STEG CEPR UKaid from the British people

STEG WORKING PAPER

MISALLOCATION ACROSS ESTABLISHMENT GENDER

Ashantha Ranasinghe

JULY 2022
STEG WP008

Misallocation across Establishment Gender

Ashantha Ranasinghe*
University of Alberta

14 July 2022

Abstract

I find substantial differences in the extent of misallocation across male and female-led establishments spanning many low and middle-income countries. In South American and South Asian countries female establishments face higher distortions to operating a business, whereas in Eastern European countries male establishments face higher distortions. These differences across gender hold when controlling for relevant establishment characteristics. Across countries, economic development is negatively related with female establishments facing higher distortions and misallocation, and is primarily from distortions on capital. Removing distortions across gender increases female market shares and average size, especially in poorer countries.

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, less understood are the particular characteristics of businesses that are most encumbered by the frictions that engender misallocation. 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 females account for a small share of all establishments and a smaller share of the market. For instance, in South America females account for 14 percent of establishments and 7 percent of sales; in Eastern Europe/Central Asia the corresponding numbers are 17 and 10 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 face higher distortions to production, particularly in poorer countries, and removing these distortions 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 imbed implicit taxes on capital and output 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 many 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, and this varies by the level of development. Female establishments face higher distortions on average in South Asian, South American, African and East Asian countries, whereas in Eastern European countries males face higher distortions. Across sub-continents the distribution of revenue productivity is more right-skewed for female establishments and exhibits higher dispersion, evidence that is consistent with females facing higher distortions and more misallocation. These results show that even among a plausibly more talented group of women that have selected in to formal sector entrepreneurship—having overcome any gender specific barriers to entry—women still face higher distortions to running a business than men.

To evaluate whether differences in the extent of distortions across male and female establishments are tied to gender, I estimate the relationship between distortions (revenue productivity) and gender controlling for relevant establishment characteristics. In South Asia, and particularly India, female establishments are associated with facing 20 percent higher distortions, whereas in South America the higher distortions are attributable to the other controls.[1] In contrast, in Eastern Europe male establishments are associated with 18 percent higher distortions. Across countries, the female estimate is negatively related to economic development (GDP per capita), implying that in poorer countries female establishments face

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

higher distortions to operating a business (relative to males). Importantly, this negative link with development is primarily from females facing higher distortions on capital. In addition, the country-level estimates, which are based on establishment-level data, are consistent 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] With respect to 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] 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 and finds that female establishments should account for a larger share of market resources than what is observed in the data, particularly in poorer countries where females face higher distortions. For example, among a fixed set of producers, a policy that eliminates misallocation in South Asian countries raises the share of female sales by over 10 percent, with larger impacts on capital and labour. This is quantitatively large since males account for over 90 percent of establishments and also (on net) increase sales from this policy. Allowing for a broader class of reallocation policies where distortions are tied to productivity, but not gender, implies much larger increases in the share of sales female establishments account for.

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 accounting for gender is an important part of the overall misallocation

[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 (2018) who examine differences in access to credit across entrepreneur gender in the U.S.

picture. In addition, I find 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 they account for a small share of the market and serves as a lower bound since selection into entrepreneurship, and hence the share of female establishments, is held fixed. More importantly, while removing misallocation raises average establishment size/sales across both genders, it has a much larger impact on average female size.

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 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 imbeds gender-specific distortions to production. Section 3 describes the data, its

[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.

advantages and limitations. Section 4 documents differences in distortions and misallocation across male and female establishments, estimates on gender, and changes in market shares from undoing misallocation. 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. Nevertheless, for completeness I include intermediate steps. In the quantitative analysis I primarily focus on a single industry (manufacturing), but here the framework is presented for multiple industries.

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 κ_{si}^j which affects the marginal product of capital relative to labour (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009).[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

[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.

[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.

(high $\bar{\kappa}_s^j$) or due to factors not tied to gender (high κ_{si}).

Establishment profit is

$$\pi_{si}^j = (1-\tau_{si}^j)p_{si}^j y_{si}^j - wn_{si} - (1+\kappa_{si}^j)rk_{si} \tag{2}$$

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 that have low productivity or face high distortions (output or capital), charge higher prices. Moreover, absent of productivity differences, female establishments charge a higher price relative to male establishments only if 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

$$\frac{\sigma-1}{\sigma}(1-\alpha_s)\frac{p_{si}^j y_{si}^j}{n_{si}^j} = \frac{w}{1-\tau_{si}^j} \equiv mrpn_{si}^j, \tag{4}$$

$$\frac{\sigma-1}{\sigma}\alpha_s\frac{p_{si}^j y_{si}^j}{k_{si}^j} = \frac{(1+\kappa_{si}^j)}{1-\tau_{si}^j}r \equiv mrpk_{si}^j. \tag{5}$$

Within an industry, a high marginal revenue product of labour $mrpn_{si}^j$ implies high output distortions (and high sales to labour ratios), and a high marginal revenue product of capital $mrpk_{si}^j$ 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

$$tfpq_{si}^{j} \equiv z_{si}^{j} = \frac{y_{si}^{j}}{k_{si}^{\alpha_s} n_{si}^{1-\alpha_s}}, \tag{6}$$

$$tfpr_{si}^{j} \equiv p_{si}^{j} z_{si}^{j} = \frac{p_{si}^{j} y_{si}^{j}}{k_{si}^{\alpha_s} n_{si}^{1-\alpha_s}} = \frac{\sigma}{\sigma-1} \Omega_s \frac{\left(1+\kappa_{si}^{j}\right)^{\alpha_s}}{\left(1-\tau_{si}^{j}\right)}, \tag{7}$$

where the expression for $tfpr_{si}^{j}$ is obtained using equations (3) and (6). Within an industry, high revenue productivity $tfpr_{si}^{j}$ is indicative that an establishment faces high barriers to production, high τ_{si}^{j} and/or κ_{si}^{j}, 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 $tfpr_{si}^{j}$ establishments) and operate on a larger scale than what is efficient. In this regard, $tfpr_{si}^{j}$ serves as a summary statistic for establishment distortions $\big($i.e., $tfpr_{si}^{j} \propto \frac{(1+\kappa_{si}^{j})^{\alpha_s}}{(1-\tau_{si}^{j})}\big)$, 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

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

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 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 $TFPR_s^j = \frac{(P_s Y_s)^j}{\left(K_s^j\right)^{\alpha_s}\left(N_s^j\right)^{1-\alpha_s}} = \frac{\sigma}{\sigma-1}\Omega_s \frac{(1+\bar{\kappa}_s^{j,avg})^{\alpha_s}}{1-\bar{\tau}_s^{j,avg}}$.[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.

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

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

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 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).

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

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

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

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 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). There is a finer disaggregation within manufacturing 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

[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.

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. Countries that have fewer than 5 female establishment observations are excluded. The final sample includes 78 countries, where about 40 percent of these countries have over 100 observations.

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. Sample weights are meant to correct for this and produce a representative sample at the country level. Nevertheless, missing observations are a concern, which is especially the case for capital in non-manufacturing (service and core) sectors, and potentially exacerbated when disaggregating by gender and industry.

To alleviate this concern, to an extent, I focus on an aggregated manufacturing sector (ISIC 15–37) so there are sufficient observations, and where missing information is less of an issue.[14] This would also mitigate biases that may arise from differential sorting across sectors, especially if females are more likely to select into non-manufacturing industries (though the proportion of female establishments does not markedly differ across manufacturing and non-manufacturing sectors in the WBES). 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. This is especially the case if there is little dispersion in marginal products within a finer disaggregation of industries but considerable dispersion when analyzing at the manufacturing level. And, if sub-industries within manufacturing that face high distortions also have a higher concentration of a specific gender. Focusing on sub-industries (2 digit ISIC level) that have more than 100 observations at the country level, neither of these issues appear to bias the results that follow. For instance, $tfpr$ dispersion in 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.[15] Related to the second point, 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 manufacturing sub-industries. 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 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.[16]

Nevertheless, I include sub-industry controls when feasible. In addition, in Section 5 I check

[14] Once missing information on capital and input costs are excluded the sample overwhelmingly reduces to manufacturing establishments (99 percent of the sample).

[15] As an example, in India the standard deviation for $tfpr$ in each sub-industry is within 12 percent of its country-level statistic. This holds across all sub-industries except for one sub-industry in China.

[16] A regression of $TFPR$ differences across gender (as defined in Figure 1) on the percentage of female establishments at the sub-industry level has a coefficient and standard error estimate of -0.005 (0.0101).

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 establishments. Thereafter, I estimate whether gender, among other establishment characteristics, can account for the observed differences in distortions across establishments. Finally, I examine the implications for market shares across gender and productivity if distortions are equalized across establishments. While I focus on establishments, the results that follow also hold for firms, though the sample is smaller. 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/managerial positions. Finally, and as noted before, I define male and female establishments based on the top manager's gender (the following section considers variants to this definition).

While the analysis is at the country level, I present results at the sub-continent level to provide a snap-shot of gender differences across broad geographic regions. I follow country sub-continent 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 briefly discussed, and are reported in the Appendix. A complete list of countries, observations and sub-continent groupings are reported in Table A.9.

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 $y_{si}^{j}=\zeta_s\left(p_{si}^{j}y_{si}^{j}\right)^{\frac{\sigma}{\sigma-1}}$. 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 (the exact value is not critical). For the manufacturing sector factor share, I set $\alpha_s = 0.406$.[17] 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. In Section 5 I focus on one sub-industry.

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.

4.2 Sample statistics, distortions and misallocation

4.2.1 Sample statistics and distortions

Table 1 reports sample statistics and measures of distortions for South America, South Asia and Eastern Europe, for all establishments and by gender (see Table A.3 in the Appendix for other sub-continents). All statistics and results are calculated using country-level sample

[17] 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).

weights.[18] Two points are worth noting before focusing on specific sub-continents. First, the proportion of female establishments in a country in the WBES is highly correlated with the proportion of female top managers in the World Development Indicators (WDI) dataset—also from the World Bank—suggesting that the proportion of establishments across gender in the WBES are a reasonable starting point.[19] Of note, the proportion of female establishments in the data 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)). Although the type and scale of entry barriers can vary across countries, they are a common hurdle females face across the range of development. Second, average establishment size in the WBES, reported in logs, is higher than what is found in other data sources (Bento and Restuccia, 2017; Poschke, 2018). The extent that larger establishment size may influence the results is addressed in Section 5.

In South America, females account for 14 percent of establishments and 7 percent of (value added) sales. They operate smaller establishments on average, as measured by full-time employees, and have about 6 fewer years working in their specific industry (experience). Importantly, female establishments face higher distortions on average as summarized by $TFPR^j$. They face higher capital distortions $\bar{\kappa}^j$—a pattern that is common in most sub-continents—and higher output distortions $\bar{\tau}^j$.[20] These broad patterns in South America are also evident in Africa, Central America and East Asia/Pacific.

