I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16530

# Allocation of Female Talent and Cross-Country Productivity Differences

Munseob Lee

OCTOBER 2023

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16530

# Allocation of Female Talent and Cross-Country Productivity Differences

**Munseob Lee**
*University of California San Diego and IZA*

OCTOBER 2023

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

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

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 nonagricultural sectors, then female workers who are better skilled at non-agricultural jobs may 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 4.3-7.6 percent, nonagricultural labor productivity decreases by 0.7-1.4 percent, and GDP per capita increases by 0.8-1.5 percent.

**JEL Classification:** O11, O13, O47

**Keywords:** misallocation, gender, occupational choice, productivity

**Corresponding author:**
Munseob Lee
9500 Gilman Dr #0519
La Jolla, California 92093-0519
USA

E-mail: munseoblee@ucsd.edu

* 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 Banternghansa, Bong Geun Choi, Kerem Cosar, Taryn Dinkelman, Kevin Donovan, Jorge L. Garcia, Joseph Kaboski, David Lagakos, Rachel Ngai, 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, Pacic Conference for Development Economics, STEG Workshop, North American Winter Meeting of the Econometric Society, and University of Alberta. The author wishes to acknowledge the statistical offices that provided the underlying data through the IPUMS project that made this research possible. The author further extends gratitude to Minki Kim and Churn Ken Lee for providing excellent research assistance.

# 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 to 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).

This study proposes a new explanation for the substantial agricultural productivity gap across countries. Typically, women are severely underrepresented in high-skilled jobs in poorer countries. Gender inequality may 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. Therefore, agricultural productivity measured by the outcome per worker decreases because of the misallocation of female talent, and the productivity in other industries increases because only women who are sufficiently talented to overcome these 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 gender 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; Ngai, Olivetti and Petrongolo, 2021; Dinkelman and Ngai, 2021, 2022) 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 infer 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 general parameters in the model to the benchmark economy. I then calibrate country specific parameters for the other 65 countries. 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 conduct a counterfactual analysis by setting gender-sector specific frictions to the US level for all countries. I find that agricultural labor productivity increases by 4.3-7.6 percent, non-agricultural labor productivity decreases by 0.7-1.4 percent, and GDP per capita increases by 0.8-1.5 percent, on average.[1] The gains in productivity and overall income are attributed mainly to poor countries. Two extensions in Section 5.1 and 5.2 show that estimated gains are higher when I allow for a positive correlation in workers' productivity draws across the two sectors, while those gains are lower when I allow country-specific comparative advantages inferred from U.S. experiences.

This model can be easily applied to understanding the historical development of a single country. I apply the same model to the historical development of two countries, the US and Turkey, from 1960 until 2010 and 2000, respectively. I find improvements in female talent allocation associated with GDP gains for both countries. For example, in Turkey, the real GDP loss from female talent misallocation decreased from 10.6 to 3.0 percent from 1960 to 2000, respectively.

[1] In the benchmark specification (Section 5.1), agricultural labor productivity increases by 7.2 percent, non-agricultural labor productivity decreases by 1.3 percent, and GDP per capita increases by 1.4 percent. In the first extension with non-zero correlation in talents (Section 5.2), agricultural labor productivity increases by 7.6 percent, non-agricultural labor productivity decreases by 1.4 percent, and GDP per capita increases by 1.5 percent. In the second extension with country-specific comparative advantages inferred from U.S. experiences (Section 5.3), agricultural labor productivity increases by 4.3 percent, non-agricultural labor productivity decreases by 0.7 percent, and GDP per capita increases by 0.8 percent.

**Related Literature.** Numerous empirical studies have examined the two-way relation between gender inequality and economic growth (Goldin, 1990; Galor and Weil, 1996; Lagerlöf, 2003; Doepke and Tertilt, 2009). These studies have reached the consensus that gender inequality hinders growth, while economic growth, in turn, improves gender equality.[2] This study relates to the first causality, and specifically 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.[3]

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 complement the one presented here, however, none consider the gender-specific frictions in the labor market as a potential channel. Additionally, Adamopoulos, Brandt, Leight and Restuccia (2022) 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, namely gender inequality, impacts the allocation of talent. This leads to a reduction in 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), and Hsieh, Hurst, Jones and Klenow (2019). The assumptions of this model regarding preference and technology are akin to Lagakos and Waugh (2013), who built a selection model with two sectors to explain the cross-country gap in agricultural productivity. I introduce separable leisure in the utility function following the approach of 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

[2] See World Bank (2011) for a comprehensive review of the literature and policy implications on this topic.

[3] 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).

and Blacks in the US.[4] The present study complements Lagakos and Waugh (2013) with an emphasis on the allocation of female talent and provides counterfactual results by setting gender-sector specific frictions to the current US level for all countries.[5]

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 propositions; in Section 4 I calibrate and estimate the model; in Section 5 I simulate the model and run counterfactual exercises including two extensions of the model; Section 6 applies the model to historical development of two countries, the US and Turkey, from 1960 until 2010 and 2000, respectively; Section 7 discusses additional concerns and provides robustness checks; and Section 8 concludes.

## 2 Motivating Evidence

This study is driven by three observations: (i) cross-country disparities in 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) and (b) of Figure 1 reproduce a finding from Caselli (2005).[6] Agricultural and non-agricultural labor productivity (defined as GDP per worker) increases with the overall income of countries, but the proportional increase in agricultural labor productivity is greater.

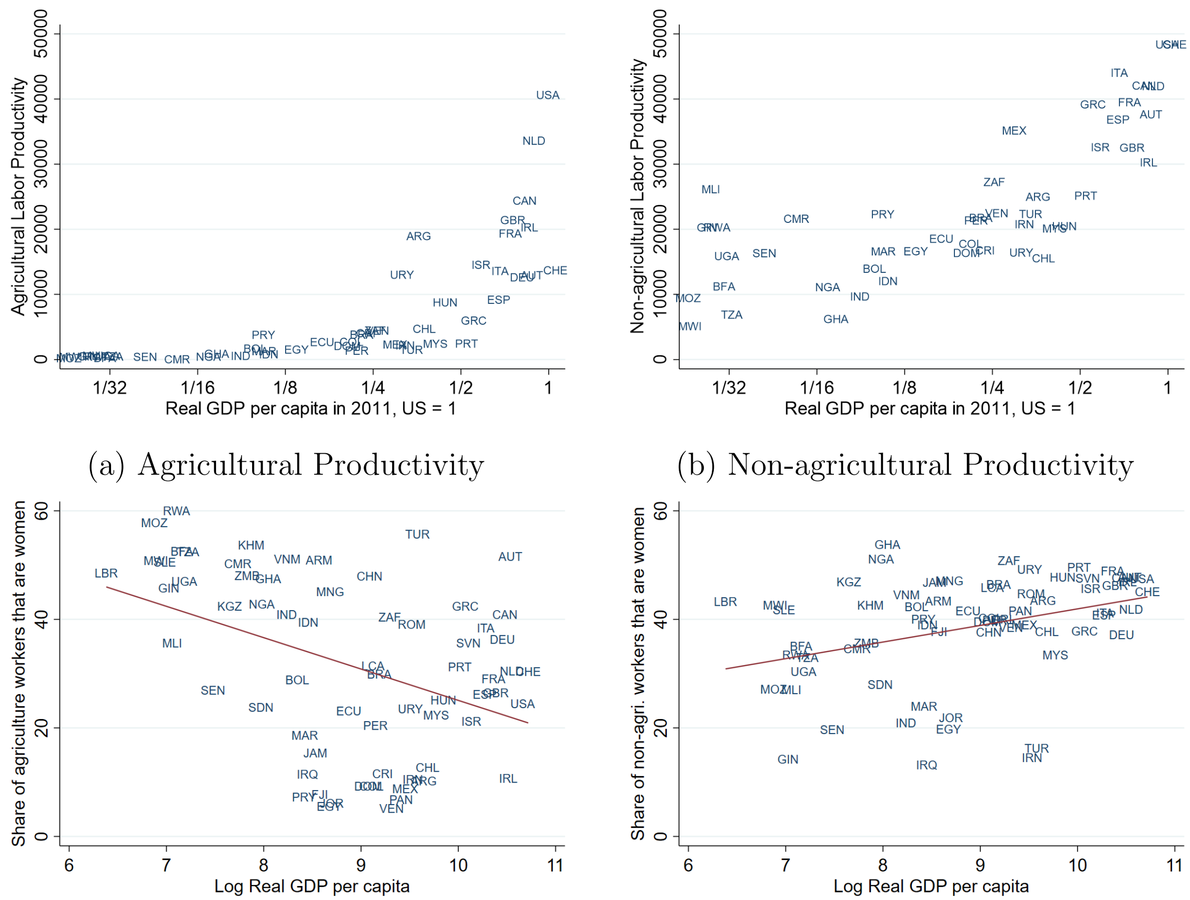
To measure the share of female workers in agriculture, I use data from the Integrated Public Use Microdata Series, International (IPUMS-International) provided by Minnesota Population Center (2020). IPUMS-International contains micro-level census data from 81 countries. For consistency, I calculate the share of female workers from the 66 countries used in our main analysis. Panel (c) of Figure 1 depicts that lower-income countries have a higher share of female workers in agricultural sectors. This raises the question of whether women are overrepresented in other occupations. Panel (d) of Figure 1 shows that lower income countries have a lower share of female workers in non-agricultural sectors. Consequently,

[4]The occupational choice model also resembles Burstein, Morales and Vogel (2019) that divide workers into multiple labor groups by gender, education, and age, and study changes in U.S. between-group inequality in a friction-less economy.

[5]The counterfactual exercise in Hsieh, Hurst, Jones and Klenow (2019) keeps friction against minorities constant at the level of 1960. The study finds that 15 to 20 percent of growth in aggregate output per worker between 1960 and 2008 can be attributed to the improved allocation of talent.

[6]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: Sectoral Labor Productivity and Share of Agricultural/Non-agricultural Workers that are Women




(a) Agricultural Productivity

(b) Non-agricultural Productivity

(c) Share of Agricultural Workers that are Women

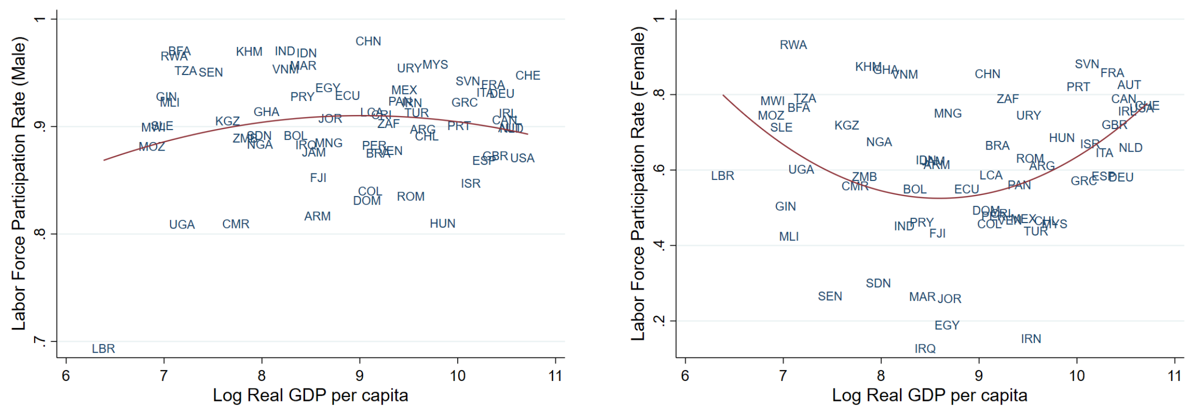
(d) Share of Non-agricultural Workers that are Women

Notes: Panel (a) and (b) show agricultural and non-agricultural labor productivity, from Caselli (2005), defined as GDP per worker. Real GDP per capita in 2011 is taken from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015) and normalized so that the United States' value is equal to one. Panel (c) 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 which is significant at the 1 percent level. Panel (d) 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.

the observed disparities in sectoral productivity and gender composition across countries motivate this study using a gender-specific occupational choice model.

