

# Allocation of Female Talent and Cross-Country Productivity Differences\*

Munseob Lee<sup>†</sup>

University of California San Diego

January 2022

## Abstract

The disparities in cross-country labor productivity are greater in agriculture than in other industries. I propose that the misallocation of female talent across sectors distorts productivity. I formalize the theory by using a general equilibrium Roy model with gender-specific frictions. If female workers experience higher frictions in non-agricultural sectors, then unproductive female workers will select into agricultural sector. From a sample of 66 countries, I find that low-income countries have higher frictions in non-agricultural industries. By setting frictions to US levels, agricultural labor productivity increases by 16.6 percent and GDP per capita increases by 1.4 percent, primarily in low-income countries.

Keywords: misallocation, gender, occupational choice, productivity

JEL Codes: O11, O13, O47

---

\*For guidance and support on this project, I thank Chang-Tai Hsieh, Erik Hurst, and Nancy Stokey. The study benefited from comments and suggestions from many others, including David Argente, Chanont Banterghansa, Bong Geun Choi, Kerem Cosar, Kevin Donovan, Jorge L. Garcia, Joseph Kaboski, David Lagakos, Ralph Ossa, Sebastian Sotelo, and seminar participants at the University of Chicago, University of Notre Dame, University of California San Diego, Korea University, IEA Annual Meeting, Midwest Macro Meeting, Econometric Society World Congress, Pacific Conference for Development Economics, and STEG Workshop. The author wishes to acknowledge the statistical offices that provided the underlying data through the IPUMS project that made this research possible. The author thanks Minki Kim and Churn Ken Lee for providing excellent research assistance.

<sup>†</sup>Email: [munseoble@ucsd.edu](mailto:munseoble@ucsd.edu). Address: 9500 Gilman Dr #0519, La Jolla, California 92093-0519.

# 1 Introduction

The disparities in labor productivity among countries are much larger in agriculture than in other industries (Caselli, 2005; Restuccia, Yang and Zhu, 2008). Moreover, poor countries allocate larger percentages of employment in agriculture. Therefore, agricultural productivity is important for understanding international income differences. One explanation for the agricultural productivity gap is that in comparison to developed countries, developing countries use much less capital per worker in agriculture, and modestly less capital per worker in other industries. However, even after controlling for these variables, large and unexplained productivity disparities in agriculture persist (Gollin, Lagakos and Waugh, 2014a,b).

This study proposes a new explanation for the substantial agricultural productivity gap across countries. Typically, women are underrepresented more severely in high-skilled jobs in poorer countries. Gender inequality might reflect various labor-market frictions that women experience, such as gender discrimination in the labor market and higher costs of labor market participation or entrepreneurship for women. If women experience higher frictions in non-agricultural versus agricultural industries, then some women who are better skilled at non-agricultural jobs might still choose to work in agriculture. Agricultural productivity measured by the outcome per worker decreases because of the misallocation of female talent, but the productivity in other industries increases because only women who are sufficiently talented to overcome frictions enter the workforce in those industries. If poor countries have higher barriers against women in non-agricultural industries, then agricultural productivity might be much lower than that in richer countries, while productivity in non-agriculture is modestly lower.

To validate this hypothesis, I formalize a theory in a two-sector general equilibrium model. The theory has three ingredients. The first is heterogeneity of workers in sectoral productivities. Workers choose where to supply their labor based on their comparative advantage. This is a Roy (1951) model of self-selection. The second is group and sector specific frictions in the labor market. I model these frictions as a gender-sector-specific tax on wage, which creates a wedge between a worker's marginal product and her earnings. The wedge can be interpreted as taste-based discrimination as in Becker (1957). If there is a disamenity value to employing minority workers, minority workers have to compensate employers by accepting a lower wage for identical productivity. The third is gender-specific labor supply. I model utility as separable in consumption and leisure, as in Ohanian, Raffo and Rogerson (2008), and I match U-shaped female labor force participation (Goldin, 1995) by calibrating gender-specific leisure preference for each country.

To solve the model analytically, I assume that workers' productivity draws follow an

extreme value distribution. This assumption allows us to have a closed-form expression on each group’s propensity to enter each sector. Workers are misallocated across sectors when frictions in the two sectors are not equal. Therefore, frictions specific to each group and sector show up in each group’s occupational choice. Using micro-level Census data from a sample of 66 countries, I back out the gender-specific frictions in agriculture and non-agriculture for each country.

The model indicates that countries with higher frictions against women in non-agricultural industries have lower labor productivity in the agricultural sector. To quantify the magnitude of productivity loss attributed to this misallocation of female talent, I calibrate and simulate a general equilibrium model. Introducing the United States as a benchmark economy, I first calibrate the parameters in the model to the benchmark economy. I then compare all of other countries to the overall productivity and frictions in the benchmark. The calibration and estimation of the sample countries indicate that poorer countries have higher frictions against women in non-agricultural industries. The results from the simulation, given the calibrated parameters, show that the misallocation of female talent explains a substantial proportion of the agricultural productivity gap in poorer countries. I run a counter-factual analysis by setting gender-specific frictions to the US level for all countries. I find that agricultural labor productivity increases by 16.6 percent and GDP per capita increases by 1.4 percent, on average, for the most preferred calibration. Gains in productivity and overall income come primarily from poor countries. The estimated gains are slightly higher when I allow for a positive correlation in workers’ productivity draws across the two sectors.

The origins of the differences in occupational choice across gender could be very diverse. There are several examples of barriers discussed in the literature and in the policy community. For example, World Bank’s 2012 World Development Report, “Gender Equality and Development” identifies four areas to fill the gaps: (i) human capital, (ii) economic opportunities, (iii) access to finance, and (iv) empowerment (World Bank, 2011). First, greater control over household resources by women leads to more investment in children’s human capital, with dynamic positive effects on economic growth (Doepke and Tertilt, 2019). Second, gender differences in access to economic opportunities are driven in part by differences in time use that result from deep-rooted norms for the distribution of responsibility for care and housework (Jayachandran, 2021). Third, access to finance and information constrains women’s opportunities to be entrepreneurs (Chiplunkar and Goldberg, 2021). Fourth, women’s empowerment helps them to transit from subsistence farming to well-paid jobs with long-term contracts.

**Related Literature.** Many empirical studies examine the two-way relation between gender inequality and economic growth (Goldin, 1990; Galor and Weil, 1996; Lagerlöf, 2003; Doepke and Tertilt, 2009). They reach the consensus that gender inequality hinders growth, and economic growth improves gender equality.<sup>1</sup> This study relates to the first causality, and especially to the recent literature that quantifies the effect of gender inequality on economic growth through improvements in female labor allocation. Esteve-Volart (2009) and Cuberes and Teignier (2016) model gender inequality as the exclusion of women from the labor market and managerial positions. These studies argue that both types of inequality hinder economic growth. However, they do not relate gender inequality to sectoral productivity. My study is the first to explain cross-country, sectoral productivity differences through misallocation of female talent.<sup>2</sup>

This study relates closely to the recent literature on cross-country agricultural productivity gaps. The explanations in the literature include distortions that limit farm size (Adamopoulos and Restuccia, 2014), intermediate input decisions (Donovan, 2021), barriers that limit specialization through trade (Tombe, 2015), capital-embodied technology (Caunedo and Keller, 2021), and the prevalence of untitled land (Chen, 2017; Gottlieb and Grobovšek, 2019). All of these explanations are complementary to the one presented here, but none consider gender-specific frictions in the labor market as a channel. Additionally, Adamopoulos, Brandt, Leight and Restuccia (2020) show that land institutions in rural China that disproportionately constrain the more productive farmers reduced aggregate agricultural productivity in China. My study also finds that an institutional constraint, gender inequality, affects two margins: misallocation and selection. This reduces aggregate agricultural productivity in developing countries.

Furthermore, the general equilibrium model that I present is closely related to the work of Lagakos and Waugh (2013), Ohanian, Raffo and Rogerson (2008), Burstein, Morales and Vogel (2019) and Hsieh, Hurst, Jones and Klenow (2019). The assumptions on preference and technology are similar to Lagakos and Waugh (2013), who build a selection model with two sectors to explain the cross-country gap in agricultural productivity. I add separable leisure in the utility function as in Ohanian, Raffo and Rogerson (2008). Then, I introduce gender-specific choice of occupation like in Hsieh, Hurst, Jones and Klenow (2019), who measure the aggregate productivity effects of the misallocation of talent among women and blacks in

---

<sup>1</sup>See World Bank (2011) for a comprehensive review of the literature and policy implications on this topic.

<sup>2</sup>See Rodríguez Mora (2009) for a nontechnical review of the sources and consequences of misallocated talent. Broadly, the current study relates to many papers on between-group inequality in labor economics. The literature identifies three channels that generate group inequality: discrimination through tastes (Becker, 1957), statistical discrimination (Phelps, 1972; Arrow, 1973), and segregation (Loury, 1977; Benabou, 1996; Durlauf, 1996).

the US. The present study complements [Lagakos and Waugh \(2013\)](#) with an emphasis on the allocation of female talent and provides counterfactual results by setting gender-specific frictions to the current US level for all countries.<sup>3</sup>

The rest of this paper is structured as follows: Section 2 presents the motivating facts on productivity and gender composition; Section 3 presents a theoretical model and proofs; in Section 4 I calibrate and estimate the model; in Section 5 I simulate the model and run counter-factual exercises; Section 6 discusses additional concerns and robustness; and Section 7 concludes.

## 2 Motivating Evidence

Three facts motivate this study: (i) the disparities in cross-country labor productivity are much larger in agricultural than in non-agricultural industries, (ii) women constitute a larger proportion of the agricultural workforce in developing countries, and (iii) female labor force participation rate first declines and then rises as countries develop.

Panel (a) of Figure 1 reproduces a finding from [Caselli \(2005\)](#).<sup>4</sup> Panel (a) of Figure 1 shows that relative agricultural labor productivity generally increases with overall income of countries. In the United States, agricultural labor productivity is 0.84, so the agricultural sector is 16 percent less productive than non-agricultural industries. In less-developed countries, agricultural productivity is much lower. For example, in Tanzania, agricultural productivity is 0.07, so agriculture is 93 percent less productive.

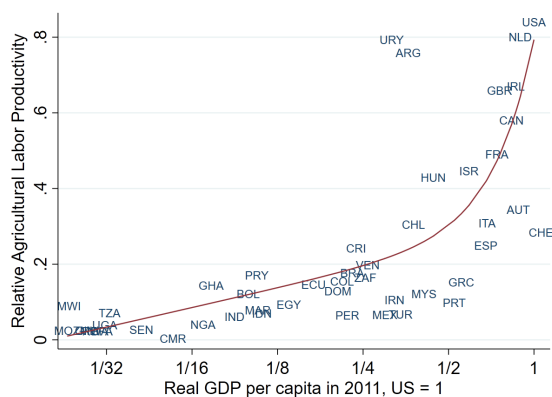
To measure the share of female workers in agriculture, I use data from the Integrated Public Use Microdata Series, International (IPUMS-International) provided by [Minnesota Population Center \(2020\)](#). The IPUMS-International contains micro-level census data from 81 countries. For consistency, I calculate the share of female workers from the 66 countries in the main analysis. Panel (b) of Figure 1 shows that low income countries have higher shares of female workers in agricultural sector. For example, in the United States 23 percent of the workers are women. One may wonder if women are over-represented in all occupations. Panel (c) of Figure 1 shows that low income countries have lower shares of female workers in non-agricultural sector. Disparities in sectoral productivity and gender composition across countries motivate the study with a gender-specific occupational choice model.

---

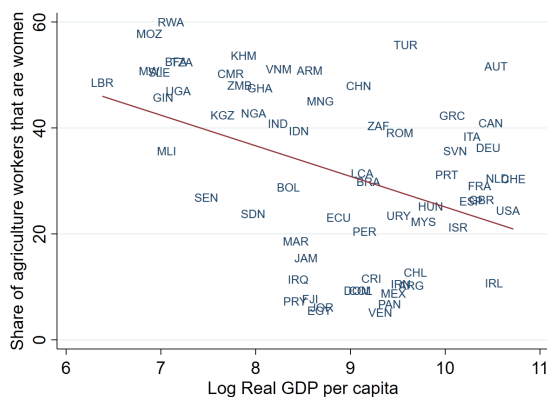
<sup>3</sup>The counterfactual exercise in [Hsieh, Hurst, Jones and Klenow \(2019\)](#) keeps friction against minorities constant at the level of 1960. They find that 15 to 20 percent of the growth in aggregate output per worker between 1960 and 2008 could be explained by the improved allocation of talent.

<sup>4</sup>[Caselli \(2005\)](#) documents this fact using a sample of 80 countries. From this sample, 47 countries match those in our sample of 66 countries and are represented in the figure, for consistency. Refer to Table A.I in Appendix A for the list of countries.

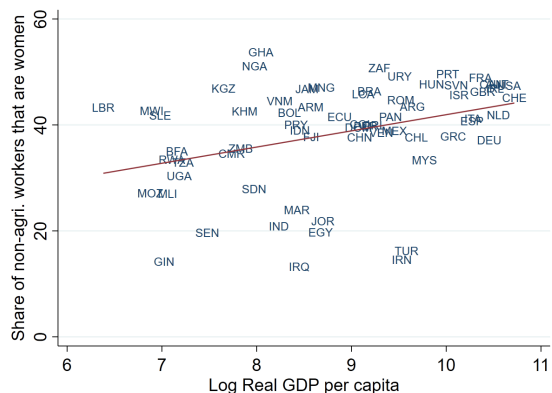
Figure 1: Agricultural Labor Productivity and Share of Agricultural/Non-agricultural Workers that are Women



(a) Relative Agricultural Labor Productivity



(b) Share of Agricultural Workers that are Women

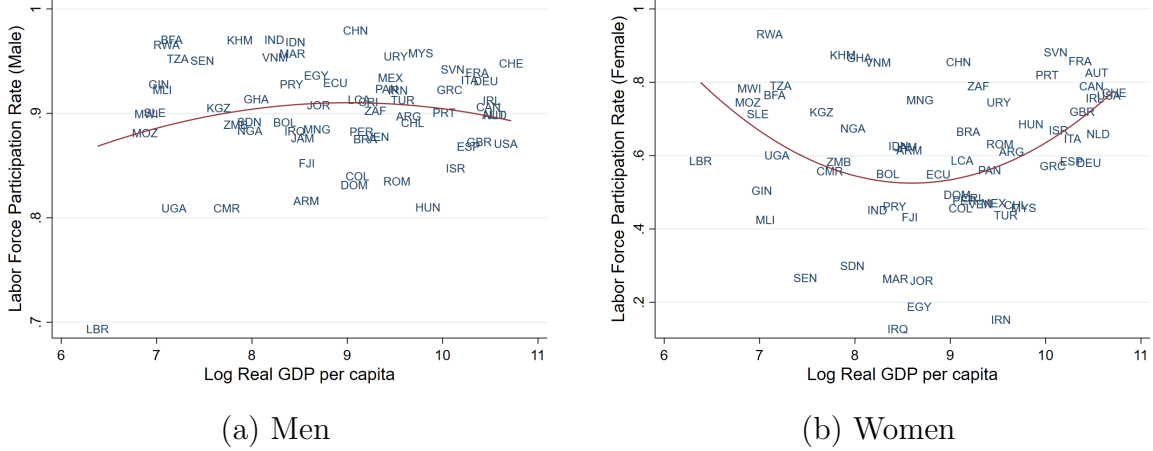


(c) Share of Non-agricultural Workers that are Women

Notes: Panel (a) shows relative agricultural labor productivity, which is labor productivity in agriculture divided by labor productivity in non-agricultural industries, from Caselli (2005). The real GDP per capita in 2011 from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015) is normalized so that the United States' value is one. The polynomial fitted line over 47 countries shows an increasing pattern in relative agricultural labor productivity. Panel (b) plots the share of agricultural workers that are women for 66 countries calculated from IPUMS-International (Minnesota Population Center, 2020). The fitted linear line has a slope of -5.77 that is significant at the 1 percent level. Panel (c) plots the share of non-agricultural workers that are women for 66 countries calculated from IPUMS-International (Minnesota Population Center, 2020). The fitted linear line has a slope of 3.05 that is significant at the 1 percent level.