In Eastern Europe, females account for 17 percent of establishments and 10 percent of sales.

[18] The broad patterns across gender hold with or without sample weights. In the WBES, establishment sample weights are at the country level. To obtain sub-continent level sample weights I further weight establishment weights by the number of observations in a country relative to its sub-continent. For instance, India accounts for about 60 percent of total observations in South Asia and Sri Lanka accounts for about 5 percent. Hence, statistics for India (Sri Lanka) receive a 60 (5) percent weight when generating statistics for South Asia. Alternatively, the WBES approach is to weight at the country level and calculate sub-continent averages by weighting each country equally.

[19] Based on the average between 2006 to 2015 reported in the WDI—where most countries have about two years of data in this time frame—the correlation across the two data sources is over 0.75 and 0.90 (for a sample of countries in the WBES that have over 100 and 500 observations respectively).

[20] The patterns for $MRPN^j$ and $MRPK^j$ are similar to $\bar{\tau}^j$ and $\bar{\kappa}^j$, respectively.

Table 1: Descriptive Statistics

	South America			South Asia			Eastern Europe		
	All	Male	Female	All	Male	Female	All	Male	Female
# of establishments	3206	2815	391	4217	3928	289	1280	1084	196
% of estabs.		0.86	0.14		0.92	0.08		0.83	0.17
% of sales		0.93	0.07		0.88	0.12		0.90	0.10
Employees: $\ln(n)$									
mean	3.31	3.40	2.81	3.40	3.39	3.60	3.06	3.06	3.05
std. dev.	1.12	1.11	1.03	1.10	1.11	0.98	1.01	1.00	0.88
Experience:									
mean	23.21	24.17	17.68	15.37	15.35	15.03	20.23	20.33	19.52
std. dev.	11.39	11.24	10.18	8.94	9.03	7.51	10.45	10.07	9.46
Distortions:									
output $\bar{\tau}^j$	0.38	0.37	0.45	0.51	0.51	0.51	0.24	0.27	-0.01
capital $\bar{\kappa}^j$	0.44	0.43	0.71	-0.06	-0.11	0.55	0.30	0.22	2.31
$TFPR^j$	1.79	1.76	2.73	2.35	2.29	2.89	1.85	1.89	1.71

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. The sample is restricted to establishments in the manufacturing sector that report values for capital and input costs; sales are value-added. Employees are permanent full-time employees and experience is the number of years the top manager has worked in the establishment's sector.

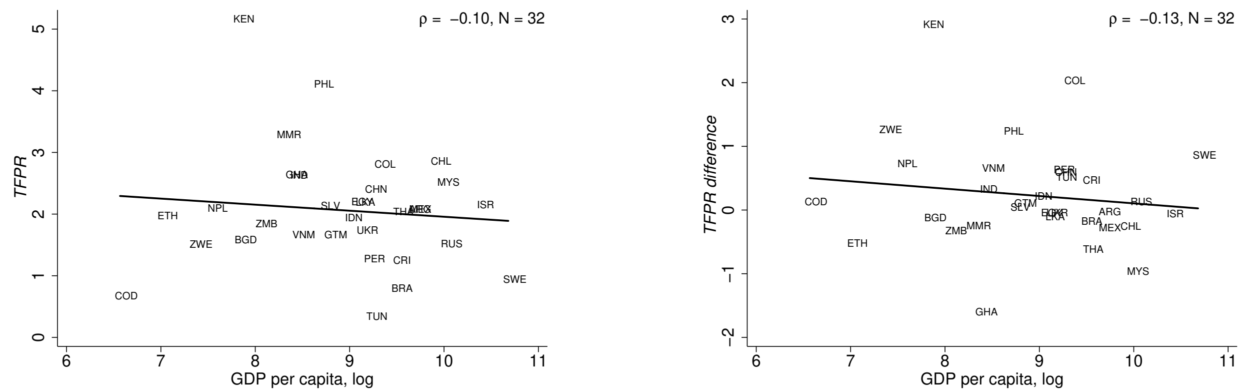
Females operate marginally smaller establishments and have about a year less experience on average.[21] Male establishments face higher distortions on production ($TFPR^j$), though there is heterogeneity across output and capital distortions. In particular, male establishments face higher output distortions but lower capital distortions. These patterns are also evident in Central America.

In South Asia, the picture is different. Females account for 8 percent of establishments and 12 percent of sales—they take a larger share of market sales relative to the proportion of female establishments. In contrast to Eastern Europe and South America, females in South Asia operate larger establishments than males on average—consistent with what Chiplunkar and Goldberg (2021) find using census-level data for India—and have similar experience working in their industry. Notably, female establishments face higher distortions on production, primarily due to higher capital distortions.[22] The fact that females operate larger estab-

[21] Average sales is lower among female establishments consistent with Sabarwal and Terrell (2008) who focus on Eastern Europe using the 2005 Business Environment and Enterprise Performance survey.

[22] Disaggregating by sub-industries in South Asia, is also consistent with females facing higher distortions on production. For instance, in 9 of the 10 sub-industries, females in South Asia face higher distortions (sample restricted to sub-industries that have at least 50 observations in a country). Results at the sub-industry

Figure 1: TFPR and economic development



(a) (b)

Notes: The sample is restricted to countries that have at least 100 observations. Panel (a) plots $TFPR$ for all establishments against (ln) GDP per capita, and Panel (b) plots $(TFPR^f - TFPR^m)/TFPR$. Country level GDP per capita is an average from 2005 to the most recent year in the Penn World Tables, 9.1.

lishments, 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.

Of interest is the link between distortions and economic development. Figure 1 Panel (a) plots country-level $TFPR$ against (log) GDP per capita, and where the sample is restricted to countries that have at least 100 observations.[23] While there is considerable variation in $TFPR$ across countries, there is no clear link to GDP. This is similar to Inklaar et al. (2017) who find that removing misallocation can generate large productivity increases in a country but it is not related to the level of development. Panel (b) plots the difference in $TFPR^j$ across gender (female less male) scaled by country-level $TFPR$, and again shows no clear pattern with GDP.

level in South America are more mixed, whereas in Eastern Europe there are inadequate observations to disaggregate at the sub-industry level.

[23]The WBES includes Sweden, which is the only high-income country in the sample. I exclude Turkey as it is an outlier for many of the estimates that follow. Results are not sensitive to including/excluding Sweden or Turkey.

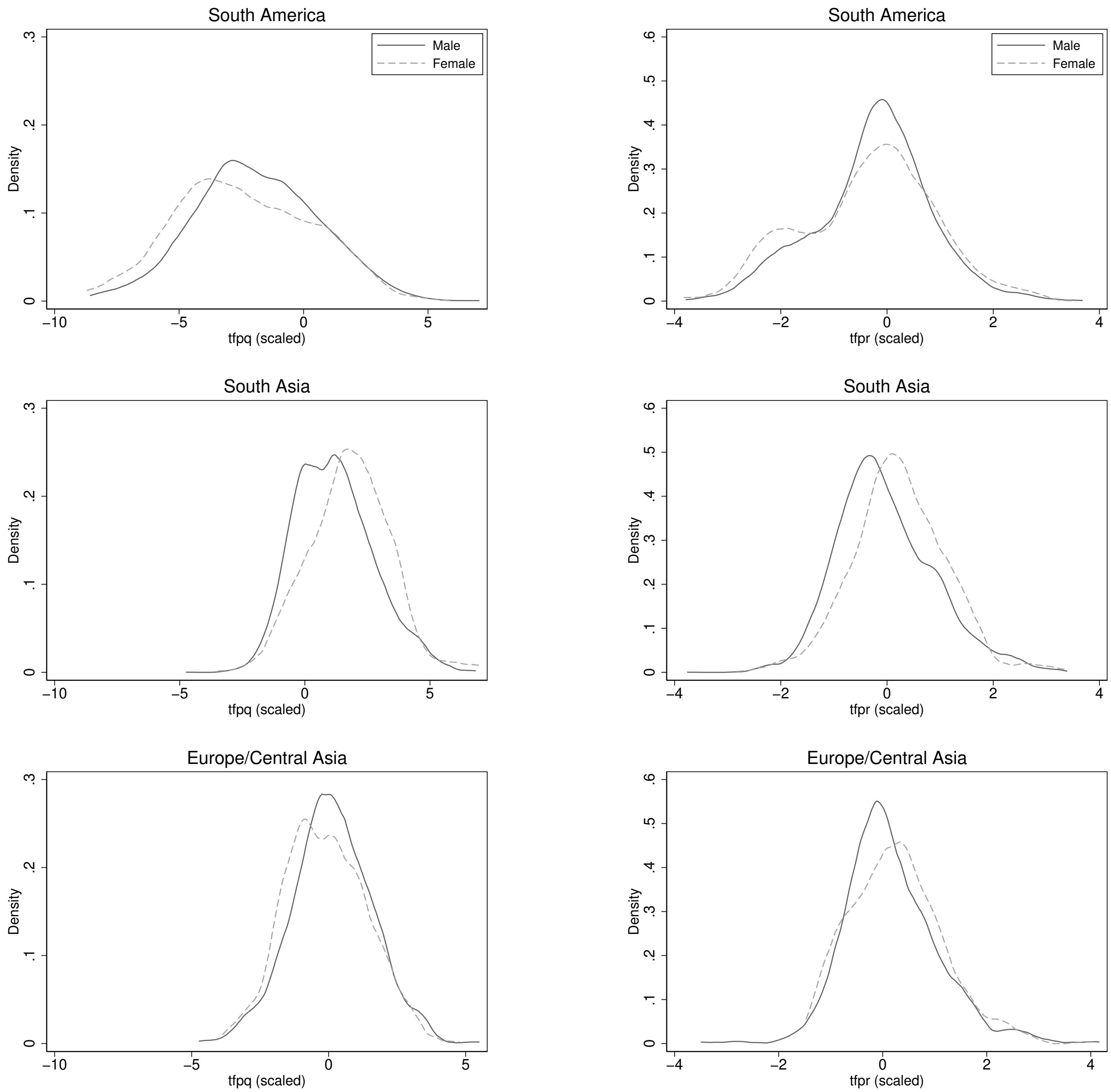
4.2.2 Misallocation

To further get at the differences across gender, I examine the dispersion in establishment productivity and distortions ($tfpr$). Figure 2 Panel (a), plots the distribution of physical productivity for male and female establishments; specifically, $ln\Big(tfpq_{si}^{j}M_s^{\frac{1}{\sigma-1}}/TFP_s^{fb}\Big)$, where M_s is the total number of establishments in the manufacturing sector. These distributions are statistically different across gender in each sub-continent. In South America and Eastern Europe, the distribution of productivity is more left-skewed for females, 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 for female establishments implying females are more productive on average. South Asia is particularly interesting because the distribution for male establishments has a long 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 are suggestive of policies that favour less productive female establishments.