Another important aspect in the labor market is the labor force participation decision, which also differs between genders. Figure 2 shows the well-known U-shaped female labor force participation and inverted U-shaped male labor force participation over the process of economic development (Goldin, 1995; Ngai, Olivetti and Petrongolo, 2021; Dinkelman

Figure 2: Labor Force Participation Rates, Men and Women




(a) Men

(b) Women

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

and Ngai, 2021, 2022). This non-monotonic pattern may not explain the monotone increase in agricultural productivity/non-agricultural productivity (Panels (a) and (b) of Figure 1). However, I will incorporate this fact on labor force participation into the model, in order to control for potential selection effects that may arise from differences in labor force participation rates by gender. At the end, this aspect plays non-negligible role in explaining the disparities in cross-country labor productivity as discussed in Section 5.4.

# 3 Baseline Model

## 3.1 Preferences and Endowments

Suppose a world where half of the population is male and another half of the population is female. 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 receive 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:

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

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 that 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 across countries (Figure 2).[7]

Each worker is endowed with one unit of time and productivity in each industry, $(\epsilon^i_{ag}, \epsilon^i_{ng})$, where $\epsilon^i_{ag}$ is the productivity in agriculture and $\epsilon^i_{ng}$ is the productivity in non-agriculture that depends on the gender $g \in \{m, w\}$ ($m$ is male and $w$ is female). She or he 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:

$$P_a c^i_a + c^i_n \leq h^i y^{i*} = (1 - \tau_{sg}) w_s h^i \epsilon^i_{sg} \tag{2}$$

where $P_a$ is the relative price of the agricultural good, and the non-agricultural good is numeraire. The derivation of the solution to the workers' optimization problem is laid out in Appendix B.

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 $\epsilon^i_{sg}$. The friction can be any type of gender-specific barrier, like preference-based discrimination by an employer. When an employer discriminates against employees, she or he acts as if she or he incurs non-pecuniary and psychological costs of production associated with 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 symmetrical such that any revenue collected from taxes in one sector is allocated 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 bi-variate Fréchet distributions. The shape parameter $\theta$ and the correlation parameter $\rho$ are assumed the same across industries and genders.[8] I assume location parameters $T_{sg}$ are dif-

[7] Note that the individual-level extensive margin of labor force participation decision does not exist; instead, the average working time in a country serves as a proxy for aggregate labor force participation.

[8] 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

ferent across industries and genders. This difference reflects the fact that men and women might have a different average productivity in each industry. For instance, a large body of literature finds women are on average less productive at agricultural work than men (Goldin and Sokoloff, 1984; Pitt, Rosenzweig and Hassan, 2012).

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

## 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:

$$Y_a = AL_a \text{ and } Y_n = AL_n \tag{3}$$

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:

$$\begin{aligned} L_a &= \sum_{g\in\{m,w\}} q_g p_{ag} E[h_{ag}\epsilon_{ag}|\text{Person chooses } a] \\ L_n &= \sum_{g\in\{m,w\}} q_g p_{ng} E[h_{ng}\epsilon_{ng}|\text{Person chooses } n] \end{aligned} \tag{4}$$

where $q_g$ is the total number of working people in gender $g$ ($q_m = q_w$, by assumption on equal population). The $p_{sg}$ is the fraction of people in gender $g$ that work in industry $s$. The total working hours in each industry are defined as:

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

$$\begin{aligned} N_n &= \sum_{g\in\{m,w\}} q_g p_{ng} E\left[h_{ng}|\text{ Person chooses } n\right] \\ &= \sum_{g\in\{m,w\}} q_g \left(1 - p_{ag}\right) E\left[h_{ng}|\text{ Person chooses } n\right] \end{aligned} \tag{5}$$

(2023). 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.

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

$$w_a = P_a A \text{ and } w_n = A \tag{6}$$

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

## 3.3 Occupational Choice

The problem of occupational choice is reduced 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:

$$\begin{aligned} y_w^{i*} &\equiv \max\{(1-\tau_{aw}) w_a \epsilon_{aw}^i, (1-\tau_{nw}) w_n \epsilon_{nm}^i\} \\ &= \max\{(1-\tau_{aw}) P_a A \epsilon_{aw}^i, (1-\tau_{nw}) A \epsilon_{nw}^i\} \end{aligned} \tag{7}$$

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:

$$\tilde{\tau}_{n|a} \equiv \frac{1-\tau_{aw}}{1-\tau_{nw}} \tag{8}$$

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). 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, counterfactual exercises are conducted incorporating the 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:*

$$p_{am} = \frac{(P_a^{\theta} T_{am})^{1/(1-\rho)}}{(P_a^{\theta} T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}} \text{ and } p_{nm} = 1 - p_{am}$$

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

*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, $\tilde{\tau}_{n|a}$.*

While the propensity of men is optimally determined by the consumer and firm's maximization problems, frictions distort the propensity for women. Women are more likely to choose an industry with lower barriers. This Roy-type model generates the average quality of workers in an industry for each group, and the following proposition demonstrates this 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:*

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

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

*Since 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 skilled 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:*

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

*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:*

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

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 infer the implied frictions from the data.

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

**Proposition 4** (Sectoral Labor Productivity): *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

Given that there are two sectors and preferences are non-homothetic, frictions are modeled as taxes, and the revenue from that "tax" is redistributed to workers to ensure that Walras's Law holds. This is different from standard models with a single final goods sector and homothetic preferences (e.g. Hsieh, Hurst, Jones and Klenow (2019)) where the sum of frictions is simply added into the aggregate demand for final goods. Under non-homothetic preferences, the sum of frictions should be redistributed to consumers who make consumption decisions between agricultural and non-agricultural goods, which govern the aggregate demands in each sector.

I introduce a government that maintains a balanced budget every period. The government imposes the gender-sector specific wage tax or subsidy on each female worker. This tax and subsidy system is my model's implementation of the gender-sector 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:

$$\sum_i \sum_s \sum_g \tau_{sg} w_s h^i \epsilon^i_{sg} = 0 \tag{9}$$

Given this set-up, Walras's Law holds: a $P_a$ that clears the agricultural output market will also clear the non-agricultural output market.

There are various ways of achieving a balanced budget. For instance, revenue from a wage tax on female workers can be allocated to male workers as a subsidy. In this case, consumption and occupational choices made by male workers will be directly affected by this subsidy. In order to minimize the distortion on the choices of male workers, I allocate wage taxes and subsidies among female workers only, across both sectors.

### 3.5 Equilibrium

Each worker takes prices and wages as given. They choose which sector to supply their labor and they maximize their utility subject to their budget constraint. Moreover, representative firms in each industry 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 workers optimize and both labor markets and output markets clear. Lastly, the government achieves a 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 ($\tau^{\text{US}}_{aw} = \tau^{\text{US}}_{nw} = 0$). Although there might be positive gender-specific frictions in the US, this assumption is simply a normalization for cross-country comparison and counterfactual exercise.[9]

I calibrate the distributional parameters $\{T_{ig}, \theta, \rho\}$ and preference parameters $\{\bar{n}, \nu\}$ from the US economy. Given those parameters, country-specific parameters, $\{A^{\text{c}}, \tau^{\text{c}}_{aw}, \tau^{\text{c}}_{nw}\, \psi^{\text{c}}_m, \psi^{\text{c}}_w\}$, where c indicates countries, are calibrated. I first calibrate parameters for individual talent distributions.

[9]Note that the relevant literature still finds evidence of gender-specific frictions in many sectors, particularly against women. 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 surgeon's after a patient's death, indicated by a sharper drop in referrals to the female surgeon.

## 4.1 Calibration of Distributional Parameters

Individual talents follow a bivariate Fréchet distribution:

$$F_g(\epsilon_a, \epsilon_n) = \exp(-\left[\Sigma_{s\in\{a,n\}}(T_{sg}\epsilon_s^{-\theta})^{1/(1-\rho)}\right]^{1-\rho}) \tag{10}$$

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

Moreover, I introduce an independent talent distribution ($\rho = 0$) as a benchmark case. Section 5.2 considers a case with non-zero correlation in talents. 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.*

Furthermore, I examine the wage dispersion within an industry by gender using micro-data from the 2010 American Community Survey (Ruggles, Flood, Foster, Goeken, Pacas, Schouweiler and Sobek, 2021). In order to reliably calculate hourly wages, I only use employed individuals between 25 and 55 years of age, and exclude people serving in the armed forces. The sample comprises both full- and part-time workers, and I follow McGrattan and Rogerson (1998) in calculating the hourly wage. Then, 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. This variance is discounted by a half to capture the permanent component.[10] The point estimate for $\theta$ is 3.5.[11]

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

[10] I follow Lagakos and Waugh (2013) to decompose permanent and transitory components using the same data. This study uses workers in 1996-2010 March CPS data who are matched across two consecutive years. The authors model log wages as the sum of a permanent component and a transitory component. By assuming that 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 the covariance of log wages from two consecutive periods for each worker. I find that about a half of the empirical dispersion can be interpreted as the permanent component.

[11] This number is higher than Hsieh, Hurst, Jones and Klenow (2019) ($\theta = 2$), because I use only the permanent component of wage variation. I assume that skills are equally dispersed by gender. If I use men and women separately in the data, I get $\theta = 3.63$ for men and $\theta = 3.38$ for women, which are close to the estimate of $\theta = 3.5$ from the pooled sample.

$$\frac{T_{am}}{T_{aw}} = (1-\tau_{aw}^{\text{US}})^{\theta}\left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{\theta}\left(\frac{p_{am}}{p_{aw}}\right)^{1-\rho} = \left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{3.5}\left(\frac{p_{am}}{p_{aw}}\right)$$
$$\frac{T_{nm}}{T_{nw}} = (1-\tau_{nw}^{\text{US}})^{\theta}\left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{\theta}\left(\frac{p_{nm}}{p_{nw}}\right)^{1-\rho} = \left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{3.5}\left(\frac{p_{nm}}{p_{nw}}\right) \tag{11}$$

where the gender wage gap $\frac{\overline{wage}_m}{\overline{wage}_w}$ and the propensity gaps $\{\frac{p_{am}}{p_{aw}}, \frac{p_{nm}}{p_{nw}}\}$ 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 higher if male workers are paid more (i.e., gender wage gap is larger) 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 infer the ratio of average talents.

Using the 2010 US census micro-data, I calculate the propensity gaps in both sectors and gender wage gap. Propensities are calculated as the probability of selecting into a certain industry conditional on being in the labor force.[12] Estimates are $\{p_{am}, p_{aw}, p_{nm}, p_{nw}\} = \{2.07\%, 0.75\%, 97.93\%, 99.25\%\}$. These estimates mean that a male worker has a 2 percent probability of entering agricutural work and a 98 percent probability of entering non-agricultural work. As such, the gender wage gap (male/female) is estimated to be 1.30039.