Another important margin in the labor market is labor force participation decision, which also differs across gender. Figure 2 shows well-known U-shaped female labor force participation and inverted U-shaped male labor force participation over the process of economic development (Goldin, 1995). This non-monotone pattern might not explain monotone increase in relative agricultural productivity (Panel (a) of Figure 1). However, I will match

Figure 2: Labor Force Participation Rates, Men and Women



Notes: The graph plots the labor force participation rate for the 66 countries, calculated from the IPUMS-International (Minnesota Population Center, 2020).

this fact on labor force participation in the model, in order to control potential selection effects that could come from differences in labor force participation rates and quantify the role of gender-specific frictions more precisely.

### 3 Baseline Model

#### 3.1 Preferences and Endowments

Suppose a world where half of population is men and another half of population is women. Men and women differ in two dimensions. First, they differ in productivity in each sector: agricultural and non-agricultural. Second, they differ in the degree of disutility they feel from supplying labor. I set utility separable in consumption and leisure as [Ohanian, Raffo and Rogerson \(2008\)](#), [Restuccia and Vandenbroucke \(2014\)](#) and [Bick, Fuchs-Schündeln and Lagakos \(2018\)](#). The preference of individual  $i$  with gender  $g$  is:

$$U^i = \log c_a^i + \nu \log(c_n^i + \bar{n}) + \psi_g \log l^i \quad (1)$$

where  $c_a^i, c_n^i, l^i$  are consumption of the agricultural good, consumption of the non-agricultural good, and leisure, respectively. Parameter  $\nu$  is the taste for non-agricultural consumption, and  $\bar{n}$  is a parameter that represents non-homotheticity in agricultural and non-agricultural consumption, where  $\bar{n} > 0$  implies the marginal utility of consuming agricultural good is always higher than the marginal utility of consuming non-agricultural good. This preference

ensures Engel’s Law holds.  $\psi_g$  represents relative utility of leisure. I make it gender-specific to match different labor force participation rates over the process of economic development (Figure 2).

Each worker is endowed with one unit of time and productivity in each industry,  $(\epsilon_{ag}^i, \epsilon_{ng}^i)$ , where  $\epsilon_{ag}^i$  is the productivity in agriculture and  $\epsilon_{ng}^i$  is the productivity in non-agriculture that depends on the gender  $g \in \{m, w\}$  ( $m$  is male and  $w$  is female). He or she chooses where to work first and then decides the amount of labor  $h^i$  to supply in the chosen industry, given the income from the chosen industry  $y^{i*}$ . The budget constraint is:

$$P_a c_a^i + c_n^i \leq h^i y^{i*} = (1 - \tau_{sg}) w_s h^i \epsilon_{sg}^i \quad (2)$$

where  $P_a$  is the relative price of an agricultural good, and the non-agricultural good is numeraire. The maximized income, after choosing industry  $s \in \{a, n\}$  ( $a$  is agriculture and  $n$  is non-agriculture), is a function of friction  $\tau_{sg}$ , wage  $w_s$ , and individual productivity  $\epsilon_{sg}^i$ . The friction can be any type of gender-specific barrier. For example, it might be preference-based discrimination by an employer. When an employer discriminates against employees, he or she acts as if he or she incurs non-pecuniary and psychic costs of production by employing them (Becker, 1957). I assume no friction against men in both industries, and therefore frictions against women in both industries represent relative barriers women experience in comparison to men. These frictions are realized as a government tax or subsidy in my model.

**Assumption 1.** (Gender-specific Frictions) Only women experience non-zero frictions in either industry, i.e.,  $\tau_{nm} = \tau_{am} = 0$  and  $\tau_{nw} \neq 0 \neq \tau_{aw}$ , and these frictions are country-specific. These frictions are set to be symmetric such that all revenue from taxes in one sector goes to the other sector as a subsidy.

Borrowing from McFadden (1974), Eaton and Kortum (2002), and Hsieh, Hurst, Jones and Klenow (2019), a worker’s idiosyncratic productivities are assumed to come from bivariate Fréchet distributions. The shape parameter  $\theta$  and the correlation parameter  $\rho$  are assumed the same across industries and genders.<sup>5</sup> I assume location parameters  $T_{sg}$  are different across industries and genders. This difference reflects the fact that men and women might have different average productivity in each industry. For example, a large body of literature finds women are less productive at agricultural work than men (Goldin and Sokoloff,

---

<sup>5</sup>After Eaton and Kortum (2002), researchers assume an independent joint Fréchet distribution because it is observationally equivalent to a joint distribution, which embeds a correlation (see footnote 14 in Eaton and Kortum (2002)). A few exceptions include Ramondo and Rodríguez-Clare (2013) and Lind and Ramondo (2021). Because individual talents across two industries might be correlated, I begin with a general case where I allow a non-zero correlation. In quantitative exercises, I begin with benchmark case  $\rho = 0$  and then assess the robustness of the results with a non-zero correlation.

1984; Pitt, Rosenzweig and Hassan, 2012).

**Assumption 2.** (Talent Distribution) Productivities come from bivariate Fréchet distributions (i.e.  $F_g(\epsilon_a, \epsilon_n) = \exp(-[\sum_{s \in \{a, n\}} (T_{sg} \epsilon_s^{-\theta})^{1/(1-\rho)}]^{1-\rho})$ ). The distribution is assumed to be the same across countries.

The solution to the consumers' utility maximization problem has two regimes that switches depending on the level of their wages in their chosen industry,  $y^{i*}$ . Intuitively, at very low levels of  $y^{i*}$ , the marginal utility of  $c_a^i$  is greater than the marginal utility of  $c_n^i$  even when  $c_n^i = 0$  and all income is spent on purchasing  $c_a^i$ . This is due to the presence of the  $\bar{n} > 0$  term. In this low wage regime, consumers spend all their income on purchasing  $c_a^i$ , and  $c_n^i = 0$ . As wages increase, the marginal utility of  $c_a^i$  decreases until it eventually drops below the marginal utility of  $c_n^i$  with  $c_n^i = 0$ . At that point, it makes sense for the consumer to purchase positive quantities of both  $c_a^i$  and  $c_n^i$ . The cutoff for  $y^{i*}$  between these two regimes is  $\frac{\bar{n}}{\nu}(1 + \psi)$ , and the optimal consumption demand and labor supply is as follows.<sup>6</sup>

**Case 1:**  $y^{i*} \leq \frac{\bar{n}}{\nu}(1 + \psi)$

$$c_a^{i*} = \frac{y^{i*}}{P_a(\psi_g + 1)} \quad (3)$$

$$c_n^{i*} = 0 \quad (4)$$

$$l^{i*} = \frac{\psi_g}{\psi_g + 1} \quad (5)$$

**Case 2:**  $y^{i*} > \frac{\bar{n}}{\nu}(1 + \psi)$

$$c_a^{i*} = \frac{\bar{n} + y^{i*}}{P_a(1 + \nu + \psi_g)} \quad (6)$$

$$c_n^{i*} = \frac{\nu y^{i*} - \bar{n}(1 + \psi_g)}{1 + \nu + \psi_g} \quad (7)$$

$$l^{i*} = \frac{\psi_g(\bar{n} + y^{i*})}{y^{i*}(1 + \nu + \psi_g)} \quad (8)$$

## 3.2 Production

I assume that representative firms in each industry produce aggregate output  $Y_a$  and  $Y_n$  by hiring labor. The technologies are:

---

<sup>6</sup>Refer to Appendix B for derivation.

$$Y_a = AL_a \text{ and } Y_n = AL_n \quad (9)$$

where  $A$  is the exogenous country-specific productivity, and  $L_a$  and  $L_n$  represent the average effective labor units employed in the two industries. These are:

$$L_a = \sum_{g \in \{m,w\}} q_g p_{ag} E[h_{ag} \epsilon_{ag} | \text{Person chooses } a] \quad (10)$$

$$L_n = \sum_{g \in \{m,w\}} q_g p_{ng} E[h_{ng} \epsilon_{ng} | \text{Person chooses } n] \quad (11)$$

where  $q_g$  is the total number of working people in gender  $g$ . The  $p_{sg}$  is the fraction of people in gender  $g$  that work in industry  $s$ . The meaning of the average effective labor units becomes clearer after introducing the total working hours in each industry, defined as:

$$N_a = \sum_{g \in \{m,w\}} q_g p_{ag} E[h_{ag} | \text{Person chooses } a]$$

$$\begin{aligned} N_n &= \sum_{g \in \{m,w\}} q_g p_{ng} E[h_{ng} | \text{Person chooses } n] \\ &= \sum_{g \in \{m,w\}} q_g (1 - p_{ag}) E[h_{ng} | \text{Person chooses } n] \end{aligned} \quad (12)$$

Firms maximize profits in competitive markets, so wages per efficiency unit of labor are:

$$w_a = P_a A \text{ and } w_n = A \quad (13)$$

Recall that  $P_a$  is the relative price of an agricultural good.

### 3.3 Occupational Choice

The problem of occupational choice reduces to choosing the industry that delivers the highest income. For men, the maximized income under full labor supply is:

$$y_m^{i*} \equiv \max\{w_a \epsilon_{am}^i, w_n \epsilon_{nm}^i\} = \max\{P_a A \epsilon_{am}^i, A \epsilon_{nm}^i\}$$

Men choose to work in non-agriculture if  $\frac{\epsilon_{nm}^i}{\epsilon_{am}^i} \geq P_a$ ; workers who do not experience frictions choose an occupation based on their own relative productivity and the price of goods in the market. For women, the maximized income under full labor supply is a function of

frictions:

$$\begin{aligned} y_w^{i*} &\equiv \max\{(1 - \tau_{aw})w_a\epsilon_{aw}^i, (1 - \tau_{nw})w_n\epsilon_{nm}^i\} \\ &= \max\{(1 - \tau_{aw})P_a A\epsilon_{aw}^i, (1 - \tau_{nw})A\epsilon_{nw}^i\} \end{aligned} \quad (14)$$

Therefore, women choose agriculture if  $\frac{\epsilon_{nw}^i}{\epsilon_{aw}^i} \geq P_a \frac{1 - \tau_{aw}}{1 - \tau_{nw}}$ . I define  $\tilde{\tau}_{n|a}$  as a relative friction in non-agricultural industries:

$$\tilde{\tau}_{n|a} \equiv \frac{1 - \tau_{aw}}{1 - \tau_{nw}} \quad (15)$$

that is a measure for the size of friction in non-agriculture relative to agriculture. Women choose non-agriculture when productivity in the industry is sufficiently high to offset the difference in price and friction. If price  $P_a$  and the relative tax in non-agriculture  $\tilde{\tau}_{n|a}$  are high, then more women enter agriculture.

I offer three propositions that relate closely to Propositions 1 through 3 from [Hsieh, Hurst, Jones and Klenow \(2019\)](#). My economy is simpler because I examine only two industries and do not incorporate the accumulation of human capital. However, this economy incorporates two outputs from both industries, so relative price  $P_a$ , which is determined endogenously in the model, is crucial to occupational choice. Proofs for the propositions appear in [Appendix B](#). Note that some comparative statics in the following propositions with respect to frictions are only strictly true in partial equilibrium. Later, counter-factual exercises will make quantitative statements incorporating general equilibrium effects.

**Proposition 1 (Propensity):** *The  $p_{sg}$  denotes the fraction of people in gender  $g$  that choose industry  $s$ . The propensities for men and women in each industry are:*

$$\begin{aligned} p_{am} &= \frac{(P_a^\theta T_{am})^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}} \text{ and } p_{nm} = 1 - p_{am} \\ p_{aw} &= \frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)}}{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}} \text{ and } p_{nw} = 1 - p_{aw} \end{aligned}$$

*Therefore, the propensities of men and women for agriculture increase with relative price  $P_a$ . The propensity of women for agriculture increases with the relative friction in non-agriculture,  $\tilde{\tau}_{n|a}$ .*

The propensity of men is determined optimally by the consumer and firm's maximization problems, but frictions distort the propensity for women. Women are more likely to choose an industry in which the barriers are low. This Roy-type model generates the average quality

of workers in an industry for each group, and the following proposition demonstrates the quality.

**Proposition 2 (Average Quality):** *The  $E[\epsilon_{sg}|\text{Person chooses } s]$  denotes the average quality of workers for gender  $g$  in industry  $s$ . For notational convenience, I omit the conditional term. The average quality of workers for gender  $g$  in industry  $s$  is:*

$$E[\epsilon_{am}] = \left(\frac{T_{am}}{p_{am}^{1-\rho}}\right)^{1/\theta} \Gamma\left(1 - \frac{1}{\theta}\right) \text{ and } E[\epsilon_{nm}] = \left(\frac{T_{nm}}{p_{nm}^{1-\rho}}\right)^{1/\theta} \Gamma\left(1 - \frac{1}{\theta}\right)$$

$$E[\epsilon_{aw}] = \left(\frac{T_{aw}}{p_{aw}^{1-\rho}}\right)^{1/\theta} \Gamma\left(1 - \frac{1}{\theta}\right) \text{ and } E[\epsilon_{nw}] = \left(\frac{T_{nw}}{p_{nw}^{1-\rho}}\right)^{1/\theta} \Gamma\left(1 - \frac{1}{\theta}\right)$$

*Because the average quality relates inversely to propensity, the comparative statistics from Proposition 1 hold inversely. The average quality of women in agriculture decreases with the relative friction in non-agriculture,  $\tilde{\tau}_{n|a}$ .*

This proposition captures a selection effect. If the friction in non-agriculture is higher than in agriculture, then women are overrepresented in agriculture; female workers who are less talented in agriculture choose that industry, which is not optimal for efficient production. The next proposition provides equations for the identification of the frictions.

