UNIVERSITY OF ALBERTA
FACULTY OF ARTS
Department of Economics

Working Paper No. 2020-02

Misallocation across Establishment Gender

Ashantha Ranasinghe
University of Alberta

Updated March 2023

Misallocation across Establishment Gender

Ashantha Ranasinghe*
University of Alberta

March 2023

Abstract

I find substantial differences in the extent of misallocation across male and female-led establishments spanning many low and middle-income countries. Across broad geographic regions, female establishments face higher distortions to operating a business and is primarily from higher capital distortions. Equalizing distortions across gender increases female sales shares implying that women should account for a larger share of the market, especially in poorer countries. When accounting for relevant establishment characteristics, higher female distortions and its negative association with development remain.

JEL: J16, O1, O4, O5.

Key Words: misallocation, gender, productivity, micro data.

*Contact: 8-14 Tory Building, University of Alberta, Edmonton, Alberta, Canada, T6G 2H4. E-mail: ranasinghe@ualberta.ca

1 Introduction

A message of the misallocation literature is that policies or implicit frictions that disadvantage some businesses to the benefit of others can account for the vast productivity differences observed across countries (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Bartelsman et al., 2013). While recent work has evaluated specific drivers that affect misallocation, the characteristics of businesses that are most encumbered by the frictions that engender misallocation are less understood. The focus of this paper is to document the extent of misallocation across male and female-led establishments and examine how it relates to economic development.

Based on the World Bank Enterprise Surveys 2008-2017 (WBES) and using the top manager's gender to define male and female-led establishments, a robust pattern across countries is that women account for a small share of all establishments and a smaller share of the market. For instance, in South America women account for 13 percent of establishments and 7 percent of sales; in Eastern Europe/Central Asia the corresponding numbers are 17 and 12 percent. The data also points to female establishments facing higher barriers to operating a business. In South America, 63 (50) percent of female (male) establishments report corruption as a major obstacle to doing business, and 26 (20) percent of female (male) establishments report access to finance as a major obstacle. The evidence is suggestive that misallocation across establishment gender is potentially large, varies across countries and important for understanding cross-country income differences. My results show that female establishments are more productive on average and face higher distortions to production, the latter which particularly holds in poorer countries. Equalizing distortions across gender can imply proportionally large increases in female market shares, average size and TFP.

I explore this formally using a variant of the framework in Hsieh and Klenow (2009, 2014). Specifically, I assume there are implicit taxes on capital and output that vary across heterogeneous producers to reflect a broad range of barriers affecting production, some of which can be gender-specific. Establishment level distortions are measured from revenue productivity, a composite of these implicit taxes, and where higher dispersion implies more misallocation. I assess the quantitative importance of misallocation across gender using publicly available data from the WBES (2008-2017), focusing on formal establishments in the manufacturing sector. The wide coverage of countries in the WBES together with information on top manager gender—often absent in more disaggregated administrative level datasets—allows me to provide a broad overview of misallocation across gender and its link to economic development spanning over 30 low and middle-income countries, regions precisely where gender discrimination is most prevalent.

I find notable differences in the extent of misallocation across establishment gender, which varies by the level of development. Female establishments face higher distortions on average across the full sample of countries, and also in South Asian, South American, African and East Asian countries; whereas in Eastern European and Central American countries male establishments face higher distortions. But a pattern that broadly holds across sub-continents is that women face substantially higher distortions on capital. In addition, across sub-continents the distribution of revenue productivity is more right-skewed for female establishments and exhibits higher dispersion, evidence that is consistent with women facing higher distortions and more misallocation. Notably, female establishments are more productive on average, implying that even among more talented women that have selected into formal entrepreneurship—having overcome any gender specific barriers to entry—women still face higher distortions to running a business than men.

That female establishments face higher distortions on production implies that women should account for a larger share of market resources than what is observed in the data, especially in poorer countries where women face higher distortions. For example, among a fixed set of producers, a policy that eliminates misallocation in South Asian countries—where women are both more productive and face higher distortions—raises the female share of sales by over 10 percent, with larger impacts among the top decile of producers (in the range of 30 percent). This is quantitatively large since men account for over 90 percent of establishments and also (on net) increase sales from this policy. Also, while removing misallocation raises average establishment size/sales across both genders, it has a larger impact on average female size. Allowing for a broader class of reallocation policies where distortions are tied to productivity, but not gender, implies much larger increases in female sales shares. A clear implication is that female establishments are inefficiently small and should expand relative to male establishments.

I also evaluate whether differences in the extent of distortions across male and female establishments are tied to gender by estimating the relationship between distortions and gender while controlling for relevant establishment characteristics. In South Asia, and particularly India, female establishments are associated with facing 20 percent higher distortions.[1] Across countries, the female estimate is negatively related to economic development (GDP per capita), implying that in poorer countries female establishments face higher distortions to operating a business (relative to males). Importantly, this negative link with development is primarily from women facing higher distortions on capital. In addition, the country-level female estimates, which are based on establishment-level data, are highly correlated with aggregate measures of gender norms found in external data sources, suggestive that the female estimates I find are broadly related to gender bias.

Few papers examine differences in the extent of misallocation across sub-groups of an economy.[2] Concerning gender, Chiplunkar and Goldberg (2021) show that barriers to female labour force participation and entrepreneurship in India can account for inequality across gender and substantial productivity losses.[3] In a similar context, Ranasinghe (2023) shows

[1] The estimates are a lower bound since it does not account for labour force participation or selection into entrepreneurship, which is male dominated.

[2] Dias et al. (2016, 2019) document that misallocation is higher in services than in manufacturing in Portugal, while Kalemli-Ozcan and Sørensen (2014) and Gorodnichenko et al. (2019) estimate the relevance of barriers and establishment characteristics for understanding distortions in Africa and Europe.

[3] In the context of the U.S., Hsieh et al. (2019) focus on the allocation of talent across occupations, and Bento (2021) examines barriers to female entrepreneurship over time. See also Morazzoni and Sy (2021) who examine differences in access to credit across entrepreneur gender in the U.S.

that differences in the distribution of distortions across gender have a larger impact on female entrepreneurship than barriers to entry. Lee (2022) focuses on gender related distortions across sectors and finds that women face higher distortions in non-agriculture sectors which can account for the large agricultural productivity gaps across countries. Also related are Cuberes and Teignier (2016, 2018) who examine the implications of labour market gender gaps using aggregate statistics in models featuring occupation choice. My paper emphasizes gender focusing on micro-level distortions across a wide range of countries and finds more broadly that women face higher distortions, notably for capital, and is a pattern that is negatively related to development.

Relatedly, Inklaar et al. (2017) find there are large productivity losses from misallocation across countries but no clear link with the level of economic development. My results show that the extent to which female establishments face higher distortions is negatively related to development and an important part of the overall misallocation picture. In addition, female establishments can account for as much as 15 percent of the TFP gains attributed to a policy that eliminates misallocation. Again, this is proportionally large as women account for a small share of establishments and serves as a lower bound since selection into entrepreneurship is fixed.

Of note, in making cross-country comparisons across gender using the WBES micro-data it is necessary to focus on a single aggregated manufacturing sector. Nevertheless, I show the central findings—that female establishments face higher distortions, particularly in poorer countries—also holds at the 2-digit sub-industry level (albeit with a smaller sample of countries). Also, while I define establishment gender based on the top manager given their importance on establishment performance (Bloom et al., 2013), the central results hold for the business owner as well.

This work also relates to the more micro-empirical papers that examine differences across establishment gender. Fairlie and Robb (2009) find female establishments are smaller and underperform relative male establishments in the U.S., and Sabarwal and Terrell (2008) find similar results for Eastern European countries. There is also work that focuses on micro-scale establishments in specific industries, often informal and with few paid employees, to identify differences across gender in observables such as revenue and investment.[4] My paper connects to this literature by measuring misallocation among formal establishments from a macro approach, providing an additional view of the barriers to production by gender among a more skilled set of the population, and for an expansive set of countries.

The rest of the paper proceeds as follows. Section 2 presents a standard model of misallocation that embeds gender-specific distortions to production. Section 3 describes the data, its advantages and limitations. Section 4 documents differences in distortions and misallocation across male and female establishments, changes in market shares from undoing misallocation, and regression estimates based on gender. Section 5 shows the results are robust to a variety of sensitivity checks, and Section 6 provides concluding remarks.

2 Model

I use a model of monopolistic competition with heterogeneous producers that differ in productivity and the distortions they face, which are modelled as implicit taxes on output and capital as in Hsieh and Klenow (2009). The model is extended to allow for differences in the extent of distortions across male and female establishments, but is otherwise a standard framework for measuring misallocation. In the quantitative analysis I primarily focus on a single industry (manufacturing), but here the framework is presented for multiple industries.

[4] See Jayachandran (2020) for a survey of the micro-entrepreneurship literature that focuses on gender. In addition, Hardy and Kagy (2018) show that micro-scale female establishments earn less profit than males in a garment sector in Ghana, and de Mel et al. (2009) show that returns are lower for female grant recipients in Sri Lanka potentially due to less control over household bargaining.

2.1 Aggregate and industry production

A representative firm produces a single final good Y in a perfectly competitive market combining output Y_s from $s \in S$ independent industries:

$$Y = \prod_{s=1}^{S} Y_s^{\theta_s}, \tag{1}$$

where Y_s is total output in industry s and $\sum_s^S \theta_s = 1$. Industry shares are $\theta_s = P_s Y_s / PY$ based on cost minimization, P_s is the price of industry s output, and P is the price of the final good (which is set equal to one).

Let $j = \{m, f\}$ denote an establishment operated by a male or female business owner or manager. (For exposition, I refer to these as male and female establishments, and use establishment and entrepreneur synonymously when convenient.) Production in industry s is

$$Y_s = \left(\sum_{i=1}^{M_s} y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} = \left(\sum_{i=1}^{M_s^m} (y_{si}^m)^{\frac{\sigma-1}{\sigma}} + \sum_{i=1}^{M_s^f} \left(y_i^f \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

where y_{si}^j is the output of establishment i of gender j in industry s, M_s^j is number of gender j establishments, and $M_s = M_s^m + M_s^f$ is the total number of establishments in industry s.

2.2 Establishment production

Establishments operate in monopolistically competitive industries. Establishment i's output (or variety i) in industry s is based on entrepreneur productivity z_{si}^j, and capital and labour inputs; $y_{si}^j = z_{si}^j \left(k_{si}^{\alpha_s} n_{si}^{1-\alpha_s} \right)$, and where factor shares α_s can be industry specific.[5]

Distortions are modelled as implicit taxes/subsidies on production: an output distortion τ_{si}^j which affects the marginal product of inputs used in production, and a capital distortion

[5]For ease, I write capital and labour inputs as k_{si} and n_{si} (without the j superscript) but it should be noted that they depend on entrepreneur productivity which is gender dependent.

κ_{si}^{j} which affects the marginal product of capital relative to labour.[6] Importantly, these distortions can be gender-specific. The output distortion is $\tau_{si}^{j} = \bar{\tau}_{s}^{j} + \tau_{si}$, where $\bar{\tau}_{s}^{j}$ is a gender-specific distortion (common across gender j establishments in an industry) and τ_{si} is an idiosyncratic part. Likewise, the distortion on capital is $\kappa_{si}^{j} = \bar{\kappa}_{s}^{j} + \kappa_{si}$, where $\bar{\kappa}_{s}^{j}$ is a gender-specific distortion on capital and κ_{si} is an idiosyncratic part.[7] For example, τ_{si}^{j} can be high if gender j establishments face a higher extent of corruption or difficulty obtaining permits (a high $\bar{\tau}_{s}^{j}$), or due to idiosyncratic factors that affect production but not related to gender (a high τ_{si}). Similarly, the distortion on capital κ_{si}^{j} can be high if gender j establishments face higher barriers to access capital (high $\bar{\kappa}_{s}^{j}$) or due to factors not tied to gender (high κ_{si}).

Establishment profit is

$$\pi_{si}^{j} = (1 - \tau_{si}^{j}) p_{si}^{j} y_{si}^{j} - w n_{si} - (1 + \kappa_{si}^{j}) r k_{si} \tag{2}$$

where p_{si}^{j} is the optimal price, w is cost of labour and r is the rental cost of capital, which are assumed common within and across industries. Values of $\tau_{si}^{j} \in (0, 1)$ and $\kappa_{si}^{j} > 0$ are implicit taxes on output and capital; $\tau_{si}^{j} \in (-1, 0)$ and $\kappa_{si}^{j} \in (-1, 0)$ are implicit subsidies; and $\tau_{si}^{j} = \kappa_{si}^{j} = 0$ implies the establishment does not face distortions. Profit maximization implies the optimal price is a fixed mark up $\sigma/(\sigma - 1)$ over the marginal cost;

$$p_{si}^{j} = \frac{\sigma}{\sigma - 1} \Omega_s \frac{\left(1 + \kappa_{si}^{j}\right)^{\alpha_s}}{z_{si}^{j}(1 - \tau_{si}^{j})}, \tag{3}$$

where $\Omega_s \equiv \left(\frac{r}{\alpha_s}\right)^{\alpha_s} \left(\frac{w}{1-\alpha_s}\right)^{1-\alpha_s}$. Establishments charge higher prices if they have low productivity or face high distortions (output or capital). Moreover, absent of productivity differences, female establishments charge a higher price relative to male establishments only if

[6] Given the specification that follows in equation (2), a tax on capital and labour cannot be separately identified. A high value for κ_{si}^{j} represents a high capital distortion or a low labour distortion (a subsidy); and vice versa. See Dias et al. (2016) for a case with three inputs in production where capital, labour and output distortions are uniquely identified.