These figures also show there is more productivity dispersion among female establishments in South America and South Asia (though some of this could be an artifact of a smaller female sample). In Europe, there is more productivity dispersion among male establishments, particularly at the tails. Across sub-continents, there is considerably more productivity dispersion in South America than in South Asia and Europe, whether looking at all establishments or by gender. Table A.1 highlights these points more formally.

Figure 2 Panel (b) plots the distribution of revenue productivity, $ln\big(tfpr_{si}^{j}/TFPR_s\big)$. Recall, in the ideal benchmark of this framework (no frictions or distortions), marginal products are equalized across all establishments and revenue productivity reduces to a constant. Hence, high revenue productivity dispersion implies a greater extent of misallocation across establishments and larger productivity losses. Moreover, high values of $ln\big(tfpr_{i}^{j}/TFPR_s\big)$ are

Figure 2: Physical and revenue productivity density plots

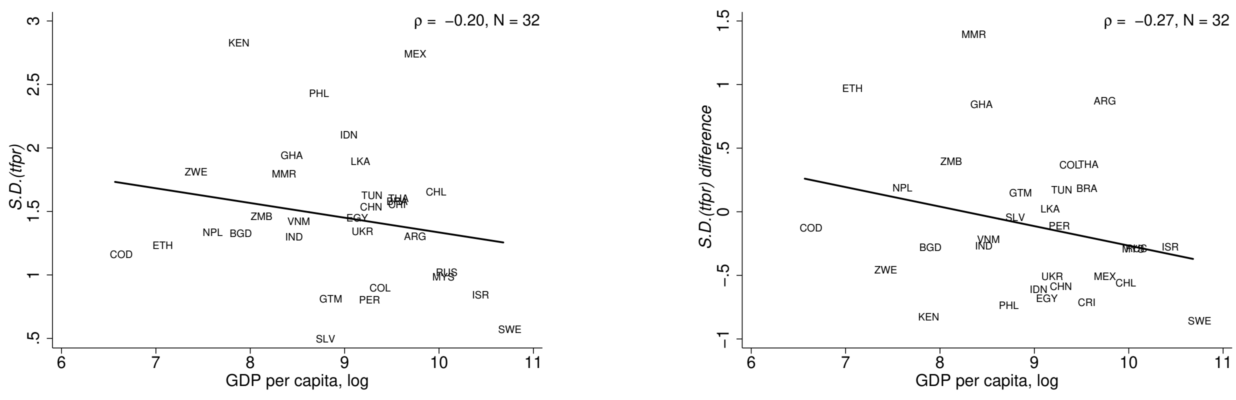


(a) Productivity

(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.

Figure 3: Misallocation and economic development



(a) (b)

Notes: The sample is restricted to countries that have at least 100 observations. Panel (a) plots the standard deviation of $tfpr$ at the country level scaled by $TFPR$, and panel (b) plots the difference in standard deviation of female and male $tfpr$'s scaled by the country level standard deviation for $tfpr$ (all establishments). GDP per capita is an average from 2005 to the most recent year in the Penn World Tables, 9.1. OLS estimates for panel (a) and (b) are −0.354 (0.297) and −0.458 (0.286).

consistent with an establishment facing high distortions on production, and vice versa.

The female revenue productivity distribution is more right-skewed in South Asia, and statistically different than the male distribution, implying they face higher distortions relative to male establishments. There is also a long left-tail for male establishments which is suggestive that a sub-group of male establishments benefit from implicit subsidies on production. The plots for South America show more revenue productivity dispersion among females, particularly at the tails, implying more misallocation among female establishments.[24] 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. Table A.2 highlights these points more formally.

A central hypothesis in the misallocation literature is that higher misallocation is a major source of cross-country income differences. To examine whether this prediction holds for the

[24]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.

WBES sample of countries, Figure 3 panel (a) plots the standard deviation of establishment $tfpr$ (scaled by country-level $TFPR$) against GDP per capita, and panel (b) plots the difference in the standard deviation of female and male establishment $tfpr$ (scaled by country-level standard deviation for $tfpr$) against GDP per capita. While not statistically significant, panel (a) and (b) depict downward slopes indicating that higher misallocation is negatively correlated with GDP per capita. Importantly, higher misallocation among females relative to male establishments is more strongly related with GDP per capita (and more amplified in the robustness checks that follow as shown in Figure 7).

4.3 Gender estimates

The evidence so far shows that female establishments face higher distortions and a greater extent of misallocation in South Asia and South America, whereas the opposite holds in Eastern Europe. I now explore whether these differences are related to gender, or if other establishment characteristics better account for them. For instance, females 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 productivity or 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 along with relevant controls in the WBES that capture some of the considerations noted above. In particular, I include controls for establishment size (small, medium 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 of 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).[25] I also include controls for sub-industry of operation (2 digit ISIC). This is particularly useful to control 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. Also included are country-level fixed effects and survey year controls.[26]

Conceptually, after accounting for these controls the coefficient on the female indicator is informative whether distortions in operating a business differ across gender, and its sign informs in which direction; a positive (negative) coefficient implies that females face higher (lower) distortions relative to males. I loosely refer to this as gender bias, i.e., establishments of a given gender are associated with facing higher distortions. Worth stressing, gender bias is a broad concept extending beyond entry into entrepreneurship/managerial positions and labour force participation, where substantial biases (most likely) against females can preclude such opportunity (Jayachandran, 2015, 2020).[27] My use of gender bias is narrower, focusing on formal businesses where labour force participation and selection into entrepreneurship has already taken place (i.e., women who have already overcome any entry barriers). As such, any association of bias against females is likely a lower bound, or put differently, only a part of the broader forms of gender bias females encounter in the workforce.

Table 2 reports regression estimates for South America, South Asia and Eastern Europe (see Table A.4 for female estimate with only industry and country controls, and Table A.5 for estimates for the other sub-continents). As my focus is on gender, I briefly highlight some of the other control variables. Establishment 'Size' is positive and highly significant in most

[25]The surveys do not include measures for education levels or productivity. Controls for whether formally registered and top manager's experience are attempts to account for this.

[26]More formally, the estimating equation is

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

where X_j and Z_c are sub-industry j and country-level c fixed effects, and W_t are survey year controls (from 2008-2017). β_1 is the coefficient of interest. The notes in Table 2 provides details on variables.

[27]That the proportion of female establishments is under 20 percent across sub-continents points to broad bias against females, whether due to lack of entrepreneurial opportunity, or low productivity attributed to nutrition/education deficits that preclude entrepreneurship or labour force participation.

Table 2: Gender estimates

	South America		South Asia		Eastern Europe	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.05	0.11	0.17**	0.12*	-0.20*	-0.19**
	(0.12)	(0.11)	(0.07)	(0.07)	(0.11)	(0.10)
Size	0.18***	0.15***	0.20***	0.19***	-0.02	-0.03
	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)
City	0.06*	0.03	-0.09***	-0.09***	0.06	0.06
	(0.04)	(0.04)	(0.03)	(0.03)	(0.05)	(0.06)
Experience	-0.05*	-0.07**	0.03	0.02	-0.05	-0.02
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)
Register	-0.04	0.01	0.08	0.06	-0.17	-0.06
	(0.09)	(0.09)	(0.07)	(0.07)	(0.27)	(0.31)
Female × Size		0.27*		0.02		0.10
		(0.17)		(0.09)		(0.14)
Female × City		0.16**		0.01		-0.08
		(0.07)		(0.09)		(0.08)
Female × Experience		0.03		0.08		-0.21**
		(0.08)		(0.06)		(0.08)
Female × Register		-0.31		0.67**		-0.11
		(0.28)		(0.27)		(0.26)
Fixed effects:						
Sub-industry	X	X	X	X	X	X
Country level	X	X	X	X	X	X
N	3206	3206	4217	4217	1157	1157
R^2	0.269	0.280	0.133	0.135	0.108	0.117

Notes: The dependent variable is $\ln\left(tfpr_{si}^{j}\right)$. Size is an indicator whether an establishment is small, medium or large; 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; 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. Variables interacted with 'Female' are differenced by its mean (male and female population). Estimates include sub-industry, country and time fixed effects. Standard errors are in parenthesis and ***, **, * denote significance at the 1, 5 and 10 percent level.

sub-continents, which is consistent with the argument in Guner et al. (2008) that distortions on production rise with establishment size. Based on column (1), moving from a small to medium-size (or medium to large) establishment in South America implies a 20 percent increase in establishment $tfpr$, or distortions on production.[28] The estimates for 'Experience' are mostly negative, and significant in some instances, implying that distortions on production fall with experience. This is consistent with the view that experienced managers have a wider network and more adept at navigating the business-politico environment. Estimates for 'City' are positive, except in South Asia where it is negative and highly significant.

Turning to gender, the coefficient on the female indicator is positive and significant in South Asia, and mainly driven by India.[29] The estimate in column (3) shows that after accounting for the various controls, a female establishment is associated with 18.5 percent higher $tfpr$ than male establishments. 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 evidence points to bias against male establishments; based on column (5), male establishments are associated with 18 percent higher $tfpr$ (this is also the case in Africa). An interpretation of this result, and consistent with Campa and Serafinelli (2019), is that gender roles are less traditional in Eastern European countries in part due to state socialism which promoted female economic inclusion. In addition, these countries are more likely to enact policies to promote female entrepreneurship, and worth exploring further using larger country-level datasets. For South America, the estimates in column (1) imply there is no evidence of bias, although the female coefficient is positive.[30]

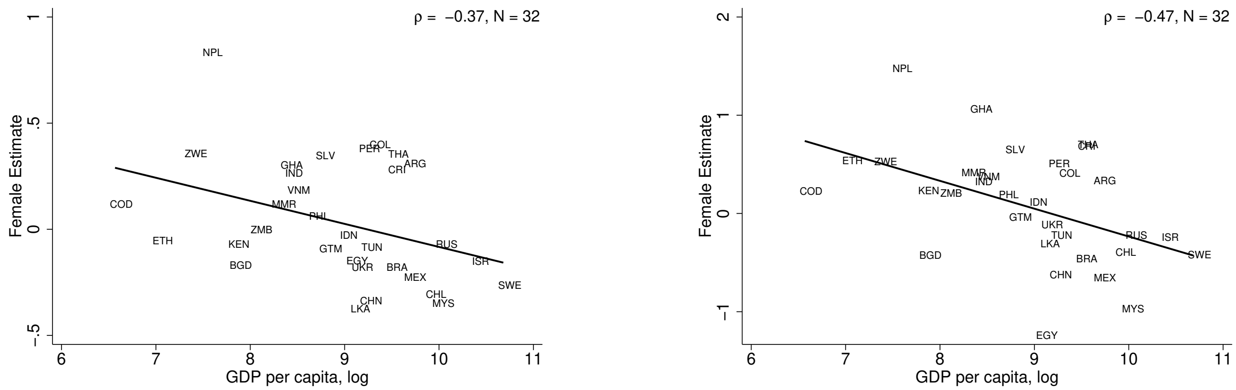
The even columns in Table 2 report estimates with interaction terms to evaluate whether the controls have differential effects across male and female establishments (the gender differential is estimated at the mean of each control). The gender bias association becomes more prevalent. Notably, in South America there is evidence of a bias against females that operate

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

[29]When India, who accounts for the majority of observations in South America, is excluded the female estimate in column (3) remains positive but is not significant, but the interaction with 'Register' in column (4) remains significant. The female estimates for South America and Eastern Europe hold when the country that accounts for the most observations, Brazil and Russia, are excluded.