**Proposition 3 (Gender Wage Gap):** *The  $\overline{wage}_{sg}$  denotes the average wage in industry  $s$  by gender  $g$ , defined as:*

$$\overline{wage}_{sg} \equiv (1 - \tau_{sg})w_s E[\epsilon_{sg}|\text{Person chooses } s]$$

*Therefore, the gender wage gaps in agriculture and non-agriculture are:*

$$\frac{\overline{wage}_{am}}{\overline{wage}_{aw}} = \frac{1}{1 - \tau_{aw}} \left(\frac{p_{am}}{p_{aw}}\right)^{-(1-\rho)/\theta} \left(\frac{T_{am}}{T_{aw}}\right)^{1/\theta}$$

$$\frac{\overline{wage}_{nm}}{\overline{wage}_{nw}} = \frac{1}{1 - \tau_{nw}} \left(\frac{p_{nm}}{p_{nw}}\right)^{-(1-\rho)/\theta} \left(\frac{T_{nm}}{T_{nw}}\right)^{1/\theta}$$

*The two wage gaps are the same as follows:*

$$\frac{\overline{wage}_{am}}{\overline{wage}_{aw}} = \frac{\overline{wage}_{nm}}{\overline{wage}_{nw}} = \frac{(P_a^\theta T_{am})^{1/(1-\rho)} + (T_{nm})^{1/(1-\rho)}}{(P_a^\theta (1 - \tau_{aw})^\theta T_{am})^{1/(1-\rho)} + ((1 - \tau_{nw})^\theta T_{nm})^{1/(1-\rho)}} \quad (16)$$

Due to the selection effect, the gender wage gaps are the same for the two industries. The overall gender wage gap and the propensity for each industry are observable for many

countries. Assuming distributional parameters, I can back out the implied frictions from the data.

The productivity of each industry can be measured by the GDP per worker. The industry’s labor productivity relates closely to gender-specific frictions as:

**Proposition 4 (Sectoral Labor Productivity):** *The labor productivity in agriculture industry  $Y_a/N_a$  increases with  $\tau_{aw}$  and decreases with  $\tau_{nw}$ . Therefore, relative labor productivity in agriculture decreases with relative friction  $\tilde{\tau}_{n|a}$ .*

### 3.4 Government

The government maintains a balanced budget every period. The government imposes on each female worker the gender-occupation specific wage tax or subsidy. This tax and subsidy system is my model’s implementation of the gender-specific frictions. A wage tax corresponds to a positive friction ( $\tau > 0$ ), while a wage subsidy corresponds to a negative friction ( $\tau < 0$ ). The government’s period-by-period budget constraint is as follows<sup>7</sup>:

$$\sum_i \sum_s \sum_g \tau_{sg} w_s h^i \epsilon_{sg}^i = 0 \quad (17)$$

### 3.5 Equilibrium

Each worker takes prices and wages as given when she first chooses which sector to supply her labor and then maximizes her utility subject to her budget constraint. Representative firms in each industry also take output prices as given and maximize profits. Therefore, an equilibrium of the economy consists of a relative agricultural price,  $P_a$ , wages per efficiency unit of labor,  $(w_a, w_n)$ , and allocations for all workers, such that all workers optimize and both labor markets and output markets clear. Lastly, the government achieves balanced budget.

## 4 Model Calibration

For cross-country comparisons, I use the US economy as a benchmark. I assume no gender-specific frictions in the United States ( $\tau_{aw}^{\text{US}} = \tau_{nw}^{\text{US}} = 0$ ).<sup>8</sup> I calibrate the distributional param-

<sup>7</sup>There are many ways of achieving a balanced budget. For example, wage tax from female workers can be allocated to male workers as a subsidy. In this case, consumption and occupation choices by male workers will be directly affected by this subsidy. In order to minimize the distortion on choices by male workers, I allocate wage taxes and subsidies only among female workers across sectors.

<sup>8</sup>There might be positive, gender-specific frictions in the United States. This assumption is just a normalization for cross-country comparison and counter-factual exercise. Note that the literature still finds

eters  $\{T_{ig}, \theta, \rho\}$  and preference parameters  $\{\bar{n}, \nu, \psi_m, \psi_w\}$  from the US economy. Given those parameters, country-specific parameters,  $\{A^c, \tau_{aw}^c, \tau_{nw}^c, \psi_m^c, \psi_w^c\}$ , where c represents countries, are calibrated. I first calibrate parameters for individual talent distributions.

## 4.1 Calibration of Distributional Parameters

I assume individual talents follow a bivariate Fréchet distribution:

$$F_g(\epsilon_a, \epsilon_n) = \exp(- [\sum_{s \in \{a,n\}} (T_{sg} \epsilon_s^{-\theta})^{1/(1-\rho)}]^{1-\rho}) \quad (18)$$

where  $T_{sg}$  are location parameters,  $\theta$  is a shape parameter, and  $\rho$  is a correlation parameter.

I assume an independent talent distribution ( $\rho = 0$ ) as a benchmark case. In Appendix D, I report results with positive correlation in talent across sectors.

Two methods are used to calibrate the shape parameter  $\theta$ . The first method employs the property of extreme value distributions. I derive the following proposition that relates the distributional parameter to the observable moments. Proof for the following proposition is in Appendix B.

**Proposition 5 (Coefficient of Variation in Wages):** *The coefficient of variation in wages within an industry is:*

$$\frac{Std.Dev.}{Mean} = \left[ \frac{\Gamma(1 - \frac{2}{\theta})}{\Gamma(1 - \frac{1}{\theta})^2} - 1 \right]^{1/2}$$

where  $\Gamma$  is a Gamma function and  $\theta$  is a shape parameter of the Fréchet distribution.

I examine the wage dispersion within an industry by gender. I use micro-data from the 2010 American Community Survey (Ruggles, Flood, Foster, Goeken, Pacas, Schouweiler and Sobek, 2021). To calculate the reliable hourly wages, I only use employed individuals between the ages of 25 and 55 and exclude people serving in the armed forces. The sample comprises both full- and part-time workers, and I follow McGrattan and Rogerson (1998) to calculate the hourly wage from the census. I take the residuals from a regression of log wages on industry and gender dummy variables and calculate the mean and variance across workers by using the exponent of the wage residuals. The variance is discounted by a half

---

evidence of gender-specific frictions, mostly against women in many sectors these days. For example, among entrepreneurs, Hebert (2020) finds that female-founded start-ups are 27% less likely to raise external equity including venture capital. However, in female-dominated sectors, female-founded start-ups are no longer at a disadvantage. Among medical doctors, Sarsons (2019) finds that physicians become more pessimistic about a female surgeon's ability than a male's after a patient death, indicated by a sharper drop in referrals to the female surgeon.

to capture the permanent component. The point estimate for  $\theta$  is 3.5.<sup>9</sup>

Because the assumption on the benchmark economy is  $\tau_{aw}^{\text{US}} = \tau_{nw}^{\text{US}} = 0$ , given  $\rho = 0$  and  $\theta = 3.5$ , the ratios of the location parameters in each industry are identified by Proposition 3:

$$\begin{aligned}\frac{T_{am}}{T_{aw}} &= (1 - \tau_{aw}^{\text{US}})^{\theta} \left( \frac{\overline{wage}_m}{\overline{wage}_w} \right)^{\theta} \left( \frac{p_{am}}{p_{aw}} \right)^{1-\rho} = \left( \frac{\overline{wage}_m}{\overline{wage}_w} \right)^{3.5} \left( \frac{p_{am}}{p_{aw}} \right) \\ \frac{T_{nm}}{T_{nw}} &= (1 - \tau_{nw}^{\text{US}})^{\theta} \left( \frac{\overline{wage}_m}{\overline{wage}_w} \right)^{\theta} \left( \frac{p_{nm}}{p_{nw}} \right)^{1-\rho} = \left( \frac{\overline{wage}_m}{\overline{wage}_w} \right)^{3.5} \left( \frac{p_{nm}}{p_{nw}} \right)\end{aligned}\quad (19)$$

where the gender wage gap  $\frac{\overline{wage}_m}{\overline{wage}_w}$  and the propensity gaps  $\left\{ \frac{p_{am}}{p_{aw}}, \frac{p_{nm}}{p_{nw}} \right\}$  can be estimated from data.

The reasoning behind these equations is the following: Without friction, the relative mean productivity of male workers in each industry is high, if male workers are paid more (i.e., gender wage gap is high) or are overrepresented in the industry (i.e., propensity gap is high). I observe the right-hand side of these equations in the data, and therefore use it to back out the ratio of average talents.

Data on the gender wage gap comes from the Human Development Report (UNDP, 2015), except for the United States. Due to a lack of gender-level disaggregated income data, especially from the less-developed countries, female and male earned income are estimated crudely based on data regarding the ratio of female non-agricultural wages to male non-agricultural wages, female and male shares of an economically active population, total female and male populations, and GDP per capita in PPP US\$. The wage ratios in this calculation are based on data for the most recent year. Although the calculation is crude, I can use this wage gap because Proposition 3 suggests that wage gaps across industries are equal in equilibrium.

I also find that the gender earnings gaps from UNDP (2015) are consistent to estimates from micro-level Census data. For example, the gender earnings gap in the United States according to UNDP (2015) is 1.46 and Goldin (2014) also estimates the gender wage gap to be 1.39 from American Community Survey micro-data.

I calculate the propensity gaps in both sectors from the 2010 US census micro-data. Propensities now calculated as conditional probabilities; probability of selecting into a certain industry conditional on being in the labor force. Estimates are  $\{p_{am}, p_{aw}, p_{nm}, p_{nw}\} = \{2.07\%, 0.75\%, 97.93\%, 99.25\%\}$ . These estimates mean that a man enters agriculture with a probability of 2 percent and enters non-agriculture with a probability of 98 percent.

---

<sup>9</sup>This number is higher than Hsieh, Hurst, Jones and Klenow (2019) ( $\theta = 2$ ) and Burstein, Morales and Vogel (2019) ( $\theta = 1.8$ ), because I use only permanent component of wage variation. I assume that skills are equally dispersed by gender. If I use only men and women in the data, I get  $\theta$  for men equals to 3.63 and  $\theta$  for women equals to 3.38, which are close to the estimate from pooled sample 3.5.

Given the gender wage and propensity gaps in the United States, I have ratios of location parameters for agriculture and non-agriculture:

$$\begin{aligned}\frac{T_{am}}{T_{aw}} &= \left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{3.5} \left(\frac{p_{am}}{p_{aw}}\right) = (1.30)^{3.5} * \frac{2.07\%}{0.75\%} = 6.9137 \\ \frac{T_{nm}}{T_{nw}} &= \left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{3.5} \left(\frac{p_{nm}}{p_{nw}}\right) = (1.30)^{3.5} * \frac{97.93\%}{99.25\%} = 2.4716\end{aligned}\quad (20)$$

The location parameter for men in agriculture is about four times higher than for women. This parameter reflects productivity disparities between male and female workers in agriculture. This value may be an overestimate because I assume no friction against female workers in the United States. The difference in non-agriculture is much smaller at 2.4716 that reflects less disparity in productivity between men and women in non-agriculture. If the wage difference were zero, the location parameter in non-agriculture would be near one. The main takeaway is a comparative (not absolute) advantage structure. Female workers have comparative advantage in non-agriculture while male workers do in agriculture (consistent to [Goldin and Sokoloff \(1984\)](#) and [Pitt, Rosenzweig and Hassan \(2012\)](#)).

## 4.2 Full Calibration for the US Economy

Given the calibrated conditions for the distributional parameters,  $\{\theta, T_{am}/T_{aw}, T_{nm}/T_{nw}\} = \{3.5, 6.9137, 2.4716\}$ , I calibrate all other parameters in the model jointly to match the moments from the US data. Given a normalization of  $T_{am} = 1$ , I need five moments to match four preference parameters and one distributional parameter,  $\{\bar{n}, \nu, \psi_m, \psi_w, T_{nm}\}$ . The first moment I target is the fraction of workers in agriculture. From the 2010 census in the IPUMS-International, the value is 1.4 percent. The second is a long-run agricultural employment share of 0.5 percent ([Restuccia, Yang and Zhu, 2008](#); [Lagakos and Waugh, 2013](#)). The third and fourth moments are male and female labor force participation rates, which are 87.3% and 76.6% in the US data. The last moment to match is the relative agricultural labor productivity at 0.84. The calibrated values for the subsistence consumption requirement, the relative taste for non-agricultural consumption, and the relative non-agricultural talent are  $\bar{n} = 783.1534$ ,  $\nu = 199$ ,  $\psi_m = 9.5633$ ,  $\psi_w = 15.1503$ , and  $T_{nm} = 119.1408$ .

A higher mean talent difference across industries is needed to replicate the relative agricultural labor productivity at 0.84. In the United States, the employment share in agriculture is only 1.4 percent. According to the model and due to a selection effect, the average productivity in agriculture must be extremely high if average talents are the same in both industries. However, agricultural labor productivity is 16 percent lower than that in non-

agriculture, which is why calibrated values for the non-agricultural location parameters are high:  $(T_{nm}, T_{nw}) = (119.1408, 48.2029)$ .

Table I summarizes calibrated parameters from the US economy. The US productivity is normalized to 100,  $A^{\text{US}} = 100$ .

Table I: Parameters Calibrated from the US Economy

Parameter	Value	Target
Distribution		
$\theta$	3.5	wage dispersion
$\rho$	0	benchmark case (no correlation)
$T_{am}$	1	normalization
$T_{aw}$	0.1446	gender wage gap & propensities
$T_{nm}$	119.1408	relative agricultural productivity = 0.84
$T_{nw}$	48.2029	gender wage gap & propensities
Preference		
$\bar{n}$	783.1534	fraction of workers in agriculture = 1.4%
$\nu$	199	long-run agricultural employment share = 0.5%
$\psi_m$	9.5633	male labor force participation rate = 0.873
$\psi_w$	15.1503	female labor force participation rate = 0.766

Notes: This table reports calibrated parameters from the US economy.