[7] I use this simple form to focus on the gender-specific components, $\bar{\tau}_{s}^{j}$ and $\bar{\kappa}_{s}^{j}$, and in particular their differences across gender. The exact composites of τ_{si}^{j} and κ_{si}^{j} can be extended to include additional structure.

they face higher distortions, and account for a smaller share of industry output (based on the equilibrium conditions of the model). The optimality conditions for establishment labour and capital are

$$\frac{\sigma-1}{\sigma}(1-\alpha_s)\frac{p^j_{si}y^j_{si}}{n^j_{si}} = \frac{w}{1-\tau^j_{si}} \ \equiv \ mrpn^j_{si}, \tag{4}$$

$$\frac{\sigma-1}{\sigma}\alpha_s\frac{p^j_{si}y^j_{si}}{k^j_{si}} = \frac{(1+\kappa^j_{si})}{1-\tau^j_{si}}r \ \equiv \ mrpk^j_{si}. \tag{5}$$

Within an industry, a high marginal revenue product of labour $mrpn^j_{si}$ implies high output distortions (and high sales to labour ratios), and a high marginal revenue product of capital $mrpk^j_{si}$ is indicative of high output or capital distortions (and high sales to capital ratios). Said differently, when output and capital distortions are uniform across establishments there is no variation in marginal revenue products (capital and labour) across establishments.

Measures for establishment physical and revenue productivity are defined as

$$tfpq^j_{si} \equiv z^j_{si} = \frac{y^j_{si}}{k^{\alpha_s}_{si}n^{1-\alpha_s}_{si}}, \tag{6}$$

$$tfpr^j_{si} \equiv p^j_{si}z^j_{si} = \frac{p^j_{si}y^j_{si}}{k^{\alpha_s}_{si}n^{1-\alpha_s}_{si}} = \frac{\sigma}{\sigma-1}\Omega_s\frac{\left(1+\kappa^j_{si}\right)^{\alpha_s}}{(1-\tau^j_{si})}, \tag{7}$$

where the expression for $tfpr^j_{si}$ is obtained using equations (3) and (6). Within an industry, high revenue productivity $tfpr^j_{si}$ is indicative that an establishment faces high barriers to production, high τ^j_{si} and/or κ^j_{si}, and that they operate on a smaller scale than what is efficient (as reflected by high marginal revenue products in equations (4) and (5)). In contrast, low revenue productivity establishments are ones that receive implicit subsidies on output and capital (at least relative to high $tfpr^j_{si}$ establishments) and operate on a larger scale than what is efficient. In this regard, $tfpr^j_{si}$ serves as a summary statistic for establishment distortions $\Big($i.e., $tfpr^j_{si} \propto \frac{\left(1+\kappa^j_{si}\right)^{\alpha_s}}{(1-\tau^j_{si})}\Big)$, which I use in the quantitative analysis to evaluate differences in distortions across gender.

2.3 Industry level marginal products and productivity

The expressions for establishment marginal products in equations (4) and (5) can be aggregated to the industry level to examine how distortions affect industry capital and labour, and particularly by gender. Aggregate capital in industry s is

$$\begin{aligned} K_s &= \sum_{i=1}^{M_s^m} k_{si}^m + \sum_{i=1}^{M_s^f} k_{si}^f \\ &= \frac{\sigma-1}{\sigma}\alpha_s P_s Y_s \left(\frac{(PY)_s^m}{P_s Y_s} \sum_{i=1}^{M_s^m} \frac{1}{mrpk_{si}^m} \frac{p_{si}^m y_{si}^m}{(PY)_s^m} + \frac{(PY)_s^f}{P_s Y_s} \sum_{i=1}^{M_s^f} \frac{1}{mrpk_{si}^f} \frac{p_{si}^f y_{si}^f}{(PY)_s^f} \right), \\ &= \frac{\sigma-1}{\sigma}\alpha_s P_s Y_s \left(\theta_s^m \frac{1}{MRPK_s^m} + \theta_s^f \frac{1}{MRPK_s^f} \right), \end{aligned} \tag{8}$$

where P_sY_s is industry s sales, $\theta_s^j \equiv (PY)_s^j / P_sY_s$ is the share of industry sales attributed to gender j establishments (i.e., $\theta_s^m + \theta_s^f = 1$), and

$$\frac{1}{MRPK_s^j} \equiv \sum_{i=1}^{M_s^j} \frac{1}{mrpk_{si}^j} \frac{p_{si}^j y_{si}^j}{(PY)_s^j} = \frac{1}{r} \sum_{i=1}^{M_s^j} \frac{1-\tau_{si}^j}{1+\kappa_{si}^j} \frac{p_{si}^j y_{si}^j}{(PY)_s^j}$$

is the average (inverse) marginal product of capital.

Following similar steps, aggregate labour in an industry is

$$N_s = \frac{\sigma-1}{\sigma}\alpha_s P_s Y_s \left(\theta_s^m \frac{1}{MRPN_s^m} + \theta_s^f \frac{1}{MRPN_s^f} \right), \tag{9}$$

where

$$\frac{1}{MRPN_s^j} \equiv \sum_{i=1}^{M_s^j} \frac{1}{mrpn_{si}^j} \frac{p_{si}^j y_{si}^j}{(PY)_s^j} = \frac{1}{w} \sum_{i=1}^{M_s^j} \left(1-\tau_{si}^j\right) \frac{p_{si}^j y_{si}^j}{(PY)_s^j},$$

is the average (inverse) marginal product of labour in industry s. Note that high average marginal revenue products ($MRPN_s^j$ and $MRPK_s^j$) lower industry labour N_s and capital K_s. In addition, the impact on N_s and K_s are especially large when the gender that accounts for a larger share of sales (θ_s^j) has a higher average marginal product (i.e., faces higher distortions on average).[8]

A natural candidate for the average distortion on output by gender in an industry is $\bar{\tau}_s^{j,avg} \equiv 1 - \sum_{i=1}^{M_s^j} \left(1 - \tau_{si}^j\right) \frac{p_{si}^j y_{si}^j}{(PY)_s^j}$; i.e., the average distortion on output weighted by sales, and is the one I use. It follows that $1 - \bar{\tau}_s^{j,avg} = \frac{\sigma}{\sigma - 1} \frac{1}{(1-\alpha_s)} \frac{wN_s^j}{(PY)_s^j}$, which says the average output distortion can be determined by gender-specific aggregates on labour (or labour costs) and sales. Similar to output, I define the average capital distortion by gender as

$$1 + \bar{\kappa}_s^{j,avg} \equiv \sum_{i=1}^{M_s^j} \left(1 + \kappa_{si}^j\right) \frac{k_{si}^j}{K_s^j} = \frac{\alpha_s}{1 - \alpha_s} \frac{wN_s^j}{rK_s^j},$$

which is weighted by the share of capital for gender j establishments, and can be backed-out by gender-specific aggregates on labour and capital costs.

Given these expressions, $|\bar{\tau}_s^{m,avg} - \bar{\tau}_s^{f,avg}|$ and $|\bar{\kappa}_s^{m,avg} - \bar{\kappa}_s^{f,avg}|$ measure the differences in average output and capital distortions by gender. Additionally, if τ_{si} and κ_{si} are truly idiosyncratic (i.e., they have the same distribution across gender) then differences in average distortions across gender reflect gender specific bias in production; that is, $|\bar{\tau}_s^{m,avg} - \bar{\tau}_s^{f,avg}| = |\bar{\tau}_s^m - \bar{\tau}_s^f|$ and $|\bar{\kappa}_s^{m,avg} - \bar{\kappa}_s^{f,avg}| = |\bar{\kappa}_s^m - \bar{\kappa}_s^f|$. If instead, and more plausibly, τ_{si}^j and κ_{si}^j include non-idiosyncratic factors that vary by gender, then $|\bar{\tau}_s^{m,avg} - \bar{\tau}_s^{f,avg}|$ and $|\bar{\kappa}_s^{m,avg} - \bar{\kappa}_s^{f,avg}|$ are biased measures of gender-specific biases in production.

Industry level productivity. I can now obtain expressions for industry level revenue and physical productivity, $TFPR_s$ and TFP_s, in terms of gender-specific distortions on output and capital. First note that industry revenue productivity is $TFPR_s = \frac{P_s Y_s}{K_s^{\alpha_s} N_s^{1-\alpha_s}} = \frac{\sigma}{\sigma-1} \Omega_s \frac{(1+\bar{\kappa}_s)^{\alpha_s}}{1-\bar{\tau}_s}$, where $\bar{\tau}_s$ and $\bar{\kappa}_s$ are industry averages across all (male and female) establishments.[9] High $TFPR_s$ is a sign of high distortions on production. Further separating by

[8]For example, if $\theta_s^m > \theta_s^f$, the impact on N_s is much bigger when $MRPN_s^m > MRPN_s^f$.

[9]Specifically, $1 - \bar{\tau}_s = \frac{\sigma}{\sigma-1} \frac{1}{(1-\alpha_s)} \frac{wN_s}{PY_s}$ and $1 + \bar{\kappa}_s = \frac{\alpha}{1-\alpha} \frac{wN_s}{rK_s}$.

gender,

$$TFPR_s = \xi_s^m \cdot TFPR_s^m + \xi_s^f \cdot TFPR_s^f \tag{10}$$

where $\xi_s^j \equiv \left(\theta_{s,k}^j\right)^{\alpha_s}\left(\theta_{s,n}^j\right)^{1-\alpha_s}$, and $\theta_{s,k}^j = K_s^j/K_s$ and $\theta_{s,n}^j = N_s^j/N_s$ are the share of capital and labour accounted for by gender j, and $TFPR_s^j = \frac{(P_sY_s)^j}{\left(K_s^j\right)^{\alpha_s}\left(N_s^j\right)^{1-\alpha_s}} = \frac{\sigma}{\sigma-1}\Omega_s\frac{(1+\bar{\kappa}_s^{j,avg})^{\alpha_s}}{1-\bar{\tau}_s^{j,avg}}$.[10] Equation (10), and similar to (8) and (9), shows that $TFPR_s$ is particularly high when the gender that accounts for a larger share of inputs in production faces higher distortions on average. Since $TFPR_s^j$ is a composite of average gender j distortions on output and capital, I use it as a summary statistic to evaluate differences in average distortions across gender.

Industry physical productivity is[11]

$$TFP_s = \frac{Y_s}{K_s^{\alpha_s}N^{1-\alpha_s}} = \frac{TFPR_s}{P_s}.$$

Noting that $P_s = \left(\sum_i (p_{si}^j)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ is a price-index over all varieties/establishments i in sector s and that $tfpr_{si}^j = p_{si}^j z_{si}^j$, industry productivity can be written as

$$TFP_s = \left(\sum_{i=1}^{M_s^m+M_s^f}\left(z_{si}^j\frac{TFPR_s}{tfpr_{si}^j}\right)^{\sigma-1}\right)^{\frac{1}{\sigma-1}}. \tag{11}$$

Worth noting, in a world where distortions are uniform across establishments, i.e., $\tau_{si}^j = \tau_s$ and $\kappa_{si}^j = \kappa_s \; \forall \; i, \; j$, industry TFP simplifies to an aggregation of establishment productivity

$$TFP_s^{fb} = \left(\sum_{i=1}^{(M_s^m+M_s^f)}(z_{si}^j)^{\sigma-1}\right)^{\frac{1}{\sigma-1}},$$

which is the first best value. That is to say, industry productivity is at its highest when

[10] An equivalent expression for industry revenue productivity as a composite of average gender specific distortions is $TFPR_s = \frac{\sigma}{\sigma-1}\Omega_s\left(\theta_{s,k}^m\frac{1+\bar{\kappa}_s^{m,avg}}{1-\bar{\tau}_s^{m,avg}} + \theta_{s,k}^f\frac{1+\bar{\kappa}_s^{f,avg}}{1-\bar{\tau}_s^{f,avg}}\right)^{\alpha_s}\left(\theta_{s,n}^m\frac{1}{1-\bar{\tau}_s^{m,avg}} + \theta_{s,n}^f\frac{1}{1-\bar{\tau}_s^{f,avg}}\right)^{1-\alpha_s}$.

[11] Or alternatively, $TFP_s = \xi_s^m \cdot TFP_s^m + \xi_s^f \cdot TFP_s^f$

distortions are common across establishments. Hence, TFP_s/TFP_s^{fb} provides a quantitative measure of productivity losses arising from non-uniform distortions across establishments.[12] In addition, it provides a benchmark to compare actual to efficient sales, capital and labour shares by gender (for a given distribution of establishments).

3 Data

I use the World Bank Enterprise Surveys (WBES) which is a publicly available, comprehensive dataset that covers formal establishments across (mostly) poor and developing countries. A feature of the WBES is that it is administered in a similar form across countries which allows for comparisons across broad geographic regions. The surveys also include a panel component which are done in four year waves for a limited set of countries, which makes tracking establishments over time and across countries not feasible. For the quantitative analysis, I therefore use a cross-section of the data from the WBES 2008-17, taking the most recent survey for each country within this time frame.

