p_j^{\omega_j}} \right]^{\eta} - D(p), \tag{11}$$

and compute occupational choices for each gender-ability-type using:

$$\Pr(oj|g, z) = \frac{\exp\left[\frac{1}{\sigma_\varepsilon} V(I_{ojgz}, p) - \frac{1}{\sigma_\varepsilon} A_{ojg} \right]}{\sum_{j'} \sum_{o'} \exp\left[\frac{1}{\sigma_\varepsilon} V(I_{o'j'gz}, p) - \frac{1}{\sigma_\varepsilon} A_{o'j'g} \right]}. \tag{12}$$

Step 2: Integrate these choice probabilities across z-types and solve for a new guess of amenities A_{ojg}^{new} to perfectly fit men's and women's observed employment shares in each occupation-sector $\Pr(oj|g)$.

Step 3: Compute average human capital in each occupation-sector and for each gender (taking into account how workers' selection into occupation-sectors is driven by their comparative advantage $z^{\kappa_{oj}}$ and other factors), using:

$$\overline{H}_{ojg} = \int_z \Pr(oj|g, z) z^{\kappa_{oj}} dF(z). \tag{13}$$

Step 4: Solve for a new guess of female wage discrimination τ_{oj}^{new} to perfectly fit observed gender wage gaps in each occupation-sector using:

$$\frac{\overline{\text{wage}}_{ojf}}{\overline{\text{wage}}_{ojm}} = (1 - \tau_{oj}) \times \frac{\overline{H}_{ojf}}{\overline{H}_{ojm}}. \tag{14}$$

Step 5: Solve for a new guess of occupation-sector-specific wage rates w_{oj}^{new} from firms' first order condition:

$$w_{oj} = \frac{\gamma_{oj} p_j B_j \prod_o H_{oj}^{\gamma_{oj}}}{H_{oj}} \tag{15}$$

Step 6: Compute aggregate sectoral expenditure shares:

$$\Phi_j = \omega_j + \nu_j \frac{P^\eta}{I} \sum_o \sum_j \sum_g \int_z I_{ojgz}^{(1-\eta)} \frac{N_g}{N} \Pr(oj|g,z), \tag{16}$$

where $P = \prod_j p_j^{\omega_j}$ and $I_{ojgz} = w_{ojg} z^{\kappa_{oj}}$.

Step 7: Solve for a new guess of sectoral prices p_j^{new} to ensure that good markets clear in each sector:

$$p_j^{new} = \frac{\Phi_j \times I}{B_j Y_j} \tag{17}$$

Iterate on the inner loop until convergence.

For each convergence of the inner loop, we proceed with the outer loop which solves for PIGL preference parameters and sectoral productivity.

Step 8: Regress observed shares of sectoral value-added (ETD data) on a constant and a model-implied measure of real income according to Equation 16 using data across multiple countries and over time. The regression constant provides a new guess for ω_j^{new} and the coefficient provides a new guess for ν_j^{new}.

Step 9: Solve for a new guess of sectoral productivity B_j^{new} to match observed growth in sectoral real value added ΔY_j in each country over time using the sector's production function (and normalizing $B_{j0} = 1$ in the first year of each country):

$$Y_j = B_j \prod_o H_{oj}^{\gamma_{oj}}. \tag{18}$$

E Data Appendix

E.1 Sample Definiton and Industry-Occupation Classifications

E.1.1 Sample Definition

1. We restrict the sample between the age of 18 to 65 years old.
2. For a small share of observations, we do observe the industry and occupation of their current/most recent job, but they are coded as "unemployed". This is mostly due to the recall periods differing on the survey. We therefore set them to "employed".
3. We drop those individuals in school, prision, disabled, ill, etc.
   keep if empstat$\geq$*1 & empstat* ≤ 3
   drop if empstatd$\geq$*320 & empstatd*$<$*370* .
4. In case a specific survey does not have such detailed information on the employment status, we adjust the number of inactive people using the share by gender of inactive individuals in school, prision, disabled, etc. from the closest survey with data available.
5. We classify all individuals who are unemployed or out of the workforce in the "home sector" i.e., *empstat == 2 — empstat == 3*
6. Table D1 provides the classification of education categories into years of education.