[30]The female estimates when $\ln(mrpk)$ and $\ln(mrpn)$ are dependent variables have the same signs and statistical significance as $\ln(tfpr)$ for these sub-continents, except that $\ln(mrpn)$ is not statistically significant for Eastern Europe.

Figure 4: Gender estimates and development



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

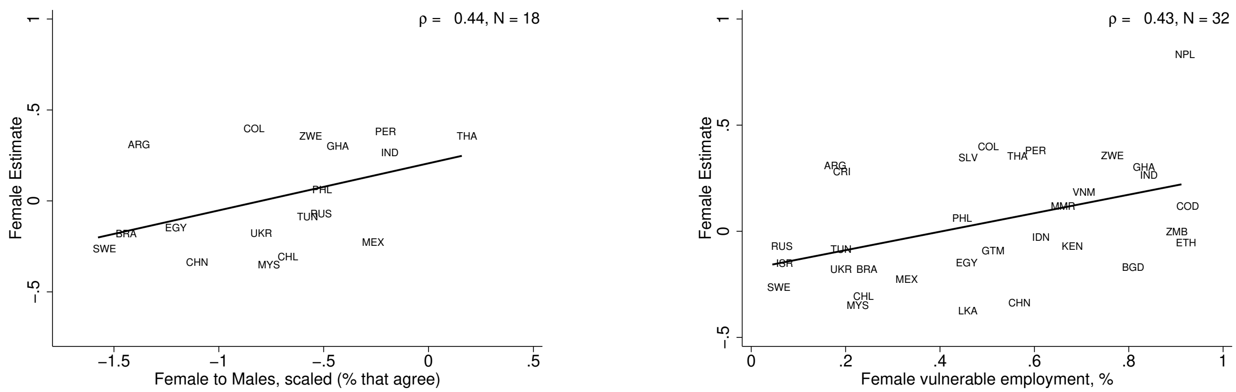
Notes: The female estimate is from a regression of establishment level $\ln(tfpr)$ (panel a) and capital distortion $\ln(mrpk)$ (panel b) on whether the top manager is female, and includes controls for establishment size, experience, city, registration status and sub-industry. OLS estimates for panel (a) and (b) are -1.278 (0.452) and -0.789 (0.221).

larger establishments or locate in bigger cities; in South Asia there is a bias against females that formally register when starting a business, and in Eastern Europe there is a bias against male establishments that have more experience.

Despite controlling for industry, female concentration in an industry (i.e., share of female establishments in an industry) can also impact differences in distortions across gender. I find that after controlling for industry, accounting for female concentration does not affect the main results, and if anything, produces larger female estimates (see Table A.4, panel B and C). This evidence supports the view that industry controls can reasonably account for any selection affects across gender.

While I have focused on sub-continents, it is useful to examine how the gender estimate at the country-level varies with economic development. Figure 4 panel (a) plots the country-level female estimate (estimated the same way as the previous regression) against GDP per capita for countries that have 100 or more observations. The relationship is negative implying that countries where females face higher distortions also have lower GDP per capita (the OLS regression estimate is -1.28 (0.45)). And this negative relationship becomes stronger

Figure 5: Gender estimates and social norms



(a) WVS (b) ILO STAT

Notes: The female estimate is from Figure 4 panel (a). The x-axis in panel (a) is the difference between the percentage of females and males that strongly agree that males are better business executives than females, scaled by its country level counter part, from the WVS. The x-axis in panel (b) is the percentage of female employment deemed vulnerable, from the ILO, based on a five year average (2010-2014). OLS estimates for panel (a) and (b) are 0.733 (0.386) and 0.426 (0.126).

when restricting the sample to establishments that have fewer than 50 employees (see Figure 8, panel a). This implies that focusing on gender among smaller establishments can be important for understanding the extent of misallocation across countries. Panel (b) plots the country-level female estimate based on the capital distortion, as measured by $\ln(mrpk)$, against GDP per capita. The estimate is more strongly correlated with GDP per capita, suggestive that on average access to capital affects females more severely in poorer countries. In contrast, the female estimate based on the output distortion, $\ln(mrpn)$, shows no relationship with development. Of note, in richer countries the female estimates in panel (a) and (b) are negative implying male establishments face higher distortions in these countries. While beyond the scope of this paper, it is useful to explore the particular features that contribute to this, notably the types of entry barriers that affect selection and policies that promote female entrepreneurship.

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) plots the difference between the percentage of women and men that hold this view scaled by the country average, 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 females face higher distortions there is a smaller relative gap in how males and females perceive men as better business executives; 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 females 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 much 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 than output distortions. 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.

What accounts for gender bias? I conclude this section by briefly exploring whether specific distortions drive high $tfpr$, and whether they impact establishments differently by gender. Figure 4 panel (b) implies capital distortions particularly affect females and is negatively related to development. The WBES can be useful to shed further insight as it reports

[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 $\rho = 0.48$ and $\rho = 0.51$ when the estimate is based on $\ln(mrpk)$, and $\rho = 0.01$ and $\rho = 0.02$ when based on $\ln(mrpn)$.

establishment responses to whether specific obstacles are a severe, major, moderate, minor or non obstacle to doing business (coded on a 0–4 scale). These are subjective measures, and I focus on several of these—corruption, crime, limited access to finance, practices of the informal sector and tax related (tax rates and how it is administered)—which are found to be highly distortionary in the literature.[33] I also include whether an inadequately educated workforce is an obstacle, which can be especially pertinent if hiring educated workers is related to establishment gender.

I regress $tfpr$ (in logs) on whether the top manager is female and indicators for whether corruption, crime, limited access to finance, the informal sector, tax related and workforce concerns are obstacles to doing business (i.e., a major or severe obstacle).[34] Table A.6 presents the results for South America, South Asia and Eastern Europe (see Tables A.7 for other continents). Crime has a significant and positive association with $tfpr$ in South Asia, and the association is negative in Eastern Europe. Informal sector and workforce concerns mostly have no link to $tfpr$, though in South America inadequate access to educated workers is associated with 12 percent higher $tfpr$. A consistent pattern across sub-continents is that finance, tax related and (to a lesser extent) corruption concerns are negatively associated with $tfpr$. This is to say, establishments that report finance, tax related concerns and corruption as obstacles to operating a business face lower distortions, which on the surface is counter intuitive. One interpretation is that these obstacles are major hinderances to business operation, but only if property rights are secure; that is once issues related to crime and informality are sufficiently addressed. To an extent, this interpretation mirrors the empirical findings in Johnson et al. (2002) and the quantitative results in Ranasinghe and Restuccia (2018) and Lopez-Martin (2019).

[33]For instance on corruption, taxation and informality see Dusha (2015), López (2017), La Porta and Shleifer (2014) and Ulyssea (2018); on crime see Ranasinghe (2017), Oguzoglu and Ranasinghe (2017) and Besley and Mueller (2018); on limited access to finance see Buera et al. (2011) and Midrigan and Xu (2014). Related to the joint effects of crime and finance, and informal sector and finance see Ranasinghe and Restuccia (2018) and Lopez-Martin (2019).

[34]The estimates are larger and in the same direction when $\ln(mrpk)$ is the dependent variable. Also, I include manager experience to account for the subjective nature of the obstacles reported, controls for city and sub-industry of operation, as well as country-level fixed effects and survey year.

Focusing on gender, the female estimate remains positive and significant in South Asia, and negative and significant for Eastern Europe. Looking at the interaction between female and the obstacles, there is no clear evidence these obstacles have a differential impact across gender, except in a few cases. Specifically, in South America taxes affect males more severely, and in South Asia corruption has a bigger effect on females. Nevertheless, there are no consistent or systematic patterns that one particular obstacle is central for understanding $tfpr$ differences across establishments. Rather, the analysis points to obstacles as more region specific, and therefore policies that aim to lower distortions need to be tailored and targeted to those regions.

4.4 Reallocation gains

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; to be consistent with the framework in Section 2 and results presented so far, I focus on the existing set of establishments and abstract from issues related to entry and exit. To the extent that barriers to entry stemming from societal norms 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. 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 South America, South Asia and Eastern Europe (see Table A.8 for other continents).

I focus on the impact of this 'first-best' policy, which in essence levels the playing field across all establishments, on female market shares and changes in average size (sales). In South Asia equalizing $tfpr$ across establishments raises female sales shares by 11 percent (from a share of 12.8 to 14.2), the share of labour by 22 percent (from a share 11.6 to 14.2), and has close to a 2-fold increase in the share of capital. These are quantitatively large increases in female market shares since males account for over 90 percent of establishments, and also (on net) expand from this policy.[35] Among the top ten percent of establishments (based on sales), females now account for close to 20 percent of sales (up from 15), though the share of female establishments falls from 17 to 15.4 percent. While there is substantial adjustment across the distribution with male and female establishments expanding and shrinking, overall the reallocation has a larger impact on female establishments; sales rise by 2.2 percent on average whereas for males it rises by 0.2 percent. The above implies that among the current set of producers, female establishments are inefficiently small in South Asian countries and should expand relative male establishments.

In Eastern Europe, the policy lowers the share of sales, capital and labour among female establishments. Recall, male establishments face higher distortions on average and account for over 80 percent of establishments (Table 1). Nevertheless, average sales rise by 12 percent for females and by 2.6 percent for males; again implying that under the policy females expand more on average.[36] In South America, the impact on female market shares are minor, as are the effects among the top ten percent of establishments. There is however a larger increase in female average sales which is consistent with the $TFPR$ differences in Table 1.

Focusing on productivity, the reallocation policy results in a three-fold increase in TFP in South America and South Asia, and a 2.5 fold increase in Europe.(Reported TFP gains are the

[35] Since females 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).

[36] An interpretation of this 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 after the reallocation.