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

$$\frac{T_{am}}{T_{aw}} = \left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{3.5}\left(\frac{p_{am}}{p_{aw}}\right) = (1.30039)^{3.5} * \frac{2.07\%}{0.75\%} = 6.9066$$
$$\frac{T_{nm}}{T_{nw}} = \left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{3.5}\left(\frac{p_{nm}}{p_{nw}}\right) = (1.30039)^{3.5} * \frac{97.93\%}{99.25\%} = 2.4744 \tag{12}$$

The location parameter for men in agriculture is nearly seven times higher than for women. This parameter reflects the 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. In contrast, the difference in non-agriculture is smaller at 2.4744, reflecting a smaller productivity disparity between men and women in non-agricultural work. If the wage difference were zero, the ratio of the location parameter in non-agriculture would be close to one.

[12] The IPUMS-International data includes all employment types. Therefore, estimated propensities include both self-employed and wage/salary workers. It is important to include self-employed workers to capture a majority of labor force in low-income countries. Appendix C shows that the calibrated frictions based on only self-employed workers or only wage/salary workers are strongly correlated with the benchmark case, using 61 countries where the employment type information is available.

The main takeaway from estimated location parameters is a comparative (not absolute) advantage structure between male and female workers. Female workers have a comparative advantage in non-agriculture while male workers have a comparative advantage in agriculture, which is consistent with Goldin and Sokoloff (1984) and Pitt, Rosenzweig and Hassan (2012). In the occupational choice model, it is important to note that comparative advantage, not absolute advantage, is a key determinant. Therefore, the underlying assumption of the model is that the comparative advantage structure is the same across countries, which we relax in Section 5.3.

## 4.2 Full Calibration for the US Economy

Given the calibrated values for the distributional parameters, $\{\theta, T_{am}/T_{aw},\ T_{nm}/T_{nw}\} = \{3.5, 6.9066, 2.4744\}$, I calibrate all other parameters in the model jointly to match moments from US data. Given a normalization of $T_{am} = 1$ and $T_{nm} = 1$, I need four moments to match the four preference parameters, $\{\bar{n}, \nu, \psi_m, \psi_w\}$. The first moment I target is the fraction of workers in agriculture. From the 2010 census in IPUMS-International, the value is 1.4 percent. The second moment I target is a long-run agricultural employment share of 0.5 percent, which has been used by others in the literature (Restuccia, Yang and Zhu, 2008; Lagakos and Waugh, 2013).[13] The method used to calibrate this parameter is to note that the relative wage must equal the expenditure share relative to the labor share. Therefore, this pins down the expenditure share, $\nu$. The third and fourth moments are male and female labor force participation rates, which are 87.3% and 76.6% in US data. The calibrated values for the subsistence consumption requirement, the relative taste for non-agricultural consumption, and the relative non-agricultural talent are $\bar{n} = 204.0414$, $\nu = 199$, $\psi_m = 9.4109$, and $\psi_w = 14.8753$.

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

## 4.3 Calibration of Country-specific Parameters

I calibrate five country-specific parameters: $\{A^{\mathrm{c}}, \psi_m^{\mathrm{c}}, \psi_w^{\mathrm{c}}, \tau_{aw}^{\mathrm{c}}, \tau_{nw}^{\mathrm{c}}\}$ 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:

$$\frac{1}{1-\tau_{aw}} = \frac{\overline{\text{wage}}_m}{\overline{\text{wage}}_w}\left(\frac{p_{am}}{p_{aw}}\right)^{(1-\rho)/\theta}\left(\frac{T_{am}}{T_{aw}}\right)^{-1/\theta} = \frac{\overline{\text{wage}}_m}{\overline{\text{wage}}_w}\left(\frac{p_{am}}{p_{aw}}\right)^{\frac{1}{3.5}}(6.9066)^{-\frac{1}{3.5}}$$

[13] This is a commonly used assumption that US agricultural employment share will decrease further in the long-run from the current level of 1.4 percent. Note that the share was 2.8 percent in 1985.

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 |
| $T_{aw}$ | 0.1448 | gender wage gap & propensities |
| $T_{nm}$ | 1 | normalization |
| $T_{nw}$ | 0.4041 | gender wage gap & propensities |
| Preference | | |
| $\bar{n}$ | 204.0414 | fraction of workers in agriculture = 1.4% |
| $\nu$ | 199 | long-run agricultural employment share = 0.5% |
| $\psi_m$ | 9.4109 | male labor force participation rate = 0.873 |
| $\psi_w$ | 14.8753 | female labor force participation rate = 0.766 |

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

$$\frac{1}{1-\tau_{nw}} = \frac{\overline{\text{wage}}_m}{\overline{\text{wage}}_w}\left(\frac{p_{nm}}{p_{nw}}\right)^{(1-\rho)/\theta}\left(\frac{T_{nm}}{T_{nw}}\right)^{-1/\theta} = \frac{\overline{\text{wage}}_m}{\overline{\text{wage}}_w}\left(\frac{p_{nm}}{p_{nw}}\right)^{\frac{1}{3.5}}(2.4744)^{-\frac{1}{3.5}} \tag{13}$$

The friction in one industry is estimated to be high when the gender wage gap is large and male workers are overrepresented in the industry. I have obtained data on the gender earning gap for 176 countries from UNDP (2015)[14], 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. Given that five countries in the IPUMS-International have no information on industry in their censuses and seven other countries are unmatched with the Penn World Tables, I have 66 matched countries for the quantitative analysis. To ensure data consistency, I selected the survey year from IPUMS-International that is closest to 2010 and contains all necessary information.[15]

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 for women as below in the 66 countries.

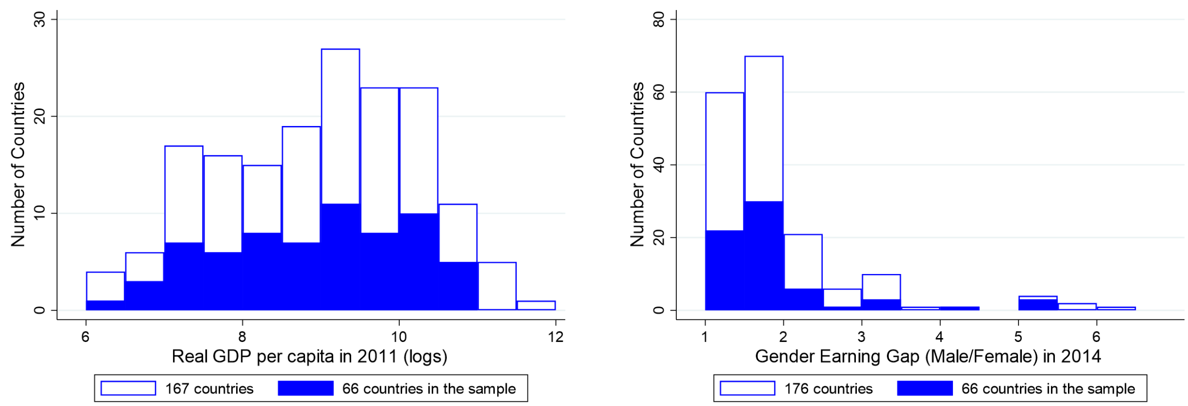
$$\tilde{\tau}_{n|a} \equiv \frac{1-\tau_{aw}}{1-\tau_{nw}}$$

The relative friction is negatively associated with real GDP per capita. Poor countries

[14] I use earning gaps instead of wage gaps due to two reasons: (i) earning gaps incorporate self-employment and (ii) wage gaps are not readily available from low-income countries.

[15] Refer to Table A.I in Appendix A for detailed sample construction.

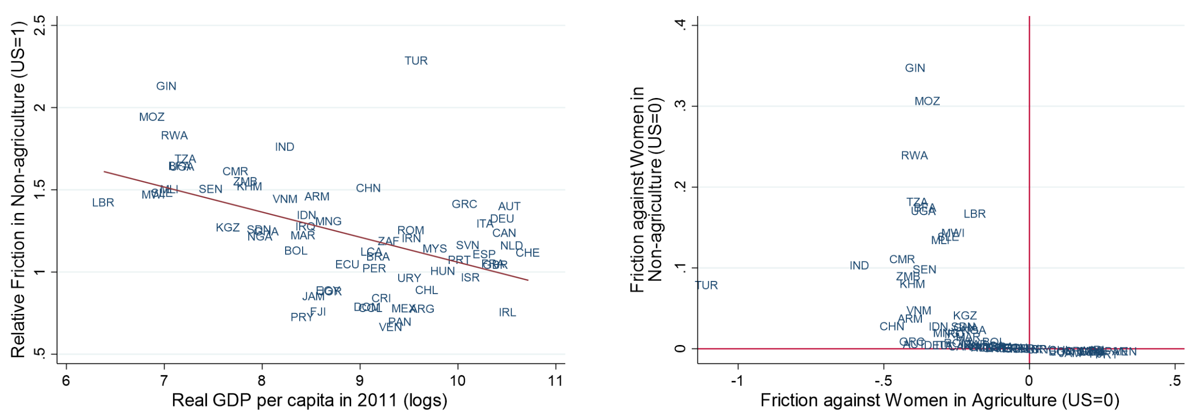
Figure 3: Representativeness of the Sample




(a) GDP per capita (b) Gender wage gap

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.

Figure 4: Estimated Frictions against Women




(a) Relative friction (b) Absolute size of friction

Notes: Panel (a) plots calibrated relative friction $\tilde{\tau}_{n|a}$ for all 66 countries. The y-axis is the relative friction and The x-axis is the real GDP per capita from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015). Mean relative friction is 1.24 with a standard deviation of 0.35. Panel (b) plots the calibrated values of $\tau_{nw}$ against $\tau_{aw}$ for all 66 countries. The horizontal and vertical lines lie on zero, which is the US level. Mean friction in non-agriculture is 0.04 with a standard deviation of 0.08. Mean friction in agriculture is -0.16 with a standard deviation of 0.26.

have a higher relative friction in non-agriculture than rich countries. From Proposition 2, *ceteris paribus*, the average quality of women in agriculture decreases as the frictions in non-agriculture decrease. Therefore, poorer countries' low labor productivity in agriculture relates to the misallocation of female talent due to higher frictions in non-agriculture. To calibrate the absolute magnitude of these frictions, I search for $\tau_{aw}$ and $\tau_{nw}$ that satisfies the government balanced budget constraint (equation 9) and is consistent with the calibrated relative frictions. Calibrated absolute frictions are shown in panel (b) of figure 4.

There is a potential measurement concern. Gollin, Lagakos and Waugh (2014b) find that the agricultural labor productivity puzzle is reduced when productivity is expressed in hours worked. In Appendix D, I use 18 countries with working hour information available and calculate relative frictions with hours-adjusted labor inputs. I find that relative frictions calculated with the number of workers and relative frictions calculated with hours-adjusted labor inputs are very close to each other. The correlation between these two frictions is 0.97 and significant at 1 percent. Therefore, even though there is a considerable variation in working hours across countries (Bick, Fuchs-Schündeln and Lagakos, 2018), taking into account working hours does not change the estimated relative friction in non-agriculture against women. Appendix D provides a detailed discussion.