The main upside of the WBES is that it reports whether the top manager of an establishment is male or female, which I use to define establishment gender. The top manager is especially useful to define establishment gender as the evidence shows they play a central role guiding establishment performance and productivity (Bloom and Van Reenen, 2007; Bloom et al., 2013). Worth emphasising, the WBES is the only dataset that reports establishment gender at the micro-level while facilitating comparisons across a broad range of low and middle-income countries.[13] There is also information related to business ownership categories—all

[12] Using the definition for the final good Y in (1), aggregate productivity losses due to distortions across all industries is

$$\frac{TFP}{TFP^{fb}} = \prod_{s=1}^{S} \left(\frac{TFP_s}{TFP_s^{fb}} \right)^{\theta_s}. \tag{12}$$

[13] Administrative level data have tighter data collection protocols but for the most part do not report establishment gender. Such datasets are also at the country level which makes it a challenge for cross-country comparisons.

men, mostly men, even split, mostly women or all women—which I use as a robustness check for establishment gender, though the sample size becomes considerably smaller.

The surveys account for establishments in manufacturing, service and other sectors (primarily transport and construction), though the manufacturing sector accounts for the majority of observations (upwards of 50 percent). It is well known that female entrepreneurs disproportionally select into the service sector and is likely an important part of the overall gender based misallocation picture. Nevertheless, I focus exclusively on the manufacturing sector because capital is not reported for service sector establishments in the WBES. Within manufacturing, there is a finer disaggregation up to a 2 digit ISIC code. About 80 percent of businesses are stand-alone (i.e., do not belong to another firm). The surveys also include general information about the establishment such as legal status, ownership structure, city of operation, initial year of operation, manager experience and information related to the production side.

To determine the level of distortions on output and capital, and hence $tfpr$, requires information on sales, labour and capital (see also Inklaar et al. (2017) who use the surveys to examine misallocation across countries). The WBES reports establishment annual sales ($d2$), and value added sales can be obtained by subtracting intermediate input costs. I follow the literature in using the total cost of labour—full-time worker salaries, wages and bonuses ($n2a$)—to account for differences in worker quality across establishments (instead of multiplying the number of full-time workers by a common wage). This has the benefit of accounting for differences in the types of workers managers hire, which can be especially relevant when examining differences between male and female establishments. For capital, I use the replacement value of machinery and equipment plus the value of land and buildings ($n7a + n7b$). I focus on establishments that have between 1 to 1000 full-time workers in the manufacturing sector (more on this below) and report positive values for sales, capital, labour and intermediate inputs. After this data cleaning, countries that have fewer than 100 observations are excluded resulting in a final sample of 32 countries.

Issues. Although the WBES's coverage of countries is expansive, an issue is the limited number of observations in several countries which can generate biased estimates and statistics. While sample weights are meant to correct for this and produce a representative sample at the country level, missing observations are a concern and potentially exacerbated when disaggregating by gender and industry. To alleviate this concern, to the extent possible, I focus on an aggregated manufacturing sector (ISIC 15–37) so there are sufficient observations and where missing information is less of an issue. I also focus on establishments whose 'figures are computed with some precision or taken directly from records'.

Of course, measuring misallocation using a coarse definition of industry (grouping at the manufacturing sector) is not ideal. For instance, the extent of misallocation across gender might be understated if there is substantial dispersion in marginal products within a finer disaggregation of industries but little dispersion when analyzing at the manufacturing level; and overstated if the reverse holds. Focusing on sub-industries (2 digit ISIC level which is the finest level of disaggregation in the WBES), $tfpr$ dispersion within a sub-industry is very close to the respective manufacturing statistic at the country level, suggestive that aggregating up to a manufacturing sector is not inflating the extent of misallocation.[14] Measures of misallocation across gender might also be affected if there is a higher concentration of a specific gender within more distorted sub-industries. This is less of a concern, however, as the manufacturing sector is primarily dominated by male establishments, in the range of 85-95 percent of all establishments, a pattern that also holds across sub-industries within manufacturing. For instance, the proportion of male establishments within sub-industries varies at the extreme by no more than 17 percent from the country statistic, and on average by no more than 10 percent. Importantly, there is no link between the percentage of female

[14]As an example, at the 2 digit level the standard deviation of $tfpr$ in each sub-industry in India is within 12 percent of its country-level statistic. This holds across all sub-industries except for one sub-industry in China. Nevertheless, this does not rule out that misallocation is over or understated based on a finer level of disaggregation, say at the 5 digit ISIC level.

establishments and average $TFPR$ differences across gender at the sub-industry level, implying there are no stark selection effects across gender within manufacturing sub-industries, at least in relation to distortions.[15]

Nevertheless, I include sub-industry controls when feasible. In addition, in Section 5 I check the sensitivity of aggregating to a manufacturing sector by restricting the analysis to the sub-industry that has the most number of observations (though the sample of countries is smaller).

A final point worth noting relates to the recent critiques that the effects of misallocation may be overstated, especially when analyzing cross-sectional data. This could be because what is observed as misallocation in a static setting may be an optimal adjustment in a dynamic setting (Asker et al., 2014), or potentially due to mis-measurement (Bils et al., 2018). This concern is less relevant when examining the extent of misallocation across subgroups such as gender. That is, as long shocks to production—that can imply static inefficiency but are dynamically efficient—are not gender-specific, then the extent of misallocation across gender should not be overstated. If shocks are gender-specific, this will point to some bias in production which will get picked up in the measure of misallocation (in what follows). The same reasoning applies to mis-measurement. As long as mis-measurement is not related to gender, differences in the extent of misallocation across gender are unlikely to be overstated (for a large enough sample), whereas if it is gender-specific this will imply some form of misallocation across gender.

4 Quantitative Analysis

This section presents the quantitative results. I outline how distortions are measured in the data, and present descriptive statistics and measures of misallocation across male and female

[15] A regression of $TFPR$ differences across gender on the percentage of female establishments at the sub-industry level has a coefficient and standard error estimate of -0.005 (0.0101).

establishments. Thereafter, I examine the implications for sales shares across gender and productivity if distortions are equalized across establishments. Finally, I estimate whether gender, among other establishment characteristics, can account for the observed differences in distortions across establishments. Throughout, I focus on a fixed set of producers in the manufacturing sector (cross-section in the data) and hence all results abstract from issues affecting selection into entrepreneurship. As already noted, male and female establishments are based on the top manager's gender (the following section considers a variant to this definition).

While the analysis is at the country level, I present results for all countries taken together and also at the sub-continent level to provide a snap-shot of gender differences across broad geographic regions. I follow sub-continent country groupings in the WBES and present results for South America, South Asia and Eastern Europe/Central Asia, where the implications across gender differences generalize to the other sub-continents. Results for other sub-continents—Africa, Central America, Middle East/North Africa, and East Asia—are reported in the Appendix. A complete list of countries and observations by sub-continents are reported in Table A.6.

4.1 Measures of distortions and productivity

The following identities are used to infer establishment output and capital distortions, as well as productivity;

$$1-\tau_{si}^{j}=\frac{\sigma}{\sigma-1}\frac{1}{(1-\alpha_s)}\frac{wn_{si}^{j}}{p_{si}^{j}y_{si}^{j}},\quad 1+\kappa_{si}^{j}=\frac{\alpha_s}{1-\alpha_s}\frac{wn_{si}^{j}}{rk_{si}^{j}},\quad z_{si}^{j}=\zeta_s\frac{\left(p_{si}^{j}y_{si}^{j}\right)^{\frac{\sigma}{\sigma-1}}}{\left(k_{si}^{j}\right)^{\alpha_s}\left(n_{si}^{j}\right)^{1-\alpha_s}},\qquad(13)$$

where $\zeta_s\equiv(P_sY_s)^{\frac{-1}{\sigma-1}}/P_s$ and s now represents an aggregated manufacturing sector. The distortion on output τ_{si}^{j} is obtained from equation (4) and the distortion on capital κ_{si}^{j} is obtained by dividing equations (5) and (4). Establishment productivity z_{si}^{j} is determined using the elasticity of demand in a CES framework, which implies $y^j_{si} = \zeta_s \left(p^j_{si} y^j_{si}\right)^{\frac{\sigma}{\sigma-1}}$. To use the language in Hsieh and Klenow (2009), price vs. quantity is inferred from sales and an assumed elasticity of demand. Also, while ζ_s is unobserved it can be set equal to one since it is common across establishments (and only affects physical productivity). Given the expressions in equation (13), data on establishment sales, capital and labour costs are sufficient to back out physical productivity, capital and output distortions, and hence revenue productivity, as well as the corresponding aggregates.

The parameters to calibrate are the cost of capital r, industry factor share α_s, and the elasticity of substitution σ. I set $r = 0.1$ assuming a 4 percent real interest rate and a 6 percent depreciation rate, although the exact values are not critical. For the manufacturing sector factor share, I set $\alpha_s = 0.406$.[16] The elasticity of substitution is set to $\sigma = 3$ as is fairly standard, and $\theta_s = 1$ since I focus on an aggregated manufacturing sector. Worth noting here, the exact values for α_s and σ are not critical for examining differences across gender since these parameters have a common scale effect on how distortions are measured. Moreover, allowing for differences in α_s across sub-industries (which is mostly infeasible due to sample size in countries) would have little effect since there are limited selection differences across sub-industries by gender, nor are they related to distortions as already noted.

Finally, I trim the 1 percent tails for $\ln(tfpq^j_{si}/TFP^{fb}_s)$ and $\ln(tfpr^j_{si}/TFPR_s)$ by country and gender to account for outliers, as is standard in the literature. Note that establishment physical and revenue productivity are scaled by a common factor, specifically TFP^{fb}_s and $TFPR_s$ (i.e., not gender-specific), to facilitate meaningful comparison.

[16] This is based on the average between 1998-2010 from the North American Industry Classification System (NAICS), as reported by the Bureau of Economic Analysis, and by mapping these NAICS factors shares to the corresponding ISIC code (15-37).

Table 1: Descriptive Statistics – All Countries

	All	Male	Female
# of establishments	14478	12620	1858
% of estabs.		0.87	0.13
% of sales		0.89	0.11
Employees: $\ln(n)$			
mean	3.27	3.31	3.08
Distortions:			
output $\bar{\tau}^j$	0.46	0.46	0.41
capital $\bar{\kappa}^j$	0.08	0.06	0.59
$TFPR^j$	2.15	2.13	2.69

Notes: Statistics are from the most recent survey for a country in the WBES 2008-17, and are weighted with the exception of '# of establishments' which is based on number of observations. Refer to the text for additional details.

4.2 Sample statistics, distortions and misallocation

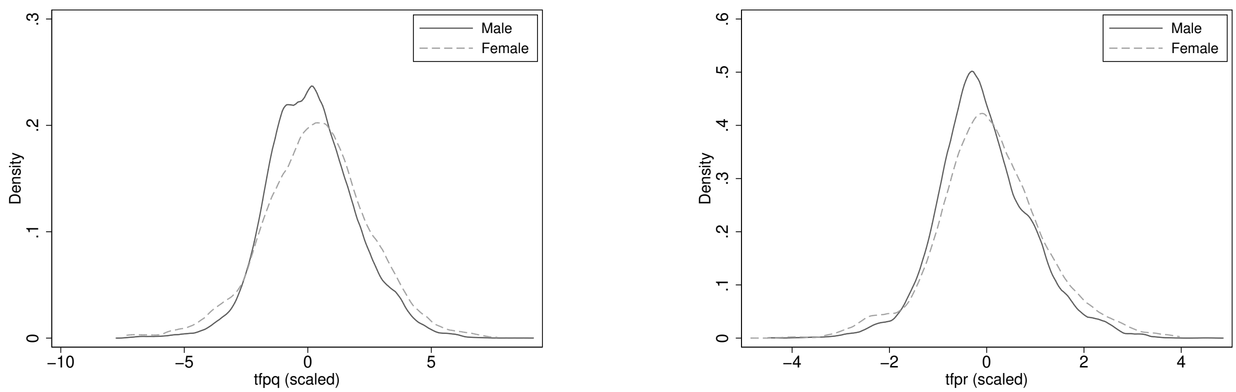
4.2.1 All Countries

Table 1 reports sample statistics and measures of distortions for all countries taken together, and broken down by gender. All statistics and results are calculated using country-level sample weights.[17] Across the full sample of countries a clear pattern is that women account for a small share of the market: about 13 percent of establishments and 11 percent of sales.[18] Women also operate smaller business on average as measured by full-time employees. Importantly, female establishments face higher distortions to operating a business as summarized by $TFPR^j$, which is 25 percent higher relative to male establishments. Of note, much of the differences in distortions across gender are from women facing higher capital distortions $\bar{\kappa}^j$, whereas output distortions $\bar{\tau}^j$ are fairly similar across gender.

[17] Reported statistics are based on calculating country level statistics based on sample weights in the WBES, and then weighting the country statistic by number of observations in a country relative to the full sample. I follow a similar approach when reporting sub-continent statistics. The broad patterns across gender hold with or without sample weights.

[18] The proportion of female establishments across all countries is low (and well below 50 percent) implying significant entry barriers, but this is not unique to the set of countries in the WBES. For instance, the share of female establishments in the U.S. is also low, below 20 percent in 2018 based on the Census Bureau (see also Fairlie and Robb (2009)).

