Informality and the Life Cycle of Plants*

Furkan Sarıkaya† M. Nazım Tamkoç‡ Jesica Torres§

September 2024

Abstract

In this paper we first exploit five waves of the Mexican establishment census to estimate the life cycle of formal and informal plants. We show that formal plants not only start their life cycle with three times more workers relative to informal plants, but also grow at a faster rate. While formal establishments more than double their size during their life cycle, informal plants increase their size only by 66%. We then develop a general equilibrium model to quantify the aggregate losses from the marked differences in both sizes at birth and growth rates during the life cycle between formal and informal plants. In our model, plants grow by investing on their productivity. Informality arises as a result of size-dependent regulations, which we capture as tax enforced only among larger firms. In equilibrium, informal plants exhibit flatter life cycle profiles to avoid detection and thus lower their tax burden. In a counterfactual experiment where effective tax rate does not vary with size and plants do not have incentives to operate informally, aggregate output increases by 8 percent relative to the baseline. Changes in the selection of entrepreneurs and steeper life cycle profiles account for these gains.

JEL classification: E23, J24, L25, O41, O33
Keywords: Informality, Life cycle Growth of Plants, Productivity, Distortions

---

*We would like to thank Edgar Avalos for his superb research assistance. 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. Tamkoç and Torres thank DECRG for providing Research Support Budget. INEGI has reviewed our results to ensure that no confidential plant-level information is disclosed.

†Arizona State University

‡The World Bank. Corresponding author: M. Nazım Tamkoç, mtamkoc@worldbank.org.

§The World Bank

1. Introduction

In this paper, we study the role of informality, and size-dependent enforcement in particular, in explaining the life cycle of plants.[1] Hsieh and Klenow (2014) have documented how plants in developed countries grow more over their life cycle compared to plants in less developed countries. In addition, informal, unregistered, plants are widespread in less developed countries. Compliance with taxes varies with the size of the plant: the effective tax rate is higher among large establishments compared to smaller plants (López and Torres (2020) for Mexico). We argue that incomplete enforcement distorts incentives to grow over the life cycle, and informal plants may optimally choose to operate at smaller scales in order to escape detection.

One of the key challenges to studying how informality distorts the life cycle of plants is the availability of plant-level panel data that contains information on the registration status of plants. We overcome this challenge by exploiting Mexico's census plant-level data administered by Instituto Nacional de Estadística, Geografía e Informática (INEGI). The data tracks plants for a period of 20 years, from 1998 to 2018, and covers whether the establishment is registered with different tax authorities. Therefore, unlike Hsieh and Klenow (2014), we are able focus on a panel of registered and unregistered plants.

In our analysis, an informal establishment is a plant which is not registered with either the central tax authority or the social security administration. In contrast, we refer to plants as formal if they are registered with either the central tax authority or the social security administration. We show first that young formal plants are bigger in terms of employment than young informal plants. Formal establishments in Mexico start operations with 6 workers on average whereas their informal counterparts report only 1.9 workers (on average). Second, formal plants grow faster over their life cycle compared to informal plants. While formal plants grow by a factor of 2.2 during their life cycle, informal plants are only 1.7 times bigger when old compared to young counterparts.

In order to quantify the losses of informality from distortions to the life cycle growth of plants, we develop a one-sector growth model where both formal and informal plants grow by investing in their productivity. Our model is based on Guner et al. (2018) which in turn builds on the Lucas (1978). Our innovation to their setup is introducing a proportional tax on output that entrepreneurs can escape by operating informally due to incomplete enforcement, in the spirit of Leal (2014). Plants can grow by investing in

[1] We use the terms establishment and plant interchangeably.

their productivity over their life cycle. Tax officials are not able enforce taxes to all plants. At the beginning of their life cycle, entrepreneurs are assigned to a tax official with some probability, which increases with their talent. Moreover, bigger plants will detected by the official if they choose to evade (regardless of whether they were assigned to an official at birth) so they choose to become formal. Therefore, in equilibrium, small, and relatively unproductive plants will operate informally whereas bigger plants and plants randomly assigned to a tax official early in their life cycle will operate in the formal sector.

We calibrate the model parameters to match some properties of the plant-size distribution and the growth rates of both formal and informal plants in Mexico. The model is able to successfully match the mean sizes of formal and informal plants as well as the concentration of employment in large plants. In addition, the model is able to capture the markedly different growth rates of formal and informal plants over their life cycle.

We then run a series of policy experiments. When the output tax is fully enforced, establishments have no incentive to remain small to avoid detection, and average size increases by 26% relative to the benchmark. Moreover, the average plant grows 20% faster during their life cycle, from a factor of 2.83 between ages 0-4 to 35-39 in the benchmark to a factor of 3.38 in the counterfactual. Aggregate output slightly decreases but by less than 1 percentage point.

In contrast, increasing the tax rate by 5 percentage points lowers aggregate output by about 8 percent relative to the benchmark. A higher tax rate generates stronger incentives to avoid it, which can only happen in this economy by investing less in managerial skill to remain small. This results in a higher rate of informality, a lower average plant size, a worse selection into entrepreneurship (lower average managerial ability), and flatter life cycle profiles for the average plant (lower growth rates).

Full enforcement of the tax while reducing the rate to keep the tax-to-output ratio constant would increase aggregate output by 8 percent relative to the benchmark. Average plant size would increase by 36%, and the average plant would grow at a pace 48% faster, from a factor of 2.8 between ages 0-4 and ages 35-39 to a factor of 4.2.

2. Establishment-level data

Our quantitative exercises exploit data from the Mexican establishment census.[2] The census is administered by Mexico's statistics agency INEGI and targets the universe of establishments in non-rural areas.[3] Enumerators conduct door-to-door visits and data collection is not restricted to registered (or formal) businesses. Establishments are included regardless of their size (even establishments without employees), as long as the economic activity takes place within delimited fixed building structures.[4]

The census is administered every five years. We exploit the five waves between 1998 and 2018. We consider only establishments in manufacturing, wholesale and retail, and non-financial services.[5] In these three broad sectors, the number of establishments in the census increased from 2.5 million in 1998 to 4.3 million in 2018. In 2018, 14% of establishments were in manufacturing, 52% were in wholesale and retail, and the remaining 34% were in (other) non-financial services. Average establishment size has hovered at about 4.5 employees (about 7.5 excluding establishments with no employees).[6] 96% of establishments employ 10 or fewer employees (a fraction that has not changed since 1998) and accounted for 44% of employment in 2018; less than 0.5% of establishments employ more than 100 workers but in 2018 accounted for 36% of employment (Table 2 below and Table A1 in the appendix).

2.1 Formal and informal plants

The 2018 questionnaire introduced two new questions where the respondent is asked whether the establishment is registered with the central tax authority SAT (*Servicio de Administración Tributaria*) and with the social security administration IMSS (*Instituto Mexicano del Seguro Social*). SAT collects corporate and personal income taxes (similar to the IRS in the US), value added taxes, and federal excise taxes. IMSS, on the other hand,

[2]Access to the census data is restricted and requires clearance from INEGI. Data was processed at INEGI's data lab in Mexico City.

[3]Rural areas have less than 2,500 residents but INEGI does include rural areas in the census when they are the capital of the municipality.

[4]For example, businesses located inside residential properties are included, but street vendors are not.

[5]Our analysis excludes agriculture (11 in the NAICS classification system); mining (21); utilities (22); construction (23); finance (52); real estate (53); management services and administrative support (55-56); education (61); healthcare and social assistance (62); postal services (491); religious and civic organizations (813); private households (814); and employment services (5613).