Other country-specific parameters are relative utility of leisure, $\psi_m^c$ and $\psi_w^c$. These 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 levels, the relative utility of leisure has a higher value when the country has a lower labor force participation rate for a specific gender. Under the separable utility function in this model, leisure preference should increase to match labor force participation rates while overall productivity (and eventually GDP) increases.

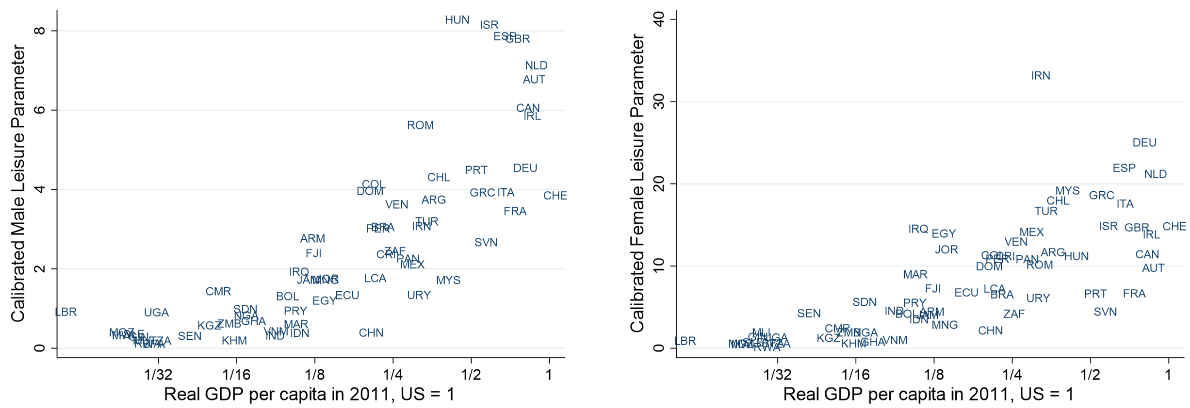
The last country-specific parameter to calibrate is productivity, $A$. Given the calibrated frictions, I calibrate productivity for the remaining 65 countries ($A$ is normalized to 100 for the United States). The target is real GDP per capita relative to the United States, from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015). The calibrated $A$ ranges from 101.9586 (Switzerland) to 1.7609 (Liberia) for targets ranging from 1.058 (Switzerland) to 0.138 (Liberia).

# 5 Quantitative Analysis

## 5.1 Counterfactual Analysis

One advantage of the model is that I can conduct a normative analysis by adjusting frictions for some countries. I conduct one counterfactual analysis in which the gender-specific

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 ($\psi_w^c$).

frictions are set to the current US value for all other countries.[16]

Table II: Summary statistics of differences between baseline and counterfactual

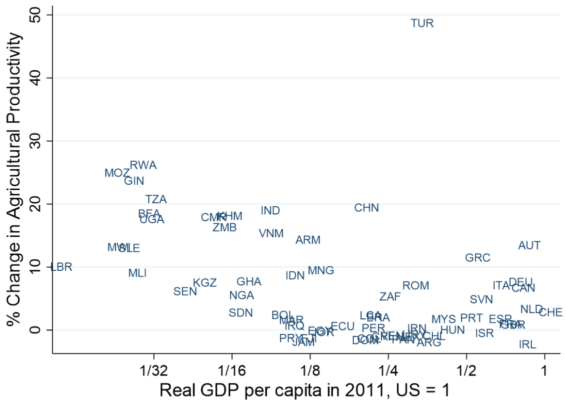
| | mean | sd | p25 | p50 | p75 |
|---|---|---|---|---|---|
| % change in ag productivity | 7.16 | 9.40 | 0.09 | 3.38 | 13.05 |
| % change in non-ag productivity | -1.32 | 2.36 | -1.53 | -0.12 | -0.00 |
| % change in real GDP per capita | 1.35 | 2.54 | 0.01 | 0.07 | 1.09 |
| % change in female labor supply | 0.89 | 1.71 | 0.02 | 0.13 | 0.83 |
| % change in male labor supply | 0.08 | 0.22 | 0.00 | 0.01 | 0.06 |
| % ag consumption equivalent welfare gain | 2.83 | 4.94 | 0.28 | 1.05 | 3.55 |

Notes: Agricultural productivity is $P_a * Y_a$ divided by total hours of labor in agriculture. Non-agricultural productivity is $Y_n$ divided by total hours of labor in non-agriculture. Real GDP is $P_a * Y_a + Y_n$. Female/male labor supply is the average number of hours of labor chosen by female/male 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 scenario.

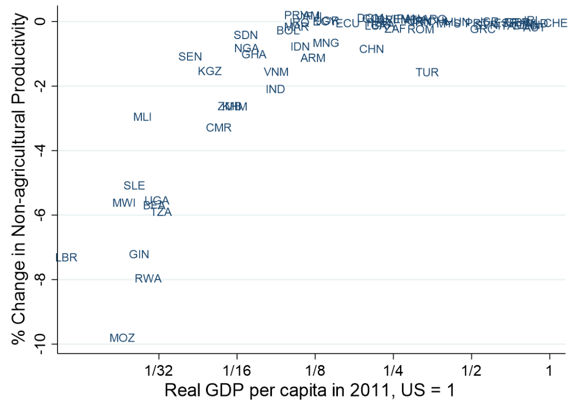
Panel (a) and (b) of Figure 6 and the first two rows of Table II show the percentage change in agricultural and non-agricultural labor productivity after setting gender specific frictions to the current US value, $\tau_{aw}^c = \tau_{nw}^c = 0$ for each country. Therefore, for the United States, the change is zero. On average, this counterfactual exercise increases agricultural

[16]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.

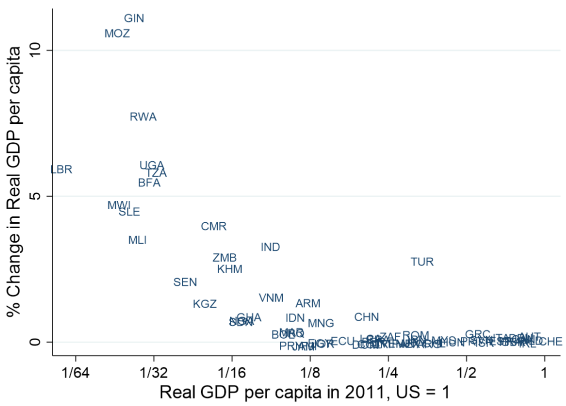
Figure 6: Counterfactual Results




(a) Agricultural Productivity




(b) Non-agricultural Productivity




(c) Real GDP per Capita

Notes: This figure plots the results from the counterfactual exercise of setting frictions to the current US level. Panel (a) and (b) shows the counterfactual percentage change in agriculture and non-agriculture labor productivity. Agricultural productivity is $P_a * Y_a/L_a$, non-agricultural productivity is $Y_n/L_n$. Aggregate productivity is $(P_a * Y_a + Y_n)/(L_a + L_n)$. $L_a$ and $L_n$ are hours of labor input in each sector. I use the modeled US equilibrium price for $P_a$ to calculate the value of agricultural output in my model, following Lagakos and Waugh (2013). Panel (c) shows the counterfactual percentage change in real GDP per capita. Real GDP is $P_a * Y_a + Y_n$. The x-axes are real GDP per capita taken from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015).

labor productivity by 7.2 percent, and decreases non-agricultural productivity by 1.3 percent. These gains are attributed mainly to poorer countries.

These counterfactual changes are driven by the selection effect in our model. In the simulation, only women who are sufficiently talented to overcome frictions enter the non-agricultural sector. When the friction is lifted to the level of the US in the counterfactual, more women enter the non-agricultural sector, which reduces the average productivity of female workers in that sector.

The changes in labor productivity say little about an economy's welfare. Panel (c) of Figure 6 and the third row of Table II show the percentage change in real GDP per capita by setting the frictions to the current US value.[17] GDP per capita increases by 1.35 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 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.[18]

The 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. Female labor supply increases by 0.89 percent on average, male labor supply remains nearly unchanged (0.08 percent increase on average), and agricultural consumption equivalent welfare gain increases on average by 2.83 percent.

## 5.2 Extension #1: 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). The 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 $\rho$. The values of the other distribution parameters are taken from the calibration with $\rho = 0$. Then, The price of agricultural goods is 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.

[17] 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 changes in quantities.

[18] 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.

Given the similarity in context, the same intuition applies here, which I replicate from Lagakos and Waugh (2013): the calibrated values of $T_{aw} = 0.1448$ and $T_{nw} = 0.4041$ that I use imply that the variance in ability draws is much higher in non-agriculture than in agriculture for women. With a 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. In turn, 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 a 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 0.2281, while the Spearman rank correlation coefficient is 0.1460.

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 value of $\rho = 0.107$. I start by calibrating the model to the US economy by matching four moments: the fraction of workers in agriculture, the long-run share of agricultural employment, and male and female labor force participation rates. The calibration produces parameters: $\bar{n} = 211.5014$, $\nu = 199$, $\psi_m = 9.1680$, $\psi_w = 14.4203$. 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.

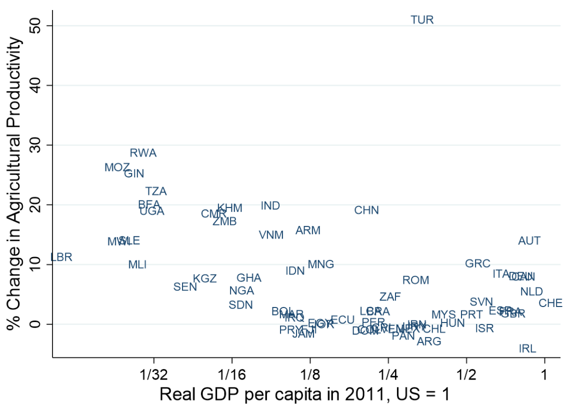
Table III: Summary statistics of differences between baseline and counterfactual, $\rho > 0$

| | mean | sd | p25 | p50 | p75 |
|---|---|---|---|---|---|
| % change in ag productivity | 7.62 | 9.88 | 0.25 | 3.82 | 13.93 |
| % change in non-ag productivity | -1.43 | 2.57 | -1.50 | -0.12 | -0.01 |
| % change in real GDP per capita | 1.49 | 2.77 | 0.01 | 0.07 | 1.16 |
| % change in female labor supply | 1.02 | 1.93 | 0.03 | 0.15 | 1.09 |
| % change in male labor supply | 0.09 | 0.22 | 0.00 | 0.01 | 0.07 |
| % ag consumption equivalent welfare gain | 3.10 | 5.59 | 0.31 | 1.02 | 3.31 |

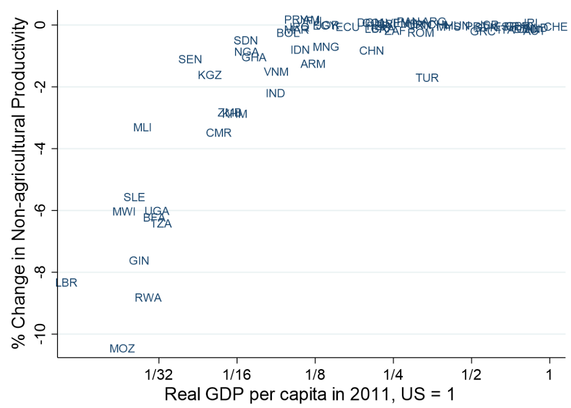
Note: Agricultural productivity is $P_a * Y_a$ divided by total hours of labor in agriculture. Non-agricultural productivity is $Y_n$ divided by total hours of labor in non-agriculture. Real GDP is $P_a * Y_a + Y_n$. Female/male labor supply is the average number of hours of labor chosen by female/male workers. The percent of 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 scenario.

