Tax collection, informal sector, and productivity[1]

Julio C. Leal-Ordonez[1]

Centro de Investigación y Docencia Económicas (CIDE).

Carretera México-Toluca 3655 Col. Lomas de Santa Fe 01210 México, D.F.

Abstract

One important determinant of informality is the tax enforcement capacity of a country which, some authors argue additionally distorts firms' decisions and creates inefficiency. In this paper, I assess the quantitative effect of incomplete tax enforcement on aggregate output and productivity using a dynamic general equilibrium framework. I calibrate the model using data for Mexico, where the informal sector is large. I then investigate the effects of improving enforcement. I find that under complete enforcement, Mexico's labor productivity and output would be 17% higher. The source of this gain is the removal of distortions induced by incomplete enforcement of taxes. These distortions affect three margins in the economy: reduce capital-labor ratios of informal establishments; allow low-productive entrepreneurs to enter; and misallocate resources towards low-productive establishments. As a result, TFP and capital accumulation are reduced, and hence output.

Keywords: tax enforcement, TFP, misallocation, informal sector

JEL: E23, E26, O17, O40

1. Introduction

Is the informal sector good or bad for an economy? Some authors have argued that firms operating in the informal sector are less regulated and less taxed than firms in the formal sector, which allows them to operate more

Email address: julio.leal@cide.edu (Julio C. Leal-Ordonez)

[1]This paper is an updated version of a chapter of my PhD dissertation. I would like to thank Richard Rogerson, Berthold Herrendorf and Ed Prescott for their advise. I would also like to thank David Lagakos, Roozbeh Hosseini, Galina Vershagina, Natalia Kovrijnykh and Manjira Datta, for their comments and suggestions.

efficiently. This, represents a positive force for the economy (see Schneider & Enste, 2002). In contrast, other authors have highlighted distortions that might arise in the presence of a large informal sector and specifically of incomplete enforcement of taxes. For example, Lewis (2004) argues that informality distorts the "natural" competitive process as informal firms enjoy of an "unfair" cost advantage through tax avoidance; Farrell (2004) reports that informal firms reduce their scale of operation in order to remain undetected by the government, which makes them less efficient; and Levy (2008) states that informality implicitly subsidizes employment in less productive activities.

In this paper, I am interested in quantifying the extent to which distortions associated with the way firms avoid taxes affect output and productivity. To do this, I develop a general equilibrium model of occupational choice and capital accumulation that includes a taxation policy with limited enforcement. In this framework, incomplete tax enforcement is the source of plant idiosyncratic distortions similar to those studied by Guner, Ventura, & Xu (2008) and Restuccia & Rogerson (2008). Individuals have heterogeneous entrepreneurial abilities (as in Lucas, 1978) and each faces a discrete occupational choice: whether to be a formal entrepreneur, an informal entrepreneur or an employee. If formal, the entrepreneur pays taxes; if informal, the entrepreneur faces a probability of being detected that depends positively on the amount of capital hired. Therefore, only small firms are able to evade taxes because it is more difficult for the government to detect them.

The novelty in this paper is that it connects informal sector data for a typical developing country with a general equilibrium model in which the consequences of incomplete enforcement of taxes can be studied. I calibrate the model using data for Mexico, an economy where 31% of the employees work in informal establishments. I then investigate the effects of improving enforcement. My main finding is that under complete enforcement, Mexico's labor productivity and output would be 17% higher. These gains come in part from a novel mechanism through which better enforcement, and therefore higher taxes (for some establishments), rises capital, as the benefits of staying small fall.

To understand the distortionary aspects of incomplete enforcement, it is important to look at two key features of the equilibrium. The first is that entrepreneurs with low productivity choose informality, while the more productive

ones choose to operate in the formal sector. The reason for this is that any firm below a certain threshold can avoid detection and so can operate informally and increase profits by avoiding taxes, at no additional cost. Since less productive entrepreneurs naturally choose a lower scale, tax avoidance acts as an implicit subsidy for less productive establishments. This feature induces two types of distortions: a misallocation of resources to establishments with low productivity; and an increase in the number of unproductive establishments in the economy.

The second feature of the equilibrium is that a significant group of informal establishments optimally reduces its scale to remain undetected by the government. This brings a distortion in the capital-labor ratio of informal establishments, because the probability of being detected rises with the amount of capital hired.

When complete enforcement is introduced, these burdens on productivity disappear. I find that removing these distortions would bring total factor productivity (TFP) and output up by 4% in the short run. Furthermore, I find that, in the long run there would be a 22% increase in the capital stock and an 11% rise in output.

There would also be a gain in the form of lower taxes. Under complete enforcement, the tax base is broadened, so revenues would remain constant even with a lower tax rate. This is precisely the core of Lewis' (2004) hypothesis, who argues that the combination of big government and incomplete enforcement creates the need to impose high taxes on the most productive part of the economy. I find that Mexico could lower taxes from a rate of 26% to one of 16% under full enforcement. This reduction would further increase output to a level 17% higher than that of the benchmark economy with informality.

Another important contribution of this paper is to document two important facts that emerge when comparing the establishment size distribution of employment in Mexico and the US. First, I document that Mexico allocates much more labor into small and large establishments than what US does. This makes looks Mexico as part of the group of countries with the "missing middle" documented by Tybout (2000). Then, I construct distributions in Mexico for both formal and informal establishments and compare them to the US distribution. When the Mexican formal sector distribution is compared to the US one it is evident that the two look very similar. Put it in other words, what this shows is that most of the differences between the establishment size distributions of employment in Mexico and the US is related to the existence of a large informal sector in Mexico. In this paper it is shown that

these differences are in part explained by incomplete enforcement of taxes.

My paper relates to the literature in the following way. First, the burdens on productivity associated with informality can be understood as a specific case of the type of idiosyncratic distortions studied in the literature on resource misallocation across heterogeneous plants and TFP, identified in the recent work of Restuccia & Rogerson (2008), Guner *et al.* (2008), and Hsieh & Klenow (2007). The first two use the US as their benchmark and impose hypothetical policies that affect the prices faced by individual establishments, while the third one studies the cases of China and India. My paper takes the Mexican case and concentrates on a specific policy that distorts the prices faced by individual producers.

Because of the link between size and detection probabilities in my model, incomplete tax enforcement is similar to a size-dependent policy of the type in Guner *et al.* (2008). Along the same lines, there is the case study by Gollin (1995) for Ghana, analyzing the impact of taxes on large establishments on productivity. One important difference between my work and that of Guner *et al.* (2008) and Gollin (1995) is that the policies considered by them do not distort the capital-labor ratios of establishments. As mentioned above, this margin is distorted by incomplete enforcement of taxes and turns out to be quantitatively important.

Second, there are models where the informal sector arises from incomplete enforcement of taxes and/or regulations: Rauch (1991), Amaral & Quintin (2006), Dabla-Norris *et al.* (2008), and de Paula & Scheinkman (2007), among others. However, these authors focus mainly on the determinants of informality rather than on the consequences of incomplete enforcement.

Finally, de V. Cavalcanti & Antunes (2007) and Moscoso-Boedo & D'Erasmo (2009) study the aggregate effects of informality within the context of GE models. However the focus of these papers is not on the way firms avoid taxes, but on other distortions associated with informality such as debt enforcement and the regulation of entry. It is worth noticing that the models in these papers (including my own) do not intend to provide a full explanation of informality and therefore each of them captures a specific aspect of informality. I focus on the taxation aspects of informality because these have been relatively less studied than the regulation and entry costs aspects emphasized originally by de Soto (1989) (see for example Herrendorf & Teixeira, 2011).

The paper is organized as follow. Following this Introduction, Section 2 presents data documenting relevant facts about the informal sector and the resource allocation in Mexico; Section **??** presents an overview of the rest of the paper and the main goals; Section 3 presents the model, while Section 4 characterizes the steady state equilibrium; in Section 5, I calibrate the model, and in 6, I present the results. The last section contains my conclusions.

2. Facts

In this section, I document the following facts: 1) the informal sector in Mexico is large, 2) the distribution of labor across establishments sizes in Mexico differs from what is found in the US, and 3) informal establishments are small. Additionally, I rely on other studies to document that informal establishments operate with smaller capital-labor ratios than their formal counterparts.

To address these inquiries, I use a number of household surveys and a census of establishments. I have access to microdata from the National Urban Employment Survey (ENEU, by its Spanish acronym), which I use more intensively. This survey is not only helpful in determining the size of the informal sector in terms of employees, but it also allows me to examine the size of the establishments referred to. Additionally, I use other surveys and a census to complement the information in ENEU and to make comparisons with US data.

2.1. The size of the informal sector

Although the ENEU's main goal is to measure unemployment, I take advantage of a question addressing whether the surveyed employee is enrolled with the Mexican Social Security Institute (IMSS) or not. As in many studies on the informal sector literature, I classify an employee as informal if she/he is not enrolled with IMSS and as formal otherwise. Using this measure, I find that 31% of employees work in the informal sector (see Table 1).The percentage corresponds to ENEU's survey in the third trimester of 2002. This percentage has not changed considerably during previous years. Furthermore, Levy (2008) reports a similar figure using a different methodology that combines establishment data from the Economic Census and IMSS registries.

Table 1: Size of the Informal Sector (ENEU 2002_3)

Sector	Share
Formal Employees	69.26
Informal Employees	30.74

2.1.1. Laws

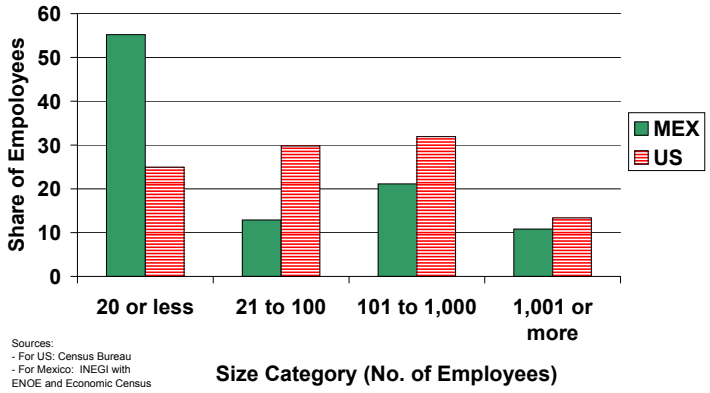
In this section, I argue that informality is closely associated with lack of tax law enforcement. According to Mexican law, all employers must be registered with IMSS and must register their employees as well. Employers who do not follow this mandate are able to avoid the payment of IMSS contributions for their employees. The advantage of using the definition above is that the ENEU asks to employees and not to employers. Since employers are the ones who bear the obligation of enrolling with IMSS, employees do not have incentives to miss report their registration status.