### 4.3 Calibration of Country-specific Parameters

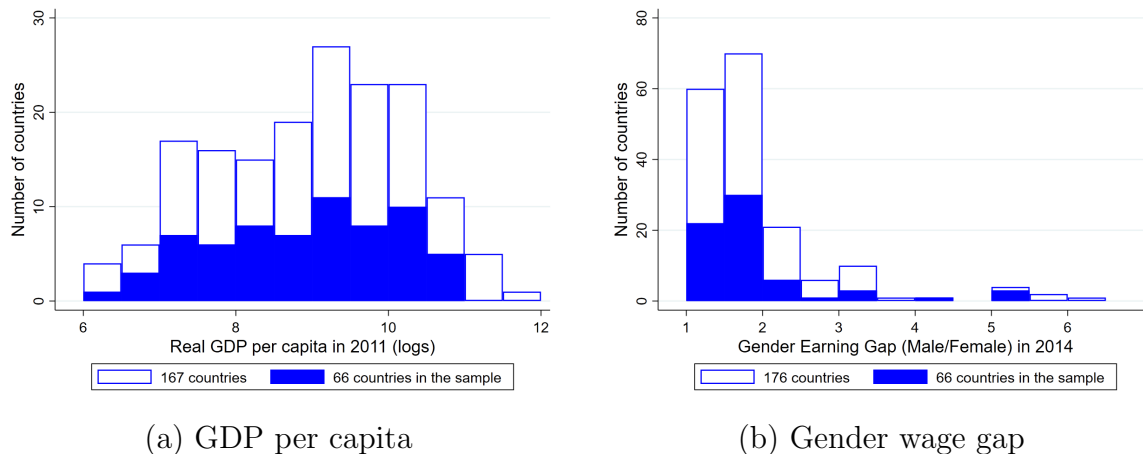
I calibrate five country-specific parameters:  $\{A^c, \psi_m^c, \psi_w^c, \tau_{aw}^c, \tau_{nw}^c\}$  where c indexes a country. According to Proposition 3, gender-specific friction is a function of the gender wage gap, propensities, and the distributional parameters:

$$\frac{1}{1 - \tau_{aw}} = \frac{\overline{\text{wage}}_m}{\overline{\text{wage}}_w} \left( \frac{p_{am}}{p_{aw}} \right)^{(1-\rho)/\theta} \left( \frac{T_{am}}{T_{aw}} \right)^{-1/\theta} = \frac{\overline{\text{wage}}_m}{\overline{\text{wage}}_w} \left( \frac{p_{am}}{p_{aw}} \right)^{\frac{1}{3.5}} (6.9137)^{-\frac{1}{3.5}}$$

$$\frac{1}{1 - \tau_{nw}} = \frac{\overline{\text{wage}}_m}{\overline{\text{wage}}_w} \left( \frac{p_{nm}}{p_{nw}} \right)^{(1-\rho)/\theta} \left( \frac{T_{nm}}{T_{nw}} \right)^{-1/\theta} = \frac{\overline{\text{wage}}_m}{\overline{\text{wage}}_w} \left( \frac{p_{nm}}{p_{nw}} \right)^{\frac{1}{3.5}} (2.4716)^{-\frac{1}{3.5}} \quad (21)$$

The friction in one industry is estimated as high when the gender wage gap is large and male workers are overrepresented in the industry. I have data on the gender wage gap for 176 countries from UNDP (2015), and micro-level census data for 81 countries from the IPUMS-International (Minnesota Population Center, 2020) to calculate the employment shares by industry and gender. Because five countries in the IPUMS-International have no information on industry in their censuses and seven other countries are unmatched with the

Figure 3: Representativeness of the Sample



Notes: In panel (a), the empty bars in the histogram are 167 countries in the PWT 8.1 (Feenstra, Inklaar and Timmer, 2015), and the solid bars are the 66 countries in the matched sample. Income distribution of the sample represents the world well. In panel (b), the empty bars in the histogram are the 176 countries in the Human Development Report (UNDP, 2015). The solid bars represent the 66 countries in the matched sample. The sample represents the distribution of gender wage gaps well.

Penn World Tables, I have 66 matched countries for the quantitative analysis. The IPUMS-International contains multiple census years so I use the survey from the most recent year where the industry code is available.<sup>10</sup> Panel (a) of Figure 3 shows that 66 countries in the sample represent the income distribution in the world well; both rich and poor countries are represented. Panel (b) of Figure 3 shows that even in terms of the gender wage gap the sample countries are distributed well.

Panel (a) of Figure 4 shows the calibrated relative non-agriculture friction

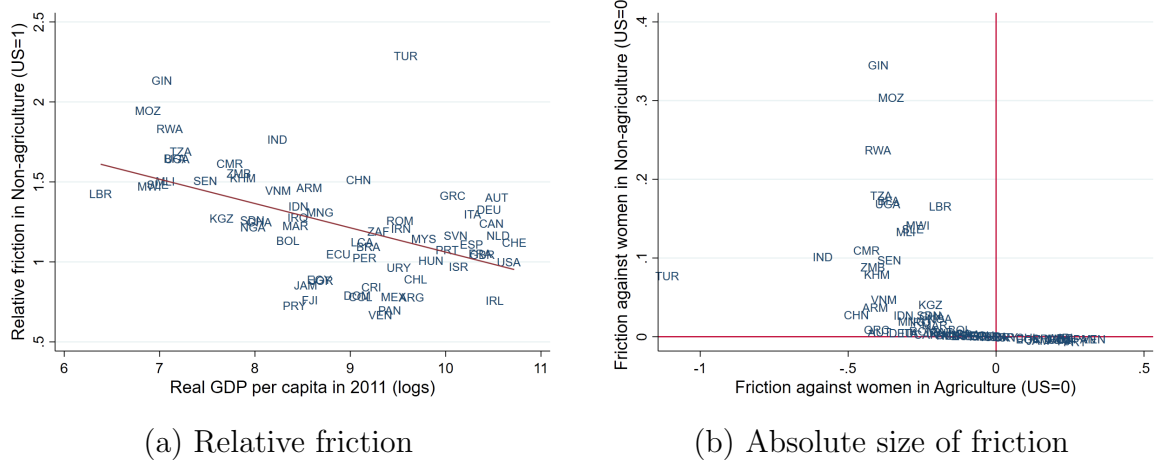
$$\tilde{\tau}_{n|a} \equiv \frac{1 - \tau_{aw}}{1 - \tau_{nw}}$$

for women in the 66 countries. The relative friction is negatively associated with real GDP per capita. Poor countries have a higher relative friction in non-agriculture than rich countries. From Proposition 2, *ceteris paribus*, the average quality of women in agriculture decreases with the frictions in non-agriculture. Therefore, the poorer countries' low labor productivity in agriculture relates to the misallocation of female talent due to higher frictions in non-agriculture.<sup>11</sup>

<sup>10</sup>Refer to Table A.I in Appendix A for detailed sample construction.

<sup>11</sup>Gollin, Lagakos and Waugh (2014b) find that the agricultural labor productivity puzzle is reduced when productivity is expressed in hours worked. In Appendix C, I use 18 countries with working hour information available and calculate relative frictions with hour-adjusted labor inputs. I find that relative frictions calculated with the number of workers and relative frictions calculated with hour-adjusted labor

Figure 4: Estimated Frictions against Women



Notes: Panel (a) plots the calibrated relative friction  $\tilde{\tau}_{n|a}$  for the 66 countries. The y-axis is the relative friction in and the x-axis is real GDP per capita from the PWT 8.1 (Feenstra, Inklaar and Timmer, 2015). A mean of relative friction is 1.23 with a standard deviation of 0.35. Panel (b) plots the calibrated values of  $\tau_{nw}$  against  $\tau_{aw}$  for the 66 countries. The horizontal and vertical lines lie on zero, which is the US level. A mean of the friction in non-agriculture is 0.04 with a standard deviation of 0.08. A mean of the friction in agriculture is -0.15 with a standard deviation of 0.25.

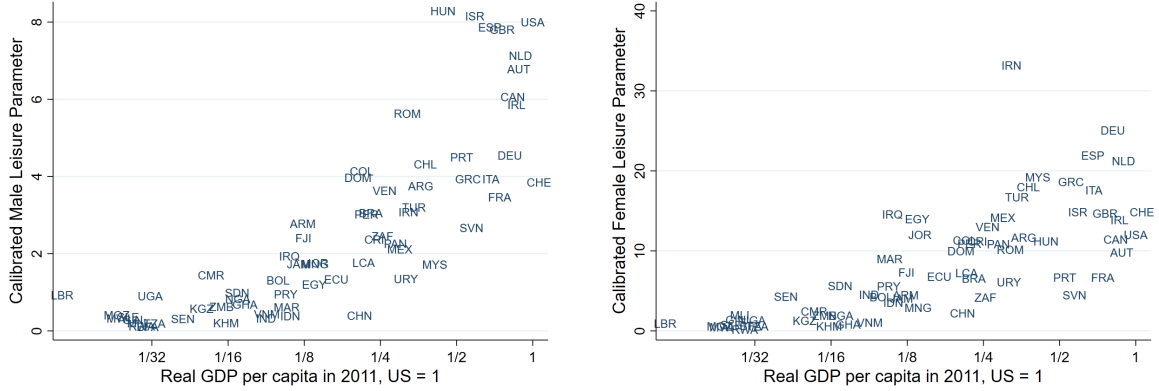
To calibrate the absolute sizes of the frictions, I search for  $\tau_{aw}$  and  $\tau_{nw}$  that satisfies the government balanced budget constraint (equation 17) and are consistent with the calibrated relative frictions. Calibrated absolute frictions are shown in panel (b) of figure 4.

Other country-specific parameters are relative utility of leisure,  $\psi_m^c$  and  $\psi_w^c$ . Those parameters are calibrated to match male and female labor force participation rates as in Figure 2. Panel (a) and (b) of Figure 5 show the estimated relative utility of leisure for men and women, respectively. Among countries with similar income level, the relative utility of leisure has a higher value when the country has a lower labor force participation rate for a specific gender.

The last country-specific parameter to calibrate is the productivity. Given the calibrated frictions, I calibrate the productivity for the 66 countries. The target is real GDP per capita relative to the United States, from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015). The calibrated  $A$  ranges from 101.9586 (Switzerland) to 1.7609 (Liberia) for targets from 1.058 (Switzerland) to 0.138 (Liberia).

input are very close to each other. Correlation is 0.97 and significant at 1 percent.

Figure 5: Estimated Relative Utility of Leisure for Men and Women



(a) Relative Utility of Leisure for Men      (b) Relative Utility of Leisure for Women

Notes: Panel (a) plots the estimated relative utility of leisure for men ( $\psi_m^c$ ). Panel (b) plots the estimated relative utility of leisure for women ( $\psi_w^c$ ).

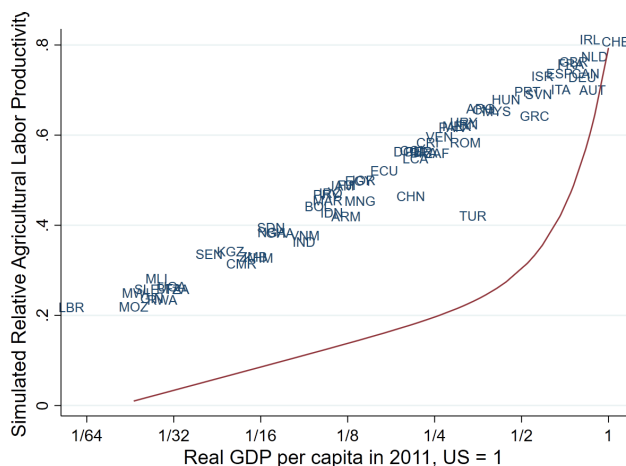
## 5 Quantitative Analysis

### 5.1 Sectoral Labor Productivity

I calculate the relative agricultural labor productivity by simulating a model for 66 countries. In the calibration, the only target is the relative agricultural labor productivity in the US at 0.84. For other countries, sectoral labor productivities are given by the worker’s occupation choice and its general equilibrium effect. As in [Lagakos and Waugh \(2013\)](#), a binding subsistence requirement induces more workers, who are unproductive at agricultural work, to select agriculture in poorer countries. At the same time, gender-specific friction distorts the allocation of female talent. Because poorer countries tend to have higher friction in non-agriculture, the agricultural labor productivity is likely low. Figure 6 shows the simulated relative agricultural labor productivity from the model, which predicts the pattern of cross-country disparities in productivity from panel (a) of Figure 1. The simulated results lie close to the polynomial fitted line from the data. I calculate to what extent the model can explain the actual agricultural productivity gap, by comparing the simulated relative productivity to the US with the data. Given that the agricultural labor productivity is 0.84 in the United States, I define the measure of model’s performance in generating relative productivities as follows:

$$\text{Explain}_c = \frac{0.84 - \text{simulated relative agricultural labor productivity}_c}{0.84 - \text{actual relative agricultural labor productivity}_c}$$

Figure 6: Simulated Relative Agricultural Labor Productivity



Notes: The graph plots the simulation results for the 66 countries. The y-axis is the relative agricultural labor productivity from the general equilibrium model simulation. The x-axis represents the data for real GDP per capita from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015). The line is a polynomial fitted from the data in panel (a) of Figure 1.

where  $c$  denotes a set of 46 countries except the United States where actual relative agricultural labor productivity data are available. Between the United States and other countries, the average size of the “Explain” variable is 0.49, implying that the model explains 49 percent of the sectoral productivity gap. Generally, the labor allocation by subsistence requirement plays a larger role, but the misallocation of females also matters. For example, Turkey has a very high relative friction in non-agriculture (2.29) in comparison to other countries with similar income. Its simulated relative agricultural productivity is 0.31, which is only 40 percent of Chile’s (0.78), where friction in non-agriculture is nearly one (0.89) and real GDP per capita is close to Turkey’s.

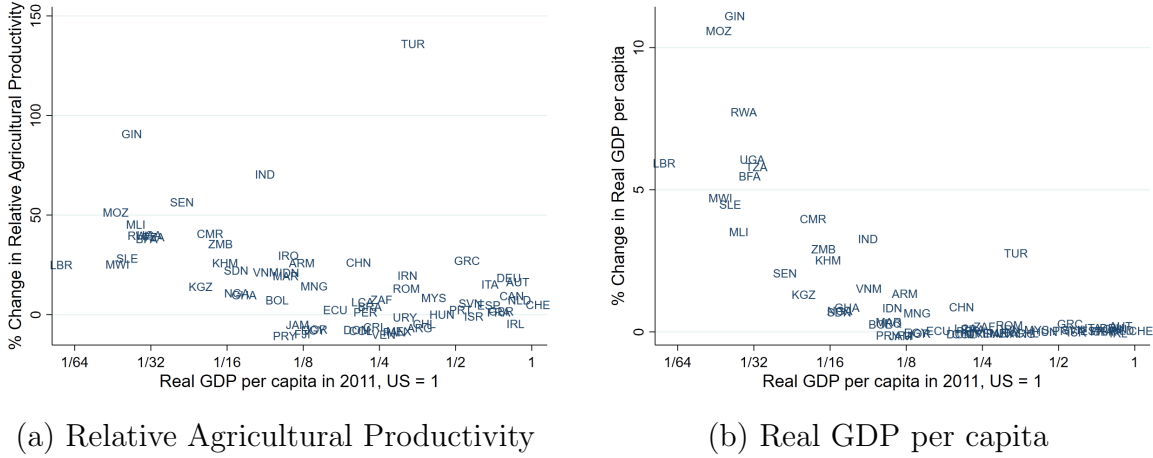
## 5.2 Counter-factual Analysis

One advantage of the model is that I can conduct a normative analysis by adjusting frictions for some countries. I run one counter-factual analysis in which the gender-specific frictions are set to the current US value for all other countries.<sup>12</sup>

Panel (a) of Figure 7 and the first row of Table II show the percentage change in relative agricultural labor productivity after setting gender-specific frictions to the current US value,  $\tau_{aw}^c = \tau_{nw}^c = 0$  for each country. Therefore, for the United States, the change is zero. On

<sup>12</sup>This experiment is similar to Hsieh and Klenow (2009), who estimate TFP gains in China and India by reallocating capital and labor hypothetically to equalize marginal products to the extent observed in the United States.