I conduct the same counterfactual analysis in which the female sector-specific frictions

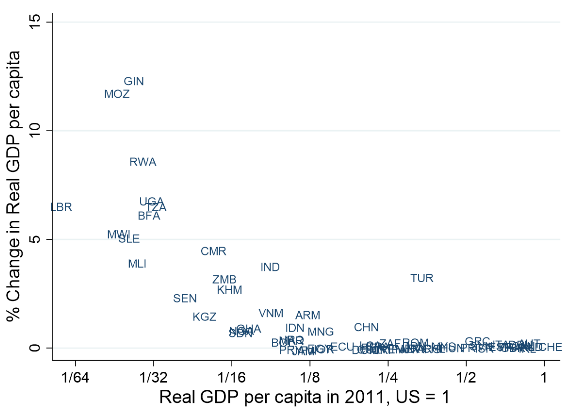
Figure 7: counterfactual Results, Correlated Talents




(a) Agricultural Productivity




(b) Non-agricultural Productivity




(c) Real GDP per Capita

Notes: This figure plots the results from the counterfactual exercise of setting frictions to the current US level. Panel (a) and (b) shows the counterfactual percentage change in agricultural and non-agricultural labor productivity. Agricultural productivity is $P_a * Y_a/L_a$, non-agricultural productivity is $Y_n/L_n$. Aggregate productivity is $(P_a * Y_a + Y_n)/(L_a + L_n)$. $L_a$ and $L_n$ are hours of labor input in each sector. I use the modeled US equilibrium price for $P_a$ to calculate the value of agricultural output in my model, following Lagakos and Waugh (2013). Panel (c) shows the counterfactual percentage change in real GDP per capita. Real GDP is $P_a * Y_a + Y_n$. The x-axes are real GDP per capita taken from PWT 8.1 (Feenstra, Inklaar and Timmer, 2015).

are set to the current US value, i.e., zero. Figure 7 and Table III show the counterfactual results. Quantitatively, the increases in 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 (2023), who find larger and more heterogeneous gains from trade under positive correlation in productivity across countries, relative to models that assume independent productivity.

## 5.3 Extension #2: Results with Country-specific Comparative Advantages

A key assumption in the baseline calibration is that the talent distribution is the same across countries, i.e., the location parameters for the Fréchet distribution, $T_{sg}$, are the same across countries. The ratio of the location parameters (comparative advantages) are identified from the ratio of wages and propensities in the benchmark economy (US in 2010), as in equation 12:

$$\frac{T_{am}}{T_{aw}} = \left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{3.5} \left(\frac{p_{am}}{p_{aw}}\right)$$

$$\frac{T_{nm}}{T_{nw}} = \left(\frac{\overline{wage}_m}{\overline{wage}_w}\right)^{3.5} \left(\frac{p_{nm}}{p_{nw}}\right)$$

So far, my approach attributed all the differences in gendered propensities between the US and other countries to gendered frictions. We may loosen this assumption by calibrating comparative advantage parameters, $T_{sm}/T_{sw}$, using historical US data, then applying the calibrated US historical comparative advantage on other countries in the present day. The idea is that poor countries today may be more similar to the US in the past than the US today in terms of gendered comparative advantage, thus lessening the role that gendered frictions play in explaining modern day propensity gaps.

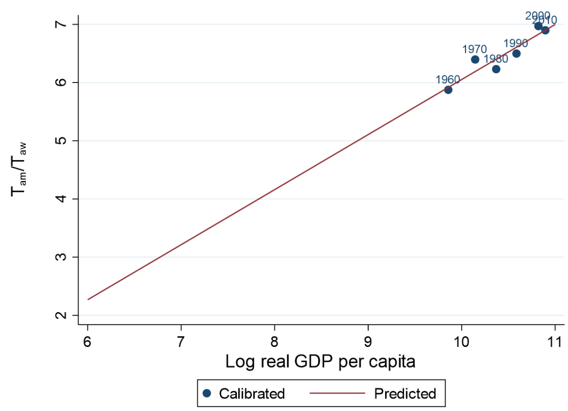
I use data from IPUMS-International for the US in 1960, 1970, 1980, 1990, 2000, and 2010[19] to calculate historical propensities $p_{sg}$, and then plug those into equation 12. Figure 8 reports US historical calibrated T-ratios from 1960 to 2010.

I make the assumption that a half of the differences in the propensities between historical and current periods are driven by the change in comparative advantages.[20] Under this assumption, the T-ratios, $T_{sm}/T_{sw}$ for $s = a, n$, are an average of the calibrated historical values and the baseline US 2010 value ($T_{am}/T_{aw} = 6.9066$, $T_{nm}/T_{nw} = 2.4744$).

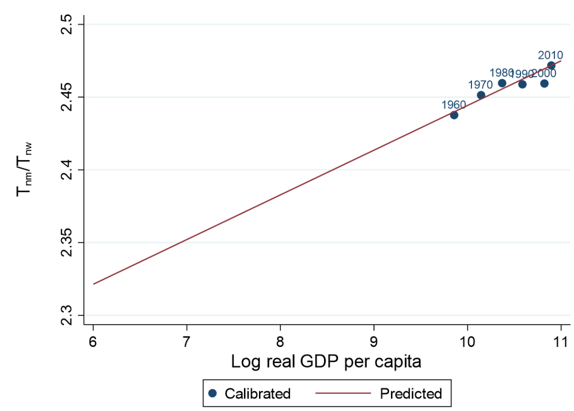
[19] Estimates for the labor force in the US prior to the 1940 Census are not consistent with our modern definitions of employment. This inconsistency is particularly severe for women, as well as younger and older men. See the extended discussion in Carter, Gartner, Haines, Olmstead, Sutch and Wright (2006), particularly the section on historical labor force estimates.

[20] Instead, I could assume that all differences in the propensities between historical and current periods are driven by the change in comparative advantages. In this case, frictions are assumed to be time-invariant, which is inconsistent with a series of papers documenting improvements in female talent allocation across sectors in recent US history (e.g., Hsieh, Hurst, Jones and Klenow (2019)). I can still conduct simulation and counterfactual exercises under this unrealistic assumption of no change in frictions in the US over time. In this case, by setting newly estimated frictions in each country to current US levels, agricultural labor productivity increases by 2.6 percent and GDP per capita increases by 0.5 percent on average. Those counterfactual gains are smaller than the gains under the assumption that differences in the propensities are driven equally by the change in frictions and the change in comparative advantages.

Figure 8: Historical comparative advantage




(a) Agriculture




(b) Non-agriculture

Notes: Values are calibrated using US historical propensities, and US 2010 wage ratios. The values presented in the plot and used for prediction are an average of calibrated historical values and the baseline values from the US 2010 calibration ($T_{am}/T_{aw} = 6.9066$, $T_{nm}/T_{nw} = 2.4744$). The prediction is obtained from a linear fit of the averaged values on log real GDP per capita. The range of log real GDP per capita in the plots reflects the range of values in our sample of 66 countries, with Liberia as the poorest country having log real GDP per capita of 6.3816.

To predict comparative advantage values for other countries in the present day, a linear fit of the averaged historical comparative advantage values for the US on historical log real GDP per capita is used. The predicted values from the linear fit are then applied on other countries based on their log real GDP per capita in the present.

Table IV reports outcomes of setting frictions to zero. The increase in agricultural productivity (4.27) is lower than the benchmark case (7.16 in Table II). The increase in real GDP per capita (0.76) is also lower than the benchmark case (1.35 in Table II). This is because I attribute parts of propensity gaps to differences in comparative advantage across countries by using US historical information. Under this assumption, in the US, female workers' comparative advantage in non-agriculture increased slightly over time, if we don't attribute all changes in propensities to changes in frictions. Taking this into account, the estimated magnitude of frictions decreases and therefore the size of counterfactual gain declines. The increase in real GDP per capita is 43% lower than the benchmark case, though it is still non-negligible.

## 5.4 Role of Non-zero Frictions and Non-uniform Leisure

In this section, I will discuss the role of non-zero frictions and non-uniform leisure across countries. To do so, I repeat the calibration procedure with different model setups. I focus

Table IV: Comparison of baseline and counterfactual, country-specific comparative advantage

| | mean | sd | p25 | p50 | p75 |
|---|---|---|---|---|---|
| % change in ag productivity | 4.27 | 7.06 | -0.47 | 1.89 | 7.16 |
| % change in non-ag productivity | -0.74 | 1.40 | -0.83 | -0.08 | 0.02 |
| % change in real GDP per capita | 0.76 | 1.51 | 0.02 | 0.06 | 0.69 |
| % change in female labor supply | 0.56 | 1.18 | 0.02 | 0.12 | 0.44 |
| % change in male labor supply | 0.04 | 0.10 | -0.00 | 0.00 | 0.04 |
| % ag consumption equivalent welfare gain | 1.78 | 3.78 | 0.15 | 0.87 | 2.10 |

Note: Agricultural productivity is $P_a * Y_a$ divided by total hours of labor in agriculture. Non-agricultural productivity is $Y_n$ divided by total hours of labor in non-agriculture. Real GDP is $P_a * Y_a + Y_n$. Female/male labor supply is the average number of hours of labor chosen by female/male 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 scenario.

on the 44 countries that have productivity data from Caselli (2005) and overlap with my sample.

Table V reports the ratios of mean agricultural productivity, non-agricultural productivity, aggregate productivity, and agricultural productivity/non-agricultural productivity of the top 20% to the bottom 20% countries. "Model 1: Baseline" is the model calibrated in Section 4, with country specific friction and leisure preference parameters. "Model 2: Zero friction" is the same model but with frictions set to US levels, i.e., zero, as described in Section 5.1. "Model 3: Uniform leisure" is the same as the baseline model except I do not calibrate country-specific leisure preference parameters and they are instead set to US values. "Model 4: Uniform leisure and zero friction" is the same as Model 3, except with frictions set to zero.

As expected, when comparing Model 1 and Model 2, setting frictions to zero reduces the gap between rich and poor countries both in terms of relative agricultural productivity and aggregate productivity. This is because frictions push women with a relative advantage in non-agriculture into agriculture, reducing their relative agricultural labor productivity, and these frictions are higher in poorer countries. When comparing Model 1 and Model 3, equalizing leisure preferences across countries reduces the gap in relative agricultural productivity but doesn't change the gap in aggregate productivity. Model 4 is comparable to the setup of Lagakos and Waugh (2013) where the model doesn't allow gender-specific comparative advantages and labor force participation rates. My baseline model explains agricutural/non-agricutural ratio by 45% ($=\frac{1.75-1.21}{1.21}$) more than Lagakos and Waugh (2013).