The ENEU survey, however is not useful in determining if the employer that avoids IMSS registration also evades other kind of taxes and regulations. However, with the help of the micro-business survey (ENAMIN), I conclude below that this is more likely than not. To make this point, however, I must first study the characteristics of establishments in the informal sector, which I do later in the next subsection.

2.2. Establishment size distribution

Here, I present data on the distribution of Mexican establishments by size, including both the formal and the informal sectors. I compare this to the distribution in the US, which I take as a relatively undistorted economy. Then, as a second step, I repeat the comparison using only information on the Mexican formal sector. For that matter, I focus on the labor allocation across establishments of different sizes in the non-agricultural sectors of the economy. The distribution of employees across similarly-sized establishments in Mexico and the US is shown in Figure 1.

Figure 1: Non-Agricultural Employment Distribution, 2003.



I have been careful to compare data in both countries that share a similar observation unit, and size categories. For the case of US, I used the data from the U.S. Census Bureau, which is based on information directly collected from employers. The 2009 Statistical Abstract reports the employment size distribution of establishments for several years. This information is only available for establishments with at least one employee.

For the case of Mexico, I used the recently created National Employment and Occupation Survey (ENOE) which is household based, and took advantage of a question that asks the size of the establishment in which the surveyed person works. The distribution obtained from this survey is comparable to the US one, to the extent that the employees report the size of the establishment with the same accuracy as their employers. This problem is somewhat mitigated by the use of broadly defined size categories. Alternatively, I could have used the Mexican Economic Census, which is based on information collected from employers. Unfortunately, that Census does not include establishments that do not operate out of fixed premises, which in the case of Mexico is not a negligible group. In contrast, the ENOE is constructed to include this type of establishment.

One final issue I had to address was the definition of size categories. The ENOE does not report size categories comparable to the US in the right tail, and these are of some importance not only for the current comparison, but especially for later exercises in this paper. Since virtually all large establishments operate out of fixed premises, I do not expect the Census and ENOE to differ much from each other in the right tail. So to obtain the full distribution

for Mexico, I used the Census information to complement the ENOE. For more details on how these two sources were combined, see the Appendix[3].

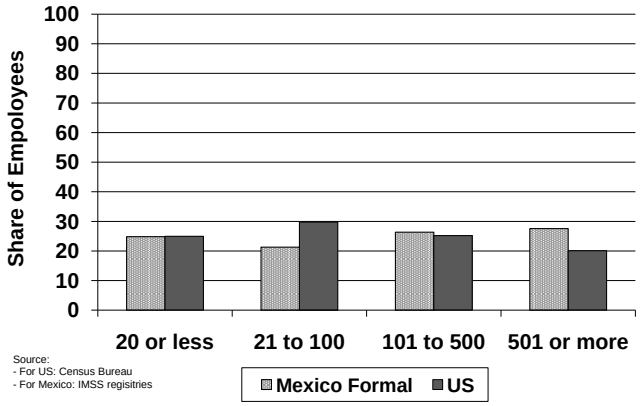
In Figure 1, the height of the bar represents the proportion of employees in each size category. It is clear that in Mexico, more labor is concentrated among establishments with less than 20 employees. While in Mexico around 55% of the employees are employed in these small establishments, the figure is only 25% in the US. The opposite is the case with larger establishments. Hence, when compared to the US, it is clear that Mexico allocates much more resources in small establishments.

To shed some light on how the existence of informal establishments affects size distribution in Mexico, I examine the distribution of labor in the formal sector alone, and compare it against the US. There are two data sets that I could use to look at the distribution in the formal sector. One is the IMSS registries (available in Levy (2008)), and the other is the ENEU household survey for which I have microdata. I present both calculations.

The advantage of using the IMSS registries is that the size categories are comparable to the categories in the US data; the disadvantage however, is that the observation unit is not the establishment. The IMSS registry units correspond to an employer ID number provided by the same institution. This ID number does not map exactly to either establishments or firms (see Appendix). Figure 2 presents the distribution of labor calculated from IMSS registries alongside distribution of employees in the US (as in Figure 1). An examination of this figure shows that the two distributions are quite similar. It follows that most of the informal employees work in small establishments, or in other words, that informal establishments are small.

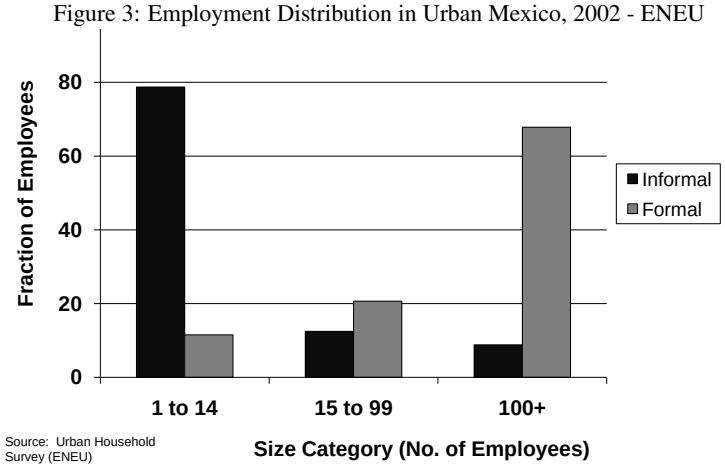
[3]In fact, the size categories in Figure 1 differ in the following way. For US the categories are: under 20, 20 to 99, 100 to 999, 1000 or more; for Mexico the categories are: 20 or less, 21 to 100, 101 to 1001, 1001 or more.

Figure 2: Non-Agricultural Employment Distribution, 2003 -IMSS registries



The size distribution of employees in the Formal sector can also be calculated from the ENEU household survey. The size categories in ENEU are not comparable with the US data, so Figure 3 compares the Formal sector distribution with the informal sector distribution. The broad picture is similar what we see using IMSS registries: most of the informal employees are working in small establishments, and the distribution within the formal sector has more mass in the upper tail as in the US case.

Figure 3: Employment Distribution in Urban Mexico, 2002 - ENEU



2.2.1. Micro-business Survey

Since I use a definition of informality that depends on the registration status with IMSS, this may raise some concerns regarding the status of informal establishments with respect to other institutions. I next address this concern using a Micro-business survey (ENAMIN).

The ENAMIN survey is a by-product of the ENEU survey. Its basic purpose to record the characteristics of Mexican micro-businesses. ENAMIN focus only on employed persons who in ENEU reported being either own-account workers or employers hiring 6 or less employees. The questionnaire asks whether the business owner is registered with the Ministry of Finance and Public Credit (SHCP), the Mexican federal tax authority. I present the proportion of businesses not registered with the SHCP by establishment size in Table 2. The same picture appears again: the smaller the establishment, the more likely it is to not be registered with tax authority.

Table 2: Under-registration with the Federal Treasury

ENAMIN Data 2002	
Size Category	% not registered
1	0.78
2	0.52
3	0.39
4	0.25
5	0.16
6	0.17

The ENAMIN is also useful because it allows us to look at other characteristics of firms avoiding registration besides their size. The literature provides evidence that informal establishments operate with a lower capital-labor ratio than their formal counterparts. For example, Thomas (1992), reports this to be the case for a survey of Peruvian establishments. Although I am unable to access data on capital per worker for Mexican informal establishments, I used the ENAMIN to examine the differences in the use of capital between informal and formal establishments. This information is summarized in Table 3. In particular, I note that 81% of businesses in the ENAMIN that report not being registered with SHCP also report the absence of fixed structures (a physical property or facility) where they perform their productive activities. Even for employers only, the percentage is still high (74%).

Table 3: Percentage w/o fixed structures (ENAMIN, 2002)

	Not Registered	Registered
All Businesses	81%	24%
Only Employers	74%	14%

3. Economic Environment

The previous section documented three facts. First, 31% of Mexican employees work in the informal sector; second, the distribution of Mexican employees by company size allocates much more mass to small establishments than the US distribution; and the third, that most informal establishments are small. Additionally, the literature has shown that informal establishments operate with smaller capital-labor ratios than their formal counterparts.

The goal of this paper is to investigate the extent to which the size distribution of establishments in Mexico is the result of a distortion induced by incomplete enforcement, and to assess the consequences of this distortion for productivity. With this in mind, I build a model with heterogeneous entrepreneurial abilities and a taxation policy with limited enforcement. This policy links the probability of detection to the amount of capital hired by the tax evader. This will lead to an endogenously determined informal sector where establishments with low productivity sort into informality. This specification captures the fact that smaller establishments are more likely to be informal and that they also show a smaller capital-labor ratio.

There is a representative household in this economy which is populated by a continuum of individuals (members) of mass 1 as in Guner, Ventura, & Xu (2008). At period zero, the household is endowed with K_0 units of capital while each member is endowed with entrepreneurial ability $z \in [\underline{z}, \bar{z}]$. This entrepreneurial ability is distributed according to pdf $g(z)$ and cdf $G(z)$ and it does not evolve over time. Additionally, individuals have 1 unit of time each period.

The household has preferences over a sequence of consumption goods defined by:

$$\sum_{t=0}^{\infty} \beta^t u(C_t) \tag{1}$$

Where C_t is the consumption in period t. The household accumulates capital by making investments X_t, and as

is standard, the accumulation is determined by the following rule:

$$K_{t+1} = X_t + (1-\delta)K_t.$$

Each household member z can have one out of three alternative occupations: entrepreneur in the formal sector, entrepreneur in the informal sector or employee in the formal or the informal sector.