Table 3: First-best policy with uniform distortions

	South America		South Asia		Eastern Europe	
	Data	Reallocation	Data	Reallocation	Data	Reallocation
Female shares:						
Sales θ^f	0.073	0.071	0.128	0.142	0.109	0.081
Capital θ^f_k	0.074	0.071	0.077	0.142	0.092	0.081
Labour θ^f_n	0.070	0.071	0.116	0.142	0.134	0.081
Female shares (top ten %):						
Sales	0.051	0.046	0.152	0.196	0.066	0.097
Establishments	0.092	0.098	0.170	0.154	0.062	0.080
Average change in sales:						
Female Estabs.	–	0.043	–	0.022	–	0.120
Male Estabs.	–	0.007	–	0.002	–	0.026
TFP gains	–	3.15	–	3.03	–	2.57
due to females (share) ≈	–	0.07	–	0.14	–	0.08

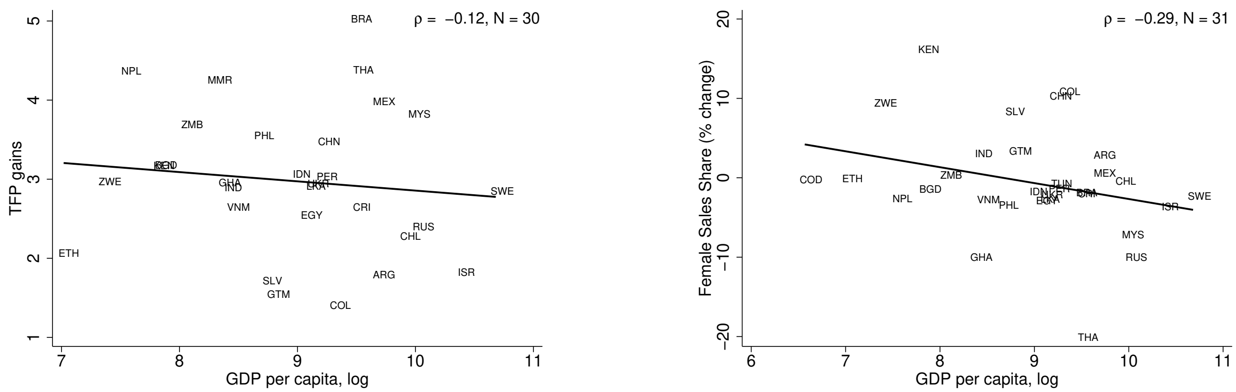
Notes: 'Data' refers to weighted statistics based on the WBES 2006-17. 'Reallocation' are statistics based on a policy that equalizes $tfpr$ across establishments in a country, and weighted up to the sub-continent level. Female shares (top 10%) refer to the proportion of females among the top ten percent of establishments 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 in the data.

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 females, whereas in South Asia 14 percent is due to females. Of note, 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.

Figure 6 shows the factor change in TFP and percentage change in female sales shares across levels of development from the reallocation policy. Panel (a) shows a fairly flat relationship between TFP and development, consistent with Inklaar et al. (2017), implying that misallocation and TFP gains are not necessarily related to development. In contrast, panel (b) shows a negative relationship where there is a larger increase in female sales shares in poorer countries, or put differently, females should account for a larger share of the market in poor countries.[37] This again supports the broader findings in this paper that misallocation across

[37]Recall, females 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. Panel (a) shows that the expansion of female (minority) relative to male (majority of) establishments is larger in poorer countries.

Figure 6: TFP gains and Female Sales Shares with Reallocation



(a) (b)

Notes: The figure shows the factor change in TFP and impact on female sales shares when $tfpr$ is uniform across establishments within a country. In panel (b), values above (below) zero indicate that female sales shares rise (fall) from the reallocation. Of note, since females account for a minority of establishments their sales shares rise only if they are *both* considerably more productive and face higher distortions. In panel (a) Congo and Tunisia, and in panel (b) Myanmar are outliers and excluded. OLS estimates for panel (a) and (b) are −0.117 (0.180) and −0.041 (0.0175) .

establishment gender is negatively related in development.

I also examine the impact of a broader reallocation policy where females face the same distribution of output and capital distortions as males based on productivity. Such a policy highlights the impact on female market shares when distortions vary with productivity ($tfpq$), and are linked to gender only through differences in establishment productivity.[38] As this policy affects aggregate resources and hence equilibrium prices, I omit statements on productivity. Table 4 shows the results. In all sub-continents female market shares rise, and by more than the policy when distortions are uniform across establishments.[39] The same holds for female shares among the top 10 percent of producers. This is because there are more male establishments at the extreme tail of the productivity distribution. And since output and capital distortions rise with productivity, females now account for a larger share of the

[38]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.

[39]In South Asia female capital shares increase by a smaller margin than in Table 3. This is because capital distortions rise faster with productivity, and females have higher productivity on average.

Table 4: Reallocation policy with common distribution of distortions

	South America		South Asia		Eastern Europe	
	Data	Reallocation	Data	Reallocation	Data	Reallocation
Female shares:						
Sales θ^f	0.073	0.094	0.128	0.160	0.109	0.112
Capital θ_k^f	0.074	0.113	0.077	0.091	0.092	0.142
Labour θ_n^f	0.070	0.104	0.116	0.181	0.134	0.131
Female shares (top ten %):						
Sales	0.051	0.068	0.152	0.231	0.066	0.103
Establishments	0.092	0.096	0.170	0.134	0.062	0.088
Average change in sales:						
Female Estabs.	–	0.032	–	0.013	–	0.062
Male Estabs.	–	0.003	–	0.001	–	0.009

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

market. The increase in average sales is larger among females but is smaller in comparison to the policy when distortions are uniform, again because distortions and productivity are positively related.

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, especially in poorer countries where female establishments face higher distortions. In South Asia, this means an 11-25 percent increase in female sales shares and a 22-56 percent increase in labour shares; in South America and Eastern Europe the results are more mixed. Across all sub-continents the reallocation policies generate a larger increase in average sales for female establishments. Again, these effects are likely a lower bound as they exclude any effects from selection into entrepreneurship.

5 Sensitivity Analysis

I now evaluate the sensitivity of the results, in particular the sensitivity to defining establishment gender based on the top manager's gender, and the sensitivity to focusing on an aggregated manufacturing sector. In regards to the former, the advantage of using the top manager to define establishment gender is that they are central to overseeing day to day operations (Bloom et al., 2013), and thus most encumbered by gender specific distortions 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 and connect to the gender of the central decision maker at the establishment, 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.[40] 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 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 1, the estimates for gender in Table 2 and its link to development (Figure 3 and 4 and) are sensitive to each of these cases. Table 5 summarizes the impact on the female estimate and Figures 7 and 8 show its link to development, all of which broadly support the central results in Section 4.

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

[40]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 5: Sensitivity of gender estimate

	South America	South Asia	Eastern Europe
(1) Restrict size to $n < 50$:			
Female	0.099	0.198**	-0.190*
	(0.124)	(0.097)	(0.103)
(2) Bus. Ownership Categories:			
Female	—	0.163	-0.166
	—	(0.105)	(0.124)
(3) Bus. Ownership Categories & $n < 50$:			
Female	—	0.164	-0.102
	—	(0.120)	(0.132)
(4) Food and Beverages (ISIC 15)			
Female	-0.054	0.204	–
	(0.185)	(0.192)	–

Notes: The female coefficient is estimated similar to those in Table 2. 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 (2) and (3), and (4) restricts to the sample to one sub-industry. The sample size for the four sensitivity checks, respectively, for South America are 1854, 82, 72, 404; for South Asia are 2721, 3957, 2583, 473; and for Eastern Europe are 792, 901, 633, 67.

fewer than 50 employees. This relies on the assumption that top managers are more likely to be the final decision maker on business operation in smaller establishments. Focusing on smaller establishments is also a useful control if there is considerable heterogeneity among top female managers across establishment size (at least more heterogeneity relative to males). For instance, large establishments may be mandated to allocate females 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 females face that vary with establishment size.

When restricting the sample to establishments that have fewer than 50 employees, the descriptive statistics in Table 1 are virtually unchanged, with female establishments having higher (lower) $TFPR$ in South Asia and South America (Eastern Europe). The female estimates have the same signs and significance as those reported in Table 2. In addition, that misallocation across gender and the female estimates are negatively related to GDP is now stronger (a correlation of -0.38 and -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, although the overall sample is smaller.

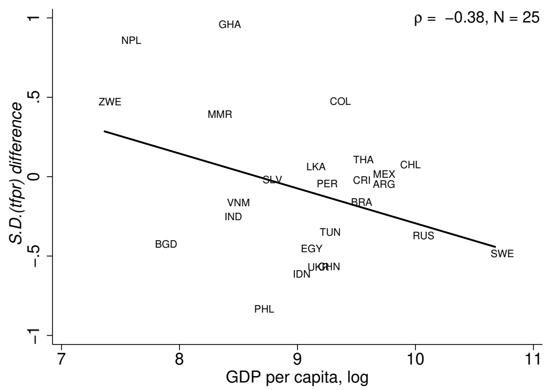
An alternate view is that females select into entrepreneurship out of necessity (Poschke, 2013), say due to discrimination that limits employment opportunity, and hence they likely operate small, low productivity establishments. To account for this, I restrict the sample to establishments that have more than 5 employees. The estimates are virtually the same in magnitude and significance as those in Table 2 (results not reported). To an extent this is not surprising since the analysis is among formal establishments, whereas selection into entrepreneurship out of necessity (or lack of outside opportunity) would more naturally apply to informal establishments.

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 1 hold for South America and South Asia, but no longer for Eastern Europe. The estimates for the female coefficient reported in Table 2 hold for South Asia and Europe in their sign but are no longer statistically significant (I do not report estimates for South America as there are fewer than 100 observations). Misallocation across gender and the female estimates, however, are now more negatively correlated with GDP.