[6]We measure employment adding white collar and blue collar employees, unpaid workers, third-party employees (who don't have an official employment relationship to the establishment), and contractors.

collects social security contributions, which employers are required both to contribute to and to withhold on behalf of their employees.[7] The two questions on registration with SAT and IMSS had yes-or-no answers, with the possibility of a refusal to respond.[8] A third of establishments in the 2018 cross-section report registration with SAT and only 0.23% have missing values in this question. In contrast, only 6.6% of plants report registration with IMSS, but the rate of non-response is 11.5% (see Table 1 below).

Table 1: Sorting of establishments into formal and informal and share of plants and employment in each category. 2018.

SAT	**IMSS**	**Formal / informal**	**Share of plants**	**Share of employment**
Yes	Yes	Formal	6.48	40.14
No	Yes	Formal	0.11	0.33
Yes	No	Formal	26.24	30.24
Yes	Non-response	Formal	0.76	0.33
Non-response	Yes	Formal	0.00	0.00
No	No	Informal	55.71	25.03
Non-response	Non-response	Informal	0.23	0.54
No	Non-response	Informal	10.48	3.39
Non-response	No	Informal	0.01	0.00

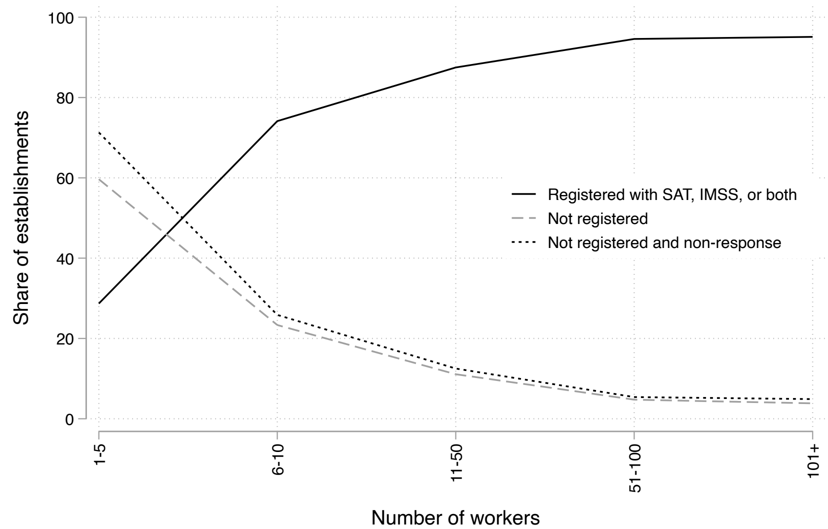
Notes: Total in each column may not add up to 100 due to rounding error.

In our analysis, establishments registered with SAT and/or IMSS (as reported by the respondent) are formal; establishments that are not formal are informal (unregistered) as we show in see Table 1. Figure 1 shows, for each size category, the fraction of formal (registered) establishments in the 2018 cross-section. The share of formal establishments markedly increases with size. In contrast, the share of plants not registered with either SAT or IMSS and the rate of non-response to the two questions is higher among small plants. Under our definition, 34% of establishments are formal, accounting for 71% of employment in the census.

[7] IMSS also delivers some social services. The share of employers' social security contributions amount to a tax of about 24% of profits. Source: World Bank Group Doing Business 2020.

[8] In Spanish, the questions read *¿Este establecimiento cuenta con registro ante el Servicio de Administración Tributaria (SAT)?* and *¿Cuenta con registro patronal ante el Instituto Mexicano del Seguro Social (IMSS)?* Follow-up questions asked the respondent to provide the identification numbers of the establishment with each tax authority. These four questions were not included in 1998-2013 waves.

Figure 1: Share of formal and informal establishments in each size category. 2018.



Notes: Formal establishments are registered with SAT, IMSS, or both. Establishments that are not formal are informal.

Table 2 shows the size distribution and the employment size distribution of registered and unregistered establishments in the 2018 cross-section. Micro-establishments (with 1-10 workers) amount to 90% of establishments and account for 26% of employment among formal plants. Among informal plants, micro-establishments amount to 99% of plants and account for 90% of employment. Large establishments (with 101 or more workers) account for almost 50% of employment among formal plants but only 4% of employment among informal establishments. Informal establishment size averages 2.03, while average formal establishment size is 10.10.

Table 2: The size distribution of formal and informal plants in the Mexican census. 2018.

Size	**Total**		**Formal plants**		**Informal plants**	
	Share of plants	Share of workers	Share of plants	Share of workers	Share of plants	Share of workers
1-10	95.89%	43.70%	89.10%	26.00%	99.32%	89.22%
11-20	2.06%	6.22%	5.28%	7.57%	0.44%	2.98%
21-50	1.19%	7.84%	3.21%	9.99%	0.18%	2.65%
51-100	0.39%	5.81%	1.10%	7.75%	0.03%	1.09%
101+	0.46%	36.42%	1.31%	48.69%	0.03%	4.06%
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Notes: Formal establishments are registered with SAT, IMSS, or both. Establishments that are not formal are informal.

2.2 The life cycle of formal and informal plants

We estimate size at birth as the average size of formal and informal plants ages 0-4 in the 2018 cross-section: 6.1 workers among formal establishments and 1.9 among informal plants (2.9 for the total average). To estimate growth rates, we exploit the panel dimension of the data. INEGI has developed identifiers that track establishments across the 2008, 2013, and 2018 waves of the census. Busso et al. (2018) developed synthetic identifiers to follow establishments from 1998 through 2008, which we leverage in our analysis.[9] Busso et al. (2019) first exploited the synthetic identifiers to compute the life cycle of formal and informal plants in Mexico, but they focused on informality on the intensive margin while our definition looks at the extensive margin. Moreover, while they restrict their analysis only to the balanced panel (the panel of surviving plants between 1998 and 2018), our estimates of the average growth rates leverage micro data from entering and exiting establishments as well.

We first sort establishments in the 2018 cross-section into formal and informal using the self-reported registration with SAT and IMSS as described above, and assume that establishments do not transition into or out of informality during their life cycle. In other words, we assume that establishments that were informal (formal) in 2018 were also informal (formal) every period back through 1998.[10] We then sort formal and informal

[9]Busso et al. (2018) use the location, the name, and the 6-digit sector to find establishments in two consecutive waves of the census.

[10]We are not able to compute in the data transition matrices across informality status using our

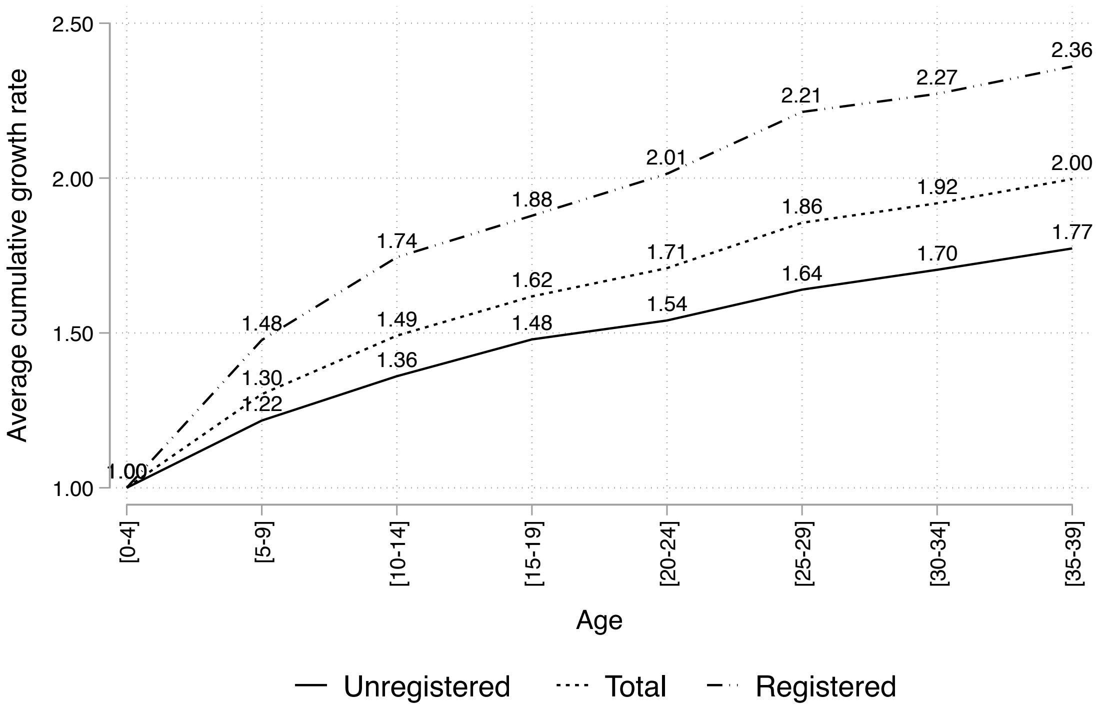
establishments in every period into age bins (0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, and 35-39). Next, we restrict the data to those establishments that were ages 0-4 in any cross-section between 1998 and 2013. That is, we restrict the analysis to establishments in the 2018 cross-section that were born between 1998 and 2013. For each establishment, we follow Eslava et al. (2022) and divide their size in every period by their size at age 0-4. We then regress these cumulative growth rates on the (double) interaction between age (in bins) and a binary indicator for whether the plant is formal or informal, 2-digit sector and cohort fixed effects, and another indicator for whether the year is before or after 2008, to control for whether the identifier that links establishments across waves is synthetic (as generated by Busso et al. (2018)) or developed by INEGI.

