Policy Research Working Paper 10956

Fading Away Informality by Development

M. Nazım Tamkoç

WORLD BANK GROUP
Development Economics
Global Indicators Group
October 2024

Policy Research Working Paper 10956

Abstract

This paper focuses on the role of development in informality through higher wages and expanded production possibilities. First, it uses informal, plant-level survey data across countries to document that on average, richer countries have smaller informal, unregistered plants in terms of employment. This negative relationship holds even after controlling for plant-level characteristics. Then, a dynamic general equilibrium model with incomplete tax enforcement is developed such that formal and informal plants coexist in equilibrium. The model allows for two groups of agents operating in the informal sector: those with lower abilities than workers, and those with abilities falling between workers and formal managers. In the model, when plants become more productive, some agents operating informally choose to be workers and some of them transition into formality due to higher wages and better production possibilities, which decreases the mean size of informal plants. The quantitative results indicate that around 30 percent of the increase in aggregate output due to higher productivity is associated with a roughly one-quarter decline in the mean size of informal plants.

This paper is a product of the Global Indicators Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The author may be contacted at mtamkoc@worldbank.org.

Fading Away Informality by Development*

M. Nazım Tamkoç†

October 2024

For the latest version, please click here

Keywords: Informality, Economic Development, Plant Size, Productivity
JEL Classification: E23, J24, L25, O41, O33

*I would like to thank Juhee Bae, Filip Jolevski, Norman Loayza, Hibret Maemir, Jorge Rodriguez Meza, Furkan Sarikaya and Gustavo Ventura for their comments and suggestions. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

†The World Bank, Enterprise Analysis Unit **E-Mail**: mtamkoc@worldbank.org

1 Introduction

Informality, economic activity outside of formal regulations and institutional frameworks, is often seen as a common feature of developing countries.[1] The literature has focused on reducing informality through strengthening enforcement of policies and lowering the costs of operating formally, such as tax rates and entry fees. However, developed countries don't have oversight of all economic activities, and their tax rates can be as high as in developing countries.[2] This paper argues that higher level of informality in poorer countries is not only the result of burdensome regulations or inadequate enforcement, but also a byproduct of underdevelopment.

In richer countries with better production technologies and economic environments, some workers prefer formal wage employment over operating an informal business, as wage rates are higher. Likewise, some managers opt for formal firm registration and operation because the income potential in richer markets makes formality more profitable despite the costs of taxes and regulations.

In this paper, I focus on the incidence of informality at the extensive margin, referring to unregistered plants. I start by documenting empirical facts related with informality in the developing world. Using plant-level data from the World Bank Informal Sector Enterprise Surveys (ISES), I show that informal plants in richer countries tend to employ fewer workers on average than informal plants in poorer countries. In other words, the informal mean size decreases with the level of development. For example, informal plants in Ghana have 2.3 workers on average, whereas in Peru, informal plants average only 1.4 workers. This negative relationship between informal plant size and GDP per capita persists even after controlling for plant characteristics such as managers' education, experience, and gender, sector of activity, and plant age. Moreover, while some informal plants achieve productivity levels, measured by value-added per worker, comparable to their formal counterparts, most operate at very low levels.

To quantify the impact of development on informality, I develop a dynamic-general equilibrium model of occupational choice where agents are optimally allocated as workers or managers according to their abilities in the spirit of Lucas (1978)'s span-of-control model based on Poschke (2018). In the model, while workers' earnings are proportional to their abilities, managerial earnings are increasing in returns to their abilities. I introduce infor-

[1] See De Soto (1989), and La Porta & Shleifer (2014), Ulyssea (2020) and Ohnsorge & Yu (2022) for a survey of informality literature.

[2] See Friedman *et al.* (2000) for the negative relationship between tax rates and informality.

mality by allowing managers to operate informally to avoid taxation due to incomplete enforcement. They risk being caught if they use capital over a certain threshold, similar to De Soto (1989) and Leal (2014). In equilibrium, agents at the lower end of the ability distribution operate as informal managers. Agents with abilities higher than those of informal managers, but not high enough to run plants, become workers. If enforcement limits or taxes are not excessively low, informality may also exist among agents with abilities exceeding those of workers. Of these higher-ability agents, the relatively less able operate informally, while the more able run formal plants. Informal managers with abilities lower than workers are referred to as operating in subsistence informality. In the case that enforcement limits or taxes are sufficiently low, all agents with abilities higher than workers operate in the formal sector.

I calibrate key model parameters to match plant size distribution, including informal plants in Ghana. The model successfully generates the observed mean plant size differences between formal and informal sectors. In the benchmark economy, while 90.6% of plants are informal, they account for 34.2% of aggregate output. Moreover, 71.1% of informal plants operate at the subsistence informality level. Hence, the average informal manager earns less than the average worker.

Motivated by better business environments and higher years of schooling in richer countries, development is introduced into the model such that all managers' productivity increases. As a result, relatively higher-ability informal plant managers start operating formally since, once more productive, they demand more labor and capital. In contrast, relatively lower-ability informal plant managers switch to become workers in the new equilibrium, enjoying higher wages. Therefore, while informality decreases with development, the share of subsistence informality increases among informal plants.

Quantitatively when the aggregate output increases by about 29.5% due to improvements in managers' productivity, the overall average plant size increases by 30% whereas the mean informal plant size declines by about 26.1%. When the model is disciplined to account for differences in informal mean plant size across countries, it can explain 32.8% of income per capita differences on average. Moreover, when only the formal sector benefits from proportional and skill-biased improvements in the managers' productivity, the share of informal plants shrinks, and the overall mean size grows faster compared to the case where all managers benefit from development. The paper concludes by comparing the gains from formalization policies, such as better enforcement and lower taxes: while formalization policies have the potential to decrease informality, their gains are limited compared to the gains from higher productivity associated with improving the business environment and improving education levels.

Background The informality literature grew extensively following De Soto (1989). First, La Porta & Shleifer (2008) and La Porta & Shleifer (2014) are among the earliest studies that utilizes ISES to document differences between formal and informal plants. Amin & Okou (2020) extends their comparison to a broader set of countries. I contribute to this literature by systematically showing that informal plants tend to be larger in poorer countries.

Second, Rauch (1991), Loayza (1996), Amaral & Quintin (2006), D'Erasmo & Boedo (2012), Leal (2014), Meghir *et al.* (2015), Ulyssea (2018) and Franjo *et al.* (2022) develop environments where informality exists due to the minimum wages, incomplete enforcement and frictions in accessing finance. I extend the model of Poschke (2018) by allowing informality to exist endogenously due to the incomplete enforcement of taxes. In the model, agents sort into being workers, informal managers, and formal managers. My contribution is providing an environment where agents with abilities lower than workers (subsistence informality) as well as those with higher abilities become informal managers.

Third, the above literature focuses on the formalization policies such as reducing the cost of being formal or taxes, motivated by Johnson *et al.* (1998), Fajnzylber *et al.* (2011), De Mel *et al.* (2013), Bruhn & McKenzie (2014), De Andrade *et al.* (2016) and Rocha *et al.* (2018) among many others; and increasing enforcement as in Kuehn (2014), Orsi *et al.* (2014), Leal (2014) and Ulyssea (2018). However, in this paper, I quantify the role of development in reducing informality. Loayza (2016) studies how growth rates and migration affect informality in a growth-accounting setup, and projects labor informality over two decades. This paper differs by developing a span-of-control model with heterogeneous production units and disciplining the model using a plant size distribution including informality.

Lastly, this paper contributes to the literature studying plant-level productivity, size, and aggregate productivity, advanced by Hsieh & Klenow (2009), Bartelsman *et al.* (2013) and Hsieh & Klenow (2014). Bento & Restuccia (2017) and Bento & Restuccia (2021) document that richer countries have larger plants in terms of employment on average. Restuccia & Rogerson (2008), Guner *et al.* (2008), Garcia-Santana & Pijoan-Mas (2014) and Gourio & Roys (2014) have focused on the role of size-dependent distortions as one of the main reasons for smaller plants in poorer countries. Tamkoç & Ventura (2024) study time taxes – rules and regulations that distort managers' productive time – as a determinant of smaller plants in poorer countries, on average. In this paper, I study the prevalence of larger informal plants in underdeveloped countries as one reason for the smaller overall mean plant size in these countries.

2 Data

The main data source of this paper is the World Bank's ISES which consists of face-to-face interviews with owners or managers of informal plants.[3] It uses an area-based sampling methodology that generates a probabilistic sample of informal firms where interviewers enumerate all plants in a randomly selected uniform blocks in a given region or city to determine registration status. If a plant is not registered in the local registration institution, it is considered informal. Within selected informal plants are randomly selected in real time for a deeper interview. Various topics such as workforce, sales, sector and background of managers are covered in the questionnaire.[4] ISES includes 25,995 plant-level observations from 78 regions/cities in 26 countries since 2008.[5] The PPP adjusted real GDP per capita, RGDP, and employment data come from World Development Indicators (WDI). The sample covers countries at different levels of development: the richest country in the sample is Argentina whereas Democratic Republic of the Congo has the lowest RGDP per capita compared to other countries in the sample.

The number of workers in an informal plant is the total number of paid and unpaid workers. Figure 1 presents the main motivating fact of the paper: richer countries tend to have smaller informal plants on average. In the figure, each dot represents a country. The y-axis and x-axis are the average number of workers in informal plants and the RGDP per capita in each country respectively. For example, the average informal plant in Argentina has 1.6 workers, while in Ghana and Nepal, the averages are 2.3 and 4.1 workers per informal plant, respectively.

The solid line in Figure 1 is the simple linear regression line where the log-informal mean size is regressed on the log-RGDP per capita where observations are weighted according to their employment size. Despite the small sample size, the elasticity of the informal mean size with respect to RGDP per capita is negative and statistically significant. It implies that doubling RGDP per capita is associated with 29.3% decline in informal mean size.

The following regression equation is estimated to test whether the negative relationship between the informal mean size and the RGDP per capita persists after controlling for plant

[3] For plant-level datasets and further information about the methodology, please visit www.enterprisesurveys.org

[4] After 2015, ISES uses an adaptive cluster sampling methodology which enables the computation of the probability of selection of a plant within a block, thus allowing the use of sampling weights to make inferences to the population of informal businesses with the region/city. See Aga *et al.* (2023) and Aberra *et al.* (2022).