Regardless of formality status, if an individual with ability z is an entrepreneur, she has access to the technology $f(z,k,l) = zk^{\theta_k}l^{\theta_l}$ and $0 < \theta_k + \theta_l < 1$, and I define $\gamma = \theta_k + \theta_l$. This technology exhibits decreasing returns to scale ensuring the coexistence of establishments with heterogeneous productivities (as in Lucas (1978)). If, on the other hand, the individual is an employee, the individual supplies 1 unit of labor which yields income w, independently of the value of z.

A government levies a tax τ_y on establishment's output, and the revenue is given back to the household as a lump sum transfer. An output tax is equivalent to simultaneously taxing labor, capital and entrepreneurial profits (before taxes). The implicit assumption is that these three margins are taxed at the same rate. Later, I analyze how deviations from this assumption affect the results of the experiments performed.

Taxes can be avoided by operating in the informal sector, but tax avoidance comes with a cost. In particular, I assume that informal entrepreneurs face a probability of detection and hence punishment. Once caught, the member will be given a fresh start in the next period and will be able to work as any of the three occupations available. The specification of the probability of detection will be referred as the enforcement function. I focus on a function that depends on the amount of capital hired in the establishment. Later in the paper I assume that the probability of detection could depend alternatively on the labor hired or the output produced. Perhaps not surprisingly, the results show that when the enforcement policy depends on capital, the negative effects of incomplete enforcement on accumulation are larger. The following is assumed:

$$p(k(z)) \quad = \quad \begin{cases} 0, & k(z) \leqslant b \\ & \\ 1, & else \end{cases}, \tag{2}$$

where $k(z)$ is the capital hired by entrepreneur z and $b > 0$.

This enforcement policy gives informal entrepreneurs the opportunity of choosing to operate with a capital level equal to b or lower, low enough not to get caught by the government, while still enjoying the benefits of tax avoidance.

The stepwise specification might at first glance seem too restrictive, so I should take a moment to comment on the advantages and disadvantages of this choice. Note first that the stepwise specification is increasing, and therefore is consistent with the facts in section 2: only small establishments have an incentive to choose informality because they can evade taxes at no cost. Additionally, some establishments might choose to reduce their scale in order to remain undetected. Second, note that there is never going to be the case that in equilibrium an individual decides to operate informally and at the same time chooses capital larger than b, otherwise it will be caught and punished. This also means that no one gets caught in equilibrium (which, some might agree describes well government skills), and furthermore, it means that we don't have to worry about the size of the punishment as long as ts level is set high enough to reduce informal profits (if detected) to a level below formal profits. For purposes of completeness, the punishment is set equal to the current period earnings.

Third, in terms of the equilibrium characterization of the occupational choices, this specification and any other that includes a strictly increasing probability of being caught are equivalent. Both will characterize occupational choices with two thresholds in the range of entrepreneurial abilities z (see section 4). The stepwise specification chosen, however, has a clear advantage in terms of computational burden; it avoids the need to numerically solve a nonlinear system of equations for each point of my grid. Finally, note that the stepwise specification do affects the distortion suffered by informal establishments in their capital-labor ratios (this will be clear in a moment). In the absence of good data on this, I use the simplest specification for convenience.

3.1. Earnings for alternative occupations

I will now analyze earnings for alternative occupations in more detail. As mentioned above an individual can have one out of the three possible occupations: employee, informal entrepreneur, formal entrepreneur. I assume employees are free to move across sectors and therefore a member working as an employee will simply earn wage w regardless of the sector in which she is employed.

A formal entrepreneur with entrepreneurial ability z maximizes profits according to:

$$\pi_F(z;w,r) = \max_{\{l_F,k_F\}} \left\{ (1-\tau_y) z k_F^{\theta_k} l_F^{\theta_l} - w l_F - r k_F \right\}, \tag{3}$$

where w is the wage rate and r is the price of capital. $k_F(z,w,r)$ and $l_F(z,w,r)$ denote the optimal choices of capital and labor respectively in the problem above.

An entrepreneur in the informal sector maximize expected profits taking into account the probability of detection mentioned in the previous section:

$$\pi_I(z;w,r) = \max_{\{l_I,k_I\}} \left\{ (1-p(k_I)) \left(z k_I^{\theta_k} l_I^{\theta_l} - w l_I - r k_I \right) \right\}.$$

$k_I(z,w,r)$ and $l_I(z,w,r)$ denote the optimal choices of capital and labor respectively. Note that, as mentioned above, it is not optimal for any informal entrepreneur to operate with capital greater than b (otherwise her profits will be zero). However she could choose to operate with capital equal to b, just low enough to prevent detection by the government, while still enjoying the benefits of tax avoidance. Therefore the profits of an entrepreneur in the informal sector can also be expressed as:

$$\pi_I(z;w,r) = \max_{\{l_I,k_I\}} \left\{ z k_I^{\theta_k} l_I^{\theta_l} - w l_I - r k_I \right\} \quad s.t. \quad k_I \leqslant b \tag{4}$$

Once occupations are defined for each z (occupational choices are described below), total household income is given by:

$$
\begin{aligned}
E(w,r) &= \int_{\underline{z}}^{\bar{z}} I(z)\pi_I(z;w,r)dG(z) \\
&+ \int_{\underline{z}}^{\bar{z}} F(z)\pi_F(z;w,r)dG(z) + \int_{\underline{z}}^{\bar{z}} w\left(1-I(z)-F(z)\right)dG(z)
\end{aligned} \tag{5}
$$

where $F(z)$ and $I(z)$ are index functions and equal 1 if the occupation for individual z is formal or informal entrepreneur, respectively and zero otherwise. Similarly, let the index function, $I^c(z)$ equal 1 for the case when an informal entrepreneur is constrained (i.e. $k_I(z,w,r) = b$), and zero otherwise.

3.2. Government

In the present model, the government obtains revenue from two different sources: tax revenues and enforcement penalties. It turns out that given the nature of the enforcement policy, penalty revenues will be zero in equilibrium. I assume a balanced budget for the government in every period so that all proceeds from government activities are given back to the household in the form of a lump sum transfer. The government budget balance condition is:

$$
R_t = T_t, \ \forall t \tag{6}
$$

where R_t is tax revenue.

3.3. Representative household problem

The household chooses sequences of consumption, capital and each member's occupation taking as given the price sequences $\{w_t, r_t\}$, taxes τ_y, transfers $\{T_t\}$ and enforcement parameter b to maximize lifetime utility. The problem is:

$$
\max_{\{C_t, K_t, I_t(z), F_t(z)\}} \left\{ \sum_{t=0}^{\infty} \beta^t u(C_t) \right\} \tag{7}
$$

Subject to the following budget constraint:

$$
C_t(z) + K_{t+1} - (1-\delta)K_t = r_t K_t + E(w_t, r_t; \tau_y, b) + T_t, \forall t \tag{8}
$$

where K_0 is given and $E(w_t, r_t; \tau_y, b)$ is the same as in 5. I use $I(z; w, r)$ and $F(z; w, r)$ to represent occupational optimal decisions.

I focus on the steady state (SS) equilibrium of this economy. As standard, the first order conditions of this problem in the steady state imply that:

$$r = \frac{1}{\beta} - (1 - \delta) \tag{9}$$

3.4. Market Clearing

The Market clearing condition for the labor market will equate the aggregate labor demand from the two sectors to labor supply:

$$\int_{\underline{z}}^{\bar{z}} I(z; w_t, r_t) l_I(z; w_t, r_t) dG(z) + \int_{\underline{z}}^{\bar{z}} F(z; w_t, r_t) l_F(z; w_t, r_t) dG(z) = \int_{\underline{z}}^{\bar{z}} W(z; w_t, r_t) dG(z) \tag{10}$$

where $W(z; w_t, r_t) = 1 - I(z; w_t, r_t) - F(z; w_t, r_t)$. Market clearing for the capital and good markets are, respectively:

$$\int_{\underline{z}}^{\bar{z}} I(z; w_t, r_t) k_I(z; w_t, r_t) dG(z) + \int_{\underline{z}}^{\bar{z}} F(z; w_t, r_t) k_F(z; w_t, r_t) dG(z) = K_t,$$

and,

$$C_t + K_{t+1} - (1 - \delta) K_t = \int_{\underline{z}}^{\bar{z}} I(z; w_t, r_t) y_I(z; w_t, r_t) dG(z) + \int_{\underline{z}}^{\bar{z}} F(z; w_t, r_t) y_F(z; w_t, r_t) dG(z).$$

3.5. Equilibrium Definition

An equilibrium for this economy is sequences $\{C_t, K_{t+1}, w_t, r_t\}$ and $\{I_t(z), F_t(z)\} \forall z \in [\underline{z}, \bar{z}]$, such that taking factor prices $\{w_t, r_t\}$, policies parameters τ_y and b, and transfers $\{T_t\}$, the household solves her problem, firms maximize profits $\forall t$, and markets clear $\forall t$.

3.6. Steady State

In the following section, I will focus on the steady state equilibrium. Because I define time-invariant taxation and enforcement policies, the dynamic part of this economy is no different from the one in the standard growth model. In the steady state, factor prices, occupational decisions, aggregate capital and output are constant over time.