Table V: Ratio of mean productivity in the top 20% to the bottom 20% of countries by income with different models

| Model | Agriculture | Aggregate | Non-ag | Ag/non-ag ratio |
|---|---|---|---|---|
| Data | 43.53 | 25.26 | 2.52 | 17.31 |
| Model 1: Baseline | 49.59 | 25.20 | 18.00 | 2.75 |
| Model 2: Zero friction | 44.06 | 23.90 | 18.94 | 2.33 |
| Model 3: Uniform leisure | 12.24 | 25.17 | 5.07 | 2.42 |
| Model 4: Uniform leisure and zero friction | 11.48 | 24.09 | 5.19 | 2.21 |

Notes: This table reports the ratios of mean agricultural productivity, non-agricultural productivity, aggregate productivity, and agricultural productivity/non-agricultural productivity for countries in the top 20% to countries in the bottom 20% of GDP per capita. "Model 1: Baseline" is the model calibrated in Section 4, with country specific friction, productivity, and leisure preference parameters. "Model 2: Zero friction" is the same model but with frictions set to US levels, i.e., zero, as discussed in Section 5.1. "Model 3: Uniform leisure" is the same as the baseline model except I do not calibrate country-specific leisure preference parameters and they are instead set to US values. "Model 4: Uniform leisure and zero friction" is the same as Model 3, except with frictions set to zero.

# 6 Time-series Analysis

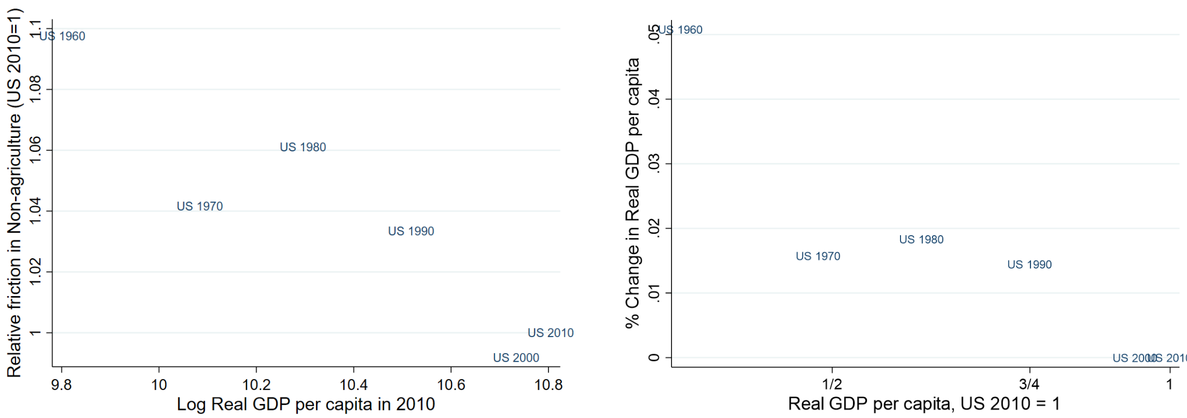
The model can be easily applied to understanding historical development of a single country. In this section, I apply the same model to historical development of two countries, the United States and Turkey.

## 6.1 US, 1960-2010

Panel (a) of Figure 9 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 (the first row of Table VI).

Panel (b) of Figure 9 and the third row of Table VI show that real GDP per capita would have increased by 0.02 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 9: Time-series Analysis, US 1960-2010




(a) Calibrated Relative Friction

(b) Counterfactual Change in Real GDP per capita

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 the real GDP per capita from Bolt, Inklaar, de Jong and van Zanden (2018). Panel (b) plots the results from the counterfactual 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 the real GDP per capita from Bolt, Inklaar, de Jong and van Zanden (2018).

Table VI: Comparison of baseline and counterfactual, US 1960-2010

| | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|
| % change in ag productivity | 0.96 | 0.56 | 0.98 | 0.63 | -0.04 |
| % change in non-ag productivity | -0.03 | -0.02 | -0.02 | -0.01 | 0.00 |
| % change in real GDP per capita | 0.02 | 0.01 | 0.01 | 0.00 | 0.00 |
| % change in female labor supply | 0.04 | 0.01 | 0.01 | 0.00 | 0.00 |
| % change in male labor supply | 0.00 | 0.00 | 0.00 | 0.00 | -0.00 |
| % ag consumption equivalent welfare gain | 0.07 | 0.05 | 0.07 | 0.01 | 0.17 |

Note: Agricultural productivity is $P_a * Y_a$ divided by total hours of labor in agriculture. Non-agricultural productivity is $Y_n$ divided by total hours of labor in non-agriculture. Real GDP is $P_a * Y_a + Y_n$. Female/male labor supply is the average number of hours of labor chosen by female/male 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 scenario.

## 6.2 Turkey, 1960-2000

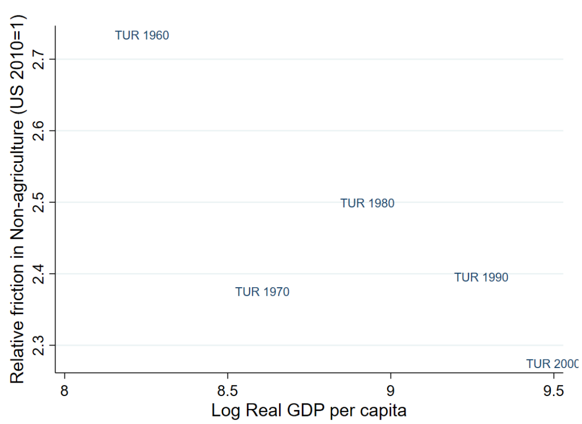
Another example is Turkey, where agricultural productivity loss and real GDP loss from female talent misallocation are large in the previous section. How does this country look like over time?

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 obtained from Statistical Indicators 1923-2009 (Göstergeler, 2010) and the GDP data are taken from the World Bank. The gender wage gap data are taken from multiple sources as follows. 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 Ilkkaracan and Selim (2007), who estimated the 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.

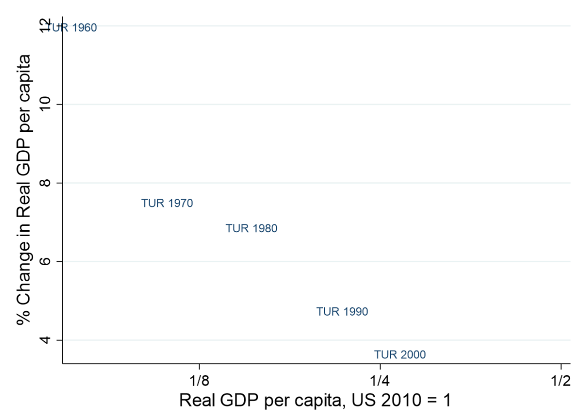
Panel (a) of Figure 10 shows the calibrated relative friction from 1960 to 2000. Female workers in Turkey faced much higher friction in non-agriculture both in 1960 and 2000 compared to the US economy, but the relative friction, and agricultural productivity loss (the first row of Table VII), has decreased over time.

As shown in panel (b) of Figure 10 and the third row of Table VII, real GDP loss from female talent misallocation was 10.59 percent in 1960, compared to 3.03 percent in 2000. This shows that there has been an improvement in female talent allocation in Turkey from 1960 to 2000, even though the gender disparity (and associated GDP loss) in Turkey today remains substantial when compared to the US.

Figure 10: Time-series Analysis, Turkey 1960-2000




(a) Calibrated Relative Friction




(b) Counterfactual Change in Real GDP per capita

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 the real GDP per capita from the World Bank. Panel (b) plots the results from the counterfactual 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 the real GDP per capita from the World Bank.

Table VII: Comparison of baseline and counterfactual, Turkey 1960-2000

| | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|
| % change in ag productivity | 63.46 | 50.75 | 56.70 | 52.93 | 49.81 |
| % change in non-ag productivity | -7.10 | -4.45 | -3.80 | -2.63 | -2.07 |
| % change in real GDP per capita | 10.59 | 6.58 | 5.96 | 4.02 | 3.03 |
| % change in female labor supply | 9.62 | 9.81 | 10.77 | 8.01 | 5.95 |
| % change in male labor supply | 0.17 | 0.28 | 0.20 | 0.16 | 0.20 |
| % ag consumption equivalent welfare gain | 54.37 | 36.49 | 45.05 | 39.09 | 33.41 |

Notes: Agricultural productivity is $P_a * Y_a$ divided by total hours of labor in agriculture. Non-agricultural productivity is $Y_n$ divided by total hours of labor in non-agriculture. Real GDP is $P_a * Y_a + Y_n$. Female/male labor supply is the average number of hours of labor chosen by female/male 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 scenario.

# 7 Discussion

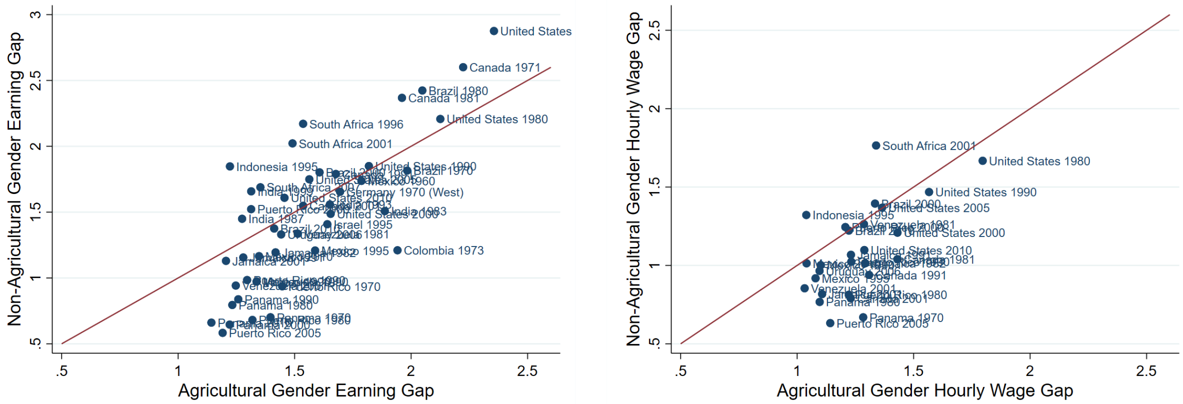
This section discusses additional concerns and provides robustness checks.

## 7.1 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 11 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 11 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 many cases.

Figure 11: Gender Gap in Agriculture and Non-agriculture




(a) Earning (b) Hourly Wage

Notes: Panel (a) plots the agricultural and non-agricultural gender earnings gaps that are calculated from 46 censuses for 15 countries from IPUMS-International (Minnesota Population Center, 2020). The solid line is a 45-degree line that is predicted by Proposition 3. Panel (b) plots the agricultural and non-agricultural gender earnings gaps that are calculated from 27 censuses for 11 countries from IPUMS-International (Minnesota Population Center, 2020). Hourly wage is calculated by dividing weekly wage income by weekly work hours, conditional on working more than zero hours. Weekly wage income is calculated by dividing annual wage income by 52 weeks. The solid line is a 45-degree line that is predicted by the Proposition 3.

## 7.2 Dispersion Parameter Estimated from Other Countries

The shape parameter of the Fréchet distribution, $\theta$, is calibrated from US data. How sensitive is it to the choice of country? To check the robustness of results, I calculate the dispersion parameter $\theta$ for four other countries where hourly wage data are available from Minnesota Population Center (2020): Brazil, Canada, Mexico, and Venezuela. I also re-estimate $\theta$ using the US data from two different years, 1990 and 2000. To be consistent with the estimate obtained from the US data, I only use employed individuals between ages of 25 and 55. I follow McGrattan and Rogerson (1998) to calculate the hourly wage from the censuse, and excluded top and bottom 1% income earners. Table VIII reports the estimated values of $\theta$ from each sample. All numbers are estimated to be close to the original estimate, 3.5.