Table D1: Classification of Education

Code	Education	Years
0	NIU (not in universe)	NA
100	Less than primary completed	2
110	No schooling	1
120	Some primary	3
130	Primary (4 years)	4
211	Primary (5 years)	5
212	Primary (6 years)	6
221	General and unspecified track	9
222	Technical track	9
311	General track completed	12
312	Some college/university	14
320	Technical track	14
321	Secondary technical degree	12
322	Post-secondary technical education	14
400	University Completed	16
999	Unknown/Missing	NA

E.1.2 Industry and Occupation Classifications

1. Tables D2 and D3 provide the classification of industries and occupations

2. We re-classify some occupations since they are sparsely represented in industries (for example, professionals in agriculture). Currently, we are using the following:

   - Professionals and Clerks in Agriculture are assigned to Services:
     ind = 4 if ind == 2 occ≥1 occ ≤ 3

   - Agri Fisheries in Manufacturing reassigned to Agriculture
     ind = 2 if ind == 3 occ==4

   - Crafts/Trade Workers & Plant & Machine Operators in Agriculture re-assigned to Manufacturing i.e., *ind = 3 if ind == 2 occ≥ 5 occ ≤ 6*

3. Some individuals report an occupation but not the industry. Where a clear mapping exists, we classify them in the correct industry.
   Agriculture: *ind = 2 if missing(indgen) occisco == 4*
   Manufacturing: *ind = 3 if missing(indgen) & (occisco == 5 — occisco == 6)*
   Services: *ind = 4 if missing(indgen) & occisco ≤ 3*

Table D2: Classification of Industry Codes in IPUMS

Code	Industry	Classification
10	Agriculture, fishing, and forestry	Agriculture (1)
20	Mining and extraction	Manufacturing (2)
30	Manufacturing	Manufacturing (2)
40	Electricity, gas, water and waste management	Manufacturing (2)
50	Construction	Manufacturing (2)
60	Wholesale and retail trade	Services (3)
70	Hotels and restaurants	Services (3)
80	Transportation, storage, and communications	Services (3)
90	Financial services and insurance	Services (3)
100	Public administration and defense	Services (3)
110	Services, not specified	Services (3)
111	Business services and real estate	Services (3)
112	Education	Services (3)
113	Health and social work	Services (3)
114	Other services	Services (3)
120	Private household services	Services (3)
130	Other industry, n.e.c.	Services (3)
998	Response suppressed	NA
999	Unknown	NA

Table D3: Classification of Occupations

Code	Occupation	Classification
1	Legislators, senior officials and managers	Professional (1)
2	Professionals	Professional (1)
3	Technicians and associate professionals	Professional (1)
4	Clerks	Clerks (2)
5	Service workers and shop and market sales	Services Workers (3)
6	Skilled agricultural and fishery workers	Skilled Agri. (4)
7	Crafts and related trades workers	Craft/Trade Wrkrs (5)
8	Plant and machine operators and assemblers	Plant & Machine (6)
9	Elementary occupations	Elementary (7)
10	Armed forces	Drop
11	Other occupations, unspecified or n.e.c.	Drop
97	Response suppressed	Drop
98	Unknown	Drop
99	NIU (not in universe)	Drop

E.2 Calculating Hourly Wages

We discuss the calculation of hourly in three steps. First, we define the different ways in which income is measured in the IPUMS data. We then discuss the measurement of hours worked. Lastly, we discuss the available of the income and hours measurement specific to the countries in our sample. Usually, within a country, the measurement does not change over time.

E.2.1 Measurement of Income Earned

There are three types of income variables available in the IPUMS:

1. **INCTOT:** reports the person's total personal income from all sources in the previous month or year.
2. **INCEARN:** reports the person's total income from their labor (from wages, a business, or a farm) in the previous month or year.
3. **INCWAGE:** reports the respondent's weekly, monthly or annual wage and salary income.

INCTOT is most commonly available across almost all country-years. Therefore, to maintain consistency across the definition of income in our sample, we use INCTOT where available, even if others are available. All income variables are reported in local currency units, with varying frequency (as we will discuss below). We remove extreme outliers in the income distribution (above 9 million LCU).

E.2.2 Measurement of Hours Worked

There are three types of weekly worked hours variables available in IPUMS:

1. **HRSWORK1:** reports the person's total worked hours per week at all jobs. It combines information of the following two variables.
2. **HRSUSUAL1:** reports the usual number of hours the person works in a typical week across all jobs.
3. **HRSACTUAL1:** reports the person's actual number of hours worked per week at all jobs.

Each variable has an corresponding categorized variable with the suffix 2 (ex.: HRSWORK2). Whenever we do not have the uncategorized variable but we do have the categorized one, we use the midpoint of the categories.

1. We use the variable HRSWORK1 or the midpoints of HRSWORK2 categories as much as possible. If none of those variables are present, we use either usual or actual hours depending on the case.
2. We trim the sample at 100 hours.