Figure 1: Physical and revenue productivity plots – All Countries



(a) Productivity

(b) Distortions

Notes: Kernel density plots for physical and revenue productivity are $\ln\left(tfpq_{si}^{j} M_s^{\frac{1}{\sigma-1}} / TFP_s^{fb}\right)$ and $\ln\left(tfpr_{si}^{j}/TFPR_s\right)$, respectively. Establishment observations are weighted at the country level by gender, and then again by the number of observations in a country relative to the full sample. The distributions are statistically different across gender.

To further get at the differences across establishment gender, I examine the dispersion in productivity and distortions. Figure 1 panel (a) plots the distribution of physical productivity for males and females; specifically $\ln\left(tfpq_{si}^{j} M_s^{\frac{1}{\sigma-1}} / TFP_s^{fb}\right)$ where M_s is the total number of establishments in the manufacturing sector. The distributions are statistically different across gender and the distribution for women exhibits higher productivity dispersion. Notably, the mean is higher for women implying that female establishments are more productive on average. Table A.2 highlights these points more formally.

Figure 1 panel (b) plots the distribution of revenue productivity, $\ln\left(tfpr_{si}^{j}/TFPR_s\right)$. Recall, higher dispersion in revenue productivity implies a greater extent of misallocation across establishments and high values of $\ln\left(tfpr_{i}^{j}/TFPR_s\right)$ are consistent with an establishment facing high distortions on production. The female revenue productivity distribution exhibits more dispersion implying more misallocation among women. The mean is also higher and the distribution more right-skewed, consistent with women facing higher distortions both on average and across the distribution (see Table A.2 as well).

Table 2: Descriptive Statistics

	South America			South Asia			Eastern Europe		
	All	Male	Female	All	Male	Female	All	Male	Female
# of establishments	2888	2536	352	4163	3879	284	744	644	100
% of estabs.		0.87	0.13		0.93	0.07		0.83	0.17
% of sales		0.93	0.07		0.88	0.12		0.88	0.12
Employees: $\ln(n)$									
mean	3.32	3.40	2.83	3.41	3.40	3.61	3.12	3.15	3.03
Distortions:									
output $\bar{\tau}^j$	0.40	0.39	0.48	0.51	0.51	0.53	0.22	0.25	0.07
capital $\bar{\kappa}^j$	0.36	0.34	0.78	-0.06	-0.11	0.51	0.25	0.14	1.41
$TFPR^j$	1.79	1.76	2.88	2.34	2.29	2.91	1.84	1.88	1.76

Notes: Statistics are from the most recent survey for a country in the WBES 2008-17, and are weighted with the exception of '# of establishments' which is based on number of observations. Refer to the text for additional details.

Taken together, Figure 1 implies the low female market shares among active producers is more a story about women facing higher distortions than one about being less productive. Consistent with the misallocation literature, this is suggestive of substantial output losses from more productive women operating smaller businesses due to higher distortions.

4.2.2 Sub-continents

While the above shows that across countries women account for a small share of the market and face higher distortions on average, I now evaluate whether these patterns hold across broad geographic regions. Table 2 reports statistics for South America, South Asia and Eastern Europe (see Table A.1 for other sub-continents). A clear pattern is that women also account for a small share of the market in these sub-continents: in the range of 7–17 percent of all establishments and 7–12 percent of total sales.

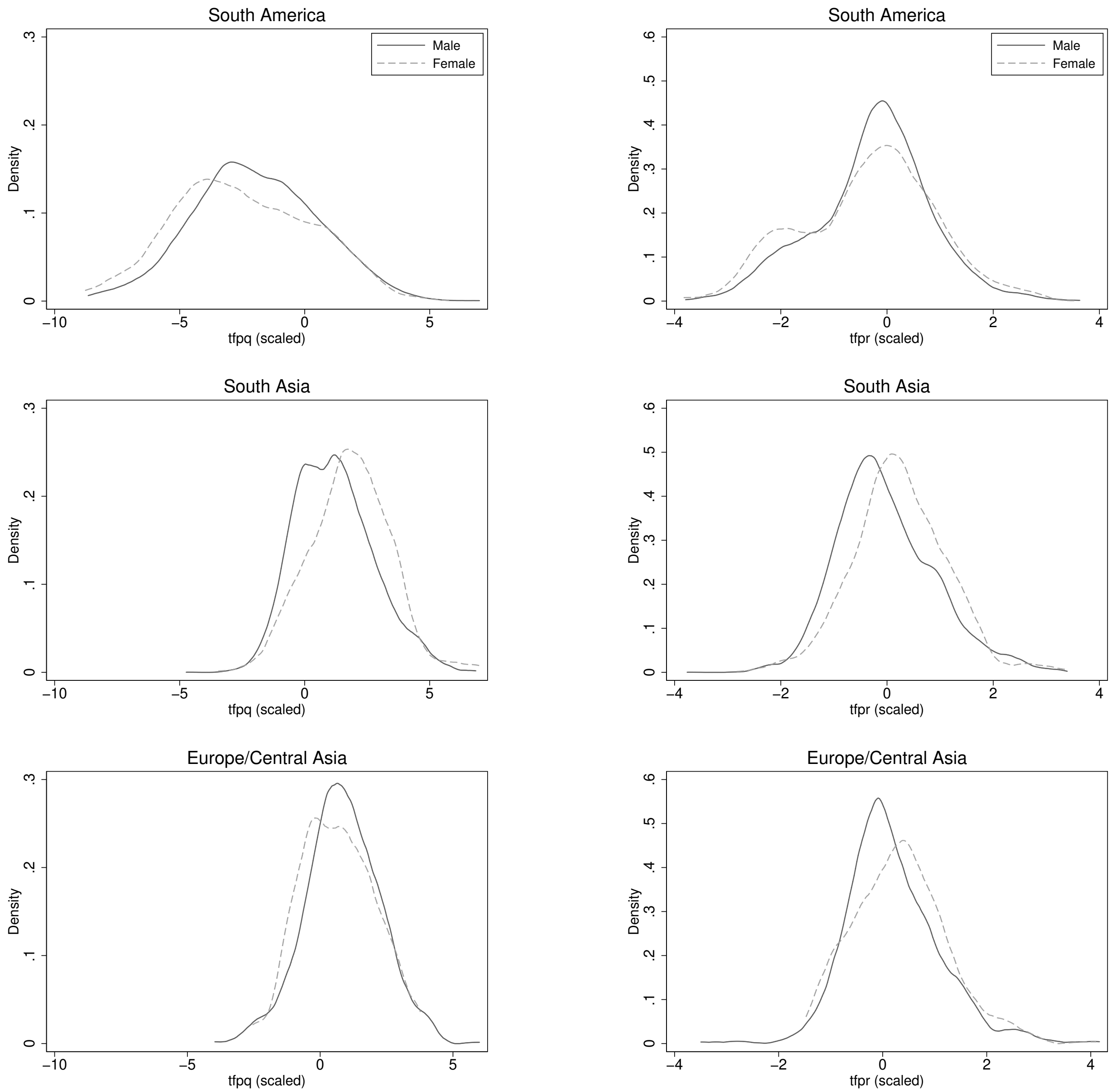
There are some differences across sub-continents that are of interest. First, female establishments face higher distortions ($TFPR^j$) in South America and South Asia, whereas in Eastern Europe male establishments face higher distortions. In South America and South Asia, the higher distortions for women is mainly from higher distortions on capital $\bar{\kappa}^j$; in fact, a broad pattern across sub-continents is that women face higher distortions on capital, including in Eastern Europe where men face higher overall distortions. Second, in contrast to Eastern Europe and South America, women operate larger establishments in South Asia and account for a larger share of sales relative to the proportion of female establishments.[19] That women operate larger establishments, account for a larger share of sales relative to their proportion, and face higher distortions is suggestive of considerable output/productivity losses in South Asia. (The broad patterns for South America are also evident in Africa and East Asia/Pacific, the patterns for Eastern Europe are evident in Central America; and North Africa is a mix of South Asia and Eastern Europe as shown in Table A.1).

Figure 2 plots the distribution of physical productivity and revenue productivity for male and female establishments by sub-continent (Table A.2 in the Appendix provides statistics). The distributions for productivity in panel (a) are statistically different across gender in each sub-continent. In South America and Eastern Europe the distribution of productivity is more left-skewed for women implying that female establishments are less productive on average relative to male establishments. In South Asia, the reverse holds: the distribution of productivity is more right-skewed implying female establishments are more productive on average. South Asia is particularly interesting because the distribution for male establishments has a longer left tail (this is evident to a lesser extent in Eastern Europe as well). On the surface, this is indicative of implicit barriers that favour less productive male establishments in South Asia that would potentially be unprofitable in the absence of such barriers. In contrast, the productivity distributions for South America and Eastern Europe are suggestive of policies that favour less productive female establishments.

Looking at distortions, panel (b) shows the female revenue productivity distribution is more right-skewed in South Asia, and statistically different than the male distribution, implying that women face higher distortions. There is also a long left-tail for male establishments, suggestive that a sub-group of male establishments benefit from implicit subsidies on production.

[19] That women operate larger establishments is consistent with what Chiplunkar and Goldberg (2021) find for India using census-level data. Also, that women operate larger firms in South Asia holds with and without India.

Figure 2: Physical and revenue productivity density plots



(a) Productivity

(b) Distortions

Notes: Kernel density plots for physical and revenue productivity are $\ln\left(tfpq_{si}^{j} M_s^{\frac{1}{\sigma-1}} / TFP_s^{fb}\right)$ and $\ln\left(tfpr_{si}^{j}/TFPR_s\right)$, respectively. A Kolmogorov-Smirnov test for equality of distributions across males and females is rejected ($p < 0.01$) for all $tfpq$ plots on the left panel, and for the $tfpr$ plot for South Asia on the right panel.

The plots for South America show more revenue productivity dispersion for women, particularly at the tails, implying more misallocation among female establishments.[20] In Eastern Europe there is more dispersion in revenue productivity among male establishments, especially at the tails. However, the distributions across gender in South America and Eastern Europe are not statistically different.

4.3 Reallocation gains

Given the extent of misallocation across gender I now quantify the potential gains from eliminating idiosyncratic and gender-specific distortions on production. I follow Hsieh and Klenow (2009) and implement a hypothetical policy that equalizes $tfpr$ across establishments within a country so that all establishments face a common, uniform distortion on capital and output. The policy is such that aggregate resources are unchanged post reallocation to abstract from equilibrium effects on prices, and I focus on the existing set of establishments and abstract from issues related to entry and exit. To the extent that higher distortions on production dissuade or preclude talented women from operating a business, the results I report should be interpreted as a lower bound for the impact of removing misallocation across gender. This is because as distortions are equalized across gender more women will select into entrepreneurship and account for a larger share of the market, a channel my results do not account for (see for instance Cuberes and Teignier (2017); Chiplunkar and Goldberg (2021); Ranasinghe (2023)). Since outliers can have a bigger impact in this hypothetical reallocation, the tails of $\ln(tfpr^j_{si}/TFPR_s)$ are trimmed at the five percent level by country and gender (instead of at the 1 percent level). While the reallocation is implemented at the country level, Table 3 presents results for all countries combined, South America, South Asia and Eastern Europe (see Table A.4 for other continents).

[20]The plots for the other sub-continents are similar to South America which exhibits more dispersion among female establishments; an exception is Central America where the distribution is more right-skewed for male establishments.

I focus on the impact of this 'first-best' policy, which in essence levels the playing field across all establishments, on female sales shares and changes in average sales.[21] For all countries and South Asia, equalizing $tfpr$ across establishments raises female sales shares by 6–11 percent (from a share of 12.8 to 14.2 in South Asia), and has larger impacts among the top ten percent of producers (in the range of 20–30 percent). That is, the reallocation policy implies that women should operate larger establishments and account for a larger share of the market. Of note, these are quantitatively large increases in female market shares since men account for roughly 90 percent of all establishments, and also (on net) expand from this policy.[22] In South America the impact on female market shares are minor as women are less productive on average (see Figure 2). In Eastern Europe, the policy lowers the share of sales among female establishments consistent with Table 2 which shows that women face lower distortions (although female sales shares among the top 10 percent of producers rise).

Table 3: First-best policy with uniform distortions

	All Countries		South America		South Asia		Eastern Europe	
	Data	Reallocation	Data	Reallocation	Data	Reallocation	Data	Reallocation
Female sales shares among:								
all estabs.	0.131	0.139	0.076	0.076	0.128	0.142	0.123	0.079
top ten %	0.114	0.137	0.053	0.050	0.152	0.196	0.051	0.070
Average change in sales:								
Female estabs.	–	0.054	–	0.039	–	0.022	–	0.093
Male estabs.	–	0.007	–	0.004	–	0.002	–	0.013
TFP gains	–	3.21	–	3.25	–	3.03	–	2.65
share due to females ≈	–	0.14	–	0.08	–	0.14	–	0.08

Notes: 'Data' refers to weighted statistics based on the WBES 2006-17. 'Reallocation' are statistics based on a policy that equalizes $tfpr$ across establishments in a country. Top 10% refers to female sales shares among the top ten percent of producers based on sales. Average change in sales, by gender, is total sales post reallocation relative pre reallocation, divided by number of establishments. TFP gains is first-best TFP relative to TFP based on data.