Figure 7: Counter-factual Results



Notes: Panel (a) plots the results from the counter-factual exercise after setting frictions to the current US level. The y-axis is the percentage change in agriculture labor productivity after setting the gender-specific frictions to the current US value. The x-axis is data for real GDP per capita from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015). Panel (b) plots the results from the counter-factual exercise after setting the frictions to the current US level. The y-axis is the percentage change in GDP per capita after setting gender-specific frictions to the current US value. The x-axis is data for real GDP per capita from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015).

Table II: Summary statistics of differences between base case and counterfactual

	mean	sd	p25	p50	p75
% change in relative ag productivity	16.60	25.47	0.08	9.83	25.96
% change in ag productivity	7.21	9.44	0.07	3.60	13.09
% change in non-ag productivity	-1.30	2.33	-1.48	-0.12	-0.00
% change in real GDP per capita	1.44	2.54	0.00	0.15	1.53
% change in women’s share of ag employment	-6.76	63.98	-48.63	-33.17	-1.32
% change in female labor supply	0.89	1.72	0.02	0.13	0.85
% ag consumption equivalent welfare gain	2.85	4.96	0.29	1.06	3.58

Notes: Relative agricultural productivity is the ratio (agricultural output per agricultural worker)/(non-agricultural output per non-agricultural worker). Real GDP is  $P_a * Y_a + Y_n$ . Women’s share of agriculture employment is just the share of agricultural workers that are women. Female labor supply is the average number of hours of labor chosen by female workers. Percent agricultural consumption equivalent welfare gain is the across-the-board percentage increase in agricultural consumption in the baseline case needed to match the level of average utility in the counterfactual world.

average, this counter-factual exercise increases the relative agricultural labor productivity by 16.6 percent. These gains come primarily from poorer countries.

However, the relative agricultural labor productivity says nothing about an economy’s welfare. Panel (b) of Figure 7 and the second row of Table II show the percentage change

in real GDP per capita by setting the frictions to the current US value.<sup>13</sup> The GDP per capita increases by 1.4 percent, on average. Thus, the gain from reallocating female talent is small in terms of the percentage change from the original value. Although removing frictions improves the GDP per capita in poorer countries considerably, a huge income gap with the United States remains. The expected gains from the reallocation of female workers across industries are too small to predict a meaningful proportion of the disparities in cross-country income.<sup>14</sup>

Last three rows of Table II show the summary statistics of the percentage change in other variables between the calibrated simulation and the counterfactual exercise. Women’s share of agricultural employment decreases by 6.76 percent on average, female labor supply increases by 0.89 percent on average, and agricultural consumption equivalent welfare gain increases by 2.84 percent on average.

## 6 Discussion

### 6.1 Auxiliary Test

One strong prediction from the model is that gender wage gaps are equalized across industries. This property is standard for the Roy-type model with the Fréchet distribution. Hsieh, Hurst, Jones and Klenow (2019) suggest that the gender wage gap in an occupation is uncorrelated with frictions in that occupation and show that it holds from the US Census data. According to Proposition 3, gender wage gaps in agriculture and non-agriculture are the same in equilibrium. Therefore, the auxiliary test is to assess whether gender wage gaps are equal across two sectors.

I collect all available censuses with any type of individual income data from the IPUMS-International. The sample consists of 46 censuses from 15 countries. Panel (a) of Figure 8 shows gender earnings gaps in agriculture and non-agriculture. The theory predicts a 45-degree line: the gender wage gaps are equal across two sectors. Generally, the dots lie close to the 45-degree line. Panel (b) of Figure 8 shows gender hourly wage gaps in agriculture and non-agriculture using 27 censuses from 11 countries where working hour information is available. Gender hourly wage gaps is in general lower than gender earning gaps because men work more on average. Still, the observations are not far from the 45-degree line in

---

<sup>13</sup>In order to calculate real GDP per capita, I fix prices of agricultural goods and non-agricultural goods at the US level. Therefore, changes in real GDP per capita come from change in quantities.

<sup>14</sup>This result is consistent with Vollrath (2009, 2014). He finds misallocation in the labor market in terms of large differences in marginal products across sectors. His analysis shows that reallocating individuals across sectors until marginal products converge yields gains in output of less than 5 percent for most countries.



Table III: Estimated  $\theta$  from other samples

Country	Census Year	Estimated $\theta$ from Proposition 5
United States	2000	3.46
United States	1990	3.33
Brazil	2000	2.85
Canada	2011	4.04
Mexico	2010	3.38
Venezuela	2001	3.24

Notes: This table reports estimated  $\theta$ 's from five different countries.  $\theta$ 's are estimated from Proposition 5.

except that I split along both gender and sector, instead of just sector. For the simulated standard deviation, I can simply compute the standard deviation of the log of realized wages in the model. Given that I do not target the gender-sector specific variation in wages, it is reassuring that the simulated variation is not too different from the empirical variation.

Table IV: Standard deviations of log wages,  $\rho = 0$ 

	Simulation	Empirical
Female, agriculture	0.3900	0.4218
Female, non-agriculture	0.3658	0.4171
Male, agriculture	0.3823	0.3625
Male, non-agriculture	0.3669	0.4915

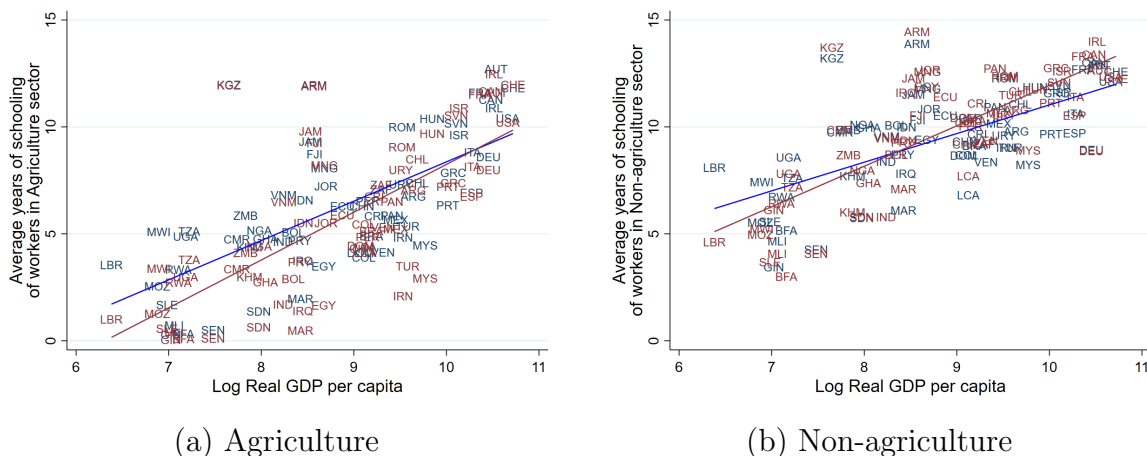
### 6.3 Role of Education

The gulf in average levels of education between rich and poor countries is huge. However, education decision is not explicitly modeled, because the gender gap in years of schooling given occupational choice is not substantially different across countries. Figure 9 shows the average years of schooling in agriculture and non-agriculture for male and female workers. Gradients of fitted lines for male and female workers are very similar between agriculture and non-agriculture. Therefore, it is less likely that education, measured by years of schooling,

the transitory component is serially uncorrelated, independent of the permanent component, and has finite variance and zero mean, the variance of the non-transitory component of wages is just the covariance of log wages from two consecutive periods for each worker.

is a fundamental source of agricultural productivity gaps.

Figure 9: Years of Schooling across Countries within Each Sector



Notes: The graphs plot the average years of schooling in agriculture and non-agriculture for 64 countries for which years of schooling data is available. The U.K. and the Netherlands are excluded. Blue lines are the fitted lines for male, and red lines are fitted lines for female. The x-axis is log real GDP per capita. The y-axis represents average years of schooling, calculated from IPUMS-International ([Minnesota Population Center, 2020](#)).

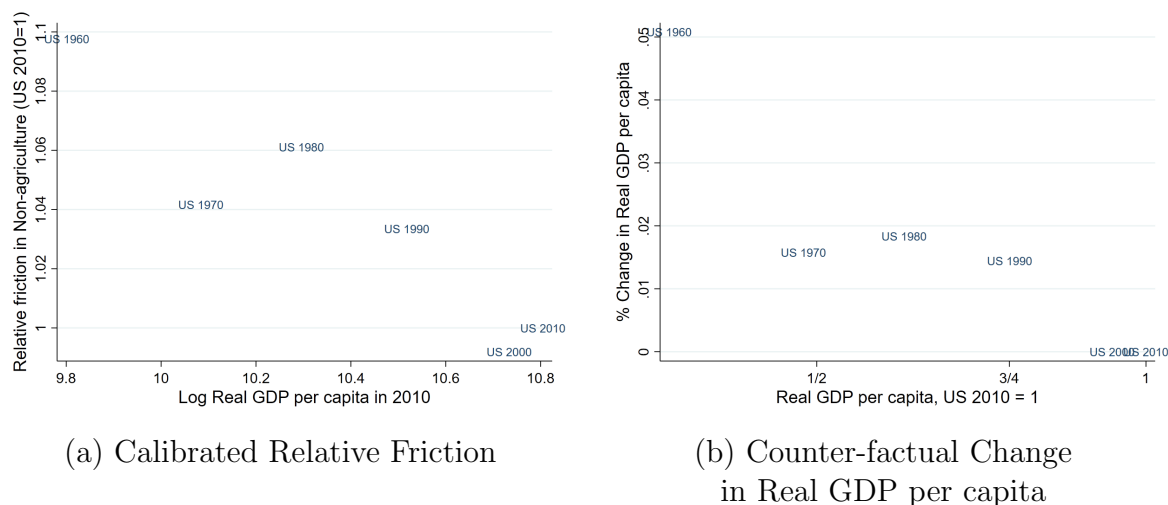
## 6.4 Time-series Analysis

The model can be easily applied to understanding historical development of a single country. In this subsection, I apply the same model to historical development of two countries, the US and Turkey. Given that the innate talent for each sector is unlikely to have changed differently across groups over time in a single country, the change in the occupational choice over time suggests that a substantial pool of innately talented women several decades ago in the US and Turkey were not pursuing their comparative advantage.

### 6.4.1 US, 1960-2010

Panel (a) of Figure 10 shows calibrated relative friction in non-agriculture from 1960 to 2010 in the US. Note that the relative friction in 2010 is normalized to one. In 1960, women faced 10 percent higher friction in non-agriculture compared to agriculture. Naturally, part of low agricultural productivity back then can be explained by this female talent misallocation. Panel (b) of Figure 10 shows that real GDP per capita would have increased by 0.05 percent if US had no relative friction against women in 1960. Both relative friction and GDP loss associated with that friction have declined over time.

Figure 10: Time-series Analysis, US 1960-2010



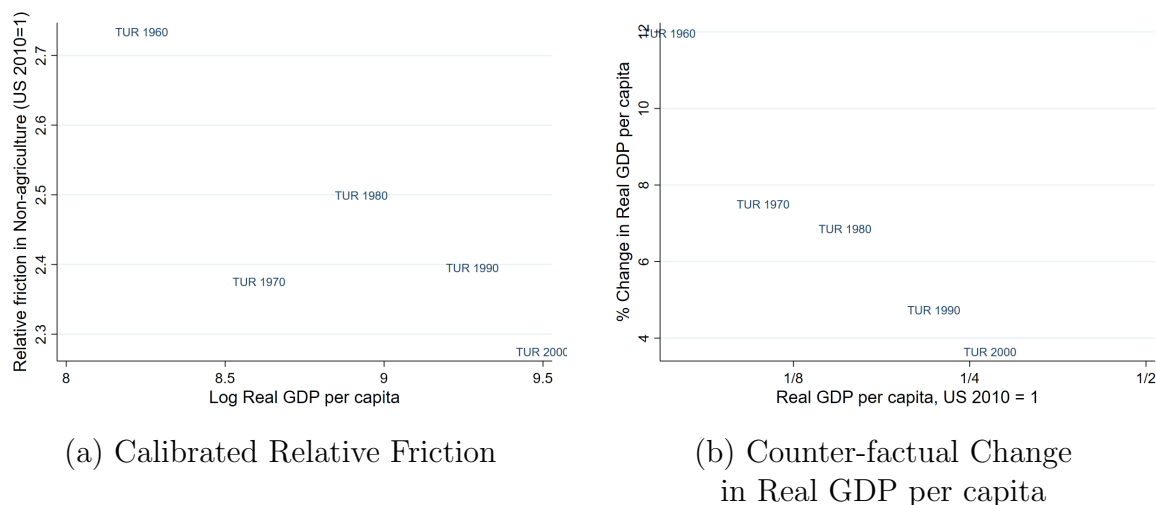
Notes: Panel (a) plots the calibration results from six time series observation for the US. The y-axis is the estimated relative friction in non-agriculture and the x-axis is real GDP per capital from [Bolt, Inklaar, de Jong and van Zanden \(2018\)](#). Panel (b) plots the results from the counter-factual exercise after setting the frictions to the 2010 US level. The y-axis is the percentage change in GDP per capita after setting gender-specific frictions to the 2010 US value. The x-axis is real GDP per capita from [Bolt, Inklaar, de Jong and van Zanden \(2018\)](#).

#### 6.4.2 Turkey, 1960-2000

Another example is Turkey, where agricultural productivity loss and real GDP loss from female talent misallocation are large. How does this country look like over time? For calibration of the model to the Turkish economy, I collect additional data. The employment and labor force participation data by sector and gender are taken from Statistical Indicators 1923-2009 ([Göstergeler, 2010](#)). The GDP data are taken from the World Bank. The gender wage gap data are taken from multiple sources. For year 2000, I use the gap in average monthly wage among salary workers, taken from [Cudeville and Gurbuzer \(2010\)](#). The data for year 1990 is taken from [Ilkcaracan and Selim \(2007\)](#), who estimated gender wage gap using the 1994 Employment and Wage Structure Survey. Due to the lack of a reliable data source, the data for 1970 and 1960 are extrapolated from the data from 1980 and 1990.

Panel (a) of Figure 11 shows the calibrated relative friction from 1960 to 2000. Female workers in Turkey faced much higher friction in non-agriculture both in 1960 and 2000 compared to the US economy, but the relative friction has decreased over time. As shown in panel (b) of Figure 11, real GDP loss from female talent misallocation was 12 percent in 1960, compared to less than 4 percent in 2000. This shows that there has been an improvement in female talent allocation in Turkey from 1960 to 2000, even though gender disparity in Turkey today is still massive compared to the US.