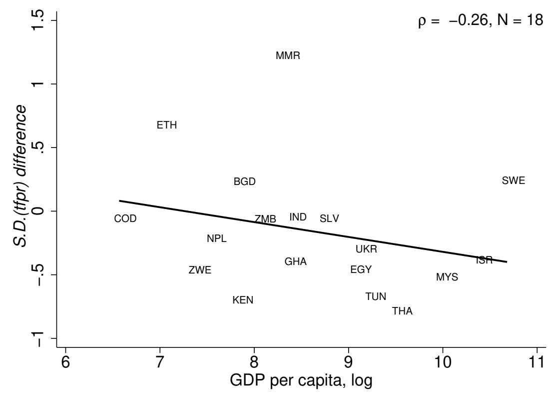
Business Ownership Categories and Establishment Size $n < 50$. A concern with using business ownership categories, as above, is that a business owner may not be actively involved in 'running' the business, and instead delegate decision making to the man-

Figure 7: Sensitivity of Misallocation across Gender and Development

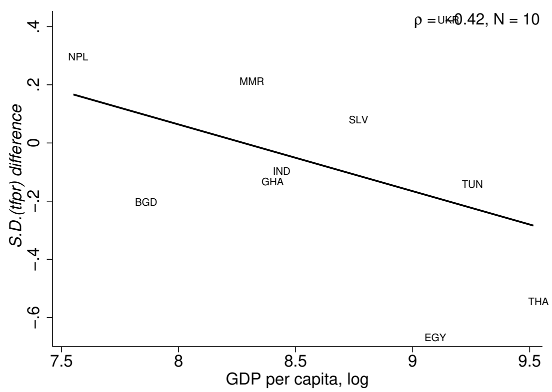
(a) Size, $n < 50$



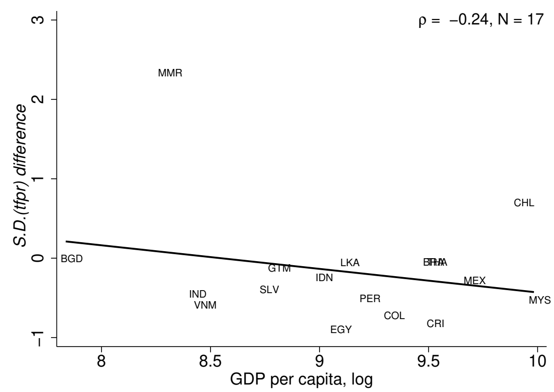
(b) Bus. Ownership Categories



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



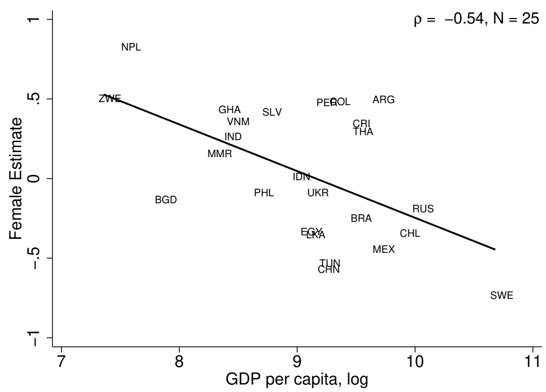
(d) ISIC 15



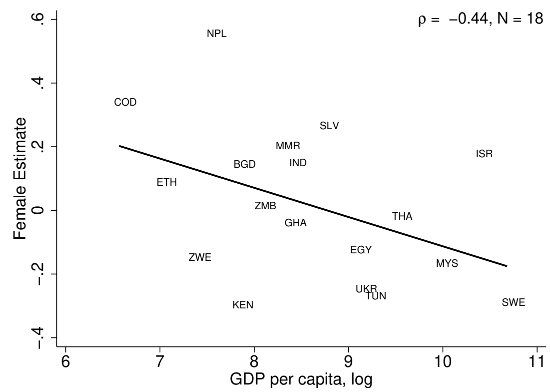
Notes: The figures are the equivalent of Figure 3 panel (b), which plots the difference in standard deviation of female and male $tfpr$'s scaled by the country level counterpart, for the four specifications. For panel (d) the sample is restricted to countries that have more than 50 observations; the pattern also holds when restricting to more than 100 observations. OLS regression estimates and standard error of log GDP per capita on the female estimate for panel (a) is −0.66 (0.338), for panel (b) is −0.60 (0.443), for panel (c) is −0.763 (0.620), and for panel (d) is −.19 (0.157).

Figure 8: Sensitivity of Female estimate and Development

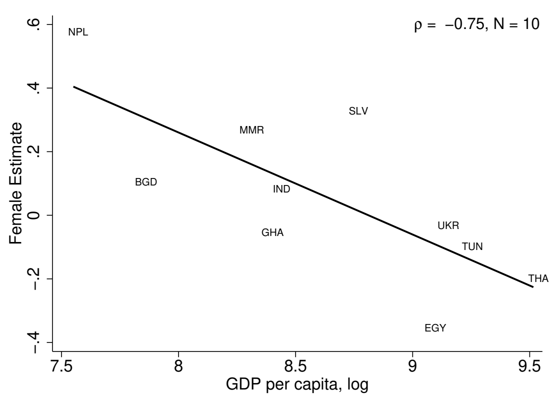
(a) Size, $n < 50$



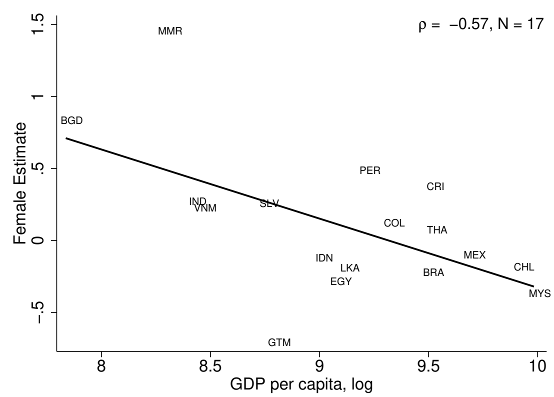
(b) Bus. Ownership Categories



(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. For panel (d) the sample is restricted to countries that have more than 50 observations; the pattern also holds when restricting to more than 100 observations. OLS regression estimates and standard error of log GDP per capita on the female estimate for panel (a) is −1.01 (0.340), for panel (b) −2.07 (0.996), for panel (c) is −1.74 (0.430), and for panel (d) −0.68 (0.264).

ager. 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 participation in decision making and management of the establishment. The descriptive statistics reported in Table 1 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 females face higher (lower) distortions to production in South Asia (Eastern Europe), and the relationship between misallocation and 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 assess whether the main results are due to aggregating across manufacturing sub-industries using common factor shares.[41] Since the value of α_s is not overly critical, I keep its value unchanged. The descriptive statistics in Table 1 mostly hold for South America and South Asia, specifically with respect to the distortion on capital. That male establishments face higher distortions in Eastern Europe cannot be verified due to insufficient observations. In Figures 7 and 8 panel (d), I focus on countries that have at least 50 observations to expand the number of countries (the main point holds when restricting to more than 100 observations as well but has fewer countries). In Figure 7 the slope is negative but is highly sensitive to two outliers (Myanmar and Chile), and absent of this the slope is virtually flat, suggestive that misallocation across gender is not related to development for this sub-industry. Concerning the gender estimate, Figure 8 shows a clear strong and negative relationship, supporting the view that female establishments face higher distortions in poorer countries. While I have not reported all specifications from Section 4, the broad patterns consistently hold for South America, South Asia and those related to development. Finally, the implications from the reallocation policy

[41] 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), but the sample of countries is fewer than 10.

in Section 4.4 also holds for this sub-industry, and its impacts on female sales shares remains negatively related to development (see Figure A.2).

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. Females face higher distortions on production, primarily on capital, in many parts of the world, notably in South America and South Asia, and males face higher distortions in Eastern Europe. Relevant is that the higher distortions females face are evident among the subset of women, and plausibly a more talented group of women, who have overcome various gender-specific entry barriers into entrepreneurship. Regression estimates show that differential distortions across gender are associated with a form of bias, against females in South Asia and against males in Eastern Europe. Importantly, this bias against females and misallocation is negatively related to economic development. Removing distortions across gender implies proportionally large increases in female market shares and TFP gains attributed to females, particularly in places where females face high distortions on production. Taken together, these results suggest that differential distortions across gender is an important factor for understanding misallocation 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.

Bento, P. and Restuccia, D. (2017). Misallocation, Establishment Size, and Productivity. *American Economic Journal: Macroeconomics*, 9(3):267–303.

Besley, T. and Mueller, H. (2018). Predation, protection, and productivity: A firm-level perspective. *American Economic Journal: Macroeconomics*, 10(2):184–221.

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.

Buera, F., Kaboski, J., and Shin, Y. (2011). Finance and Development: A Tale of Two Sectors. *American Economic Review*, 101(5):1964–2002.

Campa, P. and Serafinelli, M. (2019). Politico-economic regimes and attitudes: Female workers under state socialism. *The Review of Economics and Statistics*, 101(2):233–248.

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. (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.*

Dusha, E. (2015). Intermediated corruption. *International Economic Review*, 56(3):997–1017.

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. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics*, 7(1):63–88.

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

Johnson, S., McMillan, J., and Woodruff, C. (2002). Property Rights and Finance. *American Economic Review*, 92(5):1335–1356.

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.

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.

López, J. J. (2017). A quantitative theory of tax evasion. *Journal of Macroeconomics*, 53:107–126.

Lopez-Martin, B. (2019). Informal sector misallocation. *Macroeconomic Dynamics*, 23(8):3065–3098.

Midrigan, V. and Xu, D. Y. (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review*, 104(2):422–458.

Morazzoni, M. and Sy, A. (2018). Female Entrepreneurship, Financial Frictions and Capital Misallocation in the US. BSE Working Paper 1299.

Oguzoglu, U. and Ranasinghe, A. (2017). Crime and Establishment Size: Evidence from South America. *The B.E. Journal of Economic Analysis & Policy*, 17(4).

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

Poschke, M. (2018). The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology. *American Economic Journal: Macroeconomics*, 10(3):1–41.

Ranasinghe, A. (2017). Property Rights, Extortion and the Misallocation of Talent. *European Economic Review*, 98:86–110.

Ranasinghe, A. and Restuccia, D. (2018). Financial Frictions and the Rule of Law. *Journal of Development Economics*, 134:248–271.

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

Restuccia, D. and Rogerson, R. (2017). The causes and costs of misallocation. *Journal of Economic Perspectives*, 31(3):151–74.

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.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from brazil. *American Economic Review*, 108(8):2015–47.

A Appendix

A.1 Tables

Table A.1: Physical productivity ($tfpq$): mean and dispersion

	South America			South Asia			Eastern Europe		
	All	Male	Female	All	Male	Female	All	Male	Female
Mean									
$\ln(tfpq_{si}^{j})$	-1.66	-1.61	-1.97	0.98	0.94	1.39	-0.17	-0.11	-0.48
Dispersion									
Std. Dev.	2.39	2.31	2.81	1.52	1.51	1.56	1.50	1.50	1.47
75-25	3.37	3.31	3.82	2.07	2.04	2.20	1.89	1.87	2.19
90-10	6.18	5.94	7.22	3.93	3.85	4.34	3.56	3.57	3.25
Observations	3206	2815	391	4217	3928	289	1280	1084	196

Notes: Dispersion statistics are scaled by first-best productivity, that is, $tfpq_{si}^{j}(M_s^j)^{\frac{1}{\sigma-1}}/TFP_s^{fb}$. '75-25' is difference between the 75th and 25th percentile, and similarly for '90-10'. See text for details.

Table A.2: Revenue Productivity ($tfpr$): mean and dispersion

	South America			South Asia			Eastern Europe		
	All	Male	Female	All	Male	Female	All	Male	Female
Mean									
$\ln(tfpr_{si}^{j})$	-0.07	-0.07	-0.11	-0.10	-0.12	0.06	0.18	0.21	0.01
Dispersion									
Std. Dev.	1.00	0.95	1.24	0.91	0.91	0.92	0.86	0.87	0.75
75-25	1.10	1.04	1.57	1.20	1.20	1.22	1.05	1.13	0.75
90-10	2.52	2.35	3.73	2.27	2.26	2.44	2.22	2.25	1.91
Observations	3206	2815	391	4217	3928	289	1280	1084	196

Notes: Dispersion statistics are based on $\ln(tfpr)/TPFR_c$. '75-25' is difference between the 75th and 25th percentile, and similarly for '90-10'. See text for details.