The exercise just described only allows us to obtain average growth rates through ages 20-24 since we only observe an establishment for 5 periods at most. We therefore repeat these steps for different starting ages. For example, to obtain the average growth rate through ages 30-34, we restrict the data to establishments in the 2018 cross-section that were 10-14 at any point between 1998 and 2013. Our life cycles, which we show in Figure 2, are the average of the different estimates for the cumulative growth rates from the separate regressions using different starting ages. More details on our methodology, together with the full set of results from least squares, are offered in the appendix.

Formal plants then not only start their life cycle with more workers relative to informal plants (three times more), but also grow at a faster rate. While formal establishments more than double their size through ages 35-39, informal plants increase their size only by 66%. Note that these estimated growth rates do not differ significantly when we focus only manufacturing or services (see A3 in the appendix).

definition of formality to check the validity of this assumption, since the questions on registration with SAT and IMSS were not available in previous waves of the census. We believe, however, that our assumption is reasonable in light of the evidence in Busso et al. (2019) who document that almost 90% of informal establishments in 1998 were still informal in 2013 (using definition of informality on the intensive margin). Table A2 in the appendix shows that their definition of informality is strongly correlated with the one we use in this paper: in the 2018 cross-section, 66% of establishments that are informal on the extensive margin (they are not registered with SAT or IMSS) are also informal on the intensive margin.

Figure 2: Cumulative growth rates in the life cycle of formal and informal plants in the Mexican census.



Notes: Formal establishments are registered with SAT, IMSS, or both. Establishments that are not formal are informal.

In Figure A2 in the appendix, we compute growth rates for formal and informal plants from the 2018 cross-section only, as in Hsieh and Klenow (2014). Interestingly, the cross-sectional estimates seem to bias downward the growth of informal establishments (40% vs. 66%) but overestimate growth among formal plants (170% vs. 116%). We then show in Figures A4 and A5 the life cycles for formal and informal plants using different definitions of informality. Figure A4 looks at a strict definition: plants registered with both SAT and IMSS are formal, while plants not registered with SAT nor IMSS are informal. Figure A5 follows the formality definition of Busso et al. (2019). The estimated growth rates for informal plants do not vary across definitions, but estimated growth rates among formal establishments are significantly higher under the strict definition and the intensive margin definition of Busso et al. (2019). In our counterfactuals, we present the results using these alternative growth rates as a robustness check.

3. Model

We build on the model developed by Guner, Parkhomenko, and Ventura (2018) (GPV in what follows) where plants grow by investing on their productivity. Our innovation is introducing a tax on output that some plants can escape from by operating informally due to incomplete enforcement.

In every period, a group of agents is born with an initial amount of capital k_0 and entrepreneurial ability z. This initial endowment of entrepreneurial ability is drawn from the (exogenous) probability density function $f(z)$ with cumulative distribution function $F(z)$ and support $[\underline{z}, \overline{z}]$. Agents live for a certain number of periods, J.

Each agent maximizes the expected present value of lifetime utility derived from consumption:

$$\sum_{j=1}^{J} \beta^{j-1} u(c_j) \tag{1}$$

where $\beta \in (0, 1)$ is the discount factor, c_j denotes the consumption of an agent at age j, and $u(c) = \log(c)$.

Agents have one unit of time which they supply inelastically until they reach their (exogenous) retirement age $J_R < J$. At the beginning of their lives, they choose whether to work for someone else (workers) or to become entrepreneurs. Entrepreneurs also choose whether to operate in the formal or the informal sector. These decisions are made only once, meaning that individuals do not transition across occupations or between the formal and the informal sectors in the span of their careers.

Entrepreneurs Both formal and informal entrepreneurs have access to the same span-of-control technology in which the single final good of the economy is produced using capital k and labor n as follows:

$$f(z, k, n) = Az^{1-\gamma} \left[k^{\alpha} n^{1-\alpha}\right]^{\gamma}, \tag{2}$$

where γ is the span-of-control parameter, with $\gamma \in (0, 1)$, indicating decreasing returns to scale. The parameter A denotes the aggregate productivity term common across all entrepreneurs, which we normalize to one for the remainder of the paper.

Entrepreneurs can improve their future entrepreneurial skills by investing their income. Following GPV, the law of motion for entrepreneurial ability is assumed to take the following functional form:

$$z_{j+1} = (1 - \delta_z)z_j + g(z_j, x_j) = (1 - \delta_z)z_j + B_j z^{\theta_1} x^{\theta_2}, \tag{3}$$

where x_j denotes the income invested in entrepreneurial skill by an entrepreneur at age j; δ_z stands for the rate of depreciation of skills; and $B(j)$ represents the overall efficiency of skill investment at age j. GPV assume $\theta_2 \in (0, 1)$ and $B(j) = (1-\delta_\theta)B(j-1)$, with an initial condition of $B(1) = \theta$. With these assumptions, which we follow, the skill accumulation technology exhibits complementarity between the current entrepreneurial skill and investments in skill ($g_{zx} > 0$) and diminishing returns to skill investments ($g_{xx} < 0$). Moreover, investment in entrepreneurial skill is an essential input for the skill accumulation technology ($g(z, 0) = 0$).

Workers In the model, workers are assumed to be homogeneous. They supply their labor endowment inelastically and receive a market wage, irrespective of their initial entrepreneurial skills and of whether they work for a formal or an informal entrepreneur. The objective of a worker is to optimally allocate her lifetime earnings between consumption and savings in each period.

Taxation with imperfect enforcement Entrepreneurs face a proportional output tax denoted as $\tau \in (0, 1)$. Enforcement of this output tax, however, is imperfect for two reasons. First, at the beginning of their life cycle, agents meet a tax official only with some exogenous probability, which is a function of their entrepreneurial ability: $\phi(z)$. More precisely, we set:

$$\phi(z) = \begin{cases} \overline{\phi} & \text{if} \quad z \leq \hat{z} \\ 1 & \text{if} \quad z > \hat{z} \end{cases} \tag{4}$$

where $\overline{\phi} \in (0, 1)$. If an entrepreneur faces a tax official, the output tax is fully enforced. In other words, agents with skill above $\hat{z}$ can only operate formally if they choose to become entrepreneurs. These agents therefore have to decide only between becoming workers or formal entrepreneurs. This component of the enforcement technology does not affect the choice of capital or labor for the formal entrepreneur but does affect occupational choices.