While there is substantial adjustment across the distribution with male and female establishments expanding and shrinking, a clear pattern from this reallocation policy is the average change in sales is considerably higher for women. For instance, in South Asia average sales rise by 2.2 percent for women and by 0.2 percent for men (a factor difference of 11), and in

[21]While not reported, female labour and capital shares are identical to sales shares post reallocation.

[22]Since women comprise a minority of establishments, their sales share will rise only if they are both sufficiently more productive and face higher distortions. Or alternatively, absent of large gaps in productivity and distortions by gender, a policy that removes misallocation should raise the share of sales among the majority group (males).

Eastern Europe by 9.3 and 1.3 percent women and men.[23] Taken together the above implies that among the current set of producers, female establishments are inefficiently small and should expand relative male establishments, especially in South Asia.

Focusing on productivity, the reallocation policy results in about a three-fold increase in TFP for all countries and sub-continents, with the smallest gains in Eastern Europe (reported TFP gains are the inverse of equation (12) and is also equal to output gains.) Also reported is the approximate share of TFP gains attributed to females. In South America and Eastern Europe about 8 percent of TFP gains are due to women, whereas for all countries and South Asia 14 percent are due to women. Female establishments in South Asia account for a higher share of TFP gains (relative to their share of establishments) because they are more productive than male establishments on average, whereas the opposite holds in South America and Eastern Europe.

Table 4: Reallocation with common distribution of distortions

	All Countries		South America		South Asia		Eastern Europe	
	Data	Reallocation	Data	Reallocation	Data	Reallocation	Data	Reallocation
Female sales shares among:								
all estabs.	0.131	0.145	0.076	0.098	0.128	0.138	0.123	0.118
top ten %	0.114	0.152	0.053	0.072	0.152	0.227	0.051	0.084
Average change in sales:								
Female estabs.	–	0.019	–	0.017	–	0.012	–	0.044
Male estabs.	–	0.002	–	0.001	–	0.001	–	0.004

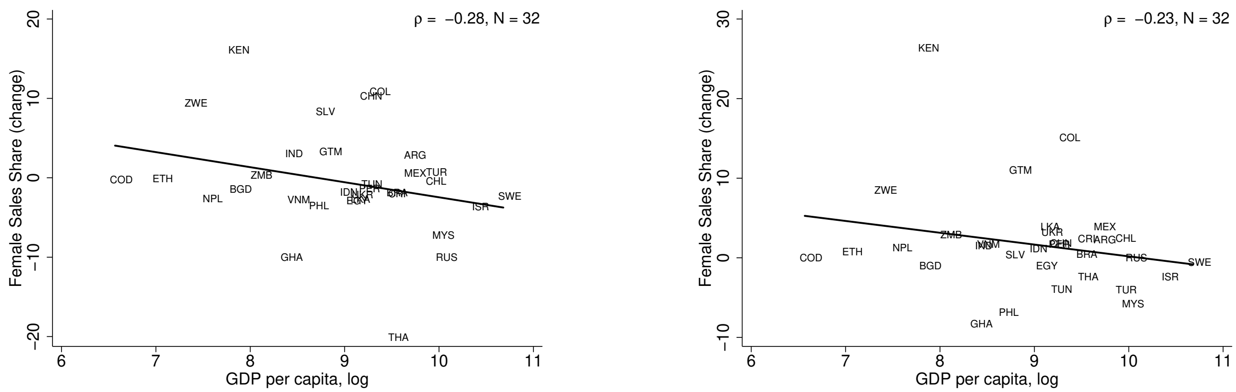
Notes: Statistics are based on a policy when the distribution of distortions are common across gender in a country. All statistics are as described in Table 3.

I also examine the impact of a broader reallocation policy where women face the same distribution of output and capital distortions as men based on productivity. Such a policy highlights the impact on female sales shares when distortions vary with productivity ($tfpq$), and are linked to gender only through differences in establishment productivity.[24] As this policy affects aggregate resources and hence equilibrium prices, I omit statements on produc-

[23] An interpretation of this for Eastern Europe is that high productivity female establishments face high distortions relative to low productivity female establishments such that on net, a policy that removes misallocation raises average female sales. This is consistent with the stronger female presence among the top ten percent of establishments post-reallocation.

[24] Specifically, I approximate a distribution for output and capital distortions, respectively, along fixed points of the $tfpq$ distribution, and apply this common distribution of distortions to evaluate female market shares.

Figure 3: Change in Female Sales Shares under Reallocation



(a) Uniform distortions

(b) Common dist. of distortions

Notes: The figure shows the change in female sales shares when $tfpr$ is uniform across establishments within a country (panel a) and when the distribution of capital and output distortions are tied to productivity (panel b). Change in female sales shares is the difference in female sales share post and pre reallocation. Values above (below) zero indicate that female sales shares rise (fall) from the reallocation. Of note, since women account for a minority of establishments their sales shares rise only if they are *both* considerably more productive and face higher distortions. OLS estimates for panel (a) and (b) are -0.040 (0.0175) and -0.034 (0.0180) .

tivity. Table 4 shows the results. A general pattern is that female market shares rise, and by more than the policy when distortions are uniform across establishments, and in particular among the top 10 percent of producers. As there are more male establishments at the extreme tail of the productivity distribution and since output and capital distortions rise with productivity, women now account for a larger share of the market. The increase in average sales is larger among women but is smaller in comparison to the policy when distortions are uniform, again because distortions and productivity are positively related.

While I have focused on broad geographic regions it is also useful to examine the impact of the reallocation policy across levels of development. Figure 3 shows the change in female sales shares from the two reallocation policies plotted against GDP per capita.[25] Across both policies there is a negative relationship where there is a larger increase in female sales shares

[25]Of interest, the factor change in TFP across countries has flat relationship with GDP per capita, consistent with Inklaar et al. (2017), implying that misallocation and TFP gains are not necessarily related to development.

in poorer countries.[26] Notably, in Colombia, Kenya and Zimbabwe the change in female sales shares is in the range of 10–30 percent.

Overall, a consistent interpretation of the two reallocation policies is that female establishments should account for a larger share of the market than what is observed in the data, and especially in poorer countries where female establishments face higher distortions. For all countries taken together and across all sub-continents the reallocation policies generate a larger increase in average sales for female establishments. Again, these effects are a lower bound as they exclude any effects from selection into entrepreneurship.

4.4 Gender estimates

The evidence so far shows that female establishments face higher distortions and should account for a larger of the market across countries and sub-continents (though Eastern Europe is an exception). I now explore whether these distortions are related to gender, or if other establishment characteristics better account for them. For instance, women may select into particular sub-industries within manufacturing, locate in specific geographic regions, or operate businesses of different size that are more prone to high distortions. Additionally, differences in manager experience will affect how an establishment navigates the business-politico climate in a country, which can influence the extent of distortions.

To this end, I regress establishment-level $tfpr$ (in logs) on whether the top manager is female and include sub-industry (2 digit ISIC) and country-level fixed effects (and use the sample of establishments in Section 4.1) . The sub-industry controls are particularly useful to account for any selection effects across gender within sub-industries and has the benefit of correcting for any measurement error from assuming common input factor shares in production (α_s) across sub-industries. I also include establishment-level controls such as size (small, medium

[26]Recall, women account for a minority of establishments so a decrease in their sales shares at the country-level should be understood as highly distorted male establishments expanding. Both panels show that the expansion of female (minority) relative to male (majority of) establishments is larger in poorer countries.

or large), characteristics of the city a business operates in (population size and whether a capital), the top manager's experience working in industry, and an indicator whether an establishment was formally registered when it began operations, which is meant to pick up the more serious and possibly more educated entrepreneurs/managers (La Porta and Shleifer, 2014; Levine and Rubinstein, 2016).[27]

Conceptually, after accounting for these controls the coefficient on the female indicator can point to whether distortions to operating a business statistically differ across gender, where a positive (negative) coefficient implies that women face higher (lower) distortions relative to men. I loosely refer to this as gender bias, i.e., establishments of a given gender are associated with facing higher distortions. Worth stressing however, the female estimate is an understatement of the extent of female gender bias for several reasons. First, and most importantly, the estimates are among women who on average are more productive, where labour force participation and selection into formal entrepreneurship has already taken place. That is, the estimates do not account for entry barriers that dissuade female entrepreneurship as revealed by the small share of female establishments in manufacturing. Second, the controls for establishment size, registration status, city of operation and experience can itself be influenced by gender bias. For instance, women may locate in specific cities or operate smaller firms to circumvent gender bias. As such, the female estimates that follow should viewed as a lower bound, or put differently, only a part of the broader forms of gender bias women encounter in the workforce.

Table 5 reports regression estimates for all countries taken together, South America, South Asia and Eastern Europe; where the odd columns report estimates with only sub-industry and country fixed effects and the even columns include the additional controls (see Table A.5

[27]More formally, the estimating equation is

$$\ln(tfpr_{ijc}) = \beta_0 + \beta_1 female_{ijc} + \beta_2 exp_{ijc} + \beta_3 register_{ijc} + \beta_4 size_{ijc} + \beta_5 city_{ijc} + X_j + Z_c + \varepsilon_{ijc}$$

where X_j and Z_c are sub-industry j and country-level c fixed effects. β_1 is the coefficient of interest. The notes in Table 5 provide details on variables.

Table 5: Gender estimates

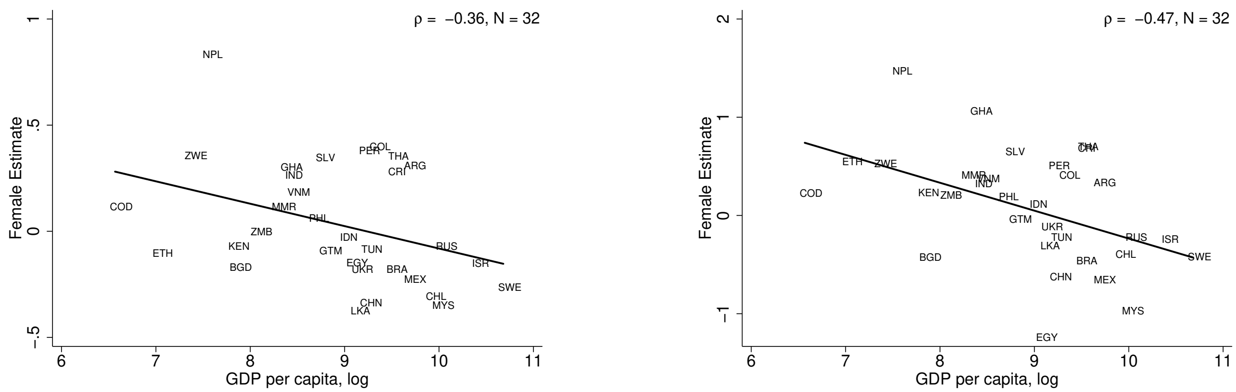
	All Countries		South America		South Asia		Eastern Europe	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.00	0.01	0.04	0.06	0.21***	0.18**	-0.10	-0.17
	(0.05)	(0.05)	(0.14)	(0.13)	(0.07)	(0.07)	(0.16)	(0.18)
Experience		-0.02*		-0.05		0.03		-0.02
		(0.01)		(0.03)		(0.02)		(0.05)
Register		0.06		-0.02		0.08		-0.20
		(0.04)		(0.10)		(0.07)		(0.33)
Size		0.16***		0.16***		0.20***		-0.05
		(0.02)		(0.05)		(0.04)		(0.08)
City		0.02		0.06*		-0.09***		0.03
		(0.02)		(0.04)		(0.03)		(0.06)
N	14478	14444	2888	2888	4163	4163	742	707
R^2	0.132	0.146	0.263	0.282	0.100	0.131	0.115	0.126

Notes: The dependent variable is ln $(tfpr^j_{si})$. All estimates include sub-industry and country fixed effects. Experience is the top manager's experience working in the industry (less than 5, 5-10, 10-15, 15-20, >20 years); Register is an indicator whether the establishment was formally registered when it began operations; Size is an indicator whether an establishment is small, medium or large; and City is an indicator whether the establishment operates in a city with a population of <50, 50-250, 250-1000, >1000 (in thousands) or is the capital. Standard errors are in parenthesis and ***, **, * denote significance at the 1, 5 and 10 percent level.

for estimates for the other sub-continents). As my focus is on gender, I briefly highlight some of the control variables. Establishment 'Size' is positive and highly significant in most sub-continents, consistent with the argument in Guner et al. (2008) that distortions on production rise with establishment size. The estimates for 'Experience' are mostly negative implying distortions on production fall with experience, and the estimates for 'City' are positive, except in South Asia where it is negative and significant.