Table A.3: 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	1331	1160	171	2107	1800	307	1168	1102	66	3632	2755	877
% of estabs.		0.88	0.12		0.85	0.15		0.95	0.05		0.76	0.24
% of sales		0.90	0.10		0.92	0.08		0.92	0.08		0.81	0.19
Employees: $\ln(n)$												
mean	2.97	3.02	2.64	2.92	2.99	2.57	2.97	2.93	3.60	3.32	3.40	3.10
std. dev.	1.14	1.15	0.89	1.13	1.14	0.90	1.08	1.08	0.87	1.08	1.08	1.06
Experience:												
mean	17.76	18.07	16.22	20.64	21.02	18.62	23.27	23.28	23.81	16.36	16.17	17.07
std. dev.	10.23	10.06	10.07	11.15	11.26	9.46	11.36	11.46	8.23	8.70	8.69	8.50
Distortions:												
output $\bar{\tau}^j$	0.62	0.63	0.40	0.28	0.27	0.09	0.53	0.52	0.55	0.48	0.50	0.40
capital $\bar{\kappa}^j$	-0.03	-0.08	1.89	0.60	0.57	1.52	-0.15	-0.14	-0.15	-0.01	0.03	0.41
$TFPR^j$	3.23	3.39	3.50	1.95	1.95	1.87	1.97	1.98	1.77	2.31	2.35	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. The sample is restricted to establishments in the manufacturing sector that report values for capital and input costs; sales are value-added. Employees are permanent full-time employees and experience is the number of years the top manager has worked in the establishment's sector.

Table A.4: Gender estimates – Alternate controls

	South America	South Asia	Eastern Europe
A: Only industry and country controls			
Female	0.03	0.20***	-0.14
	(0.13)	(0.07)	(0.10)
N	3206	4217	1279
R^2	0.247	0.102	0.101
B: Add female industry concentration (categories)			
Female	0.13	0.19***	-0.25*
	(0.12)	(0.07)	(0.14)
N	3200	4189	1141
R^2	0.246	0.138	0.148
C: Add female industry concentration (%)			
Female	0.14	0.19***	-0.30**
	(0.14)	(0.07)	(0.14)
N	3200	4189	1141
R^2	0.166	0.115	0.101

Notes: The dependent variable is $\ln\left(tfpr_{si}^{j}\right)$. Estimates include sub-industry, country and time fixed effects. Standard errors are in parenthesis and ***, **, * denote significance at the 1, 5 and 10 percent level.

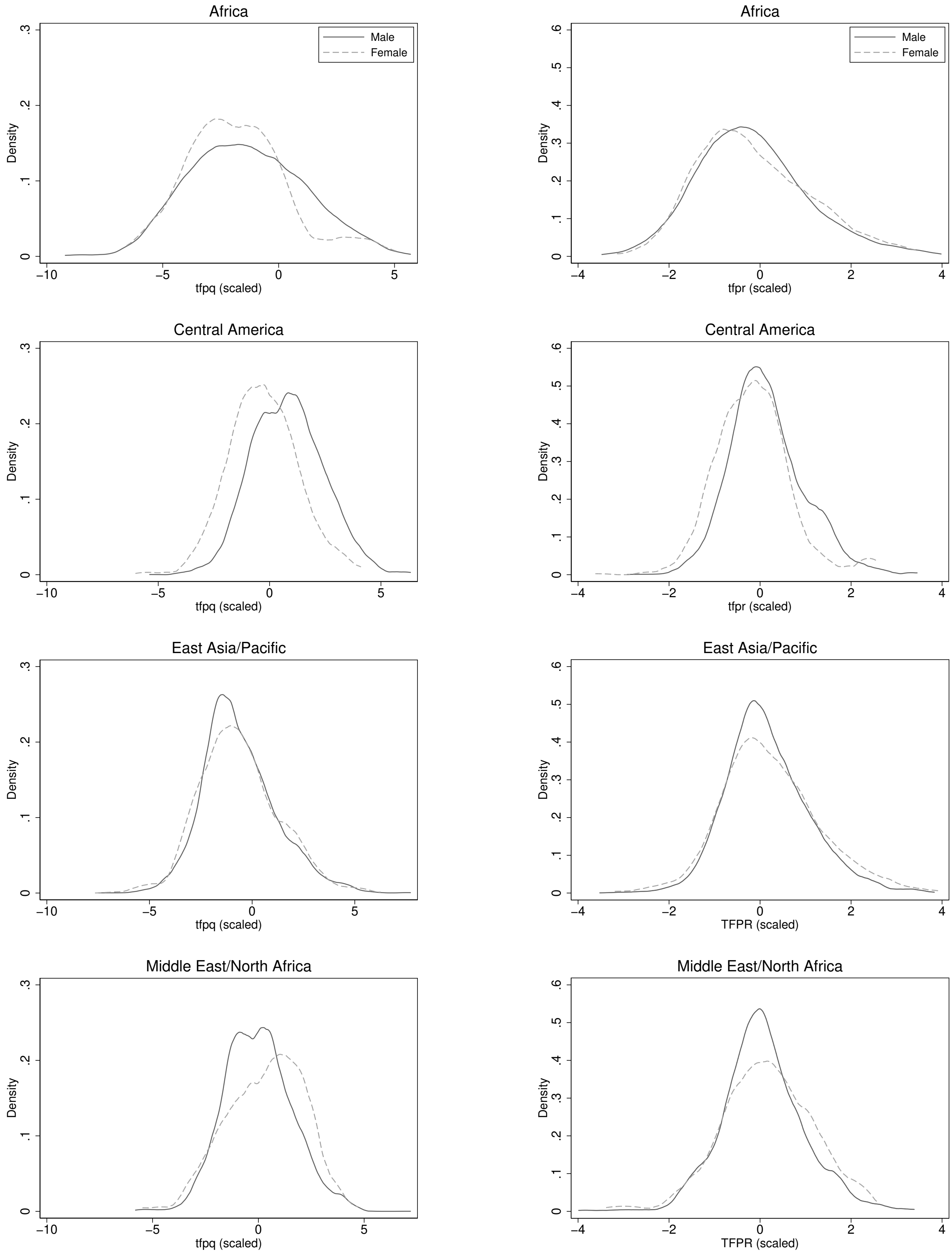
Table A.4 presents the female estimate for three specifications based off Table 2. The estimates in Panel A follows the same specification as in Table 2 but only controls for industry, country and time fixed effects (i.e., controls for size, city, experience and register are not included); Panel B has *all* the controls as Table 2 and adds female industry concentration as an additional control, where the latter is defined as the share of female establishments in an industry-country (and is a categorical variable of 7 groups); Panel C has the same specification as Table 2 but replaces the 2 digit industry control with female industry concentration by country (as continuous variable). While not reported, in Panel B and C the interaction of female with size, city, experience and register are similar to the estimates in Table 2. Panel B and C show that adding female industry concentration does not change the sign or significance of the female estimate (though the estimates are higher, especially for Eastern Europe).

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	-0.07	-0.06	-0.05	-0.02	-0.28**	-0.26**	-0.18	-0.25**
	(0.08)	(0.08)	(0.08)	(0.08)	(0.11)	(0.13)	(0.15)	(0.12)
Size	0.11***	0.11***	0.25***	0.23***	0.16***	0.14**	-0.03	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.07)	(0.07)
City	0.01	0.04	0.01	0.01	0.04	0.04	0.09**	0.09*
	(0.05)	(0.05)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)	(0.05)
Experience	-0.07***	-0.10***	-0.02	-0.00	-0.04	-0.04	0.03	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.05)	(0.06)
Register	0.16**	0.15**	-0.10	-0.05	0.12	0.19	0.08	0.05
	(0.07)	(0.07)	(0.08)	(0.08)	(0.12)	(0.13)	(0.15)	(0.15)
Female × Size		-0.03		0.11		0.15		0.11
		(0.11)		(0.10)		(0.21)		(0.17)
Female × City		-0.14		0.01		0.03		0.11
		(0.09)		(0.04)		(0.12)		(0.11)
Female × Experience		0.13**		-0.10**		0.00		0.19
		(0.06)		(0.05)		(0.08)		(0.20)
Female × Register		0.08		-0.42**		-0.47		0.76**
		(0.16)		(0.19)		(0.32)		(0.37)
N	3605	3605	2107	2107	1331	1331	1168	1168
R^2	0.091	0.097	0.209	0.218	0.383	0.385	0.178	0.183

Notes: The dependent variable is $\ln\left(tfpr_{si}^{j}\right)$. Size is an indicator whether an establishment is small, medium or large; 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; 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. Variables interacted with 'Female' are differenced by its mean (male and female population). Estimates include country, industry fixed effects and time fixed effects. Standard errors are in parenthesis and ***, **, * denote significance at the 1, 5 and 10 percent level.

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 $ln\left(tfpq^j_{si} M_s^{\frac{1}{\sigma-1}}/TFP_s^{fb}\right)$ and $ln(tfpr^j_{si}/TFPR_s)$, 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.6: Gender estimates and obstacles to doing business

	South America (1)	South America (2)	South Asia (3)	South Asia (4)	Eastern Europe (5)	Eastern Europe (6)
Female	0.02	0.03	0.18***	0.14**	-0.16	-0.17*
	(0.12)	(0.10)	(0.07)	(0.07)	(0.10)	(0.10)
Crime	0.10	0.06	0.15*	0.14	-0.33**	-0.28*
	(0.07)	(0.07)	(0.08)	(0.09)	(0.15)	(0.17)
Informal	0.02	0.01	0.04	0.03	0.01	0.02
	(0.08)	(0.07)	(0.05)	(0.06)	(0.10)	(0.11)
Workforce	0.08	0.11*	-0.01	-0.02	-0.00	-0.00
	(0.06)	(0.06)	(0.06)	(0.06)	(0.10)	(0.12)
Finance	-0.22***	-0.26***	-0.03	-0.05	-0.10	-0.04
	(0.08)	(0.08)	(0.05)	(0.05)	(0.11)	(0.10)
Tax Related	-0.13	-0.05	-0.08	-0.06	-0.47***	-0.55***
	(0.08)	(0.09)	(0.06)	(0.07)	(0.12)	(0.13)
Corruption	-0.14**	-0.14**	0.05	0.02	0.02	-0.01
	(0.06)	(0.06)	(0.05)	(0.05)	(0.08)	(0.09)
Female × Crime		0.15		0.09		-0.14
		(0.21)		(0.25)		(0.42)
Female × Informal		-0.08		0.12		-0.02
		(0.22)		(0.17)		(0.20)
Female × Workforce		-0.12		0.07		-0.02
		(0.22)		(0.17)		(0.20)
Female × Finance		0.33		0.19		-0.18
		(0.21)		(0.15)		(0.27)
Female × Tax Related		-0.62**		-0.18		0.32
		(0.24)		(0.23)		(0.26)
Female × Corruption		-0.09		0.31**		0.04
		(0.20)		(0.15)		(0.19)
N	3206	3206	4217	4217	1156	1156
R^2	0.274	0.285	0.114	0.117	0.152	0.156