4. Model Properties

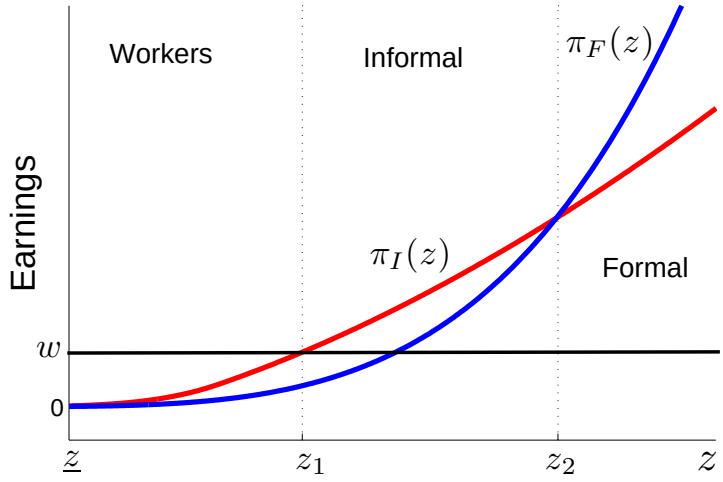
4.1. Occupational choices

In this section I analyze some properties of the model. The steady state equilibrium is characterized by three thresholds $\{z_1, z_c, z_2\}$ that summarize the occupational decisions of the agents and whether the capital choices of informal entrepreneurs are constrained or unconstrained. Figure 4 on the following page provides a graph of the optimal occupational choices. I study the determination of these thresholds next.

In a steady state equilibrium with positive formal and informal sectors, there exists thresholds $\{z_1, z_c, z_2\}$ such that:

1. $\forall z \in [\underline{z}, z_1)$ individuals decide to be employees;
2. $\forall z \in [z_1, z_2)$ individuals are informal entrepreneurs;
3. $\forall z \in [z_2, \bar{z}]$ individuals are formal entrepreneurs;
4. when $z_c > z_1$individuals $z \in (z_c, z_2)$ are constrained informal entrepreneurs; and when $z_c \leqslant z_1$ all informal entrepreneurs are constrained.

Figure 4: Characterization of Occupational Decisions



The determination of thresholds z_1 and z_2 is quite intuitive. For z_1, note that employees earnings do not vary with entrepreneurial ability (black line in Figure 4), while entrepreneurial profits increase with z (lines red and blue). Not surprisingly there exists threshold z_1 such that $w = \pi_M(z_1; w, r)$, where $\pi_M(z; w, r) = \max\{\pi_I(z; w, r), \pi_F(z; w, r)\}, \forall z$. It follows that all agents with $z < z_1$ will become employees and the rest entrepreneurs. For z_2, one must recall that informal establishments can not operate with capital larger than b, this restriction is more costly for entrepreneurs with larger z, because they naturally would like to choose a large scale according to their large productivity. It turns out that there exists a threshold z_2 such that $\pi_I(z_2; w, r) = \pi_F(z_2; w, r)$. It follows that entrepreneurs with ability $z < z_2$ would prefer the informal sector and the rest prefer the formal sector.

The optimal choices of formal entrepreneurs are quite standard:

$$k_F(z, w, r) = ((1 - \tau_y)z)^{\frac{1}{1-\gamma}} \left(\frac{\theta_l}{w}\right)^{\frac{\theta_l}{1-\gamma}} \left(\frac{\theta_k}{r}\right)^{\frac{1-\theta_l}{1-\gamma}}, \tag{11}$$

$$l_F(z, w, r) = ((1 - \tau_y)z)^{\frac{1}{1-\gamma}} \left(\frac{\theta_l}{w}\right)^{\frac{1-\theta_k}{1-\gamma}} \left(\frac{\theta_k}{r}\right)^{\frac{\theta_k}{1-\gamma}}, \tag{12}$$

and therefore maximum profits can be expressed as a function of prices and parameters:

$$\pi_F(z,w,r) = (1-\gamma)((1-\tau_y)z)^{\frac{1}{1-\gamma}} \left(\frac{\theta_l}{w}\right)^{\frac{\theta_l}{1-\gamma}} \left(\frac{\theta_k}{r}\right)^{\frac{\theta_k}{1-\gamma}}. \tag{13}$$

One less standard feature of the model is related to the presence of the informal sector. Again, some entrepreneurs in the informal sector will be better off hiring capital equal to b, just low enough to avoid detection. The threshold z_c is defined so that all informal entrepreneurs with $z < z_c$ operate unconstrained with $k_I(z,w,r) < b$ while all those $z \geq z_c$ operate constrained, i.e. $k_I(z,w,r) = b$. To illustrate this, consider an entrepreneur z in the informal sector for whom $k_I(z,w,r) < b$. The optimal capital demand for this entrepreneur will be identical to the one given by equation (11) but replacing $\tau_y = 0$. Note that, the monotonicity of this demand function with respect to z ensures the existence of the threshold z_c as defined above. Hence, the optimal informal profits are expressed in terms of prices and parameters only by:

$$\pi_I(z;w,r) = \begin{cases} (1-\gamma)z^{\frac{1}{1-\gamma}} \left(\frac{\theta_l}{w}\right)^{\frac{\theta_l}{1-\gamma}} \left(\frac{\theta_k}{r}\right)^{\frac{\theta_k}{1-\gamma}}, & k_I(z,w,r) < b \\ (1-\theta_l)z^{\frac{1}{1-\theta_l}} \left(\frac{\theta_l}{w}\right)^{\frac{\theta_l}{1-\theta_l}} b^{\frac{\theta_k}{1-\theta_l}} - rb, & k_I(z,w,r) = b \end{cases}. \tag{14}$$

With this in hand, we can provide a more formal argument for the existence of threshold z_2. Note first that both the informal and formal entrepreneurs profits are increasing convex functions of z (the exponent of the entrepreneurial ability is $\frac{1}{1-\gamma} > 1$). Second, note that by comparing equation (13) and the top case of equation (14), it is clear that at least for all $z \leqslant z_c$, informal profits are larger than formal profits. This is trivially true for other entrepreneurs to the right of z_c. Finally, note that $\frac{1}{1-\gamma} > \frac{1}{1-\theta_l}$ and hence as $z \to \infty$, $\pi_F(z;w,r) > \pi_I(z;w,r)$. This implies the existence of a threshold z_2 such that $\pi_I(z_2;w,r) = \pi_F(z_2;w,r)$ provided that $b > 0$ and $\tau_y > 0$ is not too large.

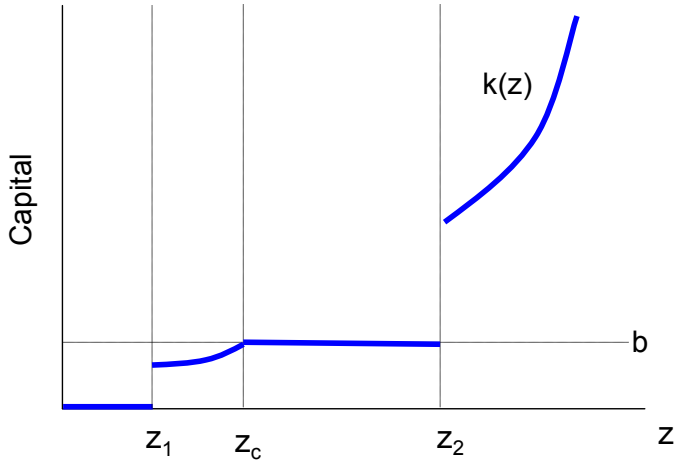
Furthermore, notice that in order to have a steady state equilibrium where both the informal and formal sectors are positive, $b > 0$ must not be too small and $\tau_y > 0$ must not be too large. When τ_y is large, the profits in the formal sector remain below the profits in the informal sector across the full range of existing entrepreneurial abilities $[\underline{z}, \bar{z}]$. This means all entrepreneurs would become informal. For example, in the case $\tau_y = 1$ formal sector profits are

zero for all $z \in [\underline{z}, \bar{z}]$, and therefore when $b > 0$, all entrepreneurs are informal. Similarly, when $b = 0$, profits in the informal sector are zero regardless of ability level, and all entrepreneurs become formal if $\tau_y < 1$. For intermediate cases, the the informal sector size will be positive provided that $b > 0$ is not too small, otherwise profits in the informal sector could remain low for all agents when compared to both employee earnings or formal profits. Finally, note that if in equilibrium both the informal and the formal sectors are positive, it must be that not all of the informal entrepreneurs are unconstrained; otherwise the threshold z_2 would not exist.

4.2. Capital-labor ratios

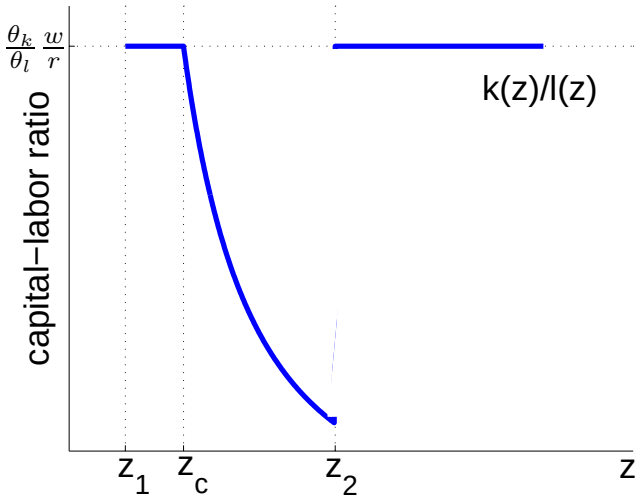
Given the nature of tax enforcement, capital-labor ratios of constrained informal entrepreneurs are distorted in equilibrium. First notice that the capital demand schedule has a discontinuity (see Figure 5). As note above, if both sectors are positive, it must be that the more capable informal entrepreneurs are constrained. To illustrate this, consider the entrepreneur z_2 who is indifferent between the two sectors. If informal, she would hire b capital, if formal, she would hire an amount strictly larger than b. To see why, note that optimal decisions of formal entrepreneur z_2 are the same as those of a hypothetical entrepreneur that operates unconstrained and pays no taxes; this entrepreneur is $z_h = (1 - \tau_y) z_2$. Entrepreneur z_2 hires capital strictly larger than b as long as $z_h > z_c$, and this inequality holds because as shown in the bottom case of equation (14), $\pi_I(z; w, r)$ is strictly increasing.