Turning to gender, the coefficient on the female indicator is essentially zero for the full sample of countries but there are differences across sub-continents. In South Asia, the female estimate is positive and significant, and mainly driven by India. The estimate in column (5), which only includes sub-industry and country fixed effects shows that female establishments are associated with 23 percent higher $tfpr$ than male establishments; when including the additional controls (column 6) female establishments are associated with having 20 percent

Figure 4: Gender estimates and development



(a) $\ln(tfpr)$ (b) $\ln(mrpk)$

Notes: The female estimate is from a regression of establishment level $\ln(tfpr)$ (panel a) and capital distortion $\ln(mrpk)$ (panel b) on whether the top manager is female and includes sub-industry controls. OLS estimates for panel (a) and (b) are -1.237 (0.462) and -0.790 (0.221).

higher $tfpr$.[28] Inasmuch as the controls are picking up key factors that affect $tfpr$, this points to a form of gender bias against female establishments in South Asia. In Eastern Europe the female estimate is negative and for South America the estimates are positive, but in both cases not significant.[29]

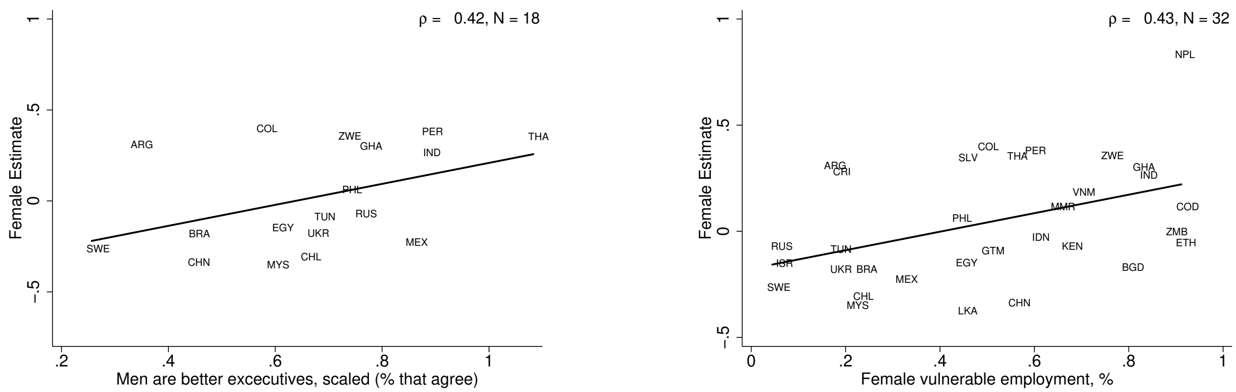
While I have focused on sub-continents, I also examine how the gender estimate at the country-level varies with economic development. Figure 4 panel (a) plots the country-level female estimate (based only on sub-industry controls) against GDP per capita.[30] The relationship is negative implying that countries where women face higher distortions also have lower GDP per capita (the OLS regression estimate is -1.24 (0.46)). And this negative relationship becomes stronger when restricting the sample to establishments that have fewer than 50 employees (see Figure 6, panel a). This is because distortions rise with size and larger establishments are mostly male owned. Panel (b) plots the country-level female es-

[28]The estimating equation is semi-log so the magnitude for coefficient β_i is $(\exp(\beta_i) - 1) \times 100$.

[29]The female estimates when $\ln(mrpk)$ and $\ln(mrpn)$ are dependent variables are similar in sign and significance as the estimates in Table 5. Also, accounting for female concentration in an industry (i.e., share of female establishments in an industry) does not have an effect on the female estimate.

[30]The patterns also hold when the female estimate is based on including controls for establishment size, city of operation, registration status and experience.

Figure 5: Gender estimates and social norms



(a) WVS (b) ILO STAT

Notes: The female estimate is from Figure 4 panel (a) based on $\ln(tfpr)$. The x-axis in panel (a) is the percentage of women that strongly agree that men are better business executives than women relative to the country average from the WVS. The x-axis in panel (b) is the percentage of female employment deemed vulnerable from the ILO and based on a five year average (2010-2014). OLS estimates for panel (a) and (b) are 0.733 (0.386) and 0.416 (0.125).

timate based on the capital distortion, as measured by $\ln(mrpk)$, against GDP per capita. The country estimates are more negatively correlated with GDP per capita, suggestive that on average access to capital affects women more severely in poorer countries. In contrast, the female estimate based on the output distortion, $\ln(mrpn)$, shows no relationship with development (and not reported). Of note, in relatively richer countries the female estimates in panel (a) and (b) are negative, however, this should not be necessarily interpreted as men facing higher distortions as the estimates are a lower bound for female gender bias as already discussed. Rather, it is the slope in Figure 4 that is relevant and points to women facing higher distortions in relatively poorer countries.

Finally, I also explore whether the country level estimate for gender, from Figure 4 panel (a), is related to gender norms in a country from other data sources. For this, I rely on the World Values Survey 2010-2013 (WVS) which reports how strongly people in a country hold to certain values, and also the percentage of female employment deemed vulnerable—measured as the share of own account and contributing family workers among females—from the International Labour Organization (ILO). From the WVS, I focus on the percentage of people that strongly agree men are better business executives than women. Figure 5, panel (a) shows the percentage of women that hold this view relative to the country average plotted against the female estimate. There is a positive and highly significant relationship in a sample of 18 countries (countries in the WBES and WVS do not perfectly overlap).[31] In countries where women face higher distortions there is a smaller relative gap in how men and women perceive men as better business executives (closer to one); alternatively, where women are more likely to agree that men make better executives. Panel (b) plots the percentage of vulnerable female employment against the female estimate, and for a broader set of countries. The correlation is also positive and significant showing that women face a harder road operating a business, especially in countries where female paid employment is lacking.

While not reported, the correlations with development in Figure 5 are stronger when the female estimate is based on $\ln(mrpk)$ and virtually flat when based on $\ln(mrpn)$.[32] Together with Figure 4, from a development perspective this suggests that in poorer countries female establishments are primarily hindered by capital distortions on production. While Kalemli-Ozcan and Sørensen (2014) find that capital distortions drive misallocation in Sub-Saharan Africa, this result shows that female establishments in particular face higher capital distortions and is a pattern that is negatively linked to development.

5 Sensitivity Analysis

I now evaluate the sensitivity of the results to defining establishment gender based on the top manager's gender and focusing on an aggregated manufacturing sector. In regards to the former, the advantage of using the top manager to define establishment gender is they are central to overseeing day to day operations and thus most encumbered by gender specific distortions

[31]The WVS also includes a question on whether men make better political leaders than women. The results from Figure 5 also hold when using this question, or taking an average of the two (better business executives and political leaders).

[32]For panel (a) and (b) respectively, the correlation with GDP (ln) is $\rho = 0.41$ and $\rho = 0.51$ when the estimate is based on $\ln(mrpk)$, and $\rho = 0.11$ and $\rho = 0.02$ when based on $\ln(mrpn)$.

on production. Nevertheless, the business owner (when different from the top manager) may be the ultimate decision maker at the establishment level and thus most affected by gender related distortions to production (e.g., securing contracts, networking). To better account for this I consider three cases: (1) focus on smaller establishments where it is more likely the top manager is also the owner/decision maker, (2) using business ownership categories by gender to define establishment gender, and (3) focusing on smaller establishments together with business owner gender.[33] Concerning the aggregated manufacturing sector, I have assumed common factor shares in production (α_s) due to data limitations. To evaluate the sensitivity of this, I focus on the 'Food products and beverage' sub-industry (ISIC 15) which is the sub-industry that has the most number of observations. Focusing on one sub-industry, albeit at the 2-digit level, can address concerns that aggregating to a manufacturing sector with common factor shares are driving the results.

To keep the analysis concise and for exposition, I evaluate whether the descriptive statistics presented in Table 2, the estimates for gender in Table 5 and its link to development (Figure 4) are sensitive to each of these cases. Table 6 summarizes the impact on the female estimate and Figure 6 shows its link to development. While I do not report all specifications from Section 4, the broad patterns and implications consistently hold for all countries and across sub-continents, including the reallocation policy in Section 4.3.

Establishment size. To better connect the top manager as the business owner and decision maker of an establishment, I now restrict the sample to establishments that have fewer than 50 employees. This relies on the assumption that top managers in small establishments are likely the final decision maker on business operations. Focusing on smaller establishments is also a useful control if there is heterogeneity among top female managers related establishment size (at least more heterogeneity relative to males). For instance, large

[33] The share of female employees in an establishment can also be used to define establishment gender (e.g., if the share of female employees is above some threshold). This however is not an ideal measure for establishment gender because hiring and personnel practices can already reflect gender specific distortions, and show up in $tfpr$.

Table 6: Sensitivity of gender estimate

	All Countries	South America	South Asia	Eastern Europe
(1) Restrict size to $n < 50$:				
Female	0.024	0.099	0.200**	-0.193
	(0.067)	(0.144)	(0.100)	(0.152)
(2) Bus. Ownership Categories:				
Female	0.06	—	0.192	-0.286
	(0.068)	—	(0.105)	(0.834)
(3) Bus. Ownership Categories & $n < 50$:				
Female	0.195**	—	0.180	-0.004
	(0.087)	—	(0.119)	(0.210)
(4) Food and Beverages (ISIC 15)				
Female	-0.047	-0.343	0.339**	–
	(0.122)	(0.221)	(0.165)	–

Notes: The female coefficient is estimated similar to the odd columns in Table 5. Four scenarios are considered: (1) size is restricted to establishments that hire less than 50 full-time workers, (2) gender is defined based on the business owner's gender, (3) combines the scenarios in (1) and (2), and (4) restricts to the sample to one sub-industry. The sample size for the four sensitivity checks, respectively, for All countries are 7966, 7130, 3888, 888; for South America are 1779, 62, 51, 237; for South Asia are 2721, 3903, 2538, 383; and for Eastern Europe are 506, 445, 202, 55.

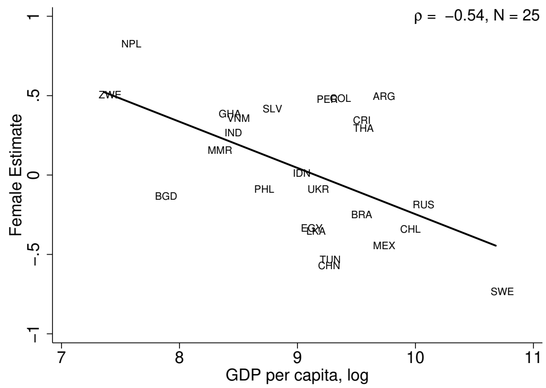
establishments may be mandated to allocate women to managerial positions, or conversely, female managers at large establishments may need to be more talented to break the 'glass ceiling'. These are examples of differences in manager quality or the types of barriers women face that vary with establishment size.

When restricting the sample to establishments that have fewer than 50 employees, the descriptive statistics in Table 2 are virtually unchanged: female establishments have higher (lower) $TFPR$ in all countries, South Asia and South America (Eastern Europe); and women consistently face higher distortions on capital across all regions. The female estimates have the same signs and significance as those reported in Table 5, and the female estimates are more negatively related to GDP (a correlation of -0.54 among a sample of 25 countries). These results also broadly hold when restricting the sample to establishments that have fewer than 30 and 10 employees.[34]

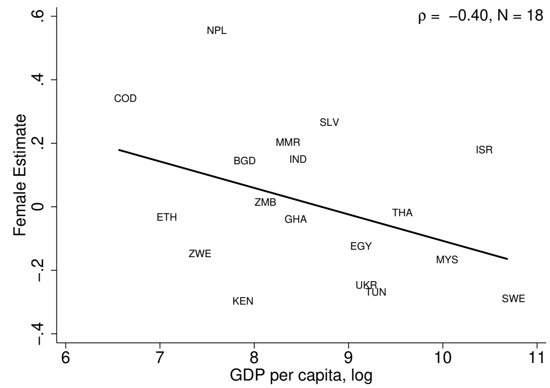
[34] An alternate view is that women select into entrepreneurship out of necessity and operate small, low productivity establishments (Poschke, 2013). When restricting the sample to establishments that have more than 5 employees, i.e., to focus on more serious entrepreneurs, the main results are unchanged.

Figure 6: Sensitivity of Female estimate and Development

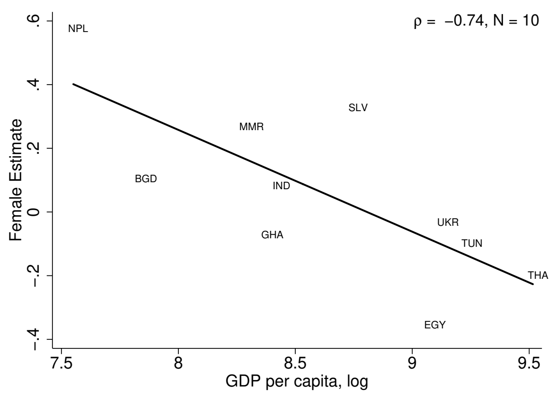
(a) Size, $n < 50$



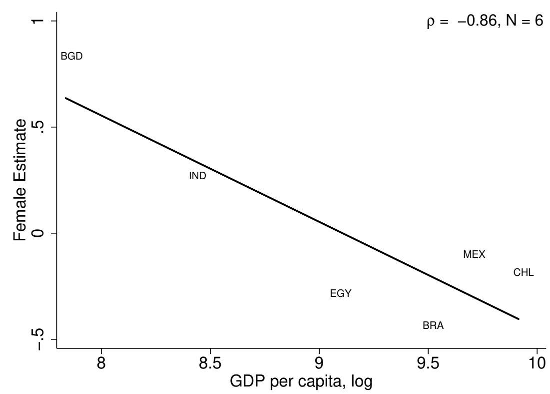
(b) Bus. Ownership Categories



(c) Bus. Ownership Categories & $n < 50$



(d) ISIC 15



Notes: The figures are the equivalent of Figure 4 panel (a), which plots the country level female estimate against GDP per capita, for the four specifications. OLS regression estimates and standard error of log GDP per capita on the female estimate for panel (a) is −1.01 (0.344), for panel (b) −1.88 (0.998), for panel (c) is −1.72 (0.431), and for panel (d) −1.49 (0.253).