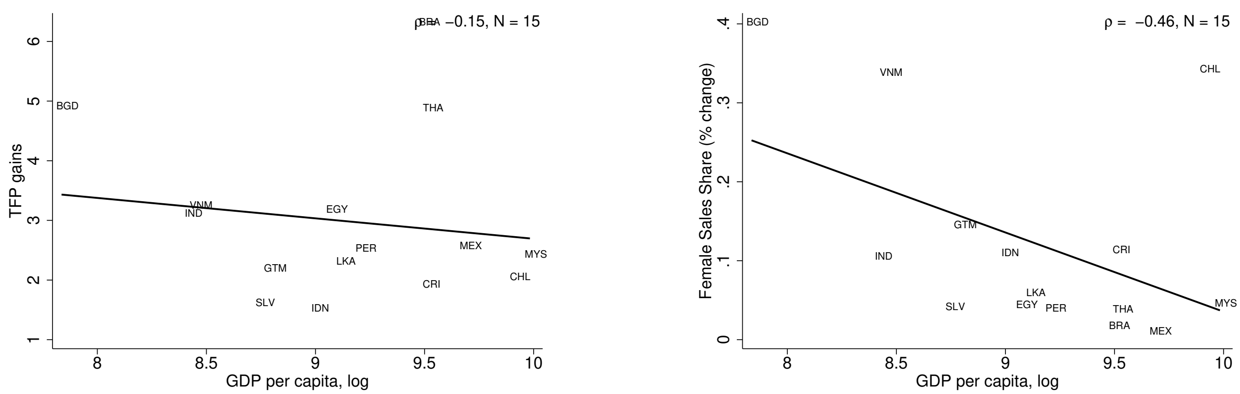
Notes: The dependent variable is $ln(tfpr_{si}^{j})$. Crime, informal, workforce, finance, tax related and corruption are indicators for whether an establishment reports these as major or severe obstacles to business operation. Also included are controls for manager experience and city of operation. Sub-industry, country and survey year fixed effects are included. Standard errors are in parenthesis and ***, **, * denote significance at the 1, 5 and 10 percent level.

Table A.7: Gender estimates and obstacles to doing business: Other continents

	East Asia/Pacific		Central America		Africa		Middle East/N. Africa	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.07	-0.05	-0.11	-0.11	-0.31***	-0.27**	-0.18	-0.09
	(0.08)	(0.09)	(0.09)	(0.08)	(0.11)	(0.11)	(0.15)	(0.12)
Crime	-0.02	0.14	-0.05	-0.00	-0.07	-0.06	-0.07	-0.11
	(0.15)	(0.20)	(0.07)	(0.07)	(0.13)	(0.13)	(0.19)	(0.21)
Informal	-0.07	0.02	-0.08	-0.16**	0.02	0.01	-0.14	-0.13
	(0.09)	(0.10)	(0.07)	(0.07)	(0.08)	(0.09)	(0.12)	(0.13)
Workforce	-0.16	-0.10	0.06	0.10*	0.06	0.12	-0.35***	-0.39***
	(0.12)	(0.12)	(0.05)	(0.06)	(0.12)	(0.13)	(0.12)	(0.13)
Finance	0.19**	0.19*	-0.12*	-0.14*	-0.02	-0.04	0.11	0.11
	(0.09)	(0.11)	(0.07)	(0.07)	(0.09)	(0.09)	(0.15)	(0.15)
Tax Related	-0.11	-0.18	-0.23***	-0.18**	0.10	0.12	-0.04	-0.02
	(0.14)	(0.15)	(0.08)	(0.08)	(0.11)	(0.12)	(0.15)	(0.15)
Corruption	-0.18	-0.13	0.11	0.10	-0.05	0.01	-0.11	-0.15
	(0.13)	(0.16)	(0.07)	(0.07)	(0.09)	(0.09)	(0.13)	(0.13)
Female × Crime		-0.33		-0.25		-0.03		-0.23
		(0.28)		(0.16)		(0.24)		(0.55)
Female × Informal		-0.33**		0.53***		-0.04		-0.43
		(0.16)		(0.19)		(0.21)		(0.39)
Female × Workforce		-0.21		-0.18		-0.35		0.36
		(0.29)		(0.15)		(0.23)		(0.31)
Female × Finance		0.02		0.08		0.10		0.25
		(0.18)		(0.16)		(0.23)		(0.30)
Female × Tax Related		0.34		-0.37*		-0.19		-0.50
		(0.33)		(0.20)		(0.29)		(0.45)
Female × Corruption		-0.18		0.10		-0.30		0.97***
		(0.28)		(0.15)		(0.22)		(0.27)
N	3605	3605	2107	2107	1330	1330	1168	1168
R^2	0.090	0.095	0.191	0.208	0.377	0.381	0.206	0.217

Notes: The dependent variable is $ln(tfpr_{si}^{j})$. Crime, informal, workforce, finance, tax related and corruption are indicators for whether an establishment reports these as major or severe obstacles to business operation. Also included are controls for manager experience and city of operation. Sub-industry, country and survey year fixed effects are included. Standard errors are in parenthesis and ***, **, * denote significance at the 1, 5 and 10 percent level.

Figure A.2: TFP gains and Female Sales Shares with Reallocation for ISIC 15



(a)

(b)

Notes: The figure shows the factor change in TFP and impact on female sales shares when $tfpr$ is uniform across establishments within a country for the 'food products and beverages' sub-industry. The correlations remain when outlier countries are removed. OLS estimates for panel (a) and (b) are −0.065 (0.138) and −2.146 (1.608) .

Table A.8: First-best policy with uniform distortions

	Africa		Central America		Middle East/North Africa		East Asia/Pacific	
	Data	Reallocation	Data	Reallocation	Data	Reallocation	Data	Reallocation
Female shares:								
Sales θ^f	0.105	0.124	0.084	0.106	0.073	0.048	0.250	0.256
Capital θ_k^f	0.139	0.124	0.052	0.106	0.107	0.048	0.197	0.256
Labour θ_n^f	0.104	0.124	0.083	0.106	0.066	0.048	0.238	0.256
Female shares (top ten %):								
Sales	0.137	0.146	0.066	0.085	0.077	0.051	0.179	0.202
Establishments	0.082	0.089	0.061	0.105	0.113	0.056	0.221	0.247
Average change in sales:								
Female Estabs.	–	0.424	–	0.081	–	0.112	–	0.045
Male Estabs.	–	0.031	–	0.017	–	0.013	–	0.015
TFP gains	–	3.09	–	2.93	–	3.73	–	3.42
due to females (share) $\approx$	–	0.12	–	0.11	–	0.05	–	0.26

Notes: Notes: 'Data' refers to weighted statistics based on the WBES 2006-17 (refer to text for details.) 'Reallocation' are statistics based on a policy that equalizes $tfpr$ across establishments in a country. Female shares (top 10%) refer to the proportion of females among the top ten percent of establishments 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.

A.2 List of Countries

Table A.9: List of countries

Country	ISO Code	Observations Total	Male	Female
Africa:				
Botswana (2010)	BWA	45	35	10
BurkinaFaso (2009)	BFA	25	19	6
Ethiopia (2015)	ETH	175	161	14
Ghana (2013)	GHA	160	139	21
Kenya (2013)	KEN	176	167	9
Lesotho (2016)	LSO	18	7	11
Liberia (2017)	LBR	47	41	6
Madagascar (2013)	MDG	95	76	19
Malawi (2014)	MWI	24	19	5
Mauritius (2009)	MUS	42	36	6
Nigeria (2014)	NGA	72	65	7
Senegal (2014)	SEN	80	74	6
Sierra Leone (2017)	SLE	53	47	6
Uganda (2013)	UGA	42	34	8
Zambia (2013)	ZMB	128	112	16
Zimbabwe (2016)	ZWE	149	128	21
Central America:				
Bahamas (2010)	BHS	28	20	8
Barbados (2010)	BRB	58	47	11
Belize (2010)	BLZ	66	58	8
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
Honduras (2016)	HND	49	44	5
Jamaica (2010)	JAM	90	76	14
Mexico (2010)	MEX	897	800	97
Nicaragua (2016)	NIC	69	52	17
Panama (2010)	PAN	16	11	5
St. Kitts & Nevis (2010)	KNA	17	12	5
St. Lucia (2010)	LCA	48	39	9
St. Vincent & Grenadines (2010)	VCT	44	31	13
Trinidad & Tobago (2010)	TTO	96	88	8
Cambodia (2016)	KHM	94	46	48
East Asia:				
China (2012)	CHN	1207	1106	101
Indonesia (2009)	IDN	515	405	110
Lao PDR (2016)	LAO	93	73	20
Malaysia (2015)	MYS	176	115	61
Mongolia (2013)	MNG	41	23	18
Myanmar (2016)	MMR	193	150	43
Philippines (2009)	PHL	323	249	74
Thailand (2016)	THA	398	119	279
Timor-Leste (2015)	TLS	52	43	9
Vietnam (2009)	VNM	540	426	114

Table A.9: List of countries

Country	ISO Code	Observations Total	Male	Female
Eastern Europe:				
Belarus (2013)	BLR	44	35	9
Bosnia & Herzegovina (2013)	BIH	60	52	8
Bulgaria (2013)	BGR	52	43	9
Croatia (2013)	HRV	75	63	12
Estonia (2013)	EST	31	21	10
Macedonia (2013)	MKD	70	57	13
Lithuania (2013)	LTU	39	34	5
Romania (2013)	ROU	67	52	15
Russia (2012)	RUS	278	226	52
Serbia (2013)	SRB	42	35	7
Slovenia (2013)	SVN	56	48	8
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
Morocco (2013)	MAR	55	49	6
Tunisia (2013)	TUN	222	208	14
West Bank & Gaza (2013)	PSE	96	91	5
South America:				
Argentina (2010)	ARG	455	428	27
Bolivia (2017)	BOL	37	31	6
Brazil (2009)	BRA	931	794	137
Chile (2010)	CHL	581	530	51
Colombia (2010)	COL	427	349	78
Ecuador (2017)	ECU	72	66	6
Guyana (2010)	GUY	57	49	8
Paraguay (2017)	PRY	38	32	6
Peru (2010)	PER	494	435	59
Suriname (2010)	SUR	75	70	5
Venezuela (2010)	VEN	39	31	8
South Asia:				
Bangladesh (2013)	BGD	957	908	49
Bhutan (2015)	BTN	54	49	5
India (2014)	IND	2769	2583	186
Nepal (2013)	NPL	197	181	16
SriLanka (2011)	LKA	240	207	33