Figure 11: Time-series Analysis, Turkey 1960-2000



Notes: Panel (a) plots the calibration results from five time series observation for Turkey. The y-axis is the estimated relative friction in non-agriculture and the x-axis is real GDP per capita from the World Bank. Panel (b) plots the results from the counter-factual exercise after setting the frictions to the 2010 US level. The y-axis is the percentage change in GDP per capita after setting gender-specific frictions to the 2010 US value. The x-axis is real GDP per capita from the World Bank.

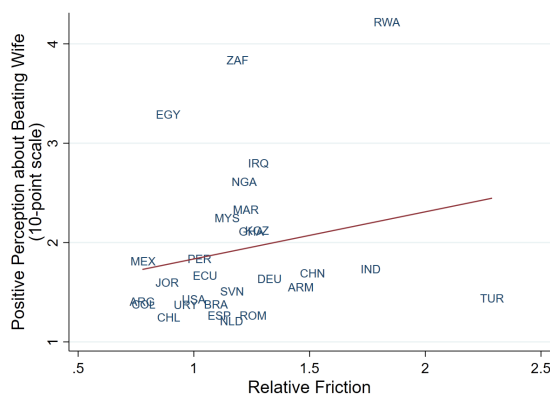
## 6.5 Potential Sources of the Friction

In this study, the relative friction in non-agriculture represents the barriers women experience when entering non-agriculture relative to agricultural industries. There are two ways of measuring social attitudes toward women: indirect/objective and direct/subjective. The indirect/objective method uses price (wage) and quantity (employment) gaps. The gender inequality index in the Human Development Report is an example, and the frictions in this study belong to this category. The direct/subjective method uses responses from survey items, of which the General Social Survey and World Values Survey are examples. I pick several survey questions from the World Values Survey ([World Values Survey Association, 2015](#)) and compare these to the frictions in this study.<sup>16</sup>

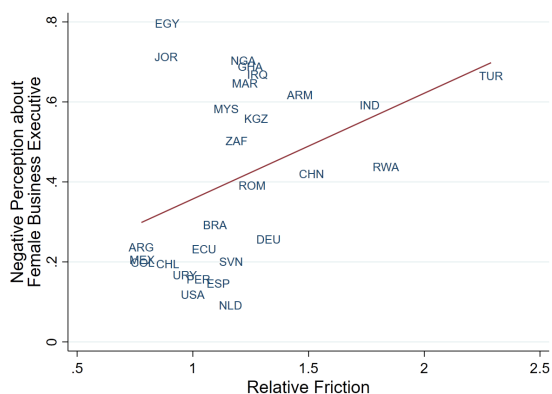
Six questions are about people's perception on beating wife, female business executives, women's university education, women earning more money, female political leader, and working mother, respectively. Among 60 countries in wave 6 (2010 through 2014) of the World Values Survey, I match 28 countries that coincide with the sample of this study. [Figure 12](#)

<sup>16</sup>Country-specific variation in gender bias has been an interest of other papers. For example, [Fernandez and Fogli \(2009\)](#) show that culture, proxied with past female labor force participation and total fertility rates from the woman's country of ancestry, have positive significant explanatory power on work and fertility choice. Refer to [Giuliano \(2020\)](#) for extensive reviews on the literature on the relevance of culture in the determination of different forms of gender gap.

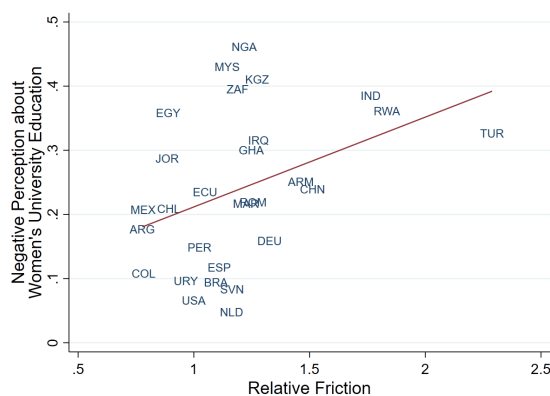
Figure 12: Perception and Relative Friction in Non-agriculture



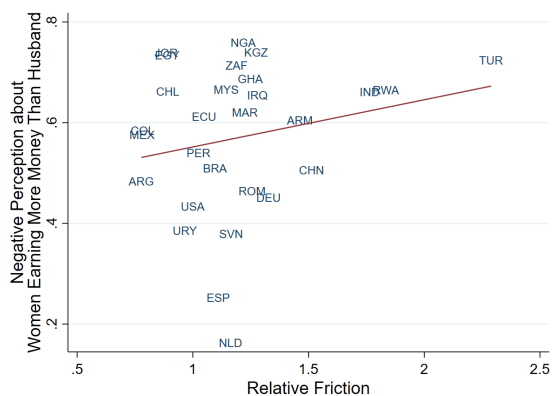
(a) Beating Wife



(b) Female Business Executives



(c) Women's University Education



(d) Women Earning More Money

Notes: The graph plots the 28 countries from the World Values Survey ([World Values Survey Association, 2015](#)). The y-axis is the proportion of respondents who agreed with the item. In panel (a), the item on the perceptions on beating wife is: “Beating wife is justified”. The slope of the fitted line is 0.474 with  $p = 0.294$ . In panel (b), the item on the perceptions on business executives is: “On the whole, men make better business executives than women do.” The slope is 0.264 with  $p = 0.037$ . In panel (c), the item on the perceptions on university education is: “A university education is more important for a boy than a girl”. The slope of the fitted line is 0.139 with  $p = 0.039$ . In panel (d), the item on the perceptions on women earning more money is: “If a woman earns more money than her husband, it is almost certain to cause problems.” The slope is 0.093 with  $p = 0.285$ .

shows the correlation between the proportion of respondents who agree with the item and the relative friction in non-agriculture. I find a positive cross-country level correlation for all six variables, and especially significant correlation for female business executives, women’s university education, and female political leader. Countries with higher frictions tend to have explicit discrimination against women in managerial/political positions and in higher education. These two barriers in business/politics and education are likely relative frictions women experience in non-agriculture.

Figure 12: Perception and Relative Friction in Non-agriculture (continued)

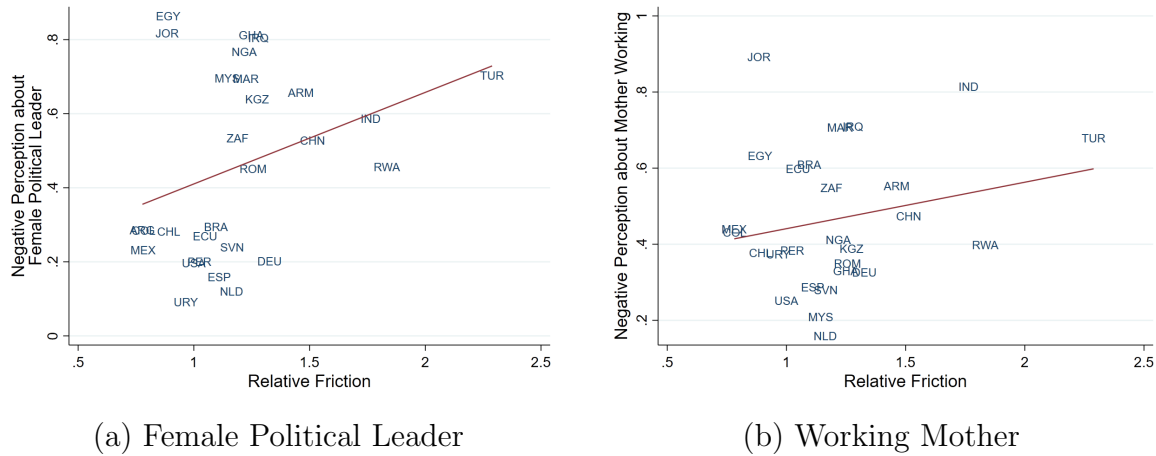
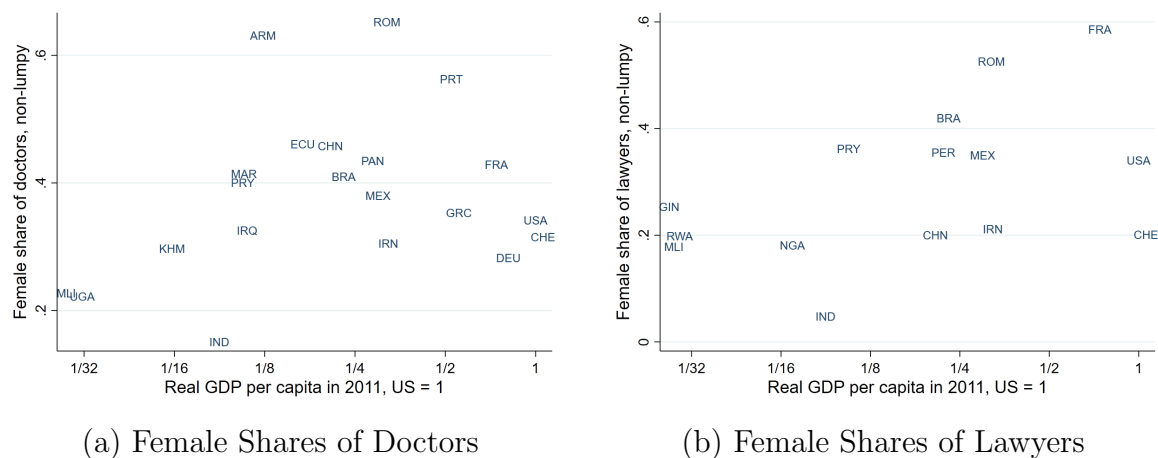


Figure 13: Female Shares of Doctors and Lawyers



Notes: Panel (a) plots female share of doctors with the set of countries where detailed occupation information is available. Panel (b) plots female share of lawyers. The x-axis is data for real GDP per capita from the World Bank.

## 7 Conclusion

I examine two important questions in the growth literature in a unified framework: (i) Why are productivity disparities in agriculture so large? and (ii) Does gender inequality hinder economic growth? I build a general equilibrium Roy model with gender-specific frictions in two sectors. I find that higher frictions against women in non-agriculture explain lower agricultural productivity in poorer countries. By setting frictions to the current US value, the other countries' agricultural productivities increase by 16.6 percent, and the GDP per capita by 1.4 percent, on average. Gains in productivity and overall income come primarily from poorer countries.

## References

- Adamopoulos, Tasso and Diego Restuccia, “The Size Distribution of Farms and International Productivity Differences,” *American Economic Review*, 2014, *104* (6), 1667–97.
- , Loren Brandt, Jessica Leight, and Diego Restuccia, “Misallocation, selection and productivity: A quantitative analysis with panel data from China,” *NBER Working Paper No. 23039*, 2020.
- Arrow, Kenneth, “The Theory of Discrimination,” in O. Ashenfelter and A. Rees, eds., *Discrimination in Labor Markets*, Princeton University Press, 1973.
- Becker, Gary S, *The Economics of Discrimination*, University of Chicago press, 1957.
- Benabou, Roland, “Equity and Efficiency in Human Capital Investment: the Local Connection,” *Review of Economic Studies*, 1996, *63* (2), 237–264.
- Bick, Alexander, Nicola Fuchs-Schündeln, and David Lagakos, “How do hours worked vary with income? Cross-country evidence and implications,” *American Economic Review*, 2018, *108* (1), 170–99.
- Bolt, Jutta, Robert Inklaar, Herman de Jong, and Jan Luiten van Zanden, *Rebasing ‘Maddison’: new income comparisons and the shape of long-run economic development*, Maddison Project Database, 2018.
- Burstein, Ariel, Eduardo Morales, and Jonathan Vogel, “Changes in between-group inequality: computers, occupations, and international trade,” *American Economic Journal: Macroeconomics*, 2019, *11* (2), 348–400.
- Caselli, Francesco, “Accounting for Cross-country Income Differences,” in P. Aghion and S. N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1A, North-Holland, 2005, chapter 9, pp. 679–741.
- Caunedo, Julieta and Elisa Keller, “Capital obsolescence and agricultural productivity,” *The Quarterly Journal of Economics*, 2021, *136* (1), 505–561.
- Chen, Chaoran, “Untitled land, occupational choice, and agricultural productivity,” *American Economic Journal: Macroeconomics*, 2017, *9* (4), 91–121.
- Chiplunkar, Gaurav and Pinelopi K Goldberg, “Aggregate Implications of Barriers to Female Entrepreneurship,” *NBER Working Paper No. 28486*, 2021.

- Cuberes, David and Marc Teignier, “Aggregate effects of gender gaps in the labor market: A quantitative estimate,” *Journal of Human Capital*, 2016, 10 (1), 1–32.
- Cudeville, Elisabeth and Leman Yonca Gurbuzer, “Gender Wage Discrimination in the Turkish Labor Market: Can Turkey be Part of Europe?,” *Comparative Economic Studies*, 2010, 52 (3), 429–463.
- Doepke, Matthias and Michele Tertilt, “Women’s Liberation: What’s in it for Men?,” *Quarterly Journal of Economics*, 2009, 124 (4), 1541–1591.
- and —, “Does female empowerment promote economic development?,” *Journal of Economic Growth*, 2019, 24 (4), 309–343.
- Donovan, Kevin, “The equilibrium impact of agricultural risk on intermediate inputs and aggregate productivity,” *The Review of Economic Studies*, 2021, 88 (5), 2275–2307.
- Durlauf, Steven N, “A Theory of Persistent Income Inequality,” *Journal of Economic Growth*, 1996, 1 (1), 75–93.
- Eaton, Jonathan and Samuel Kortum, “Technology, Geography, and Trade,” *Econometrica*, 2002, 70 (5), 1741–1779.
- Esteve-Volart, Berta, “Gender Discrimination and Growth: Theory and Evidence from India,” *mimeo*, 2009.
- Feenstra, Robert C, Robert Inklaar, and Marcel P Timmer, “The Next Generation of the Penn World Table,” *American Economic Review*, 2015, 105 (10), 3150–82.
- Fernandez, Raquel and Alessandra Fogli, “Culture: An empirical investigation of beliefs, work, and fertility,” *American economic journal: Macroeconomics*, 2009, 1 (1), 146–77.
- Galor, Oded and David N Weil, “The Gender Gap, Fertility, and Growth,” *American Economic Review*, 1996, 86 (3), 374–87.
- Giuliano, Paola, “Gender and culture,” *Oxford Review of Economic Policy*, 2020, 36 (4), 944–961.
- Goldin, Claudia, *Understanding the Gender Gap: An Economic History of American Women*, Oxford University Press, 1990.
- , “The U-Shaped Female Labor Force Function in Economic Development and Economic History,” in T. Paul Schultz, ed., *Investment in Women’s Human Capital and Economic Development*, University of Chicago Press, 1995, pp. 61–90.