If an agent does not face a tax official at the beginning of their life cycle, he or she can escape from taxation when entering entrepreneurship by operating informally (fully avoiding the output tax). Operating informally, though, carries the risk of detection and subsequent punishment (even if the individual does not face a tax official when they were born). The enforcement of taxation, in turn, follows a size-dependent technology:

$$p\left[k(z)\right] = \begin{cases} 0 & \text{if} \quad k(z) \leq b \\ 1 & \text{if} \quad \text{otherwise} \end{cases} \tag{5}$$

Informal entrepreneurs then can completely avoid paying taxes as long as they keep their capital below the threshold b. Note that no informal entrepreneur will then optimally choose to exceed this threshold as long as the fine if caught evading is strictly larger than zero. If capital exceeds b, taxes will be enforced with certainty. If the entrepreneur is informal, she will be get caught by the tax authorities. In that case, the informal entrepreneur will be required to pay a fine in addition to the output tax, which would lead to lower profits compared to bunching at $k = b$ or operating in the formal sector.

Profits of formal entrepreneurs A formal entrepreneur with ability z chooses his scale to maximize his profits:

$$\Pi^F(z) = \max_{k,n} \left(1-\tau\right) A z^{1-\gamma} \left[k^\alpha n^{1-\alpha}\right]^\gamma - wn - Rk \tag{6}$$

As we show in the appendix, factor demands are linear in managerial skill:

$$k_F(z) = \left[\gamma(1-\tau)\right]^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{R}\right)^{\frac{(\alpha-1)\gamma+1}{1-\gamma}} \left(\frac{w}{1-\alpha}\right)^{\frac{(\alpha-1)\gamma}{1-\gamma}} z \tag{7}$$

$$n_F(z) = \left[\gamma(1-\tau)\right]^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{R}\right)^{\frac{\alpha\gamma}{1-\gamma}} \left(\frac{w}{1-\alpha}\right)^{\frac{\alpha\gamma-1}{1-\gamma}} z \tag{8}$$

By substituting these optimal factor demands back into Equation 6, the earnings of a formal entrepreneur with ability z are also linear in entrepreneurial skill:

$$\Pi^F(z) = \left[(1-\gamma)(1-\tau)\right]^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{R}\right)^{\frac{\alpha\gamma}{1-\gamma}} \left(\frac{w}{1-\alpha}\right)^{\frac{(\alpha-1)\gamma}{1-\gamma}} z \tag{9}$$

Profits of informal entrepreneurs The profit maximization problem of informal entrepreneurs is slightly more complex. As described above, the enforcement technology incentivizes informal entrepreneurs to operate with a level of capital that is equal to or less than the threshold b. Some entrepreneurs then will choose a level of k strictly below b, but others will opt to adjust their capital level precisely to b (a corner solution), maintaining just enough capital to remain undetected (but potentially adjusting their level of labor).

The first group of informal entrepreneurs faces a similar maximization problem as that of entrepreneurs in the formal sector, but with the key distinction that the tax rate τ is zero, since they effectively avoid any tax payments. The second group of informal entrepreneurs adjust their level of capital to the threshold level, b, adapting their labor demand accordingly. As we show in the appendix, the profits of an informal entrepreneur with ability z are then given by:

$$\Pi^I(z) = \begin{cases} [(1-\gamma)]^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{R}\right)^{\frac{\alpha\gamma}{1-\gamma}} \left(\frac{w}{1-\alpha}\right)^{\frac{(\alpha-1)\gamma}{1-\gamma}} z & \text{if} \quad k^I(z) \leq b \\ z^{\frac{1-\gamma}{(\alpha-1)\gamma+1}} b^{\frac{\alpha\gamma}{(\alpha-1)\gamma+1}} \left(\frac{(1-\alpha)\gamma}{w}\right)^{\frac{1-\gamma}{(\alpha-1)\gamma+1}} (1-(1-\alpha)\gamma) - bR & \text{if} \quad k^I(z) = b \end{cases} \tag{10}$$

Accumulation of assets and entrepreneurial skill The optimization problem of workers in this economy remains as in GPV's original setup. A worker chooses consumption and savings at each age (c_j and a_j) to maximize the value of lifetime discounted utility:

$$\begin{aligned} W = \max_{c_j, a_{j+1}} \quad & \sum_{j=1}^{J} \beta^{j-1} u(c_j) \\ \text{s.t.} \quad & c_j + a_{j+1} = w + (1+r)a_j, \quad \forall\, 1 \leq j \leq J_R - 1 \\ & c_j + a_{j+1} = (1+r)a_j \quad \forall j \geq J_R \end{aligned}$$

Entrepreneurs choose in addition their investment in skill formation at each age x_j:

$$\begin{aligned} V^i(z_1) = \max_{c_j, x_j, a_{j+1}} \quad & \sum_{j=1}^{J} \beta^{j-1} u(c_j) \\ \text{s.t.} \quad & c_j + x_j + a_{j+1} = \Pi^i(z_j) + (1+r)a_j, \quad \forall\, 1 \leq j \leq J_R - 1 \\ & c_j + a_{j+1} = (1+r)a_j \quad \forall j \geq J_R \\ & z_{j+1} = (1-\delta_z)z_j + g(z_j, x_j), \quad \forall\, 1 \leq j \leq J_R - 1 \end{aligned}$$

where $i \in \{F, I\}$ corresponds to the formal or informal sector.

The interior solution of this problem of entrepreneurs can be characterized by two conditions. The first condition is the standard Euler equation characterizing the intertemporal decision between today's consumption and the next period's asset:

$$\frac{1}{c_j} = \beta(1+r)\frac{1}{c_{j+1}} \qquad \forall 1 \leq j \leq J-1 \tag{11}$$

The second condition is a no-arbitrage condition for asset and investment in managerial ability.

$$(1+r) = g_x(z_j, x_j)\Pi_z^i(z_{j+1}) + \frac{g_x(z_j, x_j)}{g_x(z_{j+1}, x_{j+1})}[g_z(z_{j+1}, x_{j+1}) + 1 - \delta_z] \qquad \forall 1 \leq j \leq J_R - 1 \tag{12}$$

The left-hand side of this equation represents the opportunity cost of investing in managerial skills. A manager could earn $(1+r)$ unit of consumption good by investing one unit in physical asset rather than investing one unit in managerial skills. The right-hand side of the equation is the net marginal benefit of skill investment. The first term represents the benefit deduced from the next period profit generated by additional managerial ability. An extra unit of skill investment in the current period is transformed into next period managerial skill with $g_x(z, x_j)$ and $\Pi_z^i(z, r, w)$ transforms the additional managerial ability to profit. Also, an extra unit of investment in managerial ability in the current period relaxes the skill accumulation technology constraint in the subsequent period. The second term at the right hand side stands for this relaxation. This term disappears in the last working age period of managers and the equation turns out to be the following for $j = J_R - 1$.

$$(1+r) = g_x(z_j, x_j)\Pi_z^i(z_{j+1}) \tag{13}$$

Occupational choice At age 1, individuals select the occupation that maximizes their total discounted life cycle utility given their initial entrepreneurial ability, z_1: worker, informal entrepreneur, or run a formal plant.

In the stationary equilibrium, there exists a cutoff $z_{w,e}$ that sorts individuals who meet a tax official at age 1 into workers and formal entrepreneurs: those with entrepreneurial skill below $z_{w,e}$ become workers, while those with skill above $z_{w,e}$ become formal entrepreneurs. Among those individuals who do not meet a tax official at the beginning of their life cycle, there exist two cutoffs that sort workers into workers, informal entrepreneurs, and formal entrepreneurs. Agents with skill below $z_{w,ne}$ become workers, agents with skill in $(z_{w,ne}, z_{f,ne})$ become informal entrepreneurs, and agents with skill above $z_{f,ne}$ become formal entrepreneurs. Note that $z_{w,ne} \leq z_{w,e} \leq z_{f,ne}$.