## 7.3 Role of Education

The gap in average levels of education between rich and poor countries is considerable. 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 12 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

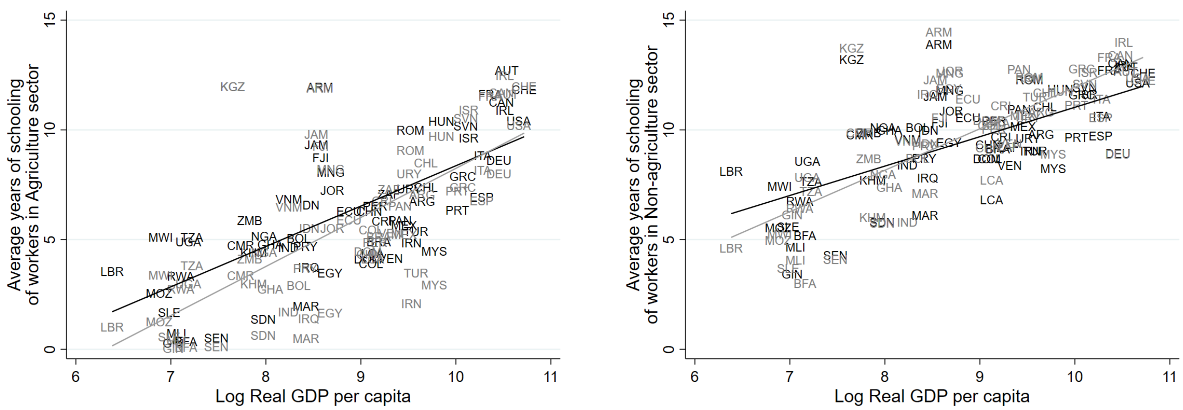
Table VIII: Estimated $\theta$ from Other Samples

| 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 four other countries and different years in the US. $\theta$'s are estimated from Proposition 5.

agriculture and non-agriculture. Therefore, it is less likely that education, measured by years of schooling, is a fundamental source of agricultural productivity gaps.

Figure 12: Years of Schooling across Countries within Each Sector




(a) Agriculture

(b) Non-agriculture

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. Black lines are the fitted lines for male, and grey 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).

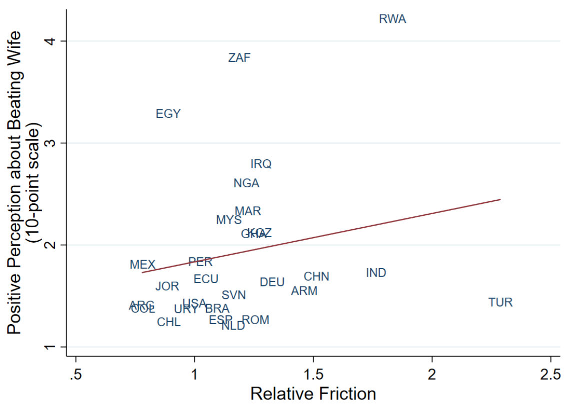
## 7.4 Potential Sources of the Friction

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.[21]

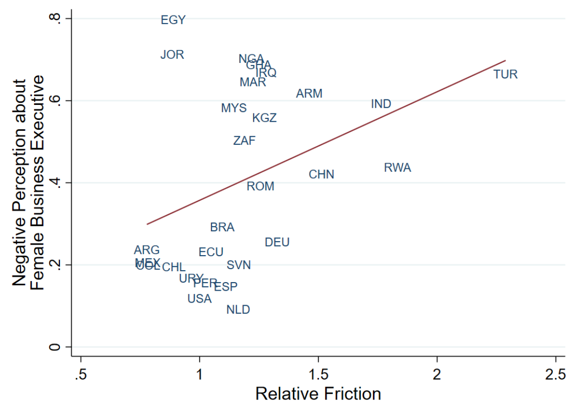
Six questions are about people's perception on beating wife, female business executives, women's university education, women earning more money, female political leader, and working mother, respectively. Among 60 countries in wave 6 (2010 through 2014) of the World Values Survey, I match 28 countries that coincide with the sample of this study. Figure 13 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.

[21] 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 gaps.

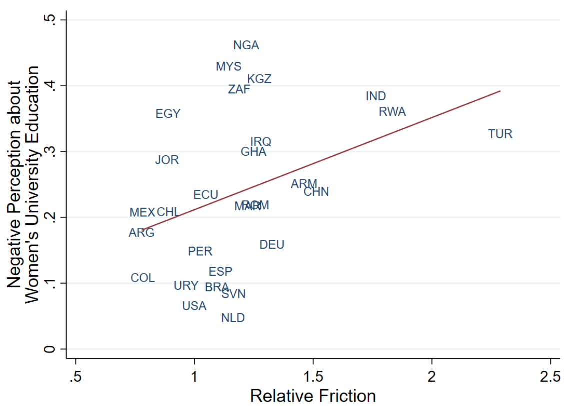
Figure 13: Perception and Relative Friction in Non-agriculture



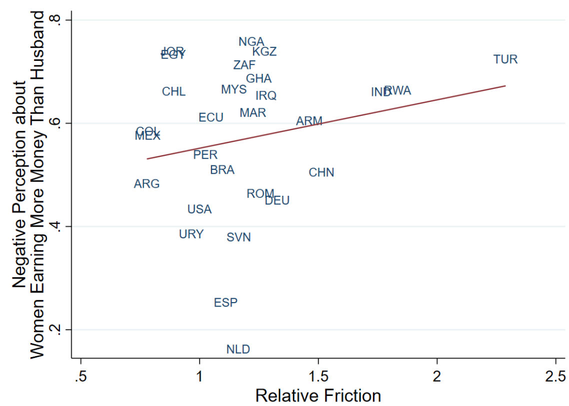

(a) Beating Wife




(b) Female Business Executives



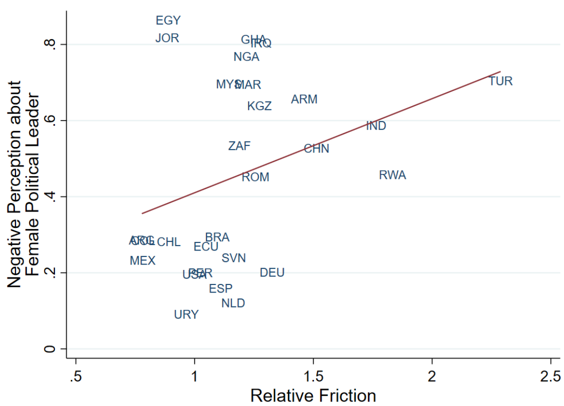

(c) Women's University Education



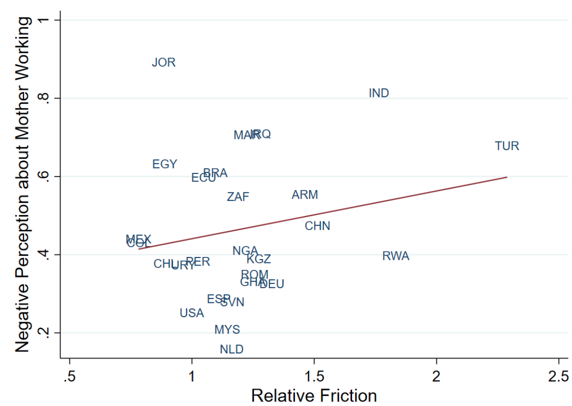

(d) Women Earning More Money

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 $p = 0.285$.

Figure 13: Perception and Relative Friction in Non-agriculture (continued)




(a) Female Political Leader




(b) Working Mother

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 female political leader is: "On the whole, men make better political leaders than women do". The slope of the fitted line is 0.247 with $p = 0.082$. In panel (b), the item on the perceptions on mother working is: "When a mother works for pay, the children suffer." The slope is 0.122 with $p = 0.274$.

# 8 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, agricultural labor productivity increases by 4.3-7.6 percent, non-agricultural labor productivity decreases by 0.7-1.4 percent, and GDP per capita increases by 0.8-1.5 percent. Gains in productivity and overall income are attributed mainly to 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," *Econometrica*, 2022, *90* (3), 1261–1282.

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.

Carter, S.B., S.S. Gartner, M.R. Haines, A.L. Olmstead, R. Sutch, and G. Wright, *The Historical Statistics of the United States Online Edition: Millennial Edition*, Cambridge University Press, 2006.

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.

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.

Dinkelman, Taryn and L Rachel Ngai, "Home Production, Women's Market Work, and Structural Transformation," *STEG Pathfinding Paper*, 2021.

_ and _ , "Time use and gender in Africa in times of structural transformation," *Journal of Economic Perspectives*, 2022, *36* (1), 57–80.

Doepke, Matthias and Michele Tertilt, "Women's Liberation: What's in it for Men?," *Quarterly Journal of Economics*, 2009, *124* (4), 1541–1591.

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.

— 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 Waugh, "Agricultural Productivity Differences across Countries," *American Economic Review*, 2014, *104* (5), 165–170.

— , — , and — , "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.

Ilkkaracan, Ipek and Raziye Selim, "The Gender Wage Gap in the Turkish Labor Market," *Labour*, 2007, *21* (3), 563–593.

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," *American Economic Review*, 2023, *113* (2), 317–353.

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.

Ngai, L Rachel, Claudia Olivetti, and Barbara Petrongolo, "Structural Transformation and U-shaped Female Employment," *mimeo*, 2021.

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 Vandenbroucke, "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

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

Solution to the problem satisfies

$$\frac{\partial \mathcal{L}}{\partial C_a} = \frac{1}{c_a} - \lambda P_a = 0 \tag{14}$$

$$\frac{\partial \mathcal{L}}{\partial C_n} = \frac{\nu}{c_n + \bar{n}} - \lambda + \mu = 0 \tag{15}$$

$$\frac{\partial \mathcal{L}}{\partial l} = \frac{\psi}{l} - \lambda y = 0 \tag{16}$$

$$\lambda(P_a c_a + c_n - (1-l)y) = 0 \tag{17}$$

$$\mu c_n = 0 \tag{18}$$

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 14 and 16 to get

$$P_a c_a = \frac{1}{\psi} y l \tag{19}$$

Plug into 17 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 19 to get optimal $c_a$:

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

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

Combine 14 and 16 to get 19 as before, and combine 15 and 16 to get

$$\begin{aligned}\frac{\nu}{c_n+\bar{n}} &= \frac{\psi}{yl}\\ c_n+\bar{n} &= \frac{\nu}{\psi}yl\\ c_n &= \frac{\nu}{\psi}yl-\bar{n} \qquad (20)\end{aligned}$$

Plug 19 and 20 into 17 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 $P_a c_a = \frac{1}{\psi} yl$ 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 20 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 $c_n >= 0$ 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 $y$.

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 $c_n^i$.