Figure 5: Capital Profile



Notice that, as a corollary of this property, the labor demand schedule is strictly increasing with respect to z, in equilibrium. Finally, notice that the discontinuity in the capital schedule translates into an informal sector that appears less capital intensive. The capital-labor ratio is smaller for constrained establishments (see Figure 6), as is the capital-output ratio.

Figure 6: Capital-labor ratios



5. Calibration

In this section I will describe the calibration strategy. Since I target a developing country (Mexico), my strategy will be different from that followed by works that focus on developed economies, like Restuccia & Rogerson (2008) and Guner, Ventura, & Xu (2008). These assume that US has small distortions and the distortion-free scenario is used as a benchmark to study how deviations affect equilibrium variables. In the case of this paper, however, the distortions characteristic of the Mexican Economy prevent us from taking the same approach.

The parameters to calibrate are the tax rate paid by the formal sector, τ_y, the technology parameters θ_k, γ and depreciation δ, the discount rate β, the enforcement policy parameter, and the entrepreneurial ability distribution parameters. The enforcement policy used as a benchmark is described in equation (2), where the probability of detection depends on capital; therefore, only one parameter needs to be calibrated (b). Later, I consider alterna-

tive specifications of this policy. Entrepreneurial ability is assumed to follow a truncated Pareto distribution with parameters z_{min}, z_{max} and s. More specifically I assume that z has cdf:

$$G(z) = \frac{1 - \left(\frac{z_{min}}{z}\right)^s}{1 - \left(\frac{z_{min}}{z_{max}}\right)^s},$$

where $s > 0$ is the shape parameter and $z \in [z_{min}, z_{max}]$, with $0 < z_{min} < z_{max}$. I make this choice for two main reasons. The first is that the firm size distribution in the US has been reported to be well described by a Pareto distribution (Axtell (2001)). The second is more practical: a truncated Pareto is fully defined on an interval that I can link directly with the model objects $\underline{z}$ and $\bar{z}$.

I start with the value of the parameters for which I am able to provide an independent calibration, which are the exponent of capital in the production function θ_k and the depreciation rate δ.

I choose $\theta_k = .33$ for the following reasons. First, because it is the standard value used by a number of studies focusing on Mexico. For example, Bergoeing *et al.* (2001) use $\theta_k = .33$ to compute TFP series for Mexico; Solimano *et al.* (2005) perform growth accounting using $\theta_k = .35$ for several Latin American economies including Mexico; and Restuccia (2008) uses a value of $\theta_k = .28$ for a production function with decreasing returns to scale. Second, this value is consistent with the estimates of Garcia-Verdu (2005).

I choose $\delta = .05$ following Solimano *et al.* (2005) and Bergoeing *et al.* (2001) who use the same value for the deprecation rate. Additionally, this value is roughly consistent with time series data on investment and consumption of fixed capital in Mexico, as I will explain below.

Given the choices of θ_k and δ, I proceed to calibrate the remaining parameters in the model. In order to do this, I solve for the equilibrium as a function of these parameters and set the value of each of them so that the model replicates a number of features of the Mexican economy. These features are the ratio of total tax revenue to GDP, various moments of the size distribution of employment, the size of the informal sector and the aggregate capital-output ratio.

The data for the moments of the size distribution of employment and the size of the informal sector was described in Section 2. The data for the other two targets (the capital-output ratio and the revenue to GDP ratio)

has not been described before.

An assessment of the magnitude of the capital-output ratio is needed. For this matter, I use data on the consumption of fixed capital (as a fraction of GNI) from Indicators (2005), and take the average since 1980 (which I call d). This average is around 10%. The model counterpart of d is $\delta K/Y$. Since $\delta$and d are known, I solve for K/Y from this equation and obtain $K/Y = d/\delta = .10/.05 = 2$.

This value of the capital-output ratio is close to the one found in two independent studies that estimate the capital stock in Mexico. Hofman (2000) performs a disaggregated estimation by type. The implied capital-output ratio in his work is around 1.7. Restuccia (2008) uses data from the Penn World Tables to estimate the capital-output ratios of a number of Latin American countries. He finds a value of around 1.9 for Mexico.

As a check, I used the capital accumulation equation in the balanced growth path combined with data on investment and capital consumption to jointly calculate the capital-output ratio and the depreciation rate. Specifically, I take yearly data on Gross fixed capital formation (%GDP) and the consumption of fixed capital (%GNI) from Indicators (2005), and take averages since 1980; then I solve the following system of equations: (1): $(1+n)(1+g)(K/Y) = (1-\delta)(K/Y) + (I/Y)$, and (2) $\delta(K/Y) = d$. Where n and g are the annual population and technology growth rates respectively, and $d = 0.105$. I set $n = .02$ and $g = .025$, again using data since 1980. The two unknowns are (K/Y) and δ. I get $K/Y = 1.9$ and $\delta = .059$.

The ratio of government tax revenue as a fraction of GDP is calculated as follows: according to OECD.stat, total tax revenue in 2003 was \$1,312,246.9183 million pesos and according to INEGI the GDP in 2003 was 6,891,992.482 million pesos. This gives a ratio of 19% of GDP.

Notice that instead of making a full characterization of the Mexican tax code, I have summarized taxes into a single output tax. Later, I perform a sensitivity analysis to determine how the model's outcomes change under alternative tax types.

Note also that the discount rate β can not be calibrated in the usual way. The usual way would be to obtain the value of r from the FOC of the firms and then using this value in the Euler equation to determine β. In principle, one would think that the FOC of the formal establishments can be used to find the value of r; but for that I would need

an estimation of the capital-output ratio in the formal sector, which is not available. Mexico's National Accounts include an estimation of the informal sector, and since I used National Accounts data to estimate the K/Y ratio, I think of it as a ratio that includes the capital and the output from both sectors.

5.1. Matching Moments

The remaining parameters are τ_y, γ, z_{min}, z_{max}, s, b and β. The choice of z_{min} is, to some extent, arbitrary. This is because all individuals with entrepreneurial ability below the z_1 threshold become identical employees (their ability is transformed into 1 unit of labor). Therefore, what matters in equilibrium is the mass of individuals to the left of z_1. Once z_{min} is set, this mass is fully determined by the parameters that describe the distribution of entrepreneurial abilities.

The rest of the parameter values are obtained by matching moments of the plant size distribution, the capital-output ratio, tax revenue and the size of the informal sector measured in terms of the proportion of employees working there. In the model there is a weakly monotonic equilibrium relationship between the size of a productive unit in terms of the labor employed and its entrepreneurial ability (see corollary 4.2). I take advantage of this feature to calibrate the parameters of the entrepreneurial ability distribution, using the employment distribution of establishments across size categories as well as information on the average size of the units in each category[4]. The moments targeted are:

1. the average size of establishments in the economy,
2. the average size of establishments with more than 100 workers,
3. the proportion of workers in establishments with more than 100 workers,
4. the size of the informal sector,
5. the capital-output ratio, and
6. tax revenue in proportion to GDP.

[4]This procedure is close to those in Guner, Ventura, & Xu (2008) and Rubini (2009)

Note that by targeting the first three moments, I can also match their complements: the share of workers and the average size of establishments with 100 workers or less. How well I can match similar moments for more disaggregated size categories will depend solely on the structure imposed by the Pareto distribution. As I show below, the calibration yields estimated parameters that replicate the data fairly well even in highly disaggregated size categories, despite the fact that I do not target such moments. I present a summary of the calibration targets in Table 4.

Table 4: Calibration Targets Summary

Parameter	Target	Source
θ_k	capital share	Gollin (2002); Garcia-Verdu (2005)
δ	gross capital formation; consumption of fixed capital;	WDI, Solimano *et al.* (2005) and Bergoeing *et al.* (2001)
zmin	arbitrary	-
γ		
β	moments of distribution;	Matching
zmax	size of informal sector;	moments
s	capital-output ratio	
τ_y	Tax revenue/GDP	
b		

As part of the sensitivity analysis, I performed calibration exercises for a number of cases varying both the type and levels of taxes and the type of enforcement policy considered. Each exercise required an independent calibration, hence, different calibrated parameters emerged in each case.

5.2. Calibration Properties

The targeted moments are well matched as can be confirmed in Table 5, which presents data and model values.

Perhaps more interesting is the fact that the calibration yields parameters that replicate well a number of moments that were not targeted explicitly. In Table 6, the model is shown to replicate the mean size and the labor shares for a number of highly disaggregated size categories. This is an important check for the methodology used, because by replicating the allocation of labor across categories that differ in average size, I am in fact replicating

Table 5: Calibration Targets

Targeted Variables	Data	Model
K/Y	2.0	2.0
mean size	5.8	5.8
Informal Size (%)	30	30
Tax revenue/GDP	.19	.1966
Mean size by employment size category:		
more than 100	364	364.9
Worker share by employment size category:		
more than 100	.32	.32

labor allocation across productivity levels.

Table 6: Calibration: Non-targeted Moments

Size Category	Labor share		Mean Size	
(total workers)	Model	Data	Model	Data
2 to 5	0.4093	0.3875	2.6	2.6
6 to 10	0.0473	0.0965	8.6	7.4
11 to 15	0.0448	0.0463	12.7	12.7
16 to 20	0.0299	0.0217	17.9	17.8
21 to 30	0.0397	0.0313	24.8	24.9
31 to 50	0.0502	0.0406	38.6	38.9
51 to 100	0.0655	0.0569	69.6	70.7
101 to 250	0.0824	0.0843	152.4	155
251 to 500	0.0624	0.0692	346.1	344
501 to 1000	0.0597	0.0578	697.4	697.1
1001 or more	0.1123	0.108	1823.9	2218.5

The calibrated parameter values are presented in Table 7. Note that the value of γ is 0.7455–relatively low compared to the what was found in studies focused in the United States. In particular, Atkeson & Kehoe (2005) estimate a value of .85 for US manufactures. This is not necessarily a bad outcome as I will explain below. Another interesting outcome of the calibration is the value of the tax rate (26%). Notice that this is below the value of the tax revenue to GDP ratio (19%). This is consistent with the fact that informal establishments do not pay taxes, but do contribute to GDP.