Equilibrium In this economy, there is no growth. Given w_t, R_t, τ, b and $\phi(z)$, the stationary equilibrium consists of value functions V_j^F, V_j^I, W_j, policy functions s_j^F, s_j^I, s_j, x_j^F, x_j^I, and occupation choices, o, such that:

1. Given prices, w and R, the tax rate, τ, and enforcement technologies b and $\phi(z)$, s_j solves the workers' problem; s_j^F and x_j^F solve the formal managers' problem; and s_j^I and x_j^I solve the informal managers' problem.
2. All agents solve their occupation choice problem such that choices are consistent with s_j's and x_j's.
3. Markets clear.

The labor market clearing condition is shown below. In the left-hand side, the first element is the demand for labor from informal plants (those that don't meet the tax official); the second element is the demand for labor from agents who don't meet the tax official but optimally choose to become formal; and the third element is the demand for labor from entrepreneurs who meet the tax official and have to operate in the formal sector. In the right-hand side, the first element is the supply of labor from workers who don't meet the tax official and become workers, whereas the second term is the supply of labor from workers who meet the tax official and choose to become workers.

$$\sum_{j=1}^{J_R-1} \int_{z_{w,ne}}^{z_{f,ne}} (1-\phi(z)) n_i(z) f_j(z) dz + \sum_{j=1}^{J_R-1} \int_{z_{f,ne}}^{\bar{z}} (1-\phi(z)) n_f(z) f_j(z) dz +$$
$$\sum_{j=1}^{J_R-1} \int_{z_{w,e}}^{\bar{z}} \phi(z) n_f(z) f_j(z) dz = \sum_{j=1}^{J_R-1} \int_{\underline{z}}^{z_{w,ne}} (1-\phi(z)) f_j(z)\, dz + \sum_{j=1}^{J_R-1} \int_{\underline{z}}^{z_{w,e}} \phi(z) f_j(z)\, dz$$

The capital and goods markets clearing conditions are in the equations that follow. Note that K denotes the capital stock, i.e. the right hand-side of capital market clearing conditions.

$$\sum_{j=1}^{J_R-1}\int_{z_{w,ne}}^{z_{f,ne}}(1-\phi(z))[k_i(z)]f_j(z)dz+\sum_{j=1}^{J_R-1}\int_{z_{f,ne}}^{\bar{z}}(1-\phi(z))[k_f(z)]f_j(z)dz+$$
$$\sum_{j=1}^{J_R-1}\int_{z_{w,e}}^{\bar{z}}\phi(z)[k_f(z)]f_j(z)dz=\sum_{j=1}^{J-1}\Bigg[\int_{\underline{z}}^{z_{w,ne}}\int_{-\bar{a}}^{\bar{a}}(1-\phi(z))s_j^W(a)d\psi_j(a,z)+$$
$$\int_{\underline{z}}^{z_{w,e}}\int_{-\bar{a}}^{\bar{a}}\phi(z)s_j^W(a)d\psi_j(a,z)+\int_{z_{w,ne}}^{z_{f,ne}}\int_{-\bar{a}}^{\bar{a}}(1-\phi(z))s_j^I(a,z)d\psi_j(a,z)+$$
$$\int_{z_{f,ne}}^{\bar{z}}\int_{-\bar{a}}^{\bar{a}}(1-\phi(z))s_j^F(a,z)d\psi_j(a,z)+\int_{z_{w,e}}^{\bar{z}}\int_{-\bar{a}}^{\bar{a}}\phi(z)s_j^F(a,z)d\psi_j(a,z)\Bigg]$$

$$\sum_{j=1}^{J}\int_{z_{w,ne}}^{z_{f,ne}}(1-\phi(z))[c_j^I(z)+s_j^I(z)+x_j^I(z)]f_j(z)dz+\sum_{j=1}^{J}\int_{z_{f,ne}}^{\bar{z}}(1-\phi(z))[c_j^F(z)+s_j^F(z)+x_j^F(z)]f_j(z)dz+$$
$$\sum_{j=1}^{J}\int_{z_{w,e}}^{\bar{z}}\phi(z)[c_j^F(z)+s_j^F(z)+x_j^F(z)]f_j(z)dz+\sum_{j=1}^{J}\int_{\underline{z}}^{z_{w,ne}}(1-\phi(z))[c_j^W+s_j^W]f_j(z)dz+$$
$$\sum_{j=1}^{J}\int_{\underline{z}}^{z_{w,e}}\phi(z)[c_j^W+s_j^W]f_j(z)dz=\sum_{j=1}^{J_R-1}\int_{z_{w,ne}}^{z_{f,ne}}(1-\phi(z))y^I(z)f_j(z)dz+$$
$$\sum_{j=1}^{J_R-1}\int_{z_{f,ne}}^{\bar{z}}(1-\phi(z))y^F(z)f_j(z)dz+\sum_{j=1}^{J_R-1}\int_{z_{w,e}}^{\bar{z}}\phi(z)y^F(z)f_j(z)dz+(1-\delta)K$$

4. Model calibration

In our calibration, the model period is 5 years. As stated above, in our version of GPV's setup we assume no population growth and no technology growth. We follow Leal (2014) and set $\delta = 0.055$, $\beta = 0.936$, and $\tau = 0.25$. We adjust the value of α so that $\alpha\gamma$ (the share of capital in the formal sector in the model economy) matches 0.33. The mean of log managerial ability (μ_z) is normalized to 0. Note we set $\theta = 1$ and $\delta_\theta = 0$, so that the skill accumulation technology is characterized by only 3 parameters (δ_z, θ_1, and θ_2).

These parametric assumptions leave eight remaining parameters to determine: $\gamma, \sigma_z, \delta_z, \theta_1, \theta_2, b, \hat{z}$, and $\overline{\phi}$. We jointly calibrate these parameters to capture the low growth rate of informal establishments in the data and other crucial features of the distribution of formal and informal establishments in the Mexican census. Our calibration matches the following moments: (i) mean size of informal establishments; (ii) mean size of formal plants; (iii) mean size; (iv) employment share of small establishments (with less than 10 workers); (v) fraction of establishments with less than 10 workers, (vi) fraction of the labor force (in the census) employed in establishments with 100 or more workers; (vii) the life cycle growth rate of formal establishments (from 0-4 to 20-24); (viii) the growth rate of informal establishments (from 0-4 to 20-24).

Table 3 shows the calibrated parameters whereas the targeted moments and their calibrated counterparts are shown in **??**. Figure 3 compares the size distribution in the Mexican census to the distribution in the calibrated economy. The benchmark model is able to match the plant size distribution in the Mexican census while it closely approximates the distribution of formal and informal plants. Figure 4 shows that the model is able to approximate the share of employment among large establishments (more than 100 workers), while Figure 5 shows that the model closely matches the fraction of formal plants among micro-establishments (less than 5 workers). The fraction of formal establishments in the model is 39%, close to the fraction observed in the data (34%).

Table 3: Parameters

Parameter	Value
Depreciation rate (δ)	0.055
Discount ractor (β)	0.936
Output tax (τ)	0.25
Capital share (α)	0.516
Mean of log managerial ability (μ_z)	0
Span-of-control (γ)	0.553
Dispersion in log managerial ability (σ_z)	2.146
Depreciation of productivity (δ_z)	0.024
Return on previous productivity (θ_1)	0.662
Return on productivity investment (θ_2)	0.492
Enforcement threshold (b)	3.619
Skill cutoff to meet official at birth ($\hat{z}$)	89.30
Probability of meeting official at birth ($\overline{\phi}$)	0.454

Figure 6 shows that the model is able to generate the life cycle of the average plant in Mexico. The model is able to generate the life cycle of informal plants observed in the data, although among formal plants, the model predicts a steeper growth profile late in the life cycle compared to the data.