[5] Please see Appendix B for the list of countries and the number of observations.

characteristics.

$$\log\left(\text{Size}_{i,c}\right) = \beta_0 + \beta_1 \log\left(\text{RGDP}_c\right) + X_{i,c} + \epsilon_{i,c} \tag{1}$$

where $\text{Size}_{i,c}$ is the number of workers in plant i in country c and RGDP_c is the RGDP per capita in country c. $X_{i,c}$ includes plant characteristic variables in plant i in country c. It consists of a manufacturing dummy; an education of a manager dummy which is 1 if the highest level of a manager is at least secondary school; experience of a manager in the sector (in years); a female dummy which takes value 1 if the manager is female; and the age of the plant (in years, as of the date of interview).

Table 1 presents the estimated coefficients of the Equation (1). The elasticity of the informal mean size with respect to RGDP per capita is -0.16 and significant even after controlling for all plant characteristics. In addition, the estimation results indicate that informal plants operating in the manufacturing sector significantly employ more workers than those operating in the services sector on average. Moreover, the experience and education of managers is positively correlated with the size of informal plants. On the other hand, plants managed by a female manager are significantly smaller compared to plants managed by a male.

Next, I compare informal plants vis-a-vis formal plants in Ghana.[6] World Bank Enterprise Surveys (WBES) conducted two other surveys in Ghana in 2013 along with ISES: Micro-WBES and WBES. Only formal (registered) plants are covered in both surveys. Micro-WBES interviews plants with fewer than 5 workers, while WBES interviews plants with at least 5 workers. I refer to plants in both Micro-WBES and WBES as formal plants.

Table 2 presents the comparison of informal and formal plants in Ghana. The mean size of informal plants is smaller than formal plants: there are 2.3 workers in informal plants whereas 3 and 34.7 workers are involved in production in the Micro-WBES and WBES respectively.[7] Moreover, informal managers have less experience in the sector they operate in compared to formal managers. While 42.8% of informal managers have at least a secondary school degree, around 65% of workers in formal plants have at least a secondary school degree. To sum up, on average, informal plants are smaller and younger compared to formal plants; their managers have fewer experience and less education; the fraction of female managers is higher among informal plants.

[6] Similar analysis is done by La Porta & Shleifer (2014) across a few countries using the same dataset, and Meghir *et al.* (2015) and Ulyssea (2018) in Brazil.

[7] As Section 4 discusses in detail, the Ghana Statistical Service's Integrated Business Establishment Survey reports the mean size of informal plants as 2.3 workers and the mean size of formal plants as 33.8. Similar mean sizes of the World Bank's plant level survey data sets with that of the Ghana Statistical Service's census build confidence in the usage of ISES and WBES to compare informal and formal plants in Ghana.

Finally, I focus on productivity differences between informal and formal plants in Ghana.[8] Figure 2 plots the distribution of Log-VA per worker in ISES, Micro-WBES and WBES. There are three observations related with the productivity differences between informal and formal plants. First, as Table 2 also shows, informal plants are less productive compared to formal plants. Secondly, while most productive informal plants are similar to most unproductive formal plants in terms of productivity levels, there are some informal plants that operate at very low levels. Third, informal plants have long left-tail productivity whereas formal plants have long right-tail productivity.

3 Model

The model is based on Poschke (2018). The innovation in this paper is that managers can operate in the informal sector to avoid distortions.

3.1 Environment

There is a household with a continuum of household members. Each household member is born with a unit of efficiency, z and a unit of time supplied inelastically. The unit of efficiency is distributed according to a cdf $G(z)$ over $[0, \bar{z}]$ and is referred as ability in this paper. The household has a preference over a consumption good and discounts the future at the rate $\beta < 1$:

$$\sum_{t=0}^{\infty} \beta^t \log(C_t) \tag{2}$$

At each period, household members are assigned to be workers or managers based on their abilities. Workers supply their ability to collect wages. Managers run plants to produce the single final good of the economy by hiring workers and renting capital. Production of the final good requires running differentiated activities i.e., using intermediate goods. Following Poschke (2018), the number of activities that a manager with ability z can perform, $M(z)$, equals to $\bar{M}^z$ where $\bar{M} > 1$ represents the aggregate technology level. I refer to the number of activities a manager can perform as productivity throughout the paper.

[8] Productivity is measured by the log of value added (Log-VA) per worker and VA is defined as sales minus expenditures on raw materials and energy following La Porta & Shleifer (2008).

The final output of a plant, y, with a manager of ability z can be defined as follows:

$$y = \left[\left(\int_0^{M(z)} \left(n_j^{\alpha} k_j^{1-\alpha} \right)^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}} \right]^{\gamma} \tag{3}$$

where j denotes different activities, n_j and k_j are the amount of ability and capital used in activity j respectively, $1 > \alpha > 0$ is the importance of labor in production, $1 > \gamma > 0$ is the span-of-control parameter and $\sigma > 1$ is the elasticity of substitution between differentiated products.

Government collects taxes from output at the rate τ. Enforcement of tax collection is incomplete in the sense that managers can operate informally to avoid taxes in line with Leal (2014). However, there is a possibility that an informal manager can get caught with probability, $p(k_j)$:

$$p(k_j) = \begin{cases} 1 & \int_0^{M(z)} k_j dj > B \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where B represents the enforcement level such that if a manager uses more capital than B in the production of the final good, she will be caught. Therefore, managers can avoid taxes by limiting their capital to less than or equal to B in total. If a manager is caught while operating informally, she loses all profits.

Problem of an Informal Manager A manager with ability z who operates informally chooses the amount of capital and labor to maximize her profit, $\pi_I(z)$:

$$\pi_I(z) = \max_{\{n_j, k_j\}_{j=0}^{M(z)}} \left(1 - p(k_j)\right) y - \int_0^{M(z)} W n_j dj - \int_0^{M(z)} R k_j dj \tag{5}$$

where W and R are wages and rental rate of capital respectively.

Problem of a Formal Manager A manager with ability z who operates formally chooses the amount of capital and labor to maximize her profit, $\pi_F(z)$. Since the manager operates formally, its output is subject to a tax, τ.

$$\pi_F(z) = \max_{\{n_j, k_j\}_j} (1 - \tau) y - \int_0^{M(z)} W n_j dj - \int_0^{M(z)} R k_j dj \tag{6}$$

The government collects taxes from formal managers and returns them to the household as a lump-sum transfer every period:

$$T_t = G_t \quad \forall t \tag{7}$$

where T_t denotes the transfers to the household and G_t is the government revenue from taxation.

Problem of the Household The household assigns its members to three occupations: workers, informal managers and formal managers, and chooses its consumption, C_t and how much capital to carry out to the next period, K_{t+1} in order to maximize its life-time utility:

$$\max_{\{S_{W,t},S_{I,t},S_{F,t},C_t,K_{t+1}\}_0^\infty} \sum_{t=0}^{\infty} \beta^t \log(C_t) \tag{8}$$

s.t.

$$C_t + K_{t+1} = I_t(S_{W,t}, S_{I,t}, S_{F,t}, W_t, R_t) + (1 - \delta + R_t)K_t + T_t$$

where $S_{W,t}, S_{I,t}, S_{F,t}$ denote the set of workers, informal managers and formal managers respectively, δ is the depreciation rate of the capital, T_t is the transfer from government and $I_t(S_{W,t}, S_{I,t}, S_{F,t}, W_t, R_t)$ is the total income of household members:

$$I_t(S_{w,t}, S_{I,t}, S_{F,t}, W_t, R_t) = W_t \int_{S_{W,t}} z g(z) dz + \int_{S_{I,t}} \pi_I(z) g(z) dz + \int_{S_{F,t}} \pi_F(z) g(z) dz \tag{9}$$

where the first item of the right-hand side is the wage income of workers and the second and the third items represent the total profit of the informal and formal managers respectively.[9]

3.2 Properties of the Equilibrium

This section focuses on a stationary equilibrium and discusses its key features. Profit of a formal manager with ability z, $\pi_F(z)$, can be written as a function of parameters and input prices using equations (A.3) and (A.4):

$$\pi_F(z) = (1-\tau)^{\frac{1}{1-\gamma}} (1-\gamma) \left(\gamma \left(\frac{\alpha}{W} \right)^{\alpha} \left(\frac{1-\alpha}{R} \right)^{1-\alpha} \right)^{\frac{\gamma}{1-\gamma}} M^{\frac{\gamma}{(\sigma-1)(1-\gamma)}} \tag{10}$$

[9] Appendix A provides the first-order conditions of managers' problems and the definition of equilibrium.

All informal managers limit their capital usage to B or less in their plants, as they would be caught operating informally otherwise. Then, the profit of an informal manager with ability z, $\pi_I(z)$, can be written as a function of parameters using equations (A.5), (A.6), (A.7) and (A.8):

$$\pi_I(z) = \begin{cases} (1-\gamma)\left(\gamma\left(\frac{\alpha}{W}\right)^{\alpha}\left(\frac{1-\alpha}{R}\right)^{1-\alpha}\right)^{\frac{\gamma}{1-\gamma}} M^{\frac{\gamma}{(\sigma-1)(1-\gamma)}} & \text{if } \int_0^{M(z)} k_j\, dj < B \\ (1-\alpha\gamma)\left(\left(\frac{\alpha\gamma}{W}\right)^{\alpha} b^{1-\alpha}\right)^{\frac{\gamma}{1-\alpha\gamma}} M^{\frac{\gamma(\sigma-\alpha(\sigma-1))}{(\sigma-1)(1-\alpha\gamma)}} - RB & \text{if } \int_0^{M(z)} k_j\, dj = B \end{cases} \tag{11}$$

I assume that $\gamma > \frac{\sigma-1}{\sigma} > \alpha$ such that profit functions of both types of managers are increasing and convex in z. This assumption also guarantees the existence and uniqueness of the equilibrium where informal and formal managers coexist (given high τ and B). If the total amount of capital used in an informal plant is less than B, the π_I is strictly greater than π_F due to $(1-\tau)^{\frac{1}{1-\gamma}} < 1$. The highest amount of capital that an informal manager use is B. I refer to informal managers who use exactly B amount of capital as constrained informal managers. They can still increase their profit by hiring more labor given their constrained capital. As a result, the slope of informal managers' profit, $\pi_I(z)$, is affected. When the optimal amount of capital for an informal manager is equal to B, the slope of the informal manager's profit with respect to z becomes smaller than that of the formal manager's profit function. Then, π_I can cross π_F at most once because both functions are strictly increasing. Since workers' earnings increase proportional to their ability, this establishes the existence and uniqueness.