Business Ownership Categories. The WBES includes a question on ownership categories by gender, which range from whether the owners are all men, majority men, equally divided across men/women, majority women and all women. I now define an establishment as female if at least 50 percent of owners are female (i.e., if the owners are not all men or majority men). This definition implies a similar proportion of female establishments as when the top manager is used to define establishment gender. A limitation is the sample size falls by 50 percent, notably for South America. The descriptive statistics reported in Table 2 continue to hold for all countries and across sub-continents except that now female establishments face higher distortions in Eastern Europe. The estimates for the female coefficient have the same sign as in Table 5 but is no longer statistically significant for South Asia (estimates for South America are not reported as there are fewer than 100 observations). The female estimates remain negatively correlated with GDP.

Business Ownership Categories and Establishment Size $n < 50$. A concern with using business ownership categories is that a business owner may not be actively involved in 'running' the business, and instead delegate decision making to the manager. To limit this possibility, I restrict the sample to establishments that have fewer than 50 employees and define an establishment as female if at least 50 percent of the owners are female (i.e., combining the two scenarios from above). This is likely to better link business ownership to active management of the establishment. The descriptive statistics reported in Table 2 hold for South America and Eastern Europe, but now males face higher distortions on average in South Asia. The signs for the female estimate continue to support the view that women face higher distortions to production in South Asia, and is now also positive and significant for all countries. The relationship between the female estimate remains negatively to development.

Food and Beverage sub-industry (ISIC 15). I now restrict the sample to the sub-industry that has the most number of observations, which allows me to focus on establishments within a narrower level of disaggregation within manufacturing.[35] I use the top manager to define establishment gender, and since the value of α_s is not overly critical I keep its value unchanged. The descriptive statistics in Table 2 mostly hold for all countries and South Asia, specifically with respect to the distortion on capital, but no longer for South America. That male establishments face higher distortions in Eastern Europe cannot be verified due to insufficient observations. Concerning the gender estimate, Figure 6 shows a strong negative relationship supporting the view that female establishments face higher distortions in poorer countries.

6 Conclusion

An avenue for understanding the vast cross-country income differences is that female business owners face discrimination in many parts of the world which prevents society from operating at its full potential. This paper has documented the extent of misallocation across male and female business owners using an established framework for measuring misallocation/distortions. Female establishments face higher distortions on production in many parts of the world, notably in South America and South Asia, and males face higher distortions in Eastern Europe. The distortions women face are primarily from higher distortions on capital, and is a pattern that consistently holds across broad geographic regions. Relevant is that the higher female distortions are evident among a plausibly more talented group of women who have overcome various gender-specific entry barriers into entrepreneurship. Removing distortions across gender implies proportionally large increases in female market shares and TFP gains attributed to females, particularly in places where women face high distortions on production. Regression estimates suggest that differential distortions by gender are associated with a form of bias, against women in South Asia and is negatively related to economic development. Taken together, these results suggest that differential distortions across gender

[35] The main results also hold for the second and third largest sub-industries in the sample, 'metal products' (ISIC 28) and 'rubber and plastic products' (ISIC 25), based on a sample of 6 countries.

is an important factor for understanding misallocation and female entrepreneurship in poor countries.

While my focus has been to present evidence for misallocation by gender across a broad range of countries and geographic regions, several areas warrant further exploration to better ground and establish discrimination across gender. The first relates to using more census level data and focusing on specific countries (say in South Asia) to more clearly establish whether female establishments face higher distortions to running a business and at a finer level of industry disaggregation. The second is to evaluate the long-run impact of removing gender specific barriers on female market shares and establishment size while allowing for movement along the entry/exit margin. In this regard, the evidence of gender discrimination and its impact on productivity that I find serves as a lower bound.

References

Asker, J., Collard-Wexler, A., and De-Loecker, J. (2014). Dynamic Inputs and Resource (Mis)Allocation. *Journal of Political Economy*, 122(5):1013–1063.

Bartelsman, E., Haltiwanger, J., and Scarpetta, S. (2013). Cross Country Differences in Productivity: The Role of Allocative Efficiency. *American Economic Review*, 103(1):305–334.

Bento, P. (2021). Female Entrepreneurship in the U.S. 1982-2012: Implications for Welfare and Aggregate Output. Manuscript.

Bils, M., Klenow, P. J., and Ruane, C. (2018). Misallocation or Mismeasurement? Manuscript.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1):1–51.

Bloom, N. and Van Reenen, J. (2007). Measuring and Explaining Management Practices Across Firms and Countries. *The Quarterly Journal of Economics*, 122(4):1351–1408.

Chiplunkar, G. and Goldberg, P. (2021). Aggregate Implications of Barriers to Female Entrepreneurship. NBER Working Paper No. 28486.

Cuberes, D. and Teignier, M. (2016). Aggregate effects of gender gaps in the labor market: A quantitative estimate. *Journal of Human Capital*, 10(1):1–32.

Cuberes, D. and Teignier, M. (2017). Gender Gaps in Entrepreneurship and their Macroeconomic Effects in Latin America. IDB-WP-848.

Cuberes, D. and Teignier, M. (2018). Macroeconomic costs of gender gaps in a model with entrepreneurship and household production. *The B.E. Journal of Macroeconomics*, 18(1):20170031.

de Mel, S., McKenzie, D., and Woodruff, C. (2009). Are Women More Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns. *American Economic Journal: Applied Economics*, 1(3):1–32.

Dias, D. A., Marques, C. R., and Richmond, C. (2016). Misallocation and productivity in the lead up to the Eurozone crisis. *Journal of Macroeconomics*, 49:46–70.

Dias, D. A., Marques, C. R., and Richmond, C. (2019). A tale of two sectors: why is misallocation higher in services than in manufacturing? *Review of Income and Wealth*.

Fairlie, R. W. and Robb, A. M. (2009). Gender differences in business performance: evidence from the characteristics of business owners survey. *Small Business Economics*, 33:375–395.

Gorodnichenko, Y., Revoltella, D., Svejnar, J., and Weiss, C. T. (2019). Resource Misallocation in European Firms: The Role of Constraints, Firm Characteristics and Managerial Decisions. NBER Working Paper No. 24444.

Guner, N., Ventura, G., and Xu, D. Y. (2008). Macroeconomic Implications of Size Dependent Policies. *Review of Economic Dynamics*, 11(4):721–744.

Hardy, M. and Kagy, G. (2018). Mind The (Profit) Gap: Why Are Female Enterprise Owners Earning Less Than Men? *AEA Papers and Proceedings*, 108:252–55.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica*, 87(5):1439–1474.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing productivity in China and India. *Quarterly Journal of Economics*, 124(4):1403–1448.

Hsieh, C.-T. and Klenow, P. J. (2014). The Life Cycle of Plants in India and Mexico. *Quarterly Journal of Economics*, 129(3):1035–1084.

Inklaar, R., Lashitew, A. A., and Timmer, M. P. (2017). The role of resource misallocation in cross-country differences in manufacturing productivity. *Macroeconomic Dynamics*, 21(3):733–756.

Jayachandran, S. (2020). Microentrepreneurship in Developing Countries. NBER Working Paper No. 26661.

Kalemli-Ozcan, S. and Sørensen, B. E. (2014). Misallocation, Property Rights, and Access to Finance: Evidence from Within and Across Africa. In *NBER Chapters, in: African Successes*, volume 3: Modernization and Development, pages 183–211.

La Porta, R. and Shleifer, A. (2014). Informality and Development. *Journal of Economic Perspectives*, 28:109–126.

Lee, M. (2022). Allocation of female talent and cross-country productivity differences. STEG Working Paper WP017.

Levine, R. and Rubinstein, Y. (2016). Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?*. *The Quarterly Journal of Economics*, 132(2):963–1018.

Morazzoni, M. and Sy, A. (2021). Female Entrepreneurship, Financial Frictions and Capital Misallocation in the US. BSE Working Paper 1299.

Poschke, M. (2013). 'Entrepreneurs out of necessity': a snapshot. *Applied Economics Letters*, 20(7):658–663.

Ranasinghe, A. (2023). Gender Specific Distortions, Entrepreneurship and Misallocation. University of Alberta, Working Paper No. 2023-01.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4):707–720.

Sabarwal, S. and Terrell, K. (2008). Does Gender Matter for Firm Performance? Evidence from Eastern Europe and Central Asia. World Bank Policy Research Working Paper 4705.

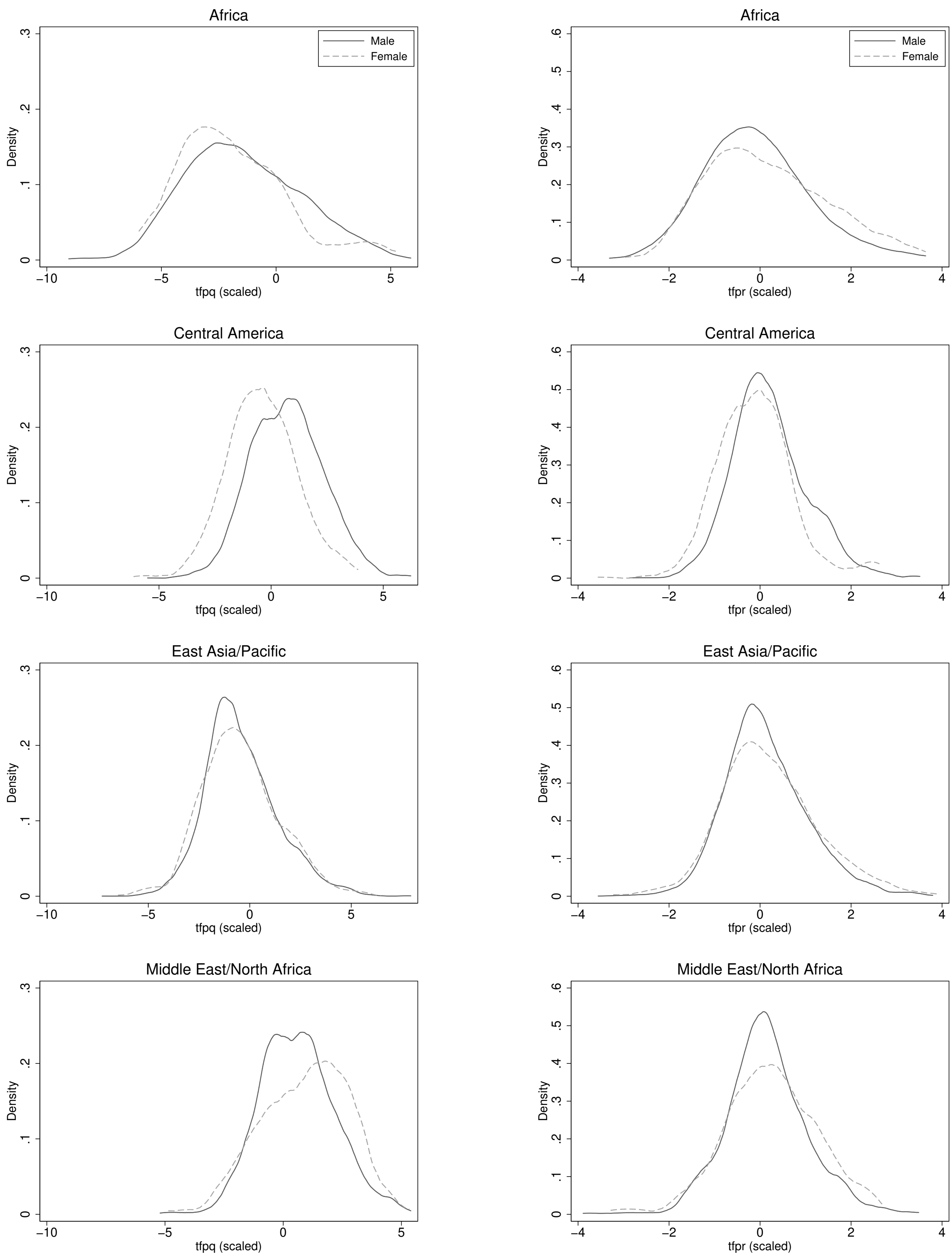
A Appendix

Table A.1: Descriptive Statistics: Other continents

	Africa			Central America			Middle East/North Africa			East Asia/Pacific		
	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
# of establishments	788	707	81	1526	1322	204	1017	962	55	3352	2570	782
% of estabs.		0.91	0.09		0.87	0.13		0.95	0.05		0.77	0.23
% of sales		0.89	0.11		0.96	0.04		0.91	0.09		0.81	0.19
Employees: $\ln(n)$												
mean	3.06	3.09	2.84	2.97	3.05	2.48	2.97	2.93	3.72	3.38	3.46	3.15
Distortions:												
output $\bar{\tau}^j$	0.66	0.66	0.45	0.25	0.26	0.08	0.52	0.51	0.59	0.50	0.52	0.46
capital $\bar{\kappa}^j$	-0.58	-0.47	-0.06	0.62	0.61	1.21	-0.29	-0.27	-0.41	-0.01	0.00	0.41
$TFPR^j$	2.71	2.92	3.59	1.85	1.87	1.66	1.79	1.79	1.69	2.39	2.39	2.90

Notes: Statistics are from the most recent survey for a country in the WBES 2008-17, and are weighted with the exception of '# of establishments' which is based on number of observations. See text for additional details.