Table 7: Parameter Values

Parameter	Calibrated Value
θ_k	.33
δ	.05
β	0.943
γ	0.7455
z_{min}	0.7
z_{max}	10.089
s	4.045
b	3.3
τ_y	.26

6. Results

Once the model is calibrated to the Mexican economy, I can investigate the negative effects of incomplete enforcement polices. To do this, I use as a benchmark this calibrated economy and perform three exercises: one that focus in short-run effects and a couple more that explore long-run effects. In these exercises, I introduce complete tax enforcement by making $b = 0$ in the model, this implies that all establishments would pay an uniform tax (τ_y).

As a first step, I look at the equilibrium in the first period, right after the new enforcement policy is introduced. In this period, the capital stock is the same as in the economy with incomplete enforcement, because accumulation has not occurred yet. Table 8 shows the value of aggregate variables in this context. Since capital and labor (mass 1) are no different from the benchmark economy, the only reason that output increases is that these resources are used more efficiently. Not surprisingly, output increases the same as TFP. Also, notice that in this first period, wages decrease.

The gains in TFP are associated with the removal of distortions present under incomplete tax enforcement. The effects of incomplete enforcement include distorting three margins: occupational optimal choices, allocation of resources across establishments, and capital-labor ratios of informal establishments. The first two distortions occur across establishments while the last one occurs within establishments. Incomplete tax enforcement distorts occupational choices because it makes more attractive entrepreneurship; also, it distorts the allocation of resources

Table 8: Short Run Effects:

First Period after change in Enforcement	
Variable	Value under Complete Enforcement relative to Benchmark
Y	1.04
τ_y	1
K	1
TFP	1.04
$Employee share$	1.11
$AvEntrepAb$	1.22
r	1.14
w	0.86

directly because it makes possible to have some establishments paying taxes and some others not; finally, it distorts the capital-labor ratio of a group of informal establishments because this is the optimal response to a probability of detection that increases with the capital in the establishment.

Note that under complete enforcement the probability of detection plays no important role and the capital-labor ratios are undistorted. Also notices that both, occupational choices and the allocation of resources across establishments are "efficient". This is true because all establishments are now formal and face the same idiosyncratic distortion (the tax). By "efficient" I mean that the allocation of resources is the same as the one in a version of the model with $\tau_y = 0$.

This last point deserves more explanation. Once every establishment is paying a uniform tax under complete enforcement, the value of the tax rate does not affect neither occupational choices nor the allocation of resources across establishments. To see this, lets write the relative labor demands for any two establishments z and z' facing the same tax rate τ_y:

$$\frac{l_F(z',w,r)}{l_F(z,w,r)} = \frac{((1-\tau_y)z')^{\frac{1}{1-\gamma}}\phi(w,r,\theta_k,\theta_l)}{((1-\tau_y)z)^{\frac{1}{1-\gamma}}\phi(w,r,\theta_k,\theta_l)} = \left(\frac{z'}{z}\right)^{\frac{1}{1-\gamma}},$$

which is independent of the tax rate and depends only on relative productivity and γ. Similarly, occupational choices under complete tax enforcement are fully described by the employee/entrepreneur threshold z_1 (remember,

that there is no informal sector in this case):

$$z_1 = \left[\frac{\theta_l}{1-\gamma} \frac{\int_{z_1}^{\bar{z}} z^{\frac{1}{1-\gamma}} g(z) dz}{G(z_1)} \right]^{1-\gamma},$$

which is also independent of the tax rate.

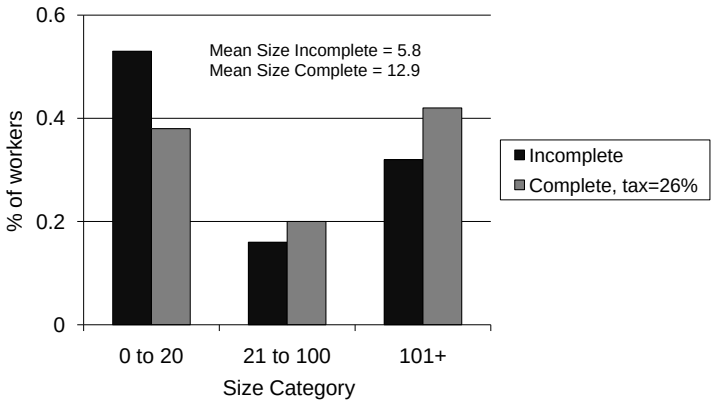
What this means is that the introduction of complete tax enforcement, makes occupational choices and resource allocation identical to the ones in a version of the model when $\tau_y = 0$. Put it differently, economies with complete enforcement look similar to the economy with $\tau_y = 0$ in two aspects: occupational choices, and allocation of resources across establishments. The aspect where these economies differ is the amount of capital accumulation, and therefore prices and output.

Therefore, the short-run gains in TFP respond to the elimination of the distortions mentioned above. Note first, that the employee/entrepreneur threshold z_1 increases because a group of low-ability individuals no longer find it attractive to be entrepreneurs, so the average entrepreneurial ability in the economy improves; second, note that resource allocation across establishments is now efficient, as explained above; and finally note that the capital-labor ratio displayed by informal establishments also improves.

In Table 8, it can be confirmed that the fraction of employees in the economy increases 11%, and the average managerial ability 22%, as a result of the change in threshold z_1. Note that consistent with this, wages decline to a level that is 0.86 of the benchmark level. Also note that the rental rate of capital increases 14% in this first period.

The effects on the efficient use of resources, are also captured by the labor reallocation across establishment size categories. This is presented in Figure 7. The improved enforcement policy reduces the allocation of resources to small establishments and increases the allocation to medium and large establishments. As a consequence, the mean size is almost doubled.

Figure 7: Labor Reallocation induced by Complete Enforcement



The next step is to look at lung-run effects of the introduction of complete tax enforcement. This exercise is presented in Table 6. Note that there is no more TFP gains beyond the ones that occur in the short-run, because occupational choices are not affected neither the allocation of resources. However, capital accumulation increases 22% relative to benchmark because capital is more productive relative to the case with incomplete enforcement. Finally, note that wages increase to 0.92 (still below the benchmark level) because workers have more capital to produce with.

The change in incentives to accumulate capital brought by complete enforcement are reflected in the 14% increase of the rental rate r in Table 8. This price change is just a consequence of the marginal productivity of capital going up, which is explained by the removal of the distortions faced by informal entrepreneurs and the better allocation of resources in the economy. Remember that an important group of establishments that used to be informal remain in operation after the enforcement change; these were using capital $k(z) = b$ and exhibited distorted capital-labor ratios; now these establishments demand capital without restrictions.

In the last exercise I perform in this section, I want to address the effects of incomplete enforcement in mind of the argument put forth in Lewis (2004): specifically, that the combination of incomplete enforcement and big

Table 9: Comparison across Steady States

Variable	Value under Complete Enforcement relative to Benchmark
Y	1.11
τ_y	1
K	1.22
TFP	1.04
$Employee\,share$	1.11
$AvEntrepAb$	1.22
w	0.92

government leads to high taxes levied on a small subset of firms. From this perspective, the relevant exercise would involve increasing enforcement levels while decreasing the tax rate to leave revenue unchanged. By increasing enforcement levels, the tax base is broadened, and therefore a lower tax rate will result in the same revenue as before. This will be a way to capture the costs of incomplete enforcement associated with the need for higher taxes.

In Table 6, I present the effects of such an exercise on the steady state values. If Mexico's present enforcement policy were complete, it would be able to reduce taxes to 64% of the current levels. This tax reduction gives an extra boost to the economy. Overall, output would increase 17%. The Table shows that this increase would be driven mainly by a 45% increase in capital accumulation, while TFP would play a smaller role with an increase of 4% which occurs fully in the short-run. Wages would increase 9%. In the long run, once accumulation of capital takes place, labor productivity is increased and the wage rate rises to a level that is 9% larger than the benchmark.

Table 10: Comparison across Steady States

Variable	Value under Complete Enforcement relative to Benchmark
Y	1.17
τ_y	0.64
K	1.45
TFP	1.04
$Employee\,share$	1.11
$AvEntrepAb$	1.22
w	1.09

6.0.1. TFP effects and development

Although not the main goal of this paper, it is interesting to look at the results from the perspective of explaining the development problem. The increase in TFP looks small (4%) for this matter. For example, Restuccia (2008) reports that in a model with human capital, one would need that US TFP be at least 60% larger than in Latin America to account for differences in output per worker of a factor of 4.

The main driver of the small gain obtained, is the calibrated value of $\gamma = .75$. This value is small when compared to what has been found in studies focusing in the US (e.g. Atkeson & Kehoe (2005), Guner *et al.* (2008)). γ controls the returns to scale at the establishment level. The closer is γ to 1 the lower is the degree of decreasing returns and the more efficient is to concentrate production in large establishments. In the limiting case of $\gamma = 1$ (constant returns to scale) the efficient output is reached by concentrating all resources in a single unit: the most productive one. The low value of γ I find, implies that the efficient allocation for Mexico is to have more workers in small units than countries where the degree of decreasing returns is smaller (i.e., larger γ), equivalently, it implies that the economy with incomplete enforcement is not too far from the efficient allocation. In fact, according to this estimate, the efficient allocation for Mexico is to have around 35% of the employees hired by small establishments (those who hire 20 worker or less). Compare this to a 25% allocation in the US in the same size category.

That the distortion-free allocation in Mexico is different than the one in the US, is not necessarily a bad result. It is not the thesis of this paper that the differences between Mexico and US distributions are due solely to tax enforcement differences. It could be argued that Mexico's distortion-free skewed distribution is merely the result of its early stage of development. A number of authors have documented the steady rise in average firm size in the US during the 19th and 20th century (for a short bibliography, see Desmet & Parente (2009)). Furthermore, when one looks at the distribution of US in the past, at a point in time during which US had the same GDP per capita as modern-day Mexico (around the 1930s), it is clear that it was not as highly concentrated in small establishments as it is in today's Mexico[5].