**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

$$\begin{aligned} p_{am} &= \Pr[\epsilon_{nm} \leq P_a \epsilon_{am}] \\ &= \int F_1^m(\epsilon, P_a \epsilon) d\epsilon \end{aligned} \tag{21}$$

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

$$\begin{aligned} p_{am} &= \int F_1^m(\epsilon, P_a \epsilon) \\ &= \frac{(P_a^\theta T_{am})^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}} \int dF^m(\epsilon) \\ &= \frac{(P_a^\theta T_{am})^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}} \end{aligned} \tag{22}$$

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.

$$p_{aw} = \frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)}}{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}} \tag{23}$$

**Proof of Proposition 2.**Quality

For both men and women, the maximum income is:

$$y^* = \max\{w_a \epsilon_a, w_n \epsilon_n\} \equiv w^* \epsilon^* \tag{24}$$

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

$$\begin{aligned} \Pr[y^* < z] &= \Pr[w_a \epsilon_a < w_n \epsilon_n] \\ &= F\left(\frac{z}{w_a}, \frac{z}{w_n}\right) \\ &= \exp\{-\bar{T} z^{-\theta}\} \end{aligned} \tag{25}$$

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:

$$G(x) \equiv \Pr[\epsilon^* < x] = \exp\{-T^* x^{-\theta}\} \tag{26}$$

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:

$$\begin{aligned} E[\epsilon_s|\text{Person chooses } s] &= \int \epsilon dG(\epsilon) \\ &= (T^*)^{1/\theta}\Gamma(1-\frac{1}{\theta}) \end{aligned} \tag{27}$$

where $\Gamma(\cdot)$ is the Gamma function. From Proposition 1, the average quality can be written as a function of the propensity. For example,

$$E[\epsilon_{am}|\text{Person chooses } a] = \left(\frac{T_{am}}{p_{am}^{1-\rho}}\right)^{1/\theta}\Gamma\left(1-\frac{1}{\theta}\right) \tag{28}$$

**Proof of Proposition 3.**Gender Wage Gap

The proof is simple algebra given the results from Propositions 1 and 2. From Proposition 1, the propensities for men and women in each sector are

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

The ratios of propensities within each sector, after slight rearrangement, are

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

The average wage in industry $s$ for gender $g$ is defined as

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

Apply the result of Proposition 2 and we get

$$\overline{\text{wage}}_{sg} = (1-\tau_{sg}) w_s \left(\frac{T_{sg}}{p_{sg}^{1-\rho}}\right)^{1/\theta} \Gamma\left(1-\frac{1}{\theta}\right).$$

Hence the gender wage ratios within each sector are

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

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

Now we plug in the previously derived propensity ratios into the gender wage ratio in the agricultural sector:

$$\begin{aligned}
\frac{\overline{\text{wage}}_{am}}{\overline{\text{wage}}_{aw}} &= \frac{1}{1-\tau_{aw}} \left[\frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}} \left(\frac{T_{am}}{T_{aw}\tilde{\tau}_{n|a}^\theta}\right)^{1/(1-\rho)}\right]^{-(1-\rho)/\theta} \left(\frac{T_{am}}{T_{aw}}\right)^{1/\theta} \\
&= \frac{1}{1-\tau_{aw}} \left(\frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}}\right)^{-(1-\rho)/\theta} \left(\frac{T_{aw}\tilde{\tau}_{n|a}^\theta}{T_{am}}\right)^{1/\theta} \left(\frac{T_{am}}{T_{aw}}\right)^{1/\theta} \\
&= \frac{\tilde{\tau}_{n|a}}{1-\tau_{aw}} \left(\frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}}\right)^{-(1-\rho)/\theta} \\
&= \frac{1}{1-\tau_{nw}} \left(\frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}}\right)^{-(1-\rho)/\theta},
\end{aligned}$$

where in the last line we use our definition of $\tilde{\tau}_{n|a} = (1-\tau_{aw})/(1-\tau_{nw})$. Repeat the same process with the gender wage ratio in the non-agricultural sector:

$$\begin{aligned}
\frac{\overline{\text{wage}}_{nm}}{\overline{\text{wage}}_{nw}} &= \frac{1}{1-\tau_{nw}} \left[\frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}} \left(\frac{T_{nm}}{T_{nw}}\right)^{1/(1-\rho)}\right]^{-(1-\rho)/\theta} \left(\frac{T_{nm}}{T_{nw}}\right)^{1/\theta} \\
&= \frac{1}{1-\tau_{aw}} \left(\frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}}\right)^{-(1-\rho)/\theta} \left(\frac{T_{nw}}{T_{nm}}\right)^{1/\theta} \left(\frac{T_{nm}}{T_{an}}\right)^{1/\theta} \\
&= \frac{1}{1-\tau_{nw}} \left(\frac{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)} + T_{nw}^{1/(1-\rho)}}{(P_a^\theta T_{am})^{1/(1-\rho)} + T_{nm}^{1/(1-\rho)}}\right)^{-(1-\rho)/\theta} \\
&= \frac{\overline{\text{wage}}_{am}}{\overline{\text{wage}}_{aw}},
\end{aligned}$$

so the wage ratios in both sectors are equal to each other. The expression for the wage ratios can be further rewritten to obtain

$$\begin{aligned}
\frac{\overline{\text{wage}}_{am}}{\overline{\text{wage}}_{aw}} &= \frac{1}{1-\tau_{nw}}\left(\frac{(P_a^\theta T_{am})^{1/(1-\rho)}+T_{nm}^{1/(1-\rho)}}{(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)}+T_{nw}^{1/(1-\rho)}}\right)^{(1-\rho)/\theta} \\
&= \left(\frac{(P_a^\theta T_{am})^{1/(1-\rho)}+T_{nm}^{1/(1-\rho)}}{(1-\tau_{nw})^{\theta/(1-\rho)}(P_a^\theta \tilde{\tau}_{n|a}^\theta T_{aw})^{1/(1-\rho)}+(1-\tau_{nw})^{\theta/(1-\rho)}T_{nw}^{1/(1-\rho)}}\right)^{(1-\rho)/\theta} \\
&= \left(\frac{(P_a^\theta T_{am})^{1/(1-\rho)}+T_{nm}^{1/(1-\rho)}}{[P_a^\theta(1-\tau_{aw})^\theta T_{aw}]^{1/(1-\rho)}+[(1-\tau_{nw})T_{nw}]^{1/(1-\rho)}}\right)^{(1-\rho)/\theta}.
\end{aligned}$$

**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

$$\begin{aligned}
\text{Mean} &= \bar{T}^{\frac{1}{\theta}}\Gamma(\frac{\theta-1}{\theta}) \\
\text{Variance} &= \bar{T}^{\frac{2}{\theta}}\{\Gamma(\frac{\theta-2}{\theta})-\Gamma(\frac{\theta-1}{\theta})^2\}
\end{aligned} \tag{29}$$

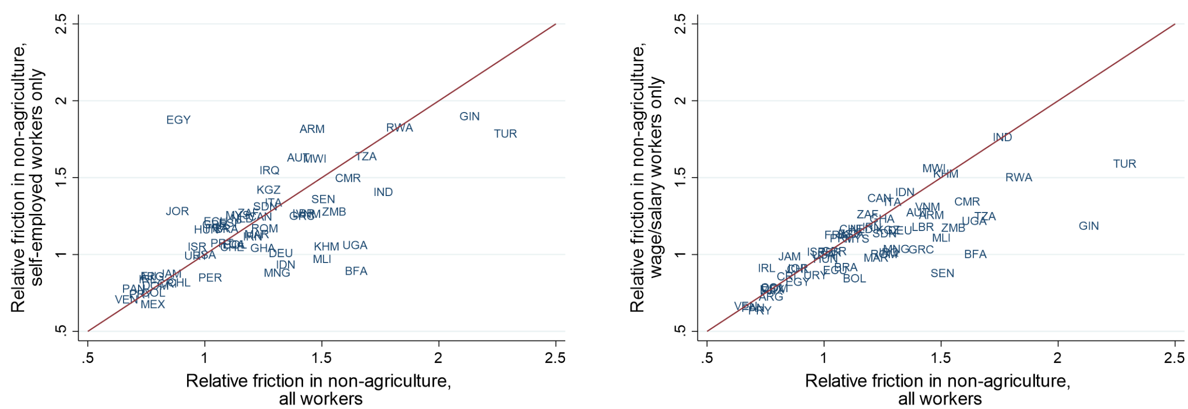
Therefore, the coefficient of variation is:

$$\frac{\text{Std.Dev.}}{\text{Mean}} = \left[\frac{\bar{T}^{\frac{2}{\theta}}\{\Gamma(\frac{\theta-2}{\theta})-\Gamma(\frac{\theta-1}{\theta})^2\}}{\{\bar{T}^{\frac{1}{\theta}}\Gamma(\frac{\theta-1}{\theta})\}^2}\right]^{1/2} = \left[\frac{\Gamma(1-\frac{2}{\theta})}{\Gamma(1-\frac{1}{\theta})^2}-1\right]^{1/2} \tag{30}$$

# C Self-employed vs. Wage/Salary Workers

The IPUMS-International data includes all employment types, i.e., both self-employed and wage/salary workers. The variable identifying employment type (CLASSWK) is available for 61 countries (66 countries minus China, Mozambique, Nigeria, Sierra Leone, and Slovenia). I calibrate relative non-agriculture friction, $\tilde{\tau}_{n|a}$, using equation 13, on different employment types and show how that affects the calibrated friction. The calibrated frictions from using only self-employed or wage/salary workers are highly correlated with the calibrated frictions from the full sample, as shown in Figure C.1. Correlation coefficient between relative frictions of self-employed workers and all workers is 0.6841, while the correlation coefficient between relative frictions of wage/salary workers and all workers is 0.7736. When either self-employed or wage/salary workers are ignored, results may hold qualitatively (given high correlation), but not quantitatively (give several countries whose calibrated frictions are far from the benchmark case).

Figure C.1: Relative friction non-agriculture: self-employed and wage/salary workers vs all workers




(a) Self-employed (b) Wage/salary workers

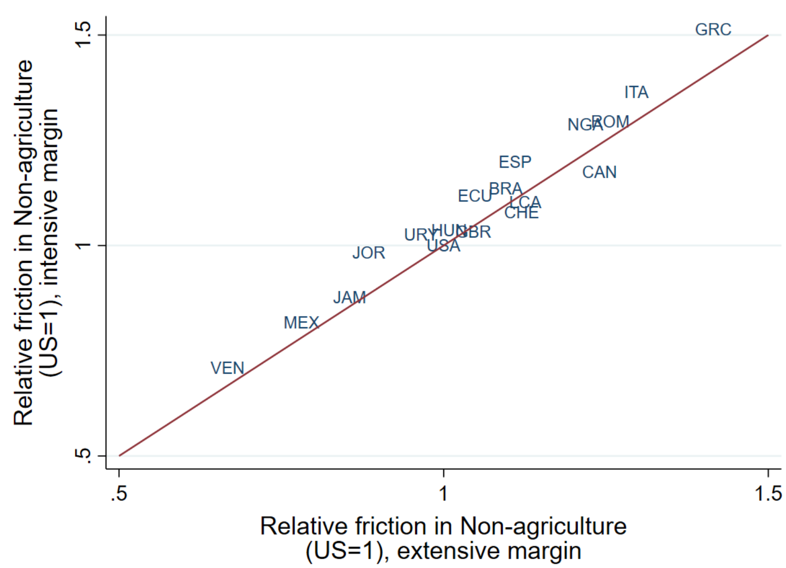
Notes: These figures compare relative frictions calibrated using self-employed workers only or wage/salary workers only with relative frictions calibrated using all workers. The red line is a 45-degree line.

# D 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 this number by the total hours of labor 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 13 in the main text to obtain a new set of calibrated relative frictions.

Figure D.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 D.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.