Figure 3 describes an equilibrium assignment of household members to different occupations. The horizontal axis represents abilities and the vertical axis is the earnings of household members. While $\pi_I(z)$, the solid line, and $\pi_F(z)$, the dash-dotted line, are increasing and convex where they cross each other only once, the earnings of workers, the dashed line, is linear and increasing with a slope of W. Household members with $z \in [0, z_1^*)$ and $z \in [z_2^*, z_3^*)$ become informal managers, household members with ability $z \in [z_3^*, \bar{z}]$ become formal managers and remaining household members with ability $z \in [z_1^*, z_2^*)$ assigned to be workers in this equilibrium. The subsistence informality exists where $z \in [0, z_1^*)$. If taxes are low or the enforcement is stricter (i.e. low B) in an economy, there may be only subsistence informality in the equilibrium. In other words, z_3^*, the ability level where the informal profit function crosses the formal profit function, can be smaller than z_2^*, the threshold level where the earnings line of workers crosses the profit of managers from above.

The solution to the managers' problem in Appendix A shows that the amount of capital

and labor devoted to each activity is the same across activities for a given manager. Therefore, I express the capital-labor ratio at the activity level without loss of generality. Using equations (A.3) and (A.4), the capital-labor ratio of formal managers can be written as follows:

$$\frac{k_f}{n_f} = \frac{1-\alpha}{\alpha}\frac{W}{R} \tag{12}$$

Similarly, the capital-labor ratio of informal managers can also be derived using equations (A.5), (A.6), (A.7) and (A.8) as follows:

$$\frac{k_I}{n_I} = \begin{cases} \dfrac{1-\alpha}{\alpha}\dfrac{W}{R} & \text{if } \int_0^{M(z)} k_j\,dj < B \\ (\alpha\gamma)^{\frac{-1}{1-\alpha\gamma}} W^{\frac{1}{1-\alpha\gamma}} b^{\frac{1-\gamma}{1-\alpha\gamma}} M^{\frac{\sigma(1-\gamma)-1}{(\sigma-1)(1-\alpha\gamma)}} & \text{if } \int_0^{M(z)} k_j\,dj = B \end{cases} \tag{13}$$

Notice that capital-labor ratios of formal and informal managers are constant and equal to each other if informal managers optimally choose less capital than B. However, more able, constrained, informal managers have a smaller capital-labor ratio compared to other managers and the ratio decreases with the ability of informal managers as $\gamma > \frac{\sigma-1}{\sigma}$.

Since managers hire efficiency units, I define the size of a plant relative to the average ability of workers as in Poschke (2018). Let F_P be the fraction of plants, i.e. the fraction of managers, which is the sum of the fraction of formal managers, F_F, and informal managers, F_I. Hence, the mean size of informal plants can be defined as the ratio of the average ability demanded in informal plants and the average ability of workers in informal plants. Using the market clearing condition, it can be written as follows

$$\text{Mean size of informal plants} = \frac{F_{WI}(1-F_P)}{F_I} \tag{14}$$

where F_{WI} is the share of informal workers among workers such that $F_{WI}(1-F_P)$ is the share of household members employed in informal plants.[10] Then, the mean size of informal plants can also be written as the ratio of average ability demanded by formal plants and the average ability of workers among formal plants:

[10] Since workers earn the same wage rate regardless of working in an informal or a formal plant, F_{WI} is the ratio of the total labor demand by informal managers to the total labor demand by all managers: $F_{WI} = \frac{\int_{S^*_{I,t}} n^*_{I,t}(z)g(z)dz}{\int_{S^*_{I,t}} n^*_{I,t}(z)g(z)dz + \int_{S^*_{F,t}} n^*_{F,t}(z)g(z)dz}$

$$\text{Mean size of formal plants} = \frac{F_{WF}(1-F_P)}{F_F} \tag{15}$$

where F_{WF} is the share of formal workers among workers such that $F_{WF}(1-F_P)$ is the share of household members employed in formal plants. In mean size calculations, labor market clearing conditions are used such that the supply of ability by workers is equal to the demand for ability by plants. Similar arguments provide the economy's overall mean size. Alternatively, we can use the above definitions to derive the mean size in equilibrium:

$$\text{Mean size} = s_F \frac{F_{WF}(1-F_P)}{F_F} + s_I \frac{F_{WI}(1-F_P)}{F_I} = \frac{1-F_P}{F_P} \tag{16}$$

where $s_i = \frac{F_i}{F_P} \quad i \in \{F, I\}$, i.e. s_i's are the share of formal and informal plants. Therefore, the overall mean size of an economy is the fraction of workers divided by the fraction of plants in equilibrium.

4 Parameter Values

In the remaining part of the paper, I quantify the gains from increasing managers' productivity exogenously. First, I calibrate model parameters to replicate key plant-size distribution and informality properties in Ghana, allowing us to quantify the effects of different policies in the benchmark economy. Some model parameters are set following the literature: the discount rate, β, depreciation rate of capital, δ, the importance of capital, $(1-\alpha)\gamma$, and the elasticity of substitution, σ, are set to 0.93, 0.072, 0.33 and 4 following Gollin (1995), Leal (2014) and Poschke (2018). The value of α is determined as 0.63 once the span-of-control parameter is calibrated below.

The rest of the parameters are calibrated jointly. I assume the abilities are distributed according to a log-normal distribution with $\mu = 0$ and standard deviation σ_z. Hence, there are five parameters to be determined: σ_z, γ, $\bar{M}$, τ and B. I use five moments from Ghana's Integrated Business Establishment Survey 2014 (IBES) which is an economic census conducted by the Ghana Statistical Service: the overall mean size, the mean size of informal plants, the mean size of formal plants, the fraction of plants with fewer than 50 workers and the employment share of plants with more than 50 workers. IBES covers all establishments regardless of size and sector. Moreover, IBES considers a plant to be formal if it is registered with the Registrar General's Department and keeps formal accounting records.[11] Otherwise,

[11] Abreha *et al.* (2022) shows that the distribution of employment similar across different definitions infor-

it is considered to be an informal plant. Table 3 displays the parameter values at the benchmark economy. Panel A of Table 4 shows the model's performance. The benchmark economy matches the targeted mean size of informal and formal plants very well while, also, it is able to capture the concentration of employment at plants with 50 workers.

While the parameters are calibrated jointly, the mean size of establishments, 5.3 workers, helps identify the span-of-control parameter, $\gamma = 0.91$. The mean size of informal, 2.3 workers, and formal plants, 33.8 workers, are targeted to match the enforcement level, B, and the output tax, τ, parameters, respectively. Finally, the fraction of plants with fewer than 50 workers, 99.1%, helps to calibrate the aggregate technology level, $\bar{M}$. The standard deviation of the ability distribution, σ_z is selected to match the concentration of employment in plants with more than 50 workers, i.e. 32.3%.

The benchmark economy performs successfully in matching not only the targeted moments but also other relevant moments, as shown in Panel B of Table 4. First, by targeting the overall mean size along with the mean sizes of informal and formal plants, the model identifies the fraction of informal plants, s_I, through Equation (16). Hence, 90.6% of all plants operate informally in the benchmark model (IBES reports 90.5% of plants are informal in Ghana). Second, since informal plants employ fewer workers compared to formal plants on average, the employment share of informal plants is 39.7% in Ghana, while in the benchmark economy, 39.8% of workers are employed by informal plants. Third, the fraction of informal plants with fewer than 5 workers is 83.7% in the data and 88.5% in the benchmark economy. Fourth, the employment share of formal plants with more than 100 workers (among all workers employed by formal plants) is 28.6% in the data, closely matching the benchmark economy's employment share of formal plants. Finally, government revenue constitutes about 14.2% of Ghana's output. Although the output tax in the benchmark economy is around 21%, government revenue in the model is approximately 14.1% of the aggregate output, since 90.6% of plants operate informally.

As the benchmark economy closely matches both the targeted and non-targeted moments related to the plant size distribution, we can use it to infer some key statistics about the Ghanaian economy. First observation is that 34.2% of the aggregate output is produced by informal plants, despite operating at smaller scales than formal plants. Second, subsistence informality exists in the benchmark economy ($z_1^* < z_2^* < z_3^*$) with 71.1% of informal plants managed by the lowest-skilled agents. This implies that bottom 10.2% of agents run informal plants in the benchmark economy. Third, the income of an average formal plant

mality.

manager is 5.2 times higher than that of an average worker. In contrast, managers of informal plants earn 8.7% less than average workers in the benchmark economy, since most of them are at subsistence level.

5 Findings

In this section, I discuss how development affects the allocation of agents into different occupations, share of informal plants, plant size distribution and aggregate output. I refer to an exogenous increase in managers' productivity, i.e. the number of activities they can perform, as development. It is motivated by two well-known observations in the development literature: a better business environment and higher years of schooling in richer countries. Better institutions in developed countries allow managers to perform more activities in business environments. Figure 4a shows that countries with a higher ease of doing business score also have smaller informal mean sizes.[12] In addition, increasing the number of activities a manager can run resembles an upward shift in the mean of their ability. In order to relate this to the data, Figure 4b shows that countries with higher years of schooling have smaller informal plants on average. In particular, I introduce development to the model using the following specification:

$$(1+\lambda)\, M(z) \tag{17}$$

where λ denotes the percentage change in the number of activities that a manager can perform regardless of whether they are formal and informal managers. Table 5 shows the effect of development on the benchmark economy with different values of λ ranging from 0 (the benchmark) to 45%. In these exercises, the parameters are fixed at the reported values in Table 3. Panel A focuses on the changes in the size distribution of plants and the aggregate output. As the value of lambda (the percentage increase in the number of activities managers can perform) increases, the overall mean size increases while the informal mean size declines. Also, the aggregate output and the employment accounted for by informal plants decline. For example, when all managers can perform 30% more activities (i.e. $\lambda = 30\%$) the share of informal plants declines to 84.7% and the contribution of informal plants to the aggregate output falls by 13.5 percentage points. Similarly, it increases the share of infor-

[12] The World Bank's Ease of Doing Business Score covers topics related to business regulations such as starting a business, paying taxes, getting credit among many. Historical data can be accessed here: https://archive.doingbusiness.org/en/data.

mal plants in subsistence informality to 93.7%. When λ is above than about 40%, only the subsistence informality exists in the economy.

The effect of development on average earnings is presented in Panel B of Table 5. Both workers and managers of formal plants benefit from higher values of λ. However, informal managers' average earnings decline as the benchmark economy develops exogenously. Informal managers' average earnings decline because higher values of λ lead to an increase in subsistence informality, where informal managers earn less than workers on average. Workers benefit the most from development as the average worker income increases by 18.7% when $\lambda = 45\%$ whereas the average formal managers income increases by 3.3%. While changes in λ affect the earning profiles, inequality in the benchmark economy increases slightly with development. When λ is 45%, the mean-to-median ratio and the Gini coefficient increases from 1.11 to 1.13 and from 0.22 to 0.23 respectively.