Table 4: Targeted Moments

Moment	Data	Model
Informal mean size	2.0	2.0
Formal mean size	10.1	9.6
Mean size	4.8	5.0
Emp. share of small (<10)	0.44	0.41
Fraction of small (<10)	0.96	0.94
Emp. share of large (+100)	0.36	0.33
Formal growth (0-4 to 20-24)	1.43	1.51
Informal growth (0-4 to 20-24)	0.53	0.52

Figure 3: Non-targeted moments. Fraction of plants in different size categories.

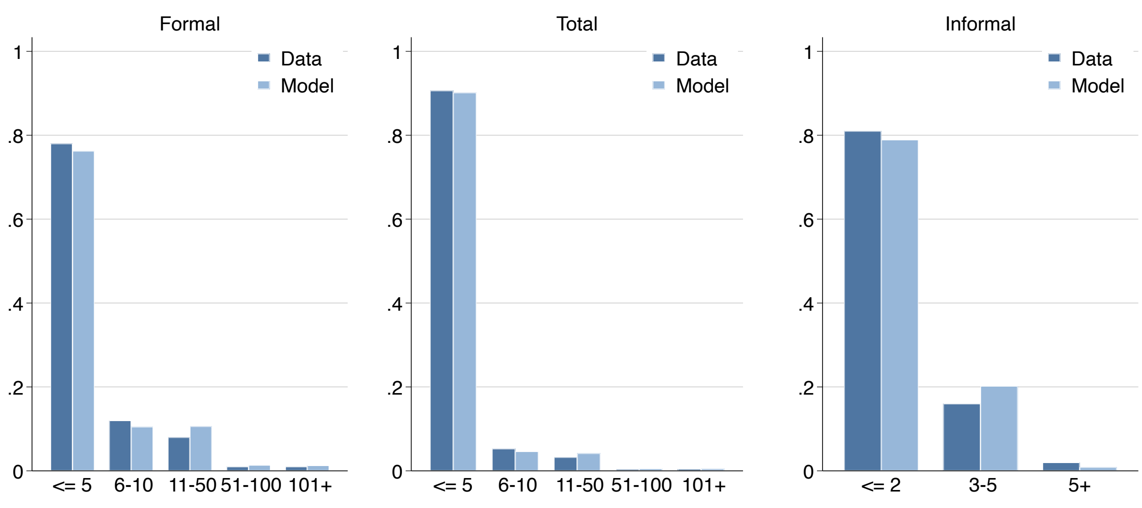


Figure 4: Non-targeted moments. Employment share of plants in different size categories.

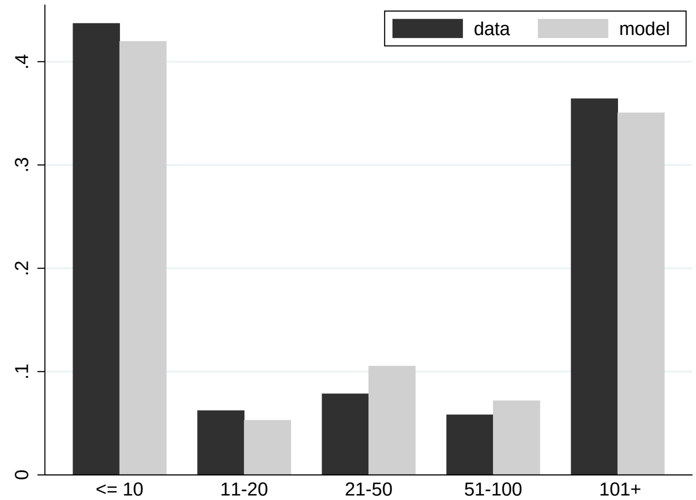


Figure 5: Non-targeted moments. Fraction of formal plants in each size bin.

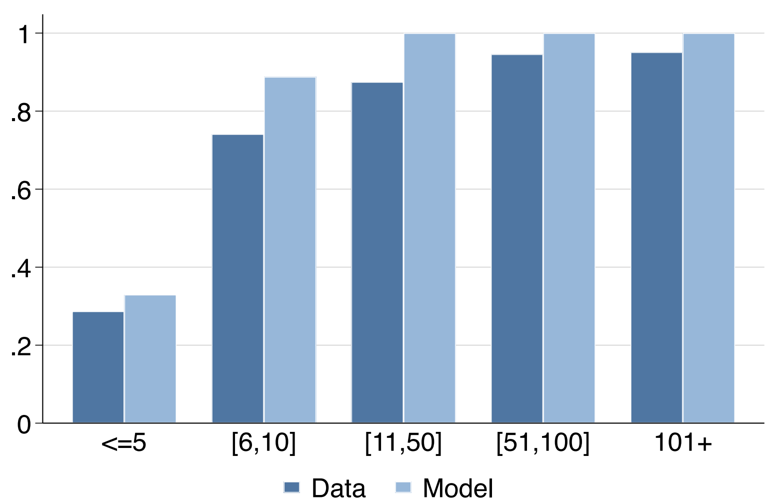
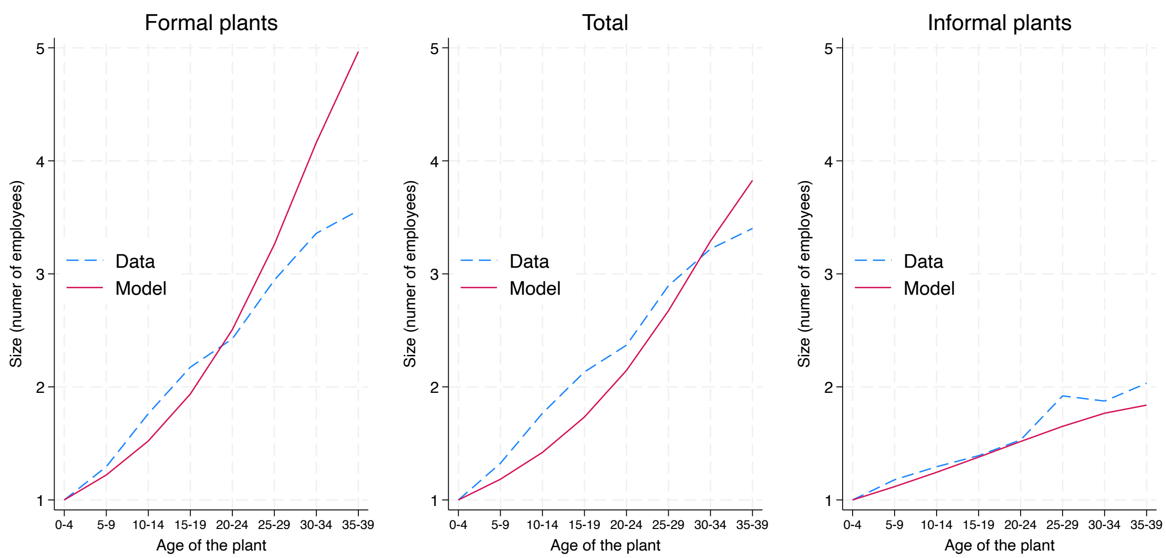


Figure 6: Non-targeted moments. The life cycle of plants: benchmark model and census data.



5. Counterfactual exercises

Table 5 shows how changes in the enforcement threshold b in this economy would affect plant sizes, aggregate output, selection into entrepreneurship (average managerial ability), and plant growth. When the output tax is fully enforced ($b = 0$), establishments have no incentive to remain small to avoid detection, and average size increases by 26% relative to the benchmark. Moreover, the average plant grows 20% faster during their life cycle, from a factor of 2.83 between ages 0-4 to 35-39 in the benchmark to a factor of 3.38 in the counterfactual. Aggregate output slightly decreases but by less than 1 percentage point.