[5]Granovetter (1984) documented the fact that the proportion of employees in US manufacturing establishments with less than 20 employees was 10% in 1933 while the proportion of employees in Mexican manufacturing establishments with less than 15 workers was 37.5% in 2005. Notice that the size category is capped at a smaller size for Mexico than for the US, but the proportion allocated is still larger.

6.0.2. Sensitivity to the tax rate

In Table 6.0.2, I present the results of improving the enforcement policy when the starting tax rate is 50% instead of 26%. I have obtained the 26% tax rate as an outcome of the calibration part by matching the tax revenue to GDP ratio. However some costs of being formal are not collected as revenue, such as periodic costs associated with sanitary, environmental, labor regulations, and the like. Therefore in this section I simply assume that $\tau_y = .50$ and perform the same matching moments exercise as before. This time I don't include the tax revenue to GDP ratio as one of the targeted moments.

The value of γ needed to match the moments of section 5.1 is larger in this case and closer to the value found by studies focusing on the US case ($\gamma = .83$).

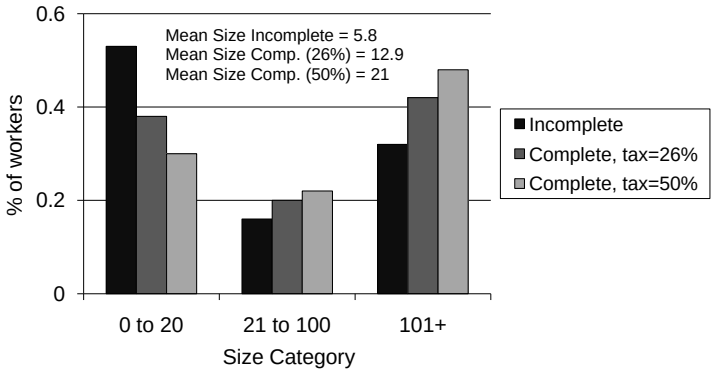
Table 11: Steady State Comparisons ($\tau_y = .50$)

Variable	Value under Complete Enforcement relative to Benchmark
Y	1.53
τ_y	0.53
K	2.6
TFP	1.12
$Employee share$	1.17
$Av Entrep Ab$	1.26
w	1.64

As can be seen in Table 6.0.2, when the implicit tax rate is doubled, the increase in productivity associated with better enforcement increases 12% that is, 3 times what it increased for a starting tax of 26% of output. More TFP, and more capital accumulation, lead to an increase of 53% in aggregate output. Given the larger value of γ (.83), more reallocation occurs. This can be inspected in Figure 8.

Similarly, for the same size categories I find that for the retail and wholesale sectors, the figures are 63.8% and 44.4% for the US in 1939, and 72% and 48% for Mexico in 2005.

Figure 8: Labor Reallocation due to Improved Enforcement (Model)



The allocation of labor in small establishments is further reduced when compared to the case when the starting tax rate is 26%. In contrast, the allocation on medium and large establishments is increased. This brings the average size of establishments from 5.8 to 21 workers.

Finally, in Table 12 I present a comparison of the hypothetical allocation of labor in Mexico under complete enforcement obtained from the model. I do this for the cases in which the starting tax rates are 26% and 50%, respectively, and compare them versus the current allocation in the US.

Table 12: Allocation of Labor under Complete Enforcement

Variable	US (data)	Undistorted Mexico (model)	
Starting τ_y	-	0.26	.50
Under 20	0.26	0.34	0.27
20 to 99	0.29	0.21	0.22
100 or more	0.46	0.46	0.51
Mean Size	18	12	21

The fraction of employees working in firms with 20 or less workers is now 27% which is closer to the 26% found in the US data. Similarly the average establishment size is increased to 21 workers, which is closer to the average size in the US.

Notice that if this sensitivity analysis is repeated for a larger value of the tax rate (say 60%), even more real-

location would occur when imposing the complete enforcement policy. Therefore, I conclude with the following conjecture:$\tau_y = .50$ is an upper bound for the actual tax rate. This is because a larger tax rate would imply an efficient allocation in Mexico that would display implausibly large establishments. It would also need of a larger value of γ than the one found in studies focusing on the US case.

7. Conclusion

The main goal of this paper was to investigate how the presence of informal establishments due to incomplete enforcement of taxes affects aggregate outcomes in developing countries. Although a long tradition (starting with Harris & Todaro (1970)) understands the informal sector simply as a symptom of early stages of development; an incrising number of authors starting with Lewis (2004), challenges this view. The Lewis hypothesis highlights the harmful effects of tax collection with limited enforcement.

I study a general equilibrium framework that includes a tax collection policy with limited enforcement. I calibrate the steady state equilibrium of this model to the case of Mexico. I then investigate the effects of improving enforcement. I find that under complete enforcement, Mexico's aggregate output would be 17% higher in the new steady state.

The first lesson learned in the paper is that informality is associated with resource misallocation. This is driven by the government inability to enforce tax and regulation policies on all firms. As a result, the tax base is small, and high taxes have to be levied on a small subset of firms, usually the most productive ones. This has a negative effect on aggregate productivity by misplacing resources into less productive establishments.

A second lesson is that incomplete enforcement not only gives existing establishments with low productivity a cost advantage; it also makes it more attractive for entrepreneurs with low ability to start new businesses. This distorts the mix of productivities of operating establishments, and therefore productivity.

A third important lesson is that the nature of enforcement policies reduces output through its effect on firms' optimal decisions. In the paper, the specification of the enforcement policy depends on the use of capital in the establishments. So a group of firms are better off by reducing their capital demands to a level low enough not to be detected by government authorities. This distorts the capital per worker of informal establishments and therefore

aggregate capital and output.

AppendixA. Employment Distribution across Establishment size categories

I have been careful to compare figures that share a similar observation unit in both countries. For Mexico I combine data from the most recent National Employment and Occupations Survey (ENOE) and the Economic Census both collected by the National Institute of Statistics (INEGI). The observation unit for both is the establishment [6]. I am in the need to combine these two sources because the Economic Census doesn't collect information of businesses that don't use a building or some kind of physical premise permanently "stick to the ground" and as I have shown, this are not negligible in the Mexican case. I also want to have a distribution of establishments with employees only, because that is the way data in US is collected. This means that I need disaggregated data for the lower tier of the distribution which is only publicly available from ENOE.

Because of the way the Census is performed I believe is a good approximation of the size distribution of medium and large establishments; using the Micro-business survey ENAMIN, I can establish that the larger the establishment, the larger the probability of being included in the Census. On the other hand, the ENOE is the best approximation of the size distribution of small establishments. Therefore, I combine the ENOE and the Census to obtain a full distribution of employment across size categories. Specifically, I use ENOE data to pin down the fraction of workers in two size categories: 2 to 5 workers, and 6 or more.

Once I have these numbers, I proceed to break down the distribution for the category of 6 or more. I do this using the census data. I obtain the distribution of workers across size categories in the census conditional on having 6 or more workers. It is important to remark that neither the Census nor the ENAMIN are able to describe fully the distribution of workers across size categories because the ENAMIN is based on owner information (not establishments) and the Census does not include establishments with out fixed structures as explained.

[6] Given the peculiarities of the economic activities in Mexico, the observation unit in the census varies with industry codes. For manufacturing, services and trade the unit of observation was the establishment. These three sectors include 97% of the surveyed units. Data for most of the remaining sectors is collected using the firm as unit of observation. An hybrid unit of observation had to be used for specific industries such as mining and fishing.

AppendixB. Sensitivity to Enforcement Specification and Tax types

AppendixB.1. Considering alternative enforcement policies for the output tax case

In this section I explore the robustness of the results to alternative enforcement policies. I keep the same step-wise enforcement policy but make the probability of being caught depend on labor and output, each case is considered independently. The results can be found in Table B.13

Table B.13: Effects of Complete Enforcement on Aggregates

Output tax				
Variable	Incomp.	Complete Enforcement		
(1)	Enfmt. (2)	k as signal	y as signal	l as signal
τ_y	0.26	.197	.186	.172
Rev/Y	1	1	1	1
$Y = Y_I + Y_F$	1	1.16	1.10	1.08
$K = K_I + K_F$	1	1.38	1.17	1.08
$E = E_I + E_F$	1	1.11	1.13	1.14
w	1	1.04	1.04	1.09
$TFP = Y/K^{\theta_k}$	1	1.04	1.04	1.05
AvManagAb	1	1.22	1.26	1.28

The effect on output is reduced to 10% when output is used as a signal to enforce tax policy and to 8% when labor is used. This compares to an increase of 16% when the enforcement depends on capital. The effect on TFP is almost the same for the three alternative policy types, 4% when enforcement depends on capital, 4% when it depends on output and 5% when on labor. The main driver of the different results on output is the effect of enforcement on capital accumulation. Accumulation is increased only 17% when enforcement depends on output and 8% when on labor while it increases 38% when on capital.

The distortion generated by the government on informal firms when enforcement depends on capital is more harmful because by affecting establishments on their capital size the aggregate demand for capital is decreased as well as the incentives to accumulate. When the enforcement depends on output this distortion is not as large because informal entrepreneurs have the ability to reduce output by decreasing both labor and capital, so more capital is hired by informal firms as a fraction of total capital demand (30%) when compared to the case in which enforcement depends entirely on capital (17%). Nonetheless the same 30% of employees is allocated on such

establishments under both policies. When the enforcement depends on labor on the other hand, informal firms have the ability to substitute the labor they can not hire with more capital, so informal firms end up hiring an even larger fraction of total capital (41%) than in the case when enforcement depends on output.

The incentives to accumulate capital increase more when enforcement is improved in the case it depends on capital. If one makes the out-of-equilibrium computation for the cases in which enforcement is a function of output and labor one will get interest rates that clear their respective capital markets of the static competitive equilibria that are 10% and 5% larger than their corresponding benchmark levels.