Mechanism As managers are able to run more activities, they demand more input. Hence, the wage rate and managerial earnings increase with higher values of λ.[13] If there would be complete enforcement, i.e. no informality, changes in λ would not have any effect on the assignment of household members into occupations as Poschke (2018) and Tamkoç & Ventura (2024) discuss. Intuitively, the increase in managers' profit would be equal to the increase in the wage rate with higher values of λ. However, when there is an opportunity to operate informally due to incomplete enforcement, changes in λ affect the sorting of agents into occupations. Because the increase in the demand for inputs by (constraint) informal plants is different from that by formal plants as Equations (A.8) and (A.4) shows. I discuss the mechanism at two stages: (i) when informality exists at both tails of ability distribution as in the benchmark and (ii) when only the subsistence informality exists.

First, the occupational assignment can be characterized by three threshold abilities, $z_1^* < z_2^* < z_3^*$ when the informality occurs at both tails of ability distribution. As shown in Equations (10) and (11), formal plants benefit more from higher values of λ compared to constrained informal plants. Hence z_3^* declines with development. At the same time, z_2^* rises due to higher wages. Therefore, the fraction of agents whose abilities are higher than workers' abilities drops with development. Some of them start operating formally (i.e. z_3^* falls) and some of them become workers (z_2^* increases) in new steady states with larger λ. On the other hand, z_1^* may increase when λ increases incrementally. However, the subsistence in-

[13] The rental rate of capital is constant across steady states as it is in other span-of-control models where the household owns the capital. See Guner *et al.* (2008), Leal (2014), Poschke (2018), Tamkoç (2022) and Tamkoç & Ventura (2024) among many.

formality (z_1^*) declines eventually with higher values of λ (in the benchmark economy, z_1^* starts declining when λ exceeds 25%) since the marginal subsistence informal managers earn more when they are workers with higher wages.

As z_2^* increases and z_3^* declines with higher λ values, eventually, there is no informal manager who has more ability than workers in the developed economy. In other words, all informality occurs at subsistence levels. At this stage, occupational assignment can be summarized by two threshold values: the small threshold and the big threshold. Agents with abilities lower than the small threshold become (subsistence) informal managers, agents with ability levels between the small and the big threshold values are assigned to be workers. Finally, agents with abilities that are higher than the big threshold run formal plants. As λ increases further, both threshold values decline. As discussed above, the increase in the wage rate would be smaller than the increase in the profits of formal plants. This is because informal plants increase their demand less than formal plants do which results in lowering the big threshold. The small threshold declines as the increase in wage rates dominates the increase in the profit of marginal subsistence level managers similar to the above mechanism.

5.1 Accounting for Informal Mean Plant Sizes across Countries

Now, Figure 5 shows the how the informal mean size, the aggregate output, the fraction of informal plants and the government revenue change for λ varying from -0.9 to 5. It is worth noting that informal plants exist in the economy regardless of how much λ increases because $M(0) > 0$. When all managers can run six times more activities than in the benchmark case, i.e., when $\lambda = 5$, the aggregate output increases by 163.9%. As a result, the fraction of informal plants falls to 2.8% and the mean size of informal plants becomes about 1 worker.

Given that higher λ values imply higher output and the smaller informal plants, Figure 6 interprets the above finding from a cross-country perspective. Each dot represents a country as discussed in the data section. The line with negative slope stands for the model's outcome with different values of λ where the aggregate output is normalized to the real GDP per capita in Ghana. For example, $\lambda = 1.47$ reduces the informal mean size to 1.4 workers what is observed in Peru. This implies a 73.9% increase in the aggregate output, accounting for 54.3% of the real GDP per capita difference between Ghana and Peru (the real GDP per capita in Peru is 136.1% higher than in Ghana). When this exercise is carried out across all countries in the sample, changes in λ can account for 32.4% of the RGDP per capita difference across countries on average.

5.2 When Only Formal Managers Benefit from Development

In this subsection, I investigate the results of improving formal managers' productivity proportionally. Then, I document the results when these improvements are skill-biased, i.e., higher-ability formal managers benefit more from development.

Table 6 shows the results of experiments where all managers experience improvements in their productivity by λ_F% except for informal managers. Aggregate output gains resulting from different λ_F values correspond to similar values from the previous exercise in Table 5. When development policies are targeted only at the formal sector, the overall mean size grows much larger than the previous case where both sectors could benefit from productivity improvements. This is because the share of the informal sector falls to around 0.2%, compared to around 81% in the previous case, when aggregate output increases by 29.5%. As a result, government revenues almost double.

Next, I consider a case where the technology level for formal managers improves exogenously by ϕ%. In other words, formal managers have access to $(1+\phi)\bar{M}$ while informal managers still produce with $\bar{M}$. This formulation is called skill-biased change in entrepreneurial technology by Poschke (2018) because higher-ability managers experience greater productivity gains when $\phi > 0$ compared to lower-ability managers. Table 7 presents the results of experiments with different ϕ levels. Again, different ϕ values are selected such that aggregate output gains are similar to those in Table 5. As a result of the skill-biased improvement in formal managers' technology, informality almost disappears when aggregate output increases 29.5%. A noticeable difference between this exercise and the previous one is the increase in employment in plants employing more than 50 workers. Since higher-ability managers runs bigger plants, the employment share of larger plants increases more in this case compared to previous exercises.

5.3 Comparing with Tax and Enforcement Policies

Reducing tax rates (the cost of being formal) and increasing tax enforcement are widely studied formalization policies in the literature. In this subsection, I discuss their effects on the benchmark economy.

Figure 7 summarizes how aggregate output, the fraction of informal plants, and government revenue change under various formalization policies. Figures in first column displays the results with different tax rates, while figures in the second column shows experiments with different enforcement levels. Due to the functional form assumptions for $M(z)$, some

informal plants exist in all exercises except when taxes are zero ($\tau = 0$) and there is full enforcement ($B = 0$). However, when τ and B are below 15% and 1.5, respectively, less than 0.01% of plants operate informally. Hence, Figure 7 displays results when τ is greater or equal to 0.15 and B is greater or equal to 1.5.

Lower taxes and stricter enforcement lead to a non-trivial decrease in the share of informal plants. However, the aggregate output gains are 21.3% and 16.7% when τ and B are 15% and 1.5, respectively. The aggregate output, indeed, increases when B is greater than around 9.5 where almost 97.8% of plants operate informally. Moreover, the benchmark economy appears to be on the right side of the Laffer curve, where reducing output taxes can generate additional government revenue. Because 90.6% of plants operate informally, any reduction in the cost of operating formally attracts productive informal plants.

While formalization policies offers smaller informality, output gains are limited to around 20% even when all plants starts operating formally in the long-run. Additionally, governments may need more resources to reduce subsistence informality, for example, through full enforcement, raising questions about the feasibility of these formalization policies.

6 Concluding Remarks

In this paper, I argue that one of the main reasons for encountering fewer informality in rich countries is development, i.e., higher wages for workers and greater production possibilities for plants.

The paper is motivated by the fact that informal plants in richer countries employ fewer workers compared to informal plants in poorer countries. This finding holds even after controlling for the sector of activity and plant characteristics such as age of the plant, experience of managers etc. Then, I develop a general equilibrium model where agents are sorted into being workers, informal managers and formal managers according to their abilities. The model is flexible to accommodate both subsistence informality, i.e., informal managers with abilities lower than workers, and informality where informal managers are more able than workers but have lower abilities than formal managers.

As the economy develops by allowing plants to operate at greater scale, some informal managers at the margin become workers thanks to higher wages, while others transition to formality to enjoy greater profits. The model matches key moments from the plant size distribution in Ghana as well as the share of informal plants and their mean size. Quantitative results show that a 73.9% increase in aggregate output, driven by exogenous improvements

in managers' productivity, reduces the informal mean size of Ghana, by about 40%, to the levels observed in Peru.

This paper focuses on informality at the extensive margin, i.e., unregistered plants. There is another form of informality: informality in the intensive margin. In this form of informality, registered plants under-report their sales or size to evade taxation. Investigating the effects of development in a model that incorporates both extensive and intensive margin informality is left for future study.

References

Aberra, Adam, Aga, Gemechu, Jolevski, Filip, & Karalashvili, Nona. 2022. Understanding Informality: Comprehensive Business-Level Data and Descriptive Findings.

Abreha, Kaleb Girma, Cirera, Xavier, Fattal Jaef, Roberto N, Maemir, Hibret Belete, & Davies, Elwyn. 2022. Deconstructing the Missing Middle: Informality and Growth of Firms in Sub-Saharan Africa. *World Bank, Washington, DC*.

Aga, Gemechu, Francis, David C, Jolevski, Filip, Rodriguez Meza, Jorge, & Wimpey, Joshua Seth. 2023. An Application of Adaptive Cluster Sampling to Surveying Informal Businesses. *Journal of Survey Statistics and Methodology*, 01.

Amaral, Pedro S, & Quintin, Erwan. 2006. A competitive model of the informal sector. *Journal of monetary Economics*, **53**(7), 1541–1553.

Amin, Mohammad, & Okou, Cedric. 2020. Casting a shadow: Productivity of formal firms and informality. *Review of Development Economics*, **24**(4), 1610–1630.

Barro, Robert J, & Lee, Jong Wha. 2013. A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, **104**, 184–198.

Bartelsman, Eric, Haltiwanger, John, & Scarpetta, Stefano. 2013. Cross-country differences in productivity: The role of allocation and selection. *American economic review*, **103**(1), 305–334.

Bento, Pedro, & Restuccia, Diego. 2017. Misallocation, establishment size, and productivity. *American Economic Journal: Macroeconomics*, **9**(3), 267–303.

Bento, Pedro, & Restuccia, Diego. 2021. On average establishment size across sectors and countries. *Journal of Monetary Economics*, **117**, 220–242.

Bruhn, Miriam, & McKenzie, David. 2014. Entry regulation and the formalization of microenterprises in developing countries. *The World Bank Research Observer*, **29**(2), 186–201.

De Andrade, Gustavo Henrique, Bruhn, Miriam, & McKenzie, David. 2016. A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms. *The World Bank Economic Review*, **30**(1), 24–54.

De Mel, Suresh, McKenzie, David, & Woodruff, Christopher. 2013. The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, **5**(2), 122–150.

De Soto, Hernando. 1989. *The other path: The invisible revolution in the third world*. Harper & Row New York.

D'Erasmo, Pablo N, & Boedo, Hernan J Moscoso. 2012. Financial structure, informality and development. *Journal of Monetary Economics*, **59**(3), 286–302.