Figure A.1: Physical and revenue productivity density plots by sub-continent



(a) Productivity

(b) Distortions

Notes: Kernel density plots for physical and revenue productivity are $ln\left(tfpq_{si}^{j} M_s^{\frac{1}{\sigma-1}} / TFP_s^{fb}\right)$ and $ln(tfpr_{si}^{j}/TFPR_s)$, respectively. A Kolmogorov-Smirnov test for equality of the $tfpq$ distributions across males and females is rejected for Africa and Central America ($p < 0.01$), and also for Middle East/North Africa ($p < 0.05$). There is no evidence the $tfpr$ distributions differ across gender in these four sub-continents.

Table A.2: Physical and revenue productivity: mean and dispersion

	All Countries			South America			South Asia			Eastern Europe		
	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
Physical productivity, $tfpq$:												
Mean												
$\ln(tfpq_{si}^{j})$	-0.06	-0.06	-0.03	-1.71	-1.66	-2.04	1.00	0.96	1.43	0.64	0.70	0.41
Dispersion												
Std. Dev.	1.96	1.93	2.14	2.43	2.35	2.88	1.52	1.51	1.55	1.22	1.22	1.20
75-25	2.56	2.51	2.75	3.44	3.33	3.95	2.06	2.05	2.10	1.53	1.60	2.13
90-10	4.90	4.85	5.35	6.30	6.10	7.22	3.93	3.86	4.34	3.06	2.97	2.65
Revenue productivity, $tfpr$:												
Mean												
$\ln(tfpr_{si}^{j})$	-0.09	-0.10	-0.02	-0.08	-0.07	-0.12	-0.10	-0.11	0.07	0.25	0.29	0.11
Dispersion												
Std. Dev.	0.96	0.94	1.07	1.01	0.96	1.26	0.90	0.90	0.91	0.84	0.86	0.74
75-25	1.17	1.14	1.23	1.13	1.05	1.62	1.20	1.20	1.23	0.95	1.09	0.57
90-10	2.30	2.26	2.54	2.58	2.46	3.85	2.27	2.26	2.44	2.15	2.15	2.05
Observations	14478	12620	1858	2888	2536	352	4163	3879	284	744	644	100

Notes: Dispersion statistics for productivity are scaled by first-best productivity, that is, $tfpq_{si}^{j}(M_s^j)^{\frac{1}{\sigma-1}}/TFP_s^{fb}$, where TFP_s^{fb} is a weighted average across the respective sample of countries. Dispersion statistics for revenue productivity are based on $\ln(tfpr)/TPFR_c$, where $TPFR_c$ is a weighted average across the respective sample of countries.'75-25' is difference between the 75th and 25th percentile, and similarly for '90-10'. See text for details.

Table A.3: Physical and revenue productivity: mean and dispersion

	Africa			Central America			Middle East/North Africa			East Asia/Pacific		
	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
Physical productivity, $tfpq$:												
Mean												
$\ln(tfpq_{si}^{j})$	-1.45	-1.40	-1.82	-0.13	-0.08	-0.46	0.27	0.26	0.59	-0.43	-0.34	-0.68
Dispersion												
Std. Dev.	2.45	2.47	2.20	1.70	1.73	1.48	1.56	1.58	1.32	1.87	1.82	1.98
75-25	3.49	3.60	2.90	2.34	2.42	1.29	1.96	1.96	1.62	2.40	2.43	2.32
90-10	6.41	6.55	4.82	4.20	4.32	3.25	3.54	3.54	2.90	4.88	4.84	5.21
Revenue productivity, $tfpr$:												
Mean												
$\ln(tfpr_{si}^{j})$	0.01	0.01	0.05	-0.04	-0.05	0.04	0.15	0.16	0.02	0.09	0.07	0.13
Dispersion												
Std. Dev.	1.11	1.10	1.27	0.82	0.82	0.84	1.03	1.03	0.96	0.95	0.92	1.06
75-25	1.52	1.49	1.68	1.14	1.12	0.96	1.09	1.11	0.71	1.11	1.03	1.27
90-10	2.71	2.67	3.39	1.91	1.93	1.61	2.21	2.30	1.46	2.32	2.27	2.52
Observations	788	707	81	1526	1322	204	1017	962	55	3352	2570	782

Notes: Dispersion statistics for productivity are scaled by first-best productivity, that is, $tfpq_{si}^{j}(M_s^j)^{\frac{1}{\sigma-1}}/TFP_s^{fb}$, where TFP_s^{fb} is a weighted average across the respective sample of countries. Dispersion statistics for revenue productivity are based on $\ln(tfpr)/TPFR_c$, where $TPFR_c$ is a weighted average across the respective sample of countries.'75-25' is difference between the 75th and 25th percentile, and similarly for '90-10'. See text for details.

Table A.4: Reallocation policies: uniform and common distribution of distortions

	Africa		Central America		Middle East/North Africa		East Asia/Pacific	
	Data	Reallocation	Data	Reallocation	Data	Reallocation	Data	Reallocation
First-best policy with uniform distortions								
Female sales shares among:								
all estabs.	0.109	0.144	0.041	0.058	0.079	0.054	0.242	0.261
top ten %	0.174	0.180	0.031	0.039	0.058	0.050	0.169	0.208
Average change in sales:								
Female estabs.	–	0.209	–	0.059	–	0.106	–	0.044
Male estabs.	–	0.015	–	0.009	–	0.008	–	0.010
TFP gains	–	2.92	–	3.38	–	3.56	–	3.43
due to females (share) ≈	–	0.14	–	0.06	–	0.05	–	0.26
Reallocation with common distribution of distortions:								
Female sales shares among:								
all estabs.	0.109	0.174	0.041	0.086	0.079	0.062	0.242	0.240
top ten %	0.174	0.236	0.031	0.056	0.058	0.061	0.169	0.187
Average change in sales:								
Female estabs.	–	0.075	–	0.022	–	0.029	–	0.006
Male estabs.	–	0.04	–	0.002	–	0.002	–	0.002

Notes: 'First best policy' are statistics based on a policy that equalizes $tfpr$ across all establishments in a country and 'Policy with common distribution of distortions' is a policy where male and female establishments face the same *distribution* of output and capital distortions.

Table A.5: Gender estimates: Other continents

	East Asia/Pacific		Central America		Africa		Middle East/N. Africa	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.03	-0.03	-0.13	-0.01	0.08	0.12	-0.17	-0.19
	(0.09)	(0.09)	(0.14)	(0.12)	(0.17)	(0.18)	(0.17)	(0.15)
Experience		-0.08***		-0.03		-0.05		0.07
		(0.02)		(0.03)		(0.05)		(0.06)
Register		0.09		-0.22**		0.16		0.09
		(0.07)		(0.10)		(0.19)		(0.17)
Size		0.11***		0.30***		0.14*		-0.01
		(0.04)		(0.04)		(0.08)		(0.08)
City		0.04		0.01		-0.08		0.10*
		(0.05)		(0.03)		(0.09)		(0.05)
N	3352	3352	1526	1526	787	787	1016	1016
R^2	0.073	0.090	0.176	0.237	0.127	0.137	0.133	0.149

Notes: The dependent variable is $\ln\left(tfpr_{si}^{j}\right)$. All estimates include sub-industry and country fixed effects. Experience is the top manager's experience working in the industry (less than 5, 5-10, 10-15, 15-20, >20 years); and Register is an indicator whether the establishment was formally registered when it began operations; Size is an indicator whether an establishment is small, medium or large; and City is an indicator whether the establishment operates in a city with a population of <50, 50-250, 250-1000, >1000 (in thousands) or is the capital. Standard errors are in parenthesis and ***, **, * denote significance at the 1, 5 and 10 percent level.

A.1 List of Countries

Table A.6: List of countries

		Observations		
Country	ISO Code	Total	Male	Female
Africa:				
Ethiopia (2015)	ETH	175	161	14
Ghana (2013)	GHA	160	139	21
Kenya (2013)	KEN	176	167	9
Zambia (2013)	ZMB	128	112	16
Zimbabwe (2016)	ZWE	149	128	21
Central America:				
Costa Rica (2010)	CRI	165	147	18
DRC (2013)	COD	101	80	21
El Salvador (2016)	SLV	185	140	45
Guatemala (2010)	GTM	178	155	23
Mexico (2010)	MEX	897	800	97
East Asia:				
China (2012)	CHN	1207	1106	101
Indonesia (2009)	IDN	515	405	110
Malaysia (2015)	MYS	176	115	61
Myanmar (2016)	MMR	193	150	43
Philippines (2009)	PHL	323	249	74
Thailand (2016)	THA	398	119	279

Table A.6: List of countries

Country	ISO Code	Observations Total	Male	Female
Vietnam (2009)	VNM	540	426	114
Eastern Europe:				
Russia (2012)	RUS	278	226	52
Sweden (2014)	SWE	159	153	6
Turkey (2013)	TUR	178	164	14
Ukraine (2013)	UKR	129	101	28
Middle East/North Africa:				
Egypt (2016)	EGY	681	647	34
Israel (2013)	ISR	114	107	7
Tunisia (2013)	TUN	222	208	14
South America:				
Argentina (2010)	ARG	455	428	27
Brazil (2009)	BRA	931	794	137
Chile (2010)	CHL	581	530	51
Colombia (2010)	COL	427	349	78
Peru (2010)	PER	494	435	59
South Asia:				
Bangladesh (2013)	BGD	957	908	49
India (2014)	IND	2769	2583	186
Nepal (2013)	NPL	197	181	16
Sri Lanka (2011)	LKA	240	207	33

Department of Economics, University of Alberta
Working Paper Series

2020-01: Does Locational Marginal Pricing Impact Generation Investment Location Decisions? An Analysis of Texas's Wholesale Electricity Market – **Brown, D.,** Zarnikau, J., Woo, C.
2019-16: A Unified Explanation of Trade Liberalization Effects Across Models of Imperfect Competition – **Alfaro, M.**, Lander, D.
2019-15: Analyzing Firm Behaviour in Restructured Electricity Markets: Empirical Challenges with a Residual Demand Analysis – **Brown, D., Eckert, A.**
2019-14: LinkedIn(to) Job Opportunities: Experimental Evidence from Job Readiness Training – **Wheeler, L.**, Garlick, R., Johnson, E., Shaw, P., Gargano, M.
2019-13: Entry Preemption by Domestic Leaders and Home-Bias Patterns: Theory and Empirics – **Alfaro, M.**
2019-12: How Much are Good Schools Worth? Evidence from School Acquisitions in Beijing – **Su, X.,** Yu, H.
2019-11: Enumerating Rights: More is Not Always Better – Ball, S., **Dave, C.,** Dodds, S.
2019-10: Motivating the Optimal Procurement and Deployment of Electric Storage as a Transmission Asset – **Brown, D.,** Sappington, D.
2019-09: Pricing Patterns in Wholesale Electricity Markets: Unilateral Market Power or Coordinated Behavior? – **Brown, D., Eckert, A.**
2019-08: Son Preference and Child Under Nutrition in the Arab Countries: Is There a Gender Bias Against Girls? – **Sharaf, M.,** Rashad, A., Mansour, E.
2019-07: How Local Economic Conditions Affect School Finances, Teacher Quality, and Student Achievement: Evidence from the Texas Shale Boom – **Marchand, J.,** Weber, J.
2019-06: Subjects in the Lab, Activists in the Field: Public Goods and Punishment – **Dave, C.,** Hamre, S., Kephart, C., Reuben, A.
2019-05: Fiscal Austerity in Emerging Market Economies – **Dave, C.,** Ghate, C., Gopalakrishnan, P., Tarafdar, S.
2019-04: Inequality and Trade Policy: Pro-Poor Bias of Contemporary Trade Restrictions – **Ural Marchand, B.**
2019-03: Determinants of Locational Patenting Behavior of Canadian Firms – **Eckert, A., Langinier, C., Zhao, L.**
2019-02: The Microeconomics of New Trade Models – **Alfaro, M.**
2019-01: Universal Childcare for the Youngest and the Maternal Labour Supply – Kunze, A., **Liu, X.**
2018-19: On the Benefits of Behind-the-Meter Rooftop Solar and Energy Storage: The Importance of Retail Rate Design – **Boampong, R., Brown, D.**
2018-18: The Value Premium During Flights – **Galvani, V.**
2018-17: Asymmetric Information, Predictability and Momentum in the Corporate Bond Market – **Galvani, V., Li, L.**
2018-16: The Momentum Effect for Canadian Corporate Bonds – **Galvani, V., Li, L.**
2018-15: Green Technology and Patents in the Presence of Green Consumers – **Langinier, C.,** Ray Chaudhuri, A.
2018-14: Subjective Performance of Patent Examiners, Implicit Contracts and Self-Funded Patent Offices – **Langinier, C.,** Marcoul, P.
2018-13: Ordered Leniency: An Experimental Study of Law Enforcement with Self-Reporting – **Landeo, C.,** Spier, K.
2018-12: Imperfect Competition in Electricity Markets with Renewable Generation: The Role of Renewable Compensation Policies – **Brown, D., Eckert, A.**