AppendixB.2. Considering alternative tax types

In this section I explore how the results are affected when I deviate from the assumption that all factors of production are taxed at the same rate. As formerly noted, a tax on output is equivalent to tax labor, capital and entrepreneurial output simultaneously at the same rate[7] and this could not be the actual case for the Mexican economy. Therefore in this section I investigate how the effects of improving enforcement levels could change if I modify the distortions faced by entrepreneurs. I focus on the case of an enforcement policy that depends on capital.

I proceed by taking deviations from the case where all tax rates are 26% (or equivalently the case of an output tax equal to 26%). I increase one of the taxes at a time and reduce the other two so that the revenue as a fraction of output remains at 26%. Every time I increase one of the taxes I reduce the other two while keeping them equal to each other. For example, consider the case in which the capital is taxed more heavily than the other two margins. A tax on capital of $\tau_r = .35$ will need tax rates on labor (τ_w) and entrepreneurial output (τ_m) of .2336 to keep the revenue as a share of formal output at .26. Given the choice of these taxes I run the SMM to match the same targets as before.

Departing from the economy just described I analyze the effects of complete enforcement by comparing it against an economy with complete enforcement but faces only an output tax set at a level that collects the same

[7] Specifically, consider the profits of a entrepreneur facing taxes τ_w, τ_r and τ_m on labor, capital and entrepreneurial output respectively. Profits are: $\pi(z,l,k) = zk^{\theta_k}l^{\theta_l} - wl - rk - \tau_w wl - \tau_r rk - \tau_m(zk^{\theta_k}l^{\theta_l} - wl - rk)$. When $\tau_w = \tau_r = \tau_m = \tau_y$, then the profits are the same as those for a entrepreneur facing only tax τ_y on output.

revenue share of the economy with incomplete enforcement. The results for the case in which I increase the tax rate of capital are presented in Table AppendixB.2.

Table B.14: Effects of Complete Enforcement. Alternative Tax Distortions.

(k as signal)				
Variable	$\tau_r = .35$		$\tau_r = .45$	
(1)	Incomp	Comp	Incomp	Comp
τ_r	.35	.194	.45	.193
τ_w	.23	.194	.21	.193
τ_m	.23	.194	.21	.193
Rev/Y	1	1	1	1
$Y = Y_I + Y_F$	1	1.19	1	1.23
$K = K_I + K_F$	1	1.53	1	1.68
$E = E_I + E_F$	1	1.11	1	1.11
w	1	1.04	1	1.05
$TFP = Y/K^{\theta_k}$	1	1.04	1	1.03
AvManagAb	1	1.22	1	1.22

The accumulation effects of better enforcement are increased when the starting situation taxes capital more heavily while the effects on TFP remain almost the same. Aggregate output is increased 19% and 23% for the cases in which the incomplete enforcement case taxes capital at 35% and 45% respectively. The main driver of the increase in output is the distance between the tax rates payed in the formal sector across enforcement scenarios. As can be seen the drop is from levels of 35 and 45 when there is incomplete enforcement to 19% when there is complete enforcement.

One might be worried about the ability to compare these scenarios given that the underlying calibrations differ across tax levels. So a word is worth about the way the calibrations differ. Two effects are important here. First notice that taxing capital at a higher rate in principle affects the capital-output ratio both in the formal sector and at the aggregate level as reveled by the first order condition $\theta_k(1-\tau_m)\left(\frac{y_F(z)}{k_F(z)}\right) = r(1+\tau_r-\tau_m)$. This implies that a smaller interest rate is needed to replicate the targeted capital-output ratio of 2.0. At the end the formal entrepreneurs are affected marginally only. The change in the interest rate however has consequences on the aggregate labor demand increasing the equilibrium wage rate and the worker/entrepreneur threshold. This in turn

reduces the size of the informal sector. If enforcement levels don't change I won't be able to replicate the informal sector size. The economies where the tax rate is larger, need a larger value of b to replicate an informal sector that hires 30% of employees. I end up with an economy almost identical to the one studied in the previous section collecting 26% of output but taxing capital at a much higher rate. When enforcement is complete, bigger savings in the cost of capital are available for the entrepreneurs.

Next I move on to study the effects of enforcement when the starting point is an economy with a larger tax rate on labor and smaller on the other two taxes. The results are presented in Table AppendixB.2.

Table B.15: Effects of Complete Enforcement on Aggregates
(k as signal)

Variable	$\tau_w = .45$		$\tau_m = .45$	
(1)	Incomp	Comp	Incomp	Comp
τ_r	.1905	.205	.13	.192
τ_w	.45	.205	.13	.192
τ_m	.1905	.205	.45	.192
Rev/Y	1	1	1	1
$Y = Y_I + Y_F$	1	1.11	1	1.12
$K = K_I + K_F$	1	1.21	1	1.26
$E = E_I + E_F$	1	1.13	1	1.10
w	1	1.10	1	0.96
$TFP = Y/K^{\theta_k}$	1	1.05	1	1.04
AvManagAb	1	1.27	1	1.16

The effect on aggregate output is 11% for the case in which the benchmark economy has a tax rate on labor of 45%. The TFP effect is slightly bigger mainly associated with the effect that large taxes on labor have on its price and therefore on the worker/entrepreneur threshold. There are more gains on TFP because a large tax on labor distorts the worker/entrepreneur threshold by creating a wedge between what entrepreneurs pay and workers receive. In fact the effect of enforcement on wages is much larger than in the case where I considered only an output tax. However, the effects of enforcement on accumulation are smaller because there is actually an increase in the tax rate on capital from about .19 to .20.

Finally I also considered the effects of having a higher tax on entrepreneurial output in the incomplete enforce-

ment economy. The results are presented in Table AppendixB.2. The results are similar in quantitative terms to the case of a labor tax. TFP effects remain small, and accumulation is smaller than in the case when the tax on capital is increased. Again, this is due to the relative costs of capital between the incomplete enforcement economy (.13) and the complete enforcement one (.19).

Amaral, P. S., & Quintin, E.: A competitive model of the informal sector, *Journal of Monetary Economics*, 53(7), (2006), 1541–1553.

Atkeson, A., & Kehoe, P. J.: Modeling and measuring organization capital, *Journal of Political Economy*, 113(5), (2005), 1026–1053.

Axtell, R.: Zipf distribution of us firm sizes, *Science*, 293(5536), (2001), 1818–1820.

Bergoeing, R., Kehoe, P., Kehoe, T., & Soto, R.: A decade lost and found: Mexico and Chile in the 1980s, *NBER Working Paper*.

Dabla-Norris, E., Gradstein, M., & Inchauste, G.: What causes firms to hide output? the determinants of informality, *Journal of Development Economics*, 85(1-2), (2008), 1–27.

de Paula, A., & Scheinkman, J.: The Informal Sector, Third Version, *PIER Working Paper Archive*.

de Soto, H.: *The Other Path: The Invisible Revolution in the Third World*, I B Tauris & Co Ltd, 1989.

de V. Cavalcanti, T. V., & Antunes, A. R.: Start up costs, limited enforcement, and the hidden economy, *European Economic Review*, 51, (2007), 203–224.

Desmet, K., & Parente, S.: The evolution of markets and the revolution of industry: A quantitative model of england's development, 1300-2000, CEPR Discussion Papers 7290, C.E.P.R. Discussion Papers, 2009.

Farrell, D.: The hidden dangers of the informal economy, *McKinsey Quarterly*, (pp. 26–37).

Garcia-Verdu, R.: Factor shares from household survey data, Working Papers 2005-05, Banco de MÃ©xico, 2005.

Gollin, D.: Do taxes on large firms impede growth? evidence from ghana, Bulletins 7488, University of Minnesota, Economic Development Center, 1995.

Gollin, D.: Getting income shares right, *Journal of Political Economy*, 110(2), (2002), 458–474.

Granovetter, M.: Small is bountiful: labor markets and establishment size, *American Sociological Review*, 49(3), (1984), 323–334.

Guner, N., Ventura, G., & Xu, Y.: Macroeconomic implications of size-dependent policies, *Review of Economic Dynamics*, 11(4), (2008), 721 – 744.

Harris, J. R., & Todaro, M. P.: Migration, unemployment and development: A two-sector analysis, *The American Economic Review*, 60(1), (1970), 126–142.

Herrendorf, B., & Teixeira, A.: Barriers to entry and development, *Forthcoming: International Economic Review*.

Hofman, A.: Standardised capital stock estimates in Latin America: a 1950-94 update, *Cambridge Journal of Economics*, 24(1), (2000), 45–86.

Hsieh, C.-T., & Klenow, P. J.: Misallocation and manufacturing tfp in china and india, Working Paper 13290, National Bureau of Economic Research, 2007.

Indicators, W. D.: World Development Indicators 2005 CD-ROM.

Levy, S.: *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*, Brookings Institution Press, 2008.

Lewis, W. W.: *The Power of Productivity: Wealth, Poverty, and the Threat to Global Stability*, University Of Chicago Press, 2004.

Lucas, R. E.: On the size distribution of business firms, *The Bell Journal of Economics*, 9(2), (1978), 508–523.

Moscoso-Boedo, H. J., & D'Erasmo, P. N.: Financial structure, informality and development, Virginia Economics Online Papers 374, University of Virginia, Department of Economics, 2009.

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

Restuccia, D.: The latin american development problem, Working Papers tecipa-318, University of Toronto, Department of Economics, 2008.

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

Rubini, L.: Innovation and the elasticity of trade volumes to tariff reductions, *unpublished*.

Schneider, F., & Enste, D.: *The shadow economy*, Springer, 2002.

Solimano, A., Soto, R., & Nations, U.: *Economic growth in Latin America in the late 20th century: evidence and interpretation*, United Nations Pubns, 2005.

Thomas, J. J.: *Informal Economic Activity*, Harvester Wheatsheaf, New York, N.Y., 1992.

Tybout, J.: Manufacturing firms in developing countries: How well do they do, and why?, *Journal of Economic literature*, 38(1), (2000), 11–44.