Fajnzylber, Pablo, Maloney, William F, & Montes-Rojas, Gabriel V. 2011. Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program. *Journal of Development Economics*, **94**(2), 262–276.

Franjo, Luis, Pouokam, Nathalie, & Turino, Francesco. 2022. Financial frictions and firm informality: A general equilibrium perspective. *The Economic Journal*, **132**(645), 1790–1823.

Friedman, Eric, Johnson, Simon, Kaufmann, Daniel, & Zoido-Lobaton, Pablo. 2000. Dodging the grabbing hand: the determinants of unofficial activity in 69 countries. *Journal of public economics*, **76**(3), 459–493.

Garcia-Santana, Manuel, & Pijoan-Mas, Josep. 2014. The reservation laws in India and the misallocation of production factors. *Journal of Monetary Economics*, **66**, 193–209.

Gollin, Douglas. 1995. Do taxes on large firms impede growth? Evidence from Ghana. *University of Minnesota, Economic Development Center*.

Gourio, François, & Roys, Nicolas. 2014. Size-dependent regulations, firm size distribution, and reallocation. *Quantitative Economics*, **5**(2), 377–416.

Guner, Nezih, Ventura, Gustavo, & Xu, Yi. 2008. Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, **11**(4), 721–744.

Hsieh, Chang-Tai, & Klenow, Peter J. 2009. Misallocation and manufacturing TFP in China and India. *The Quarterly journal of economics*, **124**(4), 1403–1448.

Hsieh, Chang-Tai, & Klenow, Peter J. 2014. The life cycle of plants in India and Mexico. *The Quarterly Journal of Economics*, **129**(3), 1035–1084.

Johnson, Simon, Kaufmann, Daniel, & Zoido-Lobaton, Pablo. 1998. Regulatory discretion and the unofficial economy. *The American economic review*, **88**(2), 387–392.

Kuehn, Zoe. 2014. Tax rates, governance, and the informal economy in high-income countries. *Economic Inquiry*, **52**(1), 405–430.

La Porta, Rafael, & Shleifer, Andrei. 2008. *The unofficial economy and economic development*. Tech. rept. National Bureau of Economic Research.

La Porta, Rafael, & Shleifer, Andrei. 2014. Informality and development. *Journal of Economic Perspectives*, **28**(3), 109–126.

Leal, Julio César Ordóñez. 2014. Tax collection, the informal sector, and productivity. *Review of Economic Dynamics*, **17**(2), 262–286.

Loayza, Norman V. 1996. The economics of the informal sector: a simple model and some empirical evidence from Latin America. *Pages 129–162 of: Carnegie-Rochester conference series on public policy*, vol. 45. Elsevier.

Loayza, Norman V. 2016. Informality in the Process of Development and Growth. *The World Economy*, **39**(12), 1856–1916.

Lucas, Robert E. 1978. On the size distribution of business firms. *The Bell Journal of Economics*, 508–523.

Meghir, Costas, Narita, Renata, & Robin, Jean-Marc. 2015. Wages and informality in developing countries. *American Economic Review*, **105**(4), 1509–1546.

Ohnsorge, Franziska, & Yu, Shu. 2022. *The long shadow of informality: Challenges and policies*. World Bank Publications.

Orsi, Renzo, Raggi, Davide, & Turino, Francesco. 2014. Size, trend, and policy implications of the underground economy. *Review of Economic Dynamics*, **17**(3), 417–436.

Poschke, Markus. 2018. The firm size distribution across countries and skill-biased change in entrepreneurial technology. *American Economic Journal: Macroeconomics*, **10**(3), 1–41.

Rauch, James E. 1991. Modelling the informal sector formally. *Journal of Development Economics*, **35**(1), 33–47.

Restuccia, Diego, & Rogerson, Richard. 2008. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics*, **11**(4), 707–720.

Rocha, Rudi, Ulyssea, Gabriel, & Rachter, Laísa. 2018. Do lower taxes reduce informality? Evidence from Brazil. *Journal of Development Economics*, **134**, 28–49.

Tamkoç, M Nazım. 2022. *Bribery, Plant Size and Size Dependent Distortions*. World Bank.

Tamkoç, M Nazım, & Ventura, Gustavo. 2024. *Rules and Regulations, Managerial Time and Economic Development*. World Bank.

Ulyssea, Gabriel. 2018. Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, **108**(8), 2015–2047.

Ulyssea, Gabriel. 2020. Informality: Causes and consequences for development. *Annual Review of Economics*, **12**, 525–546.

Tables and Figures

Table 1: Relationship between the size of informal plants and RGDP per capita

Dep. var.: Log # of workers	(1)	(2)	(3)	(4)	(5)	(6)
Log RGDP per capita	-0.244***	-0.148***	-0.142***	-0.154***	-0.164***	-0.163***
	(0.0430)	(0.0350)	(0.0332)	(0.0333)	(0.0321)	(0.0321)
Manufacturing		0.290***	0.286***	0.266***	0.285***	0.285***
		(0.0517)	(0.0538)	(0.0536)	(0.0513)	(0.0518)
Education of manager			0.0996***	0.102***	0.106***	0.106***
			(0.0265)	(0.0261)	(0.0256)	(0.0259)
Experience of manager				0.00829***	0.00765***	0.00793***
				(0.00143)	(0.00153)	(0.00229)
Female manager					-0.0868***	-0.0886***
					(0.0271)	(0.0267)
Age of plant						-0.000523
						(0.00227)
Constant	2.564***	1.634***	1.528***	1.553***	1.637***	1.632***
	(0.370)	(0.303)	(0.286)	(0.283)	(0.270)	(0.270)
# of Observation	21195	20393	20156	19337	17550	17379

Source: ISES and WDI.

Notes: This table shows that the number of workers in an informal plant plant decreases with RGDP per capita even after controlling the plant characteristics. Values in the table are the estimated coefficients of Equation (1) where the dependent variable is the natural logarithm of the number of workers in an informal plant. Huber-White robust standard errors are reported in parenthesis. Standard errors are clustered at the country and city levels. For more details see Section 2.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 2: Comparison of Informal and Formal Business in Ghana

	Informal	Micro WBES-Formal	WBES-Formal
Sampling Size		1-5	5+
Mean Size	2.3	3.0	34.7
Log-VA per worker	100	155.4	169.8
Experience of manager (years)	9.5	12.4	16.6
% of managers: at least primary sch.	88.2		
% of managers: at least secondary sch.	42.8		
% of workers: at least secondary sch.		66.0	65.2
% of female managers	62.8	26.5	14.9
Age (of plants)	8.6	9.9	15.4
# of observation	729	604	720

Notes: This tables compares characteristics of informal plants with that of formal plants in Ghana using the World Bank's plant level datasets. The second column presents summary statistics of informal plants using ISES. The third and the last columns presents summary statistics of formal plants using Micro-WBES and WBES datasets respectively. The former interviews plants with less than 5 workers whereas the latter includes plants with more than 5 workers. The average log-VA per worker of informal plants is normalized to 100. See Section 2 for details.

Table 3: Calibrated Parameter Values

Parameter	Description	Value
σ_z	Std. dev. of ability distr.	0.33
γ	Span-of-control parameter	0.91
$\bar{M}$	Level of prod. technology	2.03
τ	Output tax	0.21
B	Enforcement level	5.97

Notes: Value column shows the parameters values that generate the size distribution of plants and the mean plant size in informal sector in Ghana. For more details about parameters and the moment conditions see. Section 4 for details.

Table 4: Calibration performance: Targets and Model

	Data	Model
Panel A: Targeted Moments		
Overall mean size	5.3	5.3
Mean size of informal plants	2.3	2.3
Mean size of formal plants	33.8	33.8
Fraction of plants: less than 50 workers	99.1	97.9
Emp. share of plants: more than 50 workers	32.3	32.3
Panel B: Non-targeted Moments		
Emp. Share of Informal Plants (%)	39.7	39.8
Fraction of Informal Plants: less than 5 workers (% among Informal Plants)	83.7	88.5
Emp. Share of Formal Plants: more than 100 workers (% among Formal Plants)	28.6	33.0
Gov't Revenue (% of Agg.Output)	14.2	14.1

Data Sources: Ghana IBES and WDI

Note: This table shows the calibration performance. The data column shows the target moment conditions for the Ghana economy. The model column shows how model results with the calibrated parameters reported in Table 3. See Section 4 for description of data and the moments in detail.

Table 5: Development and Informality

	$\lambda = 0$ (Benchmark)	$\lambda = 15\%$	$\lambda = 30\%$	$\lambda = 45\%$
Panel A: Size Distribution of Plants and Aggregate Output				
Mean Size	5.3	5.6	6.2	6.9
Mean Size of Informal Plants	2.3	2.1	1.8	1.7
Agg. Output	100	110.2	120.3	129.5
Informal Output (% of Agg. Output)	34.2	27.1	20.7	15.9
Emp. Share of Informal Plants (%)	39.8	32.1	25.0	19.4
Emp. Share of 50+ (%)	32.3	34.6	36.3	37.6
Informal Plants (%)	90.6	88.0	84.7	81.0
Subsistence Informality (% of informal plants)	71.1	83.4	93.7	100.0
Tax Revenue	100	122.1	144.9	165.4
Panel B: Earnings				
Average Worker Income	100	107.3	113.6	118.7
Average Informal Man. Income	91.3	81.9	72.1	65.6
Average Formal Man. Income	521.3	525.9	529.1	538.7
Mean/Median Ratio	1.11	1.11	1.12	1.12
P90/P10 ratio	2.32	2.31	2.31	2.32

Note: This table presents results when the benchmark economy develops exogenously by improving all managers' productivity by λ%, i.e. $(1+\lambda)M(z)$. Panel A shows changes in the plant size distribution and aggregate output and Panel B displays the effects of development on earnings. The aggregate output, tax revenue and average worker income is normalized to 100 at the benchmark P90 and P10 earnings corresponds to the 90th and 10th percentile earnings. See Section 5 for details.