- , “A Grand Gender Convergence: Its Last Chapter,” *American Economic Review*, 2014, 104 (4), 1091–1119.
- and Kenneth Sokoloff, “The Relative Productivity Hypothesis of Industrialization: The American Case, 1820 to 1850,” *Quarterly Journal of Economics*, 1984, 99 (3), 461–487.
- Gollin, Douglas, David Lagakos, and Michael E Waugh, “Agricultural Productivity Differences across Countries,” *American Economic Review*, 2014, 104 (5), 165–170.
- , —, and Michael Waugh, “The Agricultural Productivity Gap,” *Quarterly Journal of Economics*, 2014, 129 (2), 939–993.
- Göstergeler, İstatistik, “Göstergeler 1923-2009,” *Türkiye İstatistik Kurumu, Ankara*, 2010.
- Gottlieb, Charles and Jan Grobovšek, “Communal land and agricultural productivity,” *Journal of Development Economics*, 2019, 138, 135–152.
- Hebert, Camille, “Gender Stereotypes and Entrepreneur Financing,” *mimeo*, 2020.
- Hsieh, Chang-Tai and Peter J Klenow, “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 2009, 124 (4), 1403–1448.
- , Erik Hurst, Charles I Jones, and Peter J Klenow, “The allocation of talent and US economic growth,” *Econometrica*, 2019, 87 (5), 1439–1474.
- Ilkcaracan, Ipek and Raziye Selim, “The Gender Wage Gap in the Turkish Labor Market,” *Labour*, 2007, 21 (3), 563–593.
- Jayachandran, Seema, “Social norms as a barrier to women’s employment in developing countries,” *IMF Economic Review*, 2021, 69 (3), 576–595.
- Lagakos, David and Michael E Waugh, “Selection, Agriculture, and Cross-Country Productivity Differences,” *American Economic Review*, 2013, 103 (2), 948–980.
- Lagerlöf, Nils-Petter, “Gender Equality and Long-run Growth,” *Journal of Economic Growth*, 2003, 8 (4), 403–426.
- Lind, Nelson and Natalia Ramondo, “Trade with correlation,” *NBER Working Paper No. 24380*, 2021.
- Loury, Glenn C, “A Dynamic Theory of Racial Income Differences,” in P. A. Wallace and A. M. LaMond, eds., *Women, Minorities and Employment Discrimination*, Boston: Lexington Books, 1977, pp. 153–186.

- McFadden, Daniel, “Conditional Logit Analysis of Qualitative Choice Behavior,” in P. Zarembka, ed., *Frontiers in Econometrics*, New York: Academic Press, 1974, pp. 105–142.
- McGrattan, Ellen R and Richard Rogerson, “Changes in Hours Worked since 1950,” *Quarterly Review*, 1998, (Win), 2–19.
- Minnesota Population Center, *Integrated Public Use Microdata Series, International: Version 7.3 [dataset]* 2020.
- Ohanian, Lee, Andrea Raffo, and Richard Rogerson, “Long-term changes in labor supply and taxes: Evidence from OECD countries, 1956–2004,” *Journal of Monetary Economics*, 2008, *55* (8), 1353–1362.
- Phelps, Edmund S, “The Statistical Theory of Racism and Sexism,” *American Economic Review*, 1972, *62* (4), 659–661.
- Pitt, Mark M, Mark R Rosenzweig, and Mohammad Nazmul Hassan, “Human Capital Investment and the Gender Division of Labor in a Brawn-based Economy,” *American Economic Review*, 2012, *102* (7), 3531–3560.
- Ramondo, Natalia and Andrés Rodríguez-Clare, “Trade, Multinational Production, and the Gains from Openness,” *Journal of Political Economy*, 2013, *121* (2), 273–322.
- Restuccia, Diego and Guillaume Vandembroucke, “Explaining educational attainment across countries and over time,” *Review of Economic Dynamics*, 2014, *17* (4), 824–841.
- , Dennis Tao Yang, and Xiaodong Zhu, “Agriculture and Aggregate Productivity: A Quantitative Cross-country Analysis,” *Journal of Monetary Economics*, 2008, *55* (2), 234–250.
- Rodríguez Mora, J. V., “The Misallocation of Talent,” *Els Opuscles del CREI*, 2009, (22).
- Roy, A.D., “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers New Series*, 1951, *3*, 135–46.
- Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek, *IPUMS USA: Version 11.0 [dataset]* 2021.
- Sarsons, Heather, “Interpreting signals in the labor market: evidence from medical referrals,” *mimeo*, 2019.

- Tombe, Trevor, “The Missing Food Problem: Trade, Agriculture, and International Productivity Differences,” *American Economic Journal: Macroeconomics*, 2015, 7 (3), 226–258.
- UNDP, *Human Development Report 2015, Work for Human Development*, United Nations Development Programme, 2015.
- Vollrath, Dietrich, “How Important are Dual Economy Effects for Aggregate Productivity?,” *Journal of Development Economics*, 2009, 88 (2), 325–334.
- , “The Efficiency of Human Capital Allocations in Developing Countries,” *Journal of Development Economics*, 2014, 108, 106–118.
- World Bank, *World Development Report 2012: Gender Equality and Development*, World Bank Publications, 2011.
- World Values Survey Association, *World Values Survey Wave 6 2010-2014 official aggregate v.20150418*, Madrid SPAIN: Aggregate File Producer: Asep/JDS, 2015.

# A Appendix Tables and Figures

Table A.I: List of 66 Matched Countries

Country	IPUMS	PWT	HDR	Sample	Year	Country	IPUMS	PWT	HDR	Sample	Year
Argentina				V	2001	Malaysia				V	2000
Austria				V	2011	Mali				V	2009
Bangladesh	X					Mexico				V	2010
Armenia				V	2011	Mongolia				V	2000
Bolivia				V	2001	Morocco				V	2004
Brazil				V	2010	Mozambique				V	2007
Belarus	X					Netherlands				V	2001
Cambodia				V	2008	Nicaragua		X			
Cameroon				V	2005	Nigeria				V	2010
Canada				V	2001	Pakistan	X				
Chile				V	2002	Panama				V	2010
China				V	1990	Paraguay				V	2002
Colombia				V	2005	Peru				V	2007
Costa Rica				V	2011	Philippines	X				2000
Cuba		X				Portugal				V	2011
Dominican Rep				V	2010	Puerto Rico		X	X		
Ecuador				V	2010	Romania				V	2002
El Salvador		X				Rwanda				V	2002
Ethiopia	X				1994	Saint Lucia				V	1991
Fiji				V	2007	Senegal				V	1988
France				V	2011	Sierra Leone				V	2004
Palestine		X	X			Vietnam				V	2009
Germany				V	1987	Slovenia				V	2002
Ghana				V	2010	South Africa				V	2007
Greece				V	2001	Spain				V	2001
Guinea				V	1983	South Sudan		X	X		
Haiti		X				Sudan				V	2008
Hungary				V	2001	Switzerland				V	2000
India				V	2004	Thailand	X				2000
Indonesia				V	2010	Turkey				V	2000
Iran				V	2006	Uganda				V	2002
Iraq				V	1997	Ukraine	X				
Ireland				V	2011	Egypt				V	2006
Israel				V	1995	United Kingdom				V	2001
Italy				V	2001	Tanzania				V	2002
Jamaica				V	2001	United States				V	2010
Jordan				V	2004	Burkina Faso				V	1996
Kenya	X					Uruguay				V	2006
Kyrgyz Rep				V	2009	Venezuela				V	2001
Liberia				V	2008	Zambia				V	2010
Malawi				V	2008						

Notes: The table lists the 81 countries where at least one year of Census micro-data is available from the Integrated Public Use Microdata Series-International ([Minnesota Population Center, 2020](#)). For four countries (Bangladesh, Belarus, Kenya, and Ukraine), the industry information for each worker is not available from the IPUMS-International. For three additional countries (Ethiopia, Philippines, Thailand), the labor force participation information for each worker is not available from the IPUMS-International. I also exclude Pakistan because the industry information is only available in the oldest survey from 1973. Seven countries do not match with the Penn World Table (PWT) for the real GDP per capita and three countries do not match with the Human Development Report (HDR) for the gender earnings gap. The final sample comprises 66 countries, marked as V in the table. The last column represents the most recent survey year I use from the IPUMS-International.

## B Proofs

This appendix contains the outlines of the proofs of the first-order conditions and the propositions reported in this study.

**Derivation of the First Order Conditions** To save notation, omit subscript  $g$  and superscript  $i$ . Consumer's utility maximization problem is:

$$\max_{c_a, c_n, l} \log c_a + \nu \log(c_n + \bar{n}) + \psi \log l$$

subject to

$$\begin{aligned} P_a c_a + c_n &\leq (1 - l)y \\ c_n &\geq 0 \end{aligned}$$

The Lagrangian is

$$\mathcal{L} = \log c_a + \nu \log(c_n + \bar{n}) + \psi \log l - \lambda \{P_a c_a + c_n - (1 - l)y\} + \mu c_n$$

Solution to the problem satisfies

$$\frac{\partial \mathcal{L}}{\partial C_a} = \frac{1}{c_a} - \lambda P_a = 0 \tag{22}$$

$$\frac{\partial \mathcal{L}}{\partial C_n} = \frac{\nu}{c_n + \bar{n}} - \lambda + \mu = 0 \tag{23}$$

$$\frac{\partial \mathcal{L}}{\partial l} = \frac{\psi}{l} - \lambda y = 0 \tag{24}$$

$$\lambda(P_a c_a + c_n - (1 - l)y) = 0 \tag{25}$$

$$\mu c_n = 0 \tag{26}$$

I only have to consider two cases:  $c_n = 0$  and  $c_n > 0$ . As the budget constraint always binds, I always have  $\lambda > 0$ .

**Case 1:**  $c_n = 0, \mu \geq 0$

Combine 22 and 24 to get

$$P_a c_a = \frac{1}{\psi} y l \tag{27}$$

Plug into 25 and solve for optimal  $l$ :

$$\begin{aligned}\frac{1}{\psi}yl &= (1-l)y \\ \frac{1}{\psi}l &= 1-l \\ l^* &= \frac{\psi}{\psi+1}\end{aligned}$$

Plug back into 27 to get optimal  $c_a$ :

$$c_a^* = \frac{y}{P_a(\psi+1)}$$

**Case 2:**  $c_n > 0, \mu = 0$

Combine 22 and 24 to get 27 as before, and combine 23 and 24 to get

$$\begin{aligned}\frac{\nu}{c_n + \bar{n}} &= \frac{\psi}{yl} \\ c_n + \bar{n} &= \frac{\nu}{\psi}yl \\ c_n &= \frac{\nu}{\psi}yl - \bar{n}\end{aligned}\tag{28}$$

Plug 27 and 28 into 25 and solve for optimal  $l$ :

$$\begin{aligned}\frac{1}{\psi}yl + \frac{\nu}{\psi}yl - \bar{n} &= (1-l)y \\ \frac{1}{\psi}yl + \frac{\nu}{\psi}yl + ly &= \bar{n} + y \\ ly \left( \frac{1}{\psi} + \frac{\nu}{\psi} + 1 \right) &= \bar{n} + y \\ ly \frac{1 + \nu + \psi}{\psi} &= \bar{n} + y \\ l^* &= \frac{\psi(\bar{n} + y)}{y(1 + \nu + \psi)}\end{aligned}$$

Plug  $l^*$  into  $P_a c_a = \frac{1}{\psi} y l$  to get optimal  $c_a$ :

$$\begin{aligned} P_a c_a &= \frac{1}{\psi} y \frac{\psi(\bar{n} + y)}{y(1 + \nu + \psi)} \\ &= \frac{\bar{n} + y}{1 + \nu + \psi} \\ c_a^* &= \frac{\bar{n} + y}{P_a(1 + \nu + \psi)} \end{aligned}$$

Plug  $l^*$  into 28 to get optimal  $c_n$ :

$$\begin{aligned} c_n &= \frac{\nu}{\psi} y \frac{\psi(\bar{n} + y)}{y(1 + \nu + \psi)} - \bar{n} \\ &= \frac{\nu(\bar{n} + y)}{1 + \nu + \psi} - \frac{\bar{n} + \bar{n}\nu + \bar{n}\psi}{1 + \nu + \psi} \\ &= \frac{\bar{n}\nu + \nu y - \bar{n} - \bar{n}\nu - \bar{n}\psi}{1 + \nu + \psi} \\ c_n^* &= \frac{\nu y - \bar{n}(1 + \psi)}{1 + \nu + \psi} \end{aligned}$$

The marginal utilities of  $c_a$  and  $c_n$  must be equal at an interior solution. At the corner solution, the marginal utility of  $c_n$  is lower than the marginal utility of  $c_a$ .

The utility in case 2 is strictly greater than the utility in case 1 when the constraint  $c_n \geq 0$  does not bind. Hence the household should always be in case 2 when  $y$  is sufficiently large to permit an interior solution. This yields a useful cutoff for  $y$  that distinguishes between case 1 and case 2, i.e., household is in case 2 if and only if

$$\begin{aligned} c_n^* &= \frac{\nu y - \bar{n}(1 + \psi)}{1 + \nu + \psi} > 0 \\ &\iff y > \frac{\bar{n}}{\nu}(1 + \psi) \end{aligned}$$

Optimal  $c_a$ ,  $c_n$ ,  $l$  are the same for both cases when exactly at the cutoff  $y$ .

### Proof of Proposition 1. Propensity

Men will choose to work in agricultural industry if  $\frac{\epsilon_{am}}{\epsilon_{nm}} \geq \frac{1}{P_a}$ . Consider the probability that a man chooses agriculture, and denote this by  $p_{am}$ . Then

$$\begin{aligned} p_{am} &= \Pr[\epsilon_{nm} \leq P_a \epsilon_{am}] \\ &= \int F_1^m(\epsilon, P_a \epsilon) d\epsilon \end{aligned} \tag{29}$$

where  $F_1^m(\cdot)$  is the derivative of the cdf with respect to its first argument. Evaluating the integral gives:

$$\begin{aligned}
p_{am} &= \int F_1^m(\epsilon, P_a \epsilon) \\
&= \frac{(P_a^\theta T_{am})^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}} \int dF^m(\epsilon) \\
&= \frac{(P_a^\theta T_{am})^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}}
\end{aligned} \tag{30}$$

Similarly, women choose to work in agriculture if:  $\frac{\epsilon_{aw}}{\epsilon_{nw}} \geq \frac{1}{P_a} \frac{1}{\tilde{\tau}_{n|a}}$ . So, by replacing  $P_a$  with  $P_a \tilde{\tau}_{n|a}$ , I have the fraction of women in agriculture.

$$p_{aw} = \frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)}}{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}} \tag{31}$$

## Proof of Proposition 2. Quality

For both men and women, the maximum income is:

$$y^* = \max\{w_a \epsilon_a, w_n \epsilon_n\} \equiv w^* \epsilon^* \tag{32}$$

From the extreme value property of the Fréchet distribution, the maximum value also has a Fréchet distribution.

$$\begin{aligned}
\Pr[y^* < z] &= \Pr[w_a \epsilon_a < w_n \epsilon_n] \\
&= F\left(\frac{z}{w_a}, \frac{z}{w_n}\right) \\
&= \exp\{-\bar{T} z^{-\theta}\}
\end{aligned} \tag{33}$$

where  $\bar{T} \equiv [(w_a^\theta T_a)^{1/(1-\rho)} + (w_n^\theta T_n)^{1/(1-\rho)}]^{(1-\rho)}$ .