3. When hours worked are missing for an individual, but available for the country-year sample, we replace it by the gender-industry-occupation average within that country-year.
4. In case no information is available on hours worked for a specific survey, we replace it by the gender-industry-occupation average of the closest year.
5. In cases where income is available at the monthly or annual frequency, we assume an individual works for 4 weeks/month and 52 months/year. In some cases (USA and Canada) we do observe the number of months worked, which we use to calculate the wages.

E.2.3 Countrywise Availability of Income and Hours

Brazil: 1970, 1980, 1991, 2000, 2010

1. Income: INCTOT and INCEARN from all sources in the previous month are available. As discussed earlier, to be consistent across countries, we use INCTOT whenever reported by the individual. If not, we replace it by INCEARN.
2. Hours worked: HRSWORK1 available for 1991, 2000 and 2010. HRSWORK2 availabe for 1980. For 1970 we use the gender-industry-occupation averages of the 1980 survey.
3. Wage = Income/(4*Hrs Worked)

Canada: 1971, 1981, 1991, 2001, 2011

1. Income: INCTOT, INCEARN, INCWAGE and INCSELF are available. INCEARN = INCWAGE + INCSELF. Like previously, we use INCTOT when reported and INCEARN in case it is missing.
2. Hours worked: HRSWORK1 available 1981 onwards, HRSUSUAL2 available for 1971. Moreover MONTHSWRK is also available, which we use to construct the number of months worked by the individual
3. Wage = Income/(4*Hrs Worked*Months Worked)

India: 1983, 1987, 1993, 1999, 2004, 2009, 2011, 2018

1. Since there is no information on hours worked for any year within IPUMS, we replace the source of data with the Indian Employment-Unemployment Survey (EUS) until 2009 and the Periodic Labor Force Survey (PLFS) for 2018. Both surveys have been harmonized the World Bnak's Global Labor Database project. (GLD)
2. Hours worked GLD: The variable records the hours of work last week for the individual's main job.
3. Wage GLD: Wage income is defined as "Last wage payment, primary job, excl. bonuses, etc (7-day ref period)". An additional variable records the frequency of payment, allowing us to convert it to weekly wages and then to hourly wages.

Indonesia: 1976, 1995, 2015, 2018

1. Income: INCWAGE from the previous month is available for both the years.
2. Hours worked: HRSWORKED1 available for 1995 and HRSACTUAL1 available for 1976.
3. Wage = Income/(4*Hrs Worked)
4. For 2015 and 2018, we complement the IPUMS data for Indonesia with data from the SAKERNAS survey. We pool data from 2013-2016 (2014 missing) and 2017-2019 respectively to increase the sample size.
5. Wages and hours worked in SAKERNAS: Wage income is available for 2013, 2015, 2016, 2017, 2018, 2019 and is defined as "Last wage payment, primary job, excl. bonuses, etc (7-day ref period)". The SAKERNAS further reports hours worked in the last week, which allows us to compute hourly wages.

Mexico: 1960, 1970, 1990, 1995, 2000, 2005, 2010, 2015

1. Income: INCTOT and INCEARN in the previous month. Not available for 2005.
2. Hours worked: HRSWORKED1 available for 1990, 1995, 2000, 2010. For 1960 and 1970 we use the gender-industry-occupation averages of the 1990 survey. For 2015 we use the gender-industry-occupation average of the 2010 survey.
3. Wage = Income/(4*Hrs Worked)

USA: 1960, 1970, 1980, 1990, 2000, 2005, 2010, 2015

1. Income: INCTOT and INCWAGE in the previous year available for all years. INCWAGE and INCSELF are also available, but I have checked that INCEARN = INCWAGE + INCSELF, so we don't need these two variables separately.
2. Hours worked: HRSWORK1 Available 1980 onwards. HRSWORK2 available for 1960 and 1970. MONTHSWRK is also available for all years
3. Wage = Income/(4*Hrs Worked*Months Worked)

E.2.4 Home Sector & Real Wages

1. We trim the wage earnings <u>within</u> each country-year-industry-occupation-gender at the 1st and 95th percentile.
2. We impute the gender-specific wages for the "home sector" using the average wages in elementary occupations in the services sector <u>within</u> each country-year.

3. We set the returns to ability (κ) at home to be equal to 1 across both men and women.
4. We use the exchange rates (LCU/USD) and real GDP at current and constant prices from the World Bank data to convert all wages in LCU to real 2010 USD as follows: $w_{USD} = \frac{w_{LCU}}{ExchangeRate} \times \frac{GDP_{Constant\,2010}}{GDP_{Current}}$.
5. Lastly, while aggregating, we use the person weights provided by the sample surveys to make the estimates representative of the population in that country-year.