Table 6: Improving only the Formal Sector

	$\lambda_F = 0$ (Benchmark)	$\lambda_F = 5.9\%$	$\lambda_F = 11.4\%$	$\lambda_F = 45\%$
Mean Size	5.3	8.1	15.3	30.1
Mean Size of Informal Plants	2.3	2.2	1.7	1.4
Agg. Output	100	110.2	120.3	129.5
Informal Output (% of Agg. Output)	34.2	17.8	4.7	0.01
Emp. Share of Informal Plants (%)	39.8	21.6	5.9	0.01
Emp. Share of 50+ (%)	32.3	37.4	40.6	41.7
Informal Plants (%)	90.6	79.8	51.9	0.2
Subsistence Informality (% of informal plants)	71.1	77.0	100.0	100.0
Tax Revenue	100	137.7	174.2	196.8

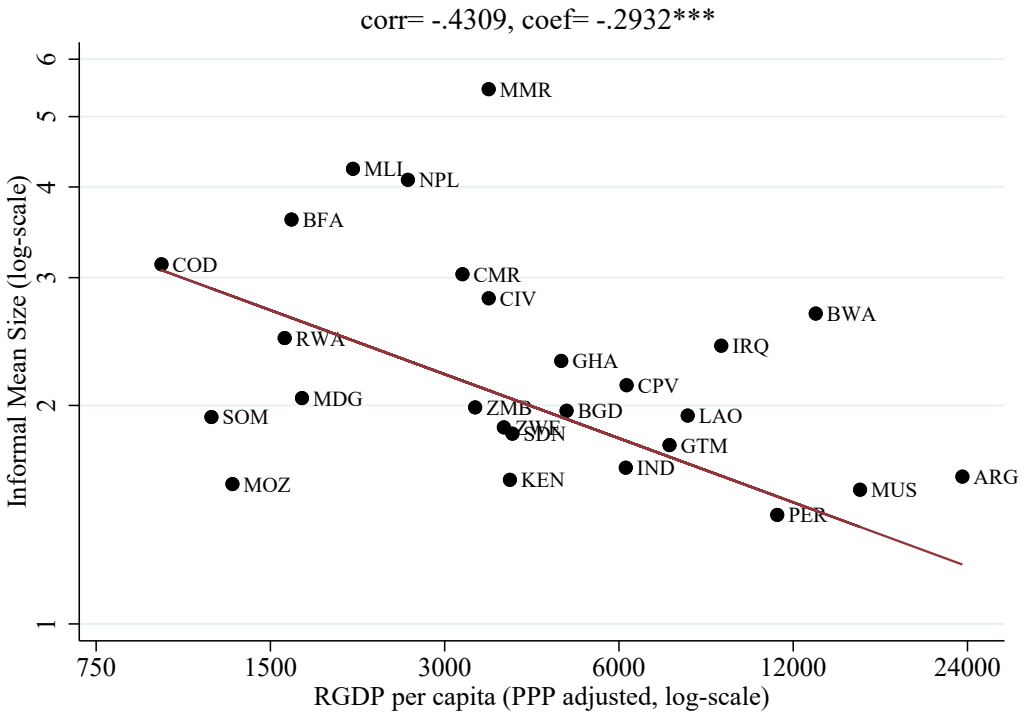
Note: This table presents results when the benchmark economy develops exogenously by improving only formal managers' productivity by λ_F%. The aggregate output, tax revenue and average worker income is normalized to 100 at the benchmark P90 and P10 earnings corresponds to the 90th and 10th percentile earnings. See Section 5 for details.

Table 7: Skill-Biased Change in Formal Managers' Technology

	$\phi = 0$ (Benchmark)	$\phi = 2.6\%$	$\phi = 4.9\%$	$\phi = 9.6\%$
Mean Size	5.3	8.1	15.3	36.4
Mean Size of Informal Plants	2.3	2.2	1.8	1.5
Agg. Output	100	110.2	120.3	129.5
Informal Output (% of Agg. Output)	34.2	18.3	5.5	0.04
Emp. Share of Informal Plants (%)	39.8	22.1	6.9	0.1
Emp. Share of 50+ (%)	32.3	39.5	45.5	52.1
Informal Plants (%)	90.6	81.1	57.3	1.3
Subsistence Informality (% of informal plants)	71.1	75.8	94.6	100.0
Tax Revenue	100	136.9	172.8	196.7

Note: This table presents results when the benchmark economy develops exogenously by improving only formal managers' productivity such that higher able formal managers benefit more compared to lower able formal managers. In other words, the technology level available to formal managers changes to $(1+\phi)\bar{M}$. The aggregate output, tax revenue and average worker income is normalized to 100 at the benchmark P90 and P10 earnings corresponds to the 90th and 10th percentile earnings. See Section 5 for

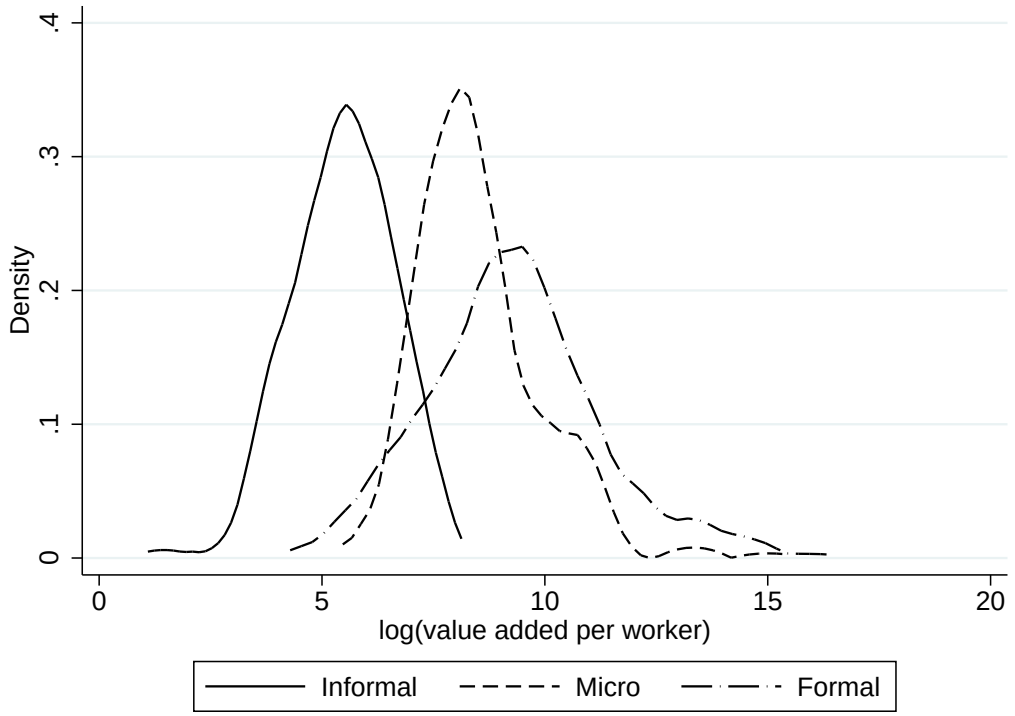
Figure 1: Informal Plants in Richer Countries are Smaller.



Sources: ISES and WDI

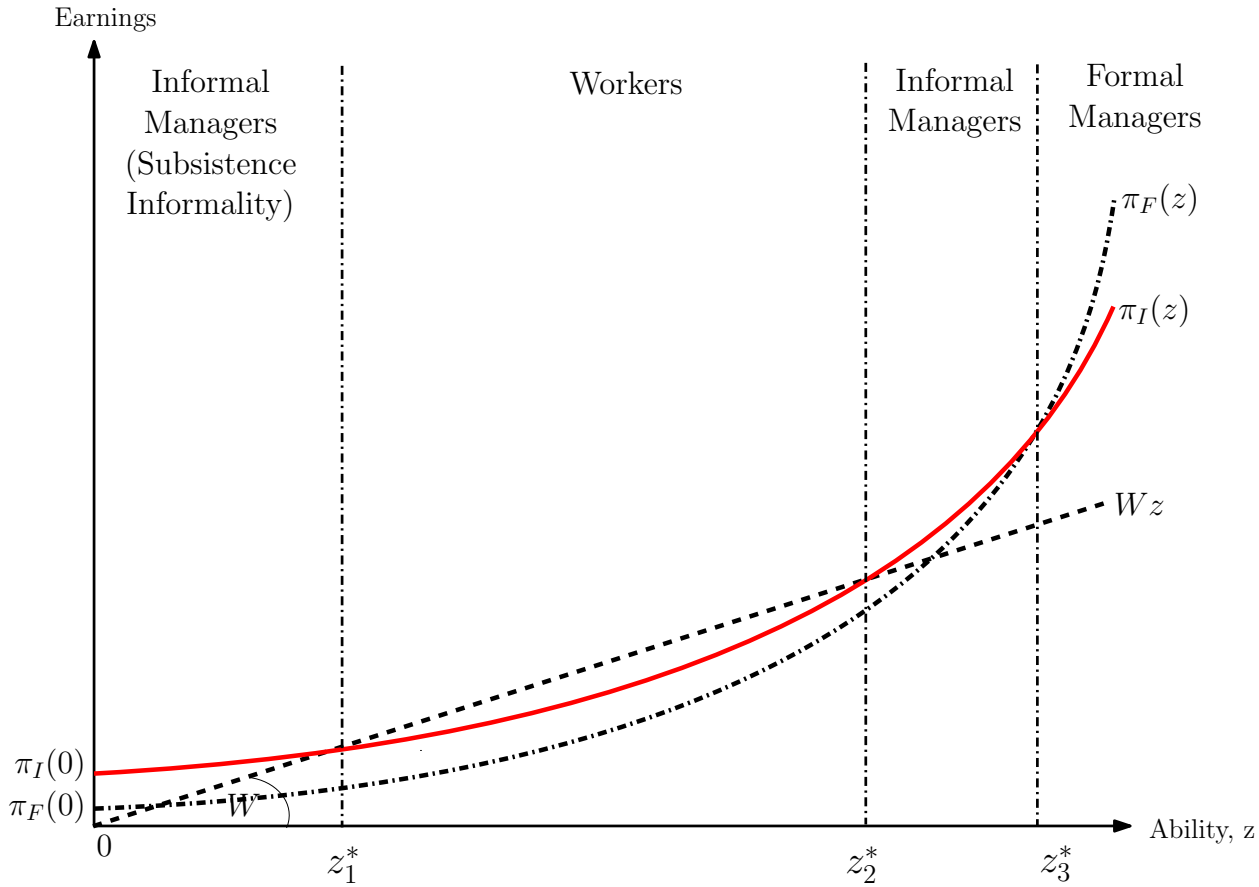
Notes: This figure shows that richer countries tend to have smaller mean size in informal plants. The mean size is the number of workers an informal plant have on average. For more details see Section 2.

Figure 2: Comparison of Productivity of informal, micro and formal plants in Ghana



Sources: ISES, Micro WBES, WBES
Notes: This figure shows that productivity of informal plants are very low compared to that of formal plants. For more details see Section 2.