Similarly, the distribution of  $\epsilon^*$ , the ability of people in their chosen industry, is also Fréchet:

$$G(x) \equiv \Pr[\epsilon^* < x] = \exp\{-T^* x^{-\theta}\} \tag{34}$$

where  $T^* \equiv [((\frac{w_a}{w^*})^\theta T_a)^{1/(1-\rho)} + ((\frac{w_n}{w^*})^\theta T_n)^{1/(1-\rho)}]^{(1-\rho)}$ .

Given this distribution, I have a conditional expectation:

$$\begin{aligned}
E[\epsilon_s | \text{Person chooses } s] &= \int \epsilon dG(\epsilon) \\
&= (T^*)^{1/\theta} \Gamma\left(1 - \frac{1}{\theta}\right)
\end{aligned} \tag{35}$$

where  $\Gamma(\cdot)$  is the Gamma function. From Proposition 1, the average quality can be written

as a function of the propensity. For example,

$$E[\epsilon_{am} | \text{Person chooses } a] = \left( \frac{T_{am}}{p_{am}^{1-\rho}} \right)^{1/\theta} \Gamma \left( 1 - \frac{1}{\theta} \right) \quad (36)$$

### Proof of Proposition 3. Gender Wage Gap

The proof is simple algebra given the results from Proposition 2.

$$\begin{aligned} \frac{\overline{\text{wage}}_{am}}{\overline{\text{wage}}_{aw}} &= \frac{w_a E[\epsilon_{am} | \text{Person chooses } a]}{(1 - \tau_{aw}) w_a E[\epsilon_{aw} | \text{Person chooses } a]} \\ &= \frac{1}{1 - \tau_{aw}} \left( \frac{p_{am}}{p_{aw}} \right)^{-(1-\rho)/\theta} \left( \frac{T_{am}}{T_{aw}} \right)^{1/\theta} \end{aligned} \quad (37)$$

Two wage gaps are the same as follows:

$$\begin{aligned} \frac{\overline{\text{wage}}_{am}}{\overline{\text{wage}}_{aw}} &= \frac{(P_a^\theta T_{am})^{1/(1-\rho)} + (T_{nm})^{1/(1-\rho)}}{(P_a^\theta (1 - \tau_{aw})^\theta T_{am})^{1/(1-\rho)} + ((1 - \tau_{nw})^\theta T_{nm})^{1/(1-\rho)}} \\ &= \frac{\overline{\text{wage}}_{nm}}{\overline{\text{wage}}_{nw}} \end{aligned} \quad (38)$$

### Proof of Proposition 4. Sectoral Labor Productivity

Given the comparative statistics from Propositions 1 and 2, the fraction in agriculture increases and the average quality in agriculture decreases in relative friction.

### Proof of Proposition 5. Coefficient of Variation of Wage

From the proof of Proposition 2, the maximum income and wage follows the Fréchet distribution,  $\exp\{-\bar{T}z^{-\theta}\}$ . The mean and variance of the distribution are

$$\begin{aligned} \text{Mean} &= \bar{T}^{\frac{1}{\theta}} \Gamma\left(\frac{\theta-1}{\theta}\right) \\ \text{Variance} &= \bar{T}^{\frac{2}{\theta}} \left\{ \Gamma\left(\frac{\theta-2}{\theta}\right) - \Gamma\left(\frac{\theta-1}{\theta}\right)^2 \right\} \end{aligned} \quad (39)$$

Therefore, the coefficient of variation is:

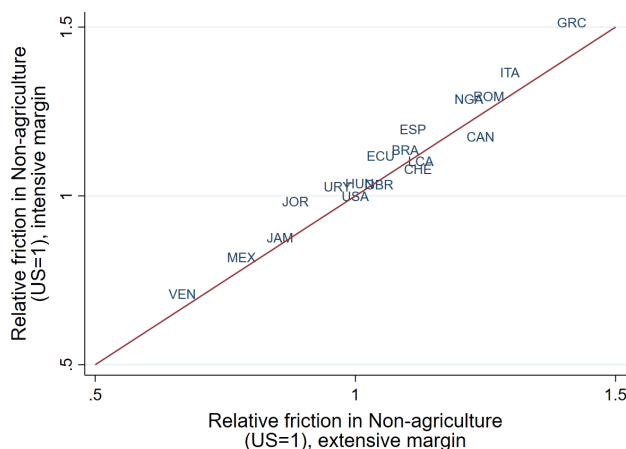
$$\frac{\text{Std.Dev.}}{\text{Mean}} = \left[ \frac{\bar{T}^{\frac{2}{\theta}} \{ \Gamma(\frac{\theta-2}{\theta}) - \Gamma(\frac{\theta-1}{\theta})^2 \}}{\{ \bar{T}^{\frac{1}{\theta}} \Gamma(\frac{\theta-1}{\theta}) \}^2} \right]^{1/2} = \left[ \frac{\Gamma(1 - \frac{2}{\theta})}{\Gamma(1 - \frac{1}{\theta})^2} - 1 \right]^{1/2} \quad (40)$$

## C Hours-adjusted Labor Input and Relative Friction

Gollin, Lagakos and Waugh (2014b) find that the agricultural labor productivity puzzle is reduced when productivity is expressed in hours worked. There are 18 countries with working hour information available among the sample of 66 countries: Brazil, Canada, Ecuador, Greece, Hungary, Italy, Jamaica, Jordan, Mexico, Nigeria, Romania, Saint Lucia, Spain, Switzerland, United Kingdom, United States, Uruguay, and Venezuela. Labor hours are measured as the number of hours worked per week.

I can incorporate hours of work by recalculating gender-specific propensities for entering into either sectors using hours information. For example, in the benchmark extensive margin propensities, the propensity of men going into agriculture was calculated by summing up the total number of men working in agriculture, and dividing it by the total number of working men. Instead, I now calculate the intensive margin propensity of men going into agriculture by summing up the total number of labor hours supplied by men working in agriculture, and dividing it by the total number of labor hours supplied by men. I do the same for intensive margin propensities of men going into non-agriculture, and women going into agriculture and non-agriculture. I then plug these intensive margin propensities into equations 21 in the main text to obtain a new set of calibrated relative frictions.

Figure C.1: Relative Friction in Non-agriculture: Intensive vs. Extensive



Notes: The graph plots the relative friction in non-agriculture with 18 countries where working hour information is available from the Census micro-data. The x-axis is relative friction calculated with the number of workers without considering working hour differences. The y-axis is relative friction calculated with hour-adjusted labor input. The line is a 45 degree reference line.

Figure C.1 compares relative frictions calculated with the number of workers (extensive margin) and relative frictions calculated with hour-adjusted labor input (intensive margin).

Within 18 countries where working hour information is available, estimated relative frictions are very close to each other. Correlation is 0.97 and significant at 1 percent level. Even though the variation in working hours across countries is massive (Bick, Fuchs-Schündeln and Lagakos, 2018), taking into account working hours do not change estimated relative friction in non-agriculture against women.

## D Results with Non-Zero Correlation in Talents

I assess the robustness of the main results using a non-zero correlation for the talent distribution. The calibration of parameters with correlation parameter  $\rho > 0$  closely follows the procedure in [Lagakos and Waugh \(2013\)](#). Calibration proceeds in two stages. The correlation parameter  $\rho$  is calibrated in the first stage, while the remaining parameters are calibrated in the second stage.

In the first stage,  $\rho$  is calibrated by targeting the ratio of average agricultural to non-agricultural wages for women. I use data from the US Current Population Survey (CPS) for 1996 through 2010, and calculate the ratio of average wages in exactly the same manner as in [Lagakos and Waugh \(2013\)](#), i.e., I use estimated non-transitory components of wages. This yields a target value of 1.060.

In order to calibrate  $\rho$ , I simulate the economy with a given value of  $\rho$ , and then optimally choosing their sector. The values of the other distribution parameters are taken from the calibration with  $\rho = 0$ . The price of agricultural goods is then adjusted so that the fraction of workers choosing the agricultural sector is 1.4%. I then compute the simulated value for the ratio of average agricultural to non-agricultural wage for women, and adjust the value of  $\rho$  until the simulated wage ratio is equal to the target of 1.060. This procedure is similar to how [Lagakos and Waugh \(2013\)](#) calibrate their correlation parameter for agricultural and non-agricultural ability draws.

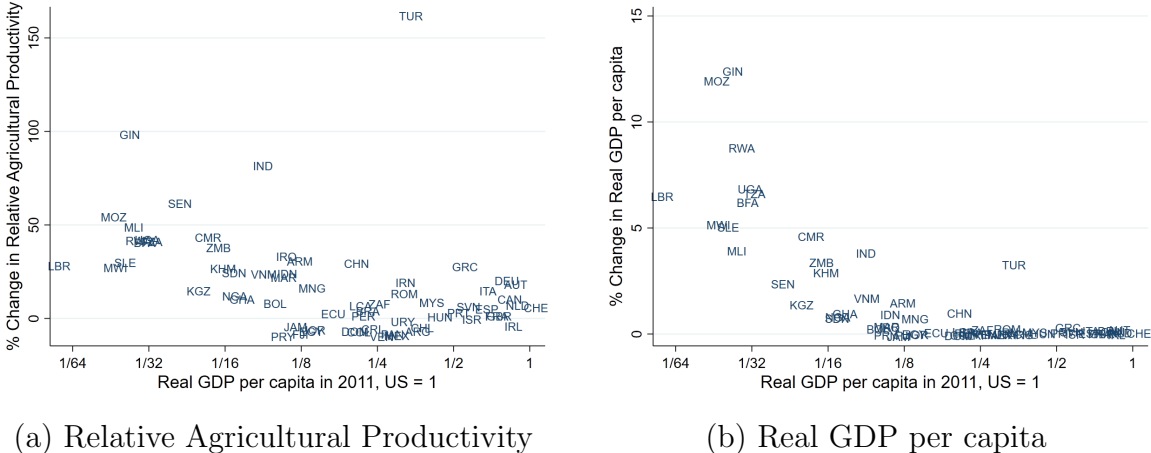
Given the similarity in context, the same intuition applies here, which I repeat from [Lagakos and Waugh \(2013\)](#): the calibrated values of  $T_{aw} = 0.1446$  and  $T_{nw} = 48.2029$  that I use imply that the variance in ability draws is much higher in non-agricultural than in agriculture for women. With high correlation in their ability draw in the two sectors, workers who get a high draw in non-agriculture are likely to get a high draw in agriculture as well. However, due to the higher dispersion in draws in non-agriculture, these workers are likely to have a comparative advantage in non-agriculture. This leads to high skill draw workers, and thus high wage workers, to be predominantly in non-agriculture. Thus the ratio of average wages of agriculture to non-agriculture decreases with  $\rho$ . With low correlation in skill draws, workers who get high agricultural skill draws are less likely to get high non-agricultural skill draws, so each sector employs workers that have high draws in each respective sector. This increases the ratio of average wages of agriculture to non-agriculture.

This calibration procedure yields a value of  $\rho = 0.107$ . The Pearson correlation coefficient of women's agricultural and non-agricultural ability draws is 0.2277, while the Spearman rank correlation coefficient is 0.1457.

In the second stage, I calibrate the model in the same manner as in the case with  $\rho = 0$ ,

except with positively correlated skill draws governed by a new value of  $\rho = 0.107$ . I start by calibrating the model to the US economy by matching three moments: the fraction of workers in agriculture, the long-run share of agricultural employment, and relative agricultural labor productivity. The calibration produces parameters:  $\bar{n} = 761.6520$ ,  $T_{nm} = 89.7594$ ,  $\psi_m = 9.2089$ ,  $\psi_w = 14.5459$ . Note that  $T_{nm} = 119.1408$  with  $\rho = 0$ , so the disparity in talent is smaller with positively correlated skill draws. This is because the selection effect is bigger with the correlation, so a lower average disparity in talent across industries can match the same target of productivity differences. Given these parameters, the country-specific productivities,  $A^c$ , and disutility parameters,  $\psi_m^c$  and  $\psi_w^c$ , are calibrated for the other 66 countries.

Figure D.1: Counter-factual Results, Correlated Talents



Notes: Panel (a) plots the results from the counter-factual exercise after setting frictions to the current US level, with correlated talent draws. The y-axis is the percentage change in agriculture labor productivity after setting the gender-specific frictions to the current US value. The x-axis is data for real GDP per capita from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015). Panel (b) plots the results from the counter-factual exercise after setting the frictions to the current US level, with correlated talent draws. The y-axis is the percentage change in GDP per capita after setting gender-specific frictions to the current US value. The x-axis is data for real GDP per capita from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015).

I run the same counter-factual analysis in which the female sector-specific frictions are set to the current US value, i.e., zero. Figure D.1 and Table D.I show counter-factual results. Quantitatively, the increases in relative agricultural productivity and real GDP per capita are slightly higher. The gains from removing the misallocation of female talent are higher when the sectoral talents are positively correlated. This is consistent with Lind and Ramondo (2021), who find larger and more heterogeneous gains from trade under positive correlation in productivity across countries, relative to models that assume independent productivity.

Table D.I: Summary statistics of differences between base case and counterfactual,  $\rho > 0$

	mean	sd	p25	p50	p75
% change in relative ag productivity	18.28	28.12	0.15	10.24	27.61
% change in ag productivity	7.67	10.09	0.08	3.60	13.78
% change in non-ag productivity	-1.36	2.43	-1.51	-0.12	-0.00
% change in real GDP per capita	1.60	2.83	0.00	0.15	1.65
% change in women's share of ag employment	-4.39	77.14	-52.84	-37.92	-2.34
% change in female labor supply	0.99	1.93	0.02	0.13	0.88
% ag consumption equivalent welfare gain	3.27	5.93	0.30	1.01	4.10

Note: Relative agricultural productivity is the ratio (agricultural output per agricultural worker)/(non-agricultural output per non-agricultural worker). Real GDP is  $P_a * Y_a + Y_n$ . Women's share of agriculture employment is just the share of agricultural workers that are women. Female labor supply is the average number of hours of labor chosen by female workers. Percent agricultural consumption equivalent welfare gain is the across-the-board percentage increase in agricultural consumption in the baseline case needed to match the level of average utility in the counterfactual world.