Table 5: The role of the enforcement threshold.

			Benchmark		
	b=0	b=2	b=3.6	b=6	b=8
Mean size (overall)	6.30	4.95	5.00	5.07	5.12
Formal mean size	6.30	9.96	9.63	9.25	9.24
Informal mean size	0	1.65	2.00	2.32	2.46
Agg. Output	99.11	99.10	100.00	100.58	100.79
Emp. Share of informal plants	0	0.20	0.24	0.27	0.29
Avg. managerial ability	118.15	99.79	100.00	100.60	100.80
Informal growth (0-4 to 35-39)	-	0.62	0.84	1.06	1.20
Formal growth (0-4 to 35-39)	3.38	4.15	3.97	3.80	3.72
Average growth (0-4 to 35-39)	3.38	2.98	2.83	2.76	2.74

Figure 7 shows how the level of the tax affects different features of the size distribution of plants, the size of the informal sector in this economy, and aggregate output. Relative to the benchmark, increasing the tax rate by 5 percentage points (from 25 to 30%) lowers aggregate output by about 8 percent. A higher tax rate generates stronger incentives to avoid it, which can only happen in this economy by investing less in managerial skill to remain small. This results in a higher rate of informality, a lower average plant size, a worse selection into entrepreneurship (lower average managerial ability), and flatter life cycle profiles for the average plant (lower growth rates).

Figure 7: The role of the size of the output tax.

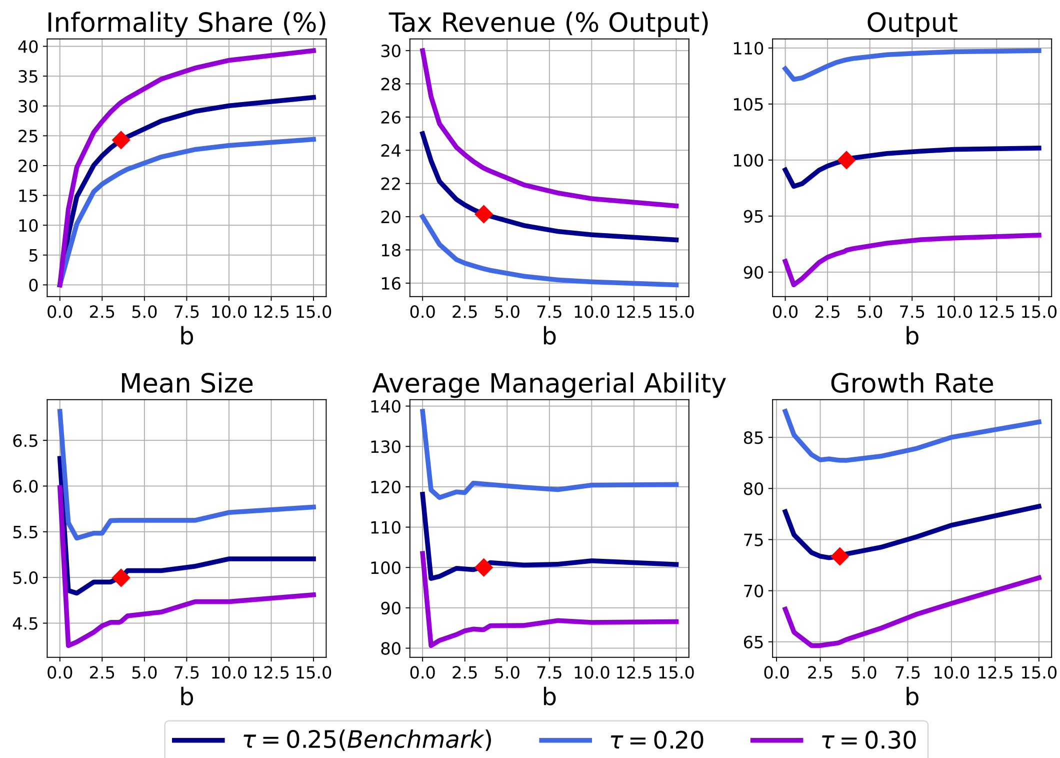


Table 6 shows the results of a policy experiment where we remove the enforcement threshold b (so that all establishments meet the tax official and there is no informal sector) but we lower the tax rate relative to the benchmark to keep the tax-to-output ratio constant at 20%. Average size increases 36% (from 5 to 6.8). Average managerial ability increases by 38%, and the average life cycle growth rate increases by 48%, from a factor of 2.8 between ages 0-4 and ages 35-39 to a factor of 4.2. Aggregate output increases by 8%, mainly due to the lower tax rate since the losses from the removal of the informal sector (through the change in the threshold) are negligible.

Table 6: A tax-to-output-neutral policy experiment.

	Benchmark	Fully Formal Economy
τ	0.25	0.202
b	3.62	0
Informality Share(%)	24.3	0
Output	100	107.9
Mean Size	5.0	6.8
Average Managerial Ability	100	138.3
Skill Investment (as share of output)	2.9	3.2
Growth rate (%) (0-4 to 20-24)	1.1	1.5
Growth rate (%) (0-4 to 35-39)	2.8	4.2
Tax to Output Ratio	0.202	0.202

6. Concluding Remarks

In this paper, we exploit data from Mexico's census plant-level data to quantify the output losses that arise from the marked differences in growth rates between formal and informal plants, which we document. Unlike Hsieh and Klenow (2014), we are able to focus on a panel of registered and unregistered plants to estimate their life cycle employment profiles. We build on the model developed by Guner et al. (2018) where plants grow by investing on their productivity. Our innovation is introducing a tax on output that some plants can escape from by operating informally due to incomplete enforcement. Our policy experiments suggest that reducing the tax rate while improving the enforcement technology to keep the tax-to-output ratio constant would lead to output gains of 8% relative to the benchmark.

References

Busso, M., O. Fentanes, and S. Levy (2018): "The longitudinal linkage of Mexico's economic census 1999-2014," *InterAmerican Development Bank*.

Busso, M., S. Levy, and J. Torres (2019): "Establishment dynamics and the persistence of resource misallocation in Mexico: An analysis of longitudinal data for 1998-2013," *InterAmerican Development Bank*.

Eslava, M., J. Haltiwanger, and Á. Pinzón (2022): "Job Creation in Colombia Versus the USA:'Up-or-out Dynamics' Meet 'The Life Cycle of Plants'," *Economica*, 89, 511–539.

Guner, N., A. Parkhomenko, and G. Ventura (2018): "Managers and productivity differences," *Review of Economic Dynamics*, 29, 256–282.

Hsieh, C.-T. and P. J. Klenow (2014): "The life cycle of plants in India and Mexico," *The Quarterly Journal of Economics*, 129, 1035–1084.

Leal, J. C. O. (2014): "Tax collection, the informal sector, and productivity," *Review of Economic Dynamics*, 17, 262–286.

López, J. J. and J. Torres (2020): "Size-dependent policies, talent misallocation, and the return to skill," *Review of Economic Dynamics*, 38, 59–93.

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

Appendix

A1 Estimating growth in the life cycle of formal and informal plants

To estimate average growth rates for formal and informal plants, we proceed as follows:

1. We sort establishments alive in 2018 into formal and informal, and assume that establishments do not transition into or out of informality during their life cycle.

2. In every period (census wave), we sort formal and informal establishments into 8 age bins: 0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, and 35-39. We then restrict the data to those that were in age bin 1 in any of the four waves prior to 2018. That is, we restrict the analysis in this first stage to establishments both (*i*) alive in 2018 and (*ii*) that were 0-4 in any period between 1998 and 2013 (prior to 2018).