Figure 3: Earnings and the Occupational Assignments in Equilibrium

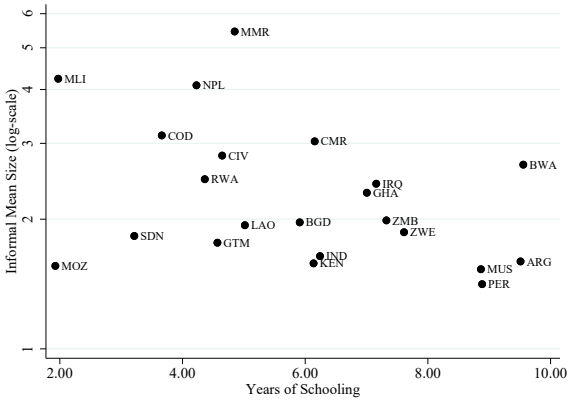


Notes: This figure shows the equilibrium assignment of household members in three different occupations. Earnings of workers increase with their ability level. π_F and π_I denote the profit functions of formal and informal managers, respectively. Occupational assignment is characterized by three threshold ability levels, z_1^*, z_2^* and z_3^*. For more details see Section 3.

Figure 4: Development: Ease of Doing Business and Years of Schooling

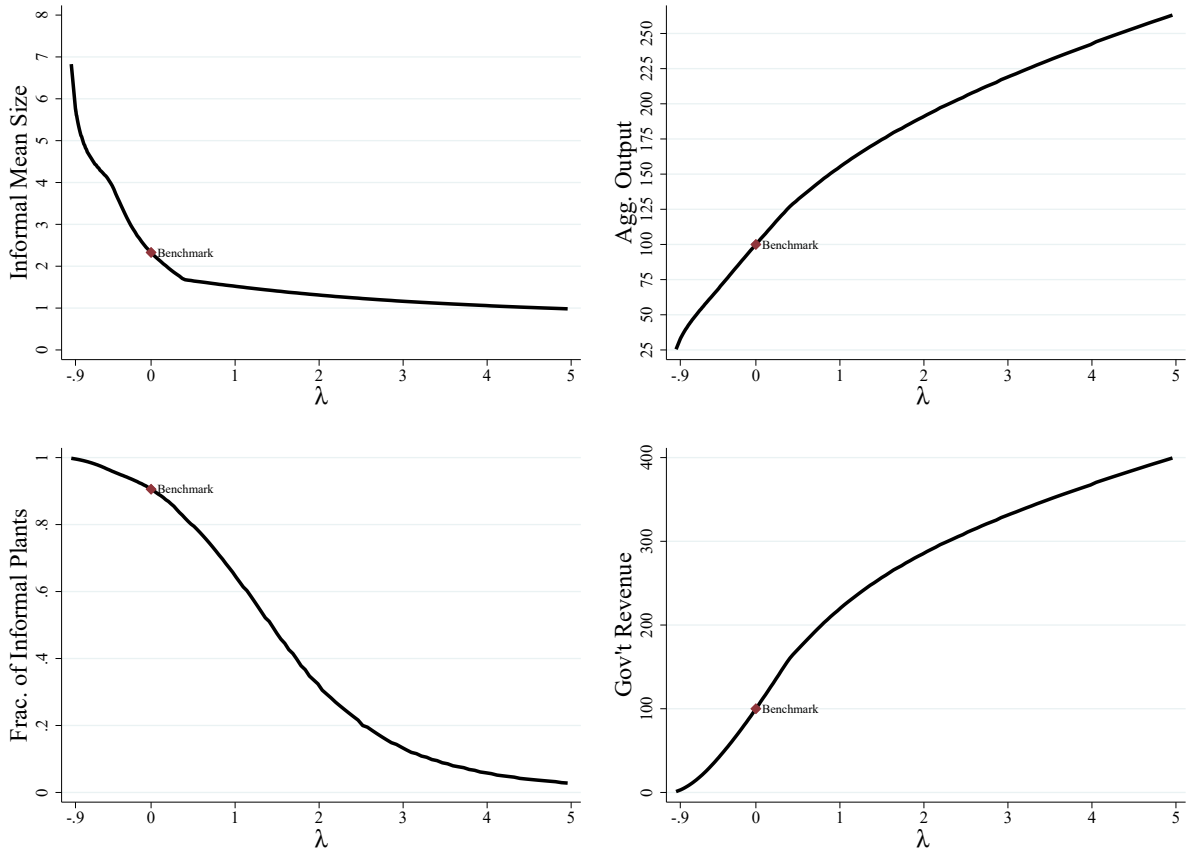
(a) Doing Business Score

(b) Years of Schooling



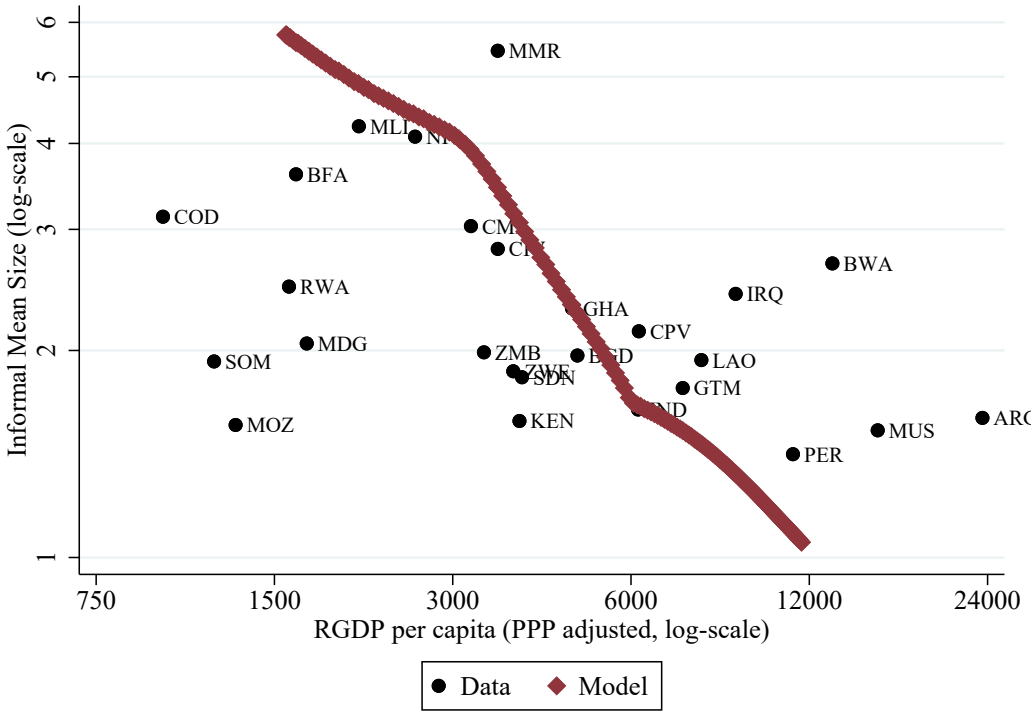
Sources: ISES, Doing Business and Barro & Lee (2013)
Notes: This figure compares the informal mean size with the ease of doing business score (Panel (a)) and the average years of schooling (Panel (b)). For more details see Section 5.

Figure 5: Changes in λ and Aggregate Variables



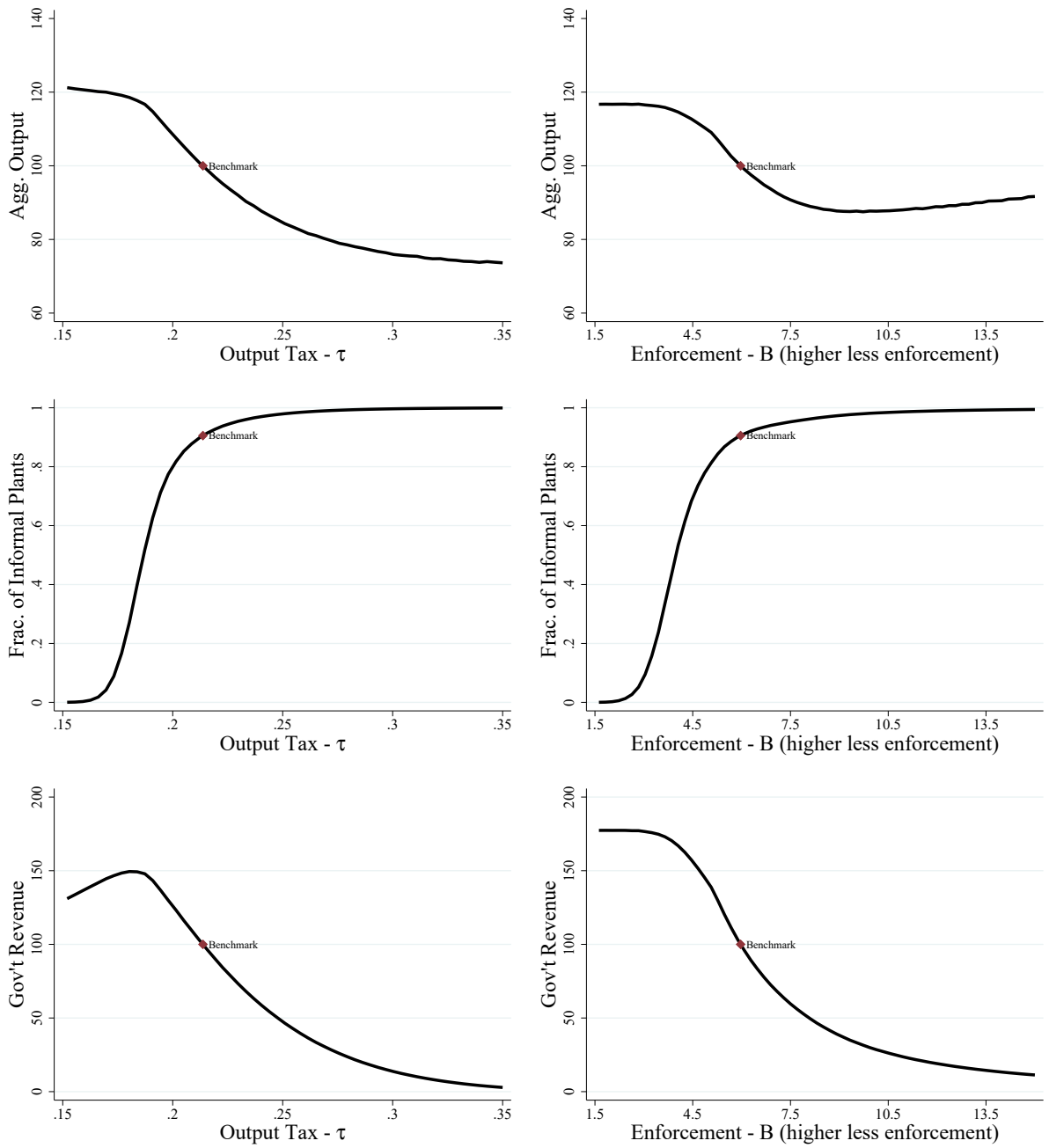
Notes: These figures show the effects of development on the informal mean size, the aggregate output, the fraction of informal plants. The aggregate output and the government revenue are normalized to 100 at the benchmark. For more details see Section 5.

Figure 6: Accounting Differences in the Mean Size of Informal Plants with Development



Sources: ISES and the author's own calculations.
Notes: This figure compares the model's prediction on the informal mean size as countries develop with the data. Each circle represents a country. Each diamond shows the model's prediction of the aggregate output and informal mean size with different λ values. For more details see Section 5.

Figure 7: Effects of Output Tax and Enforcement



Notes: This figures compares the effects of changes in output tax and enforcement levels. Figures in the first column presents the experiment results when with output tax, τ. Figures in the second column shows the results when the enforcement level (B) changes from the benchmark value. Smaller enforcement values indicates stricter enforcement policies. The aggregate output and the government revenue are normalized to 100 at the benchmark in all figures. For more details see Section 5.

Appendix

A Solving Managers' problem and the Equilibrium

In this section, I, first, present the first-order conditions (FOCs) of managers' and the household's problems. Then, I provide the definition of the equilibrium.

A.1 Solving the problem of a formal manager

Consider a formal manager's problem (6). FOCs with respect to n_j and k_j follow:

$$(1-\tau)\gamma\left[\left(\int_0^{M(z)}\left(n_j^{\alpha}k_j^{1-\alpha}\right)^{\frac{\sigma-1}{\sigma}}dj\right)^{\frac{\sigma}{\sigma-1}}\right]^{\gamma-1}\frac{\sigma}{\sigma-1}\left(\int_0^{M(z)}\left(n_j^{\alpha}k_j^{1-\alpha}\right)^{\frac{\sigma-1}{\sigma}}dj\right)^{\frac{1}{\sigma-1}}\frac{\sigma-1}{\sigma}\left(n_j^{\alpha}k_j^{1-\alpha}\right)^{\frac{-1}{\sigma}}\alpha n_j^{\alpha-1}k_j^{1-\alpha}=W \tag{A.1}$$

$$(1-\tau)\gamma\left[\left(\int_0^{M(z)}\left(n_j^{\alpha}k_j^{1-\alpha}\right)^{\frac{\sigma-1}{\sigma}}dj\right)^{\frac{\sigma}{\sigma-1}}\right]^{\gamma-1}\frac{\sigma}{\sigma-1}\left(\int_0^{M(z)}\left(n_j^{\alpha}k_j^{1-\alpha}\right)^{\frac{\sigma-1}{\sigma}}dj\right)^{\frac{1}{\sigma-1}}\frac{\sigma-1}{\sigma}\left(n_j^{\alpha}k_j^{1-\alpha}\right)^{\frac{-1}{\sigma}}(1-\alpha)n_j^{\alpha-1}k_j^{-\alpha}=R \tag{A.2}$$

It is straightforward to establish that $n_j = n_i$ and $k_j = k_i\ \forall\, j, i \in [0, M(z)]$ by comparing equations (A.1) and (A.2) with the FOCs with respect to n_i and k_i. Let n_F and k_F be the optimal labor and capital demand per activity of an informal manager. Using above FOCs, we can write k_F and n_F as a function of parameters and the informal manager's ability:

$$k_F = \left((1-\tau)\gamma\right)^{\frac{1}{1-\gamma}}\left(\frac{\alpha}{W}\right)^{\frac{\alpha\gamma}{1-\gamma}}\left(\frac{1-\alpha}{R}\right)^{\frac{1-\alpha\gamma}{1-\gamma}}M^{\frac{1-\sigma(1-\gamma)}{(\sigma-1)(1-\gamma)}} \tag{A.3}$$

$$n_F = \left((1-\tau)\gamma\right)^{\frac{1}{1-\gamma}}\left(\frac{\alpha}{W}\right)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}}\left(\frac{1-\alpha}{R}\right)^{\frac{\gamma(1-\alpha)}{1-\gamma}}M^{\frac{1-\sigma(1-\gamma)}{(\sigma-1)(1-\gamma)}} \tag{A.4}$$

A.2 Solving the problem of an informal manager

Let n_I and k_I be the optimal labor and capital demand per activity of an informal managers whose optimal capital demand is less than B. These type of informal managers' are not

going to get caught and their problem can be solved following the same way of solving formal manager's problem as above.

$$k_I = \gamma^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{W}\right)^{\frac{\alpha\gamma}{1-\gamma}} \left(\frac{1-\alpha}{R}\right)^{\frac{1-\alpha\gamma}{1-\gamma}} M^{\frac{1-\sigma(1-\gamma)}{(\sigma-1)(1-\gamma)}} \tag{A.5}$$

$$n_I = \gamma^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{W}\right)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}} \left(\frac{1-\alpha}{R}\right)^{\frac{\gamma(1-\alpha)}{1-\gamma}} M^{\frac{1-\sigma(1-\gamma)}{(\sigma-1)(1-\gamma)}} \tag{A.6}$$

On the other hand, there may be some informal managers who optimally choose exactly B amount of capital in total. I call them constrained informal managers. Similar to above managers' problems capital and labor demand is same across activities for constraint informal manager as well. Let k_C and n_{CI} denotes the optimal capital and labor demand for each activity by a constraint informal manager Then we can write the capital and labor demand for each activity as follow:

$$k_C = \frac{B}{M} \equiv b \tag{A.7}$$

$$n_C = \left(\alpha\gamma\right)^{\frac{1}{1-\alpha\gamma}} W^{\frac{-1}{1-\alpha\gamma}} b^{\frac{\gamma(1-\alpha)}{1-\alpha\gamma}} M^{\frac{1-\sigma(1-\gamma)}{(\sigma-1)(1-\alpha\gamma)}} \tag{A.8}$$

A.3 Equilibrium

Given prices, tax rate, enforcement level and transfers, $\{W_t^*, R_t^*, \tau, B, T_t^*\}_{t=0}^{\infty}$, the household maximizes her utility by choosing $\{C_t^*, K_{t+1}^*, S_{W,t}^*, S_{I,t}^*, S_{F,t}^*\}_{t=0}^{\infty}$ such that the allocation of the household members into occupations are consistent with solution of the top managers' problem, government budget is balanced and all markets clear:

- Labor market clearing condition:

$$\int_{S_{W,t}^*} z g(z) dz = \int_{S_{I,t}^*} n_{I,t}^*(z) g(z) dz + \int_{S_{F,t}^*} n_{F,t}^*(z) g(z) dz \tag{A.9}$$

  where $n_{I,t}^*$ and $n_{F,t}^*$ are demand for worker's ability by an informal and a formal manager respectively.

- Capital market clearing condition:

$$K_t^* = \int_{S_{I,t}^*} k_{I,t}^*(z) g(z) dz + \int_{S_{F,t}^*} k_{F,t}^*(z) g(z) dz \tag{A.10}$$

where K_t^* is the supply of capital and $k_{I,t}^*$ and $k_{F,t}^*$ are capital demand by an informal and a formal manager respectively.

- Goods market clears:

$$C_t^* + K_{t+1}^* = \int_{S_{I,t}^*} y_{I,t}^*(z) g(z) dz + \int_{S_{F,t}^*} y_{F,t}^*(z) g(z) dz + (1 - \delta + R_t) K_t^* \tag{A.11}$$

where $y_{I,t}^*$ and $y_{F,t}^*$ are output produced by an informal and a formal manager respectively and $\delta \in (0,1)$ is the deprecation rate of capital.

B List of Countries

Table B.1: Survey Years in ISES

country, year	Region	Registration	# of Obs.
Argentina, 2010	Buenos Aires, Resistencia (Chaco)	Administración Federal de Ingresos Públicos	384
Bangladesh,2022	Dhaka, Chittagong, Cox's Bazaar	National Board of Revenue	3,375
Botswana,2010	Gaborone	Central Statistics Office of Botswana	99
Burkina Faso,2009	Bobo-Dioulasso, Ouagadougou	Chambre de Commerce	120
Cameroon,2009	Douala, Yaounde, Bafoussam	Registre du Commerce	122
Cabo Verde,2009	Sal, Santiago, Sao Vicente	Registry of commerce or holding a Municipal License	129
Cote d'Ivoire,2008	Abidjan, San Perdo,Yamoussoukro	Chambre du Commerce et de lIndustrie	129
Congo, Dem. Rep., 2013	Western, Eastern, Center, Southern	Numero de Registre de Commerce and Identification Nationale	480
Ghana, 2013	Accra, Tema, Tekoradi, Tamale	Registrar's General Department	729
Guatemala, 2010	Ciudad de Guatemala,Quetzaltenango	Superintendencia de Administracion Tributaria	303
India, 2021	Tezpur,Mumbai,Hyderabad,Surat, Jaipur,Kochi,Varanasi,Ludhiana,Sehore	a business PAN or a GST number	10,672
Iraq, 2019	Baghdad,Basrah,Sulaymaniyah,Najaf	Directorate of Companies Registration of Federal Iraq	1,996
Kenya,2013	Nairobi, Nyanza,Mombasa,Nakuru,Central	Kenya Revenue Authority	533
Lao PDR,2019	Vientiane,Pakse	Department of Enterprise Registration	361
Madagascar,2008	Antananarivo, Mahajanga,Toamasina,Antsiranana	INSTAT or EDBM	127
Mali,2010	Bamako,Mopti,Segou,Sikasso	List of Registrar	120
Mauritius,2009	Mauritius	Register of Companies	132
Mozambique,2018	Maputo,Beira,Nampula	BAU or NUIT	554
Myanmar,2014	Yangon,Mandalay,Bago,Taunggyi,Monywa	DICA,Ministry of Industry,Department of Development Affairs	300
Nepal,2009	Central, West, East	Inland Revenue Department	120
Peru,2022	Lima,Trujillo	SUNAT	1,640
Rwanda,2011	Kigali,Butare	Rwanda Development Board	240
Somalia,2019	Mogadishu, Bosaso	Ministry of Commerce,Municipality	491
Sudan,2022	Khartoum, Omdurman, Port Sudan	Commercial Registrar	1,410
Zambia,2019	Lusaka,Kitwe,Ndola	PACRA,ZRA	914
Zimbabwe,2016	Harare	Registrar of Companies	515
Total: 25,995 observations in 26 countries			