3. We compute cumulative growth rates for each establishment by dividing their size in every period by the size at birth (when they were in the first age bin). For example, for an establishment ages 0-4 in 2003, the cumulative growth rate in 2003 will be 1; the cumulative growth rate between ages 0-4 (age bin 1) and 5-9 (age bin 2) will be the size in 2008 divided by the size in 2003; the cumulative growth rate between ages 0-4 (age bin 1) and 10-14 (age bin 3) will be the size in 2013 divided by the size in 2003; and the cumulative growth rate between ages 0-4 (age bin 1) and 15-19 (age bin 4) will be the size in 2018 relative to the size in 2003.

4. We regress the cumulative growth rates on 2-digit sector and cohort fixed effects, a binary indicator for whether the cumulative rate corresponds to the period before or after 2008 (since before 2008, we only observe the synthetic identifiers generated by Busso et al. (2018)), and the (double) interaction of the indicator for whether the plant is formal or informal with the fixed effects for the age bins. We use the package *reghdfe* in STATA to absorb the sector and cohort fixed effects.

5. To obtain the life cycle of formal and informal plants and the average (or combined) life cycle, we take the derivative of the estimated equation using the STATA command *margins*.

Notice that by construction this exercise will only allow us to obtain the cumulative growth rates through ages 20-24 (since we only observe an establishment for 5 periods at most). We therefore repeat steps 1-5 above for different starting ages:

- We restrict the data to establishments alive in 2018 that were in age bin 2 (5-9) at any point between 1998 and 2013. This way, we can compute growth rates from ages 5-9 through 25-29.

- To obtain the life cycle through ages 30-34, we restrict the data to establishments alive in 2018 that were in age bin 3 (10-14) at any point between 1998 and 2013.

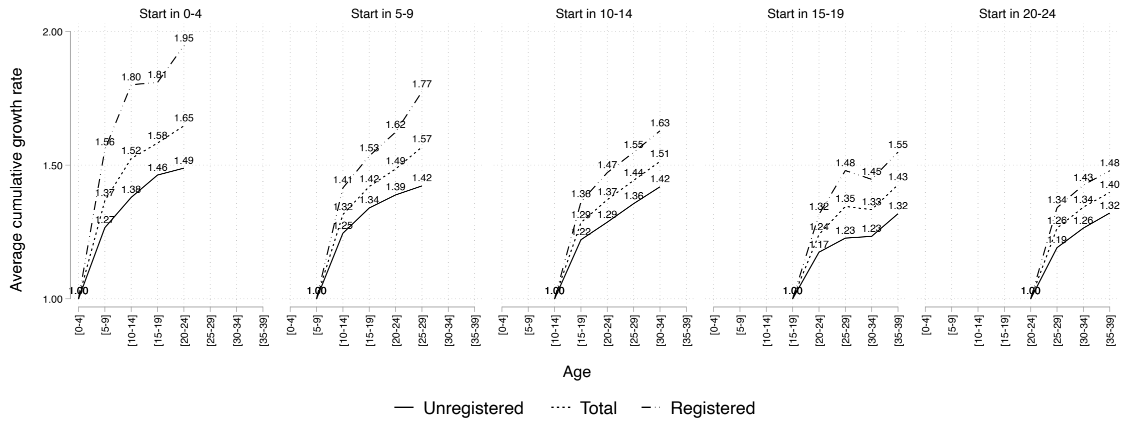
- To obtain the life cycle through ages 35-39, we restrict the data to establishments alive in 2018 that were in age bin 4 (15-19) at any point between 1998 and 2013.

- We also repeat steps 1-5 above restricting the data to establishments alive in 2018 that were in age bin 5 (20-24) at any point between 1998 and 2013.

The tables below show the results from textitreghdfe and *margins*, including the number of unique establishments in each exercise. Note that data from the same establishment might be exploited in every separate regression if the establishment is observed for at least two periods (in addition to 2019).

Repeating the strategy in steps 1-5 using different starting ages will result in different estimates for the cumulative growth rate through the same age, as shown in figure A1. For example, for the average growth rate from ages 5-9 to ages 10-14, we can use the rates from the first panel (with starting ages 0-4) or from the second one (with starting ages 5-9). We average across these results by running the estimated life cycles on dummies for age and dummies for the starting age, separately for formal, informal, and the average plant. Figure 2 plots the results.

Figure A1: Estimated life cycles from...



Notes: Formal establishments are registered with SAT, IMSS, or both. Informal establishments are not registered with SAT or IMSS, did not respond to both questions, or report no registration with one of the two authorities and refused to answer the second question.

A2 Additional tables and figures

Table A1: The establishment and employment size distribution in the census. 2018.

Size	Total number of establishments	Share	Total number of workers	Share of employment	Average size
1	1,818,816	42.26%	1,818,816	8.88%	1.00
2	1,223,455	28.42%	2,446,910	11.95%	2.00
3-5	848,306	19.71%	3,029,170	14.79%	3.57
6-10	226,043	5.25%	1,655,406	8.08%	7.32
11-20	88,567	2.06%	1,274,736	6.22%	14.39
21-50	51,263	1.19%	1,606,644	7.84%	31.34
51-100	16,745	0.39%	1,190,401	5.81%	71.09
101-250	12,411	0.29%	1,908,394	9.32%	153.77
251-500	4,296	0.10%	1,513,206	7.39%	352.24
501-1000	1,966	0.05%	1,359,090	6.64%	691.30
1001+	1,241	0.03%	2,678,432	13.08%	2,158.29
Total	4,293,109	100%	20,481,205	100%	4.77

Table A2: Correlation between formality on the extensive margin and formality as defined by Busso et al. (2019). 2018.

		Registration with SAT and/or IMSS	
		Formal	Informal
Levy	Formal	8.93	0.86
	Informal	24.64	65.57

Table A3: Reg table

	(1) [0-4]	(2) [5-9]	(3) [10-14]	(4) [15-19]	(5) [20-24]
[5-9]	0.253***				
	(0.016)				
[10-14]	0.361***	0.231***			
	(0.024)	(0.005)			
[15-19]	0.440***	0.320***	0.208***		
	(0.038)	(0.007)	(0.006)		
[20-24]	0.464***	0.366***	0.269***	0.170***	
	(0.046)	(0.010)	(0.009)	(0.026)	
[25-29]		0.398***	0.334***	0.222***	0.185***
		(0.015)	(0.014)	(0.042)	(0.010)
[30-34]			0.394***	0.228***	0.256***
			(0.022)	(0.067)	(0.017)
[35-39]				0.311***	0.310***
				(0.099)	(0.026)
Formal	−0.023	−0.020***	−0.014***	−0.022	−0.006
	(0.014)	(0.004)	(0.005)	(0.020)	(0.007)
[5-9]xFormal	0.264***				
	(0.023)				
[10-14]xFormal	0.382***	0.151***			
	(0.028)	(0.006)			
[15-19]xFormal	0.311***	0.172***	0.128***		
	(0.037)	(0.008)	(0.008)		
[20-24]xFormal	0.414***	0.208***	0.167***	0.131***	
	(0.048)	(0.009)	(0.009)	(0.033)	
[25-29]xFormal		0.316***	0.173***	0.236***	0.142***
		(0.016)	(0.013)	(0.044)	(0.012)
[30-34]xFormal			0.188***	0.198***	0.153***
			(0.025)	(0.070)	(0.017)
[35-39]xFormal				0.213*	0.150***
				(0.115)	(0.026)
Sector	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
Synth. panel	Yes	Yes	Yes	Yes	Yes
Observations	6,004,576	4,458,683	3,045,703	2,047,579	1,227,890
Unique obs.	3,006,254	2,088,920	1,450,693	1,042,205	706,722
R^2	0.000	0.004	0.003	0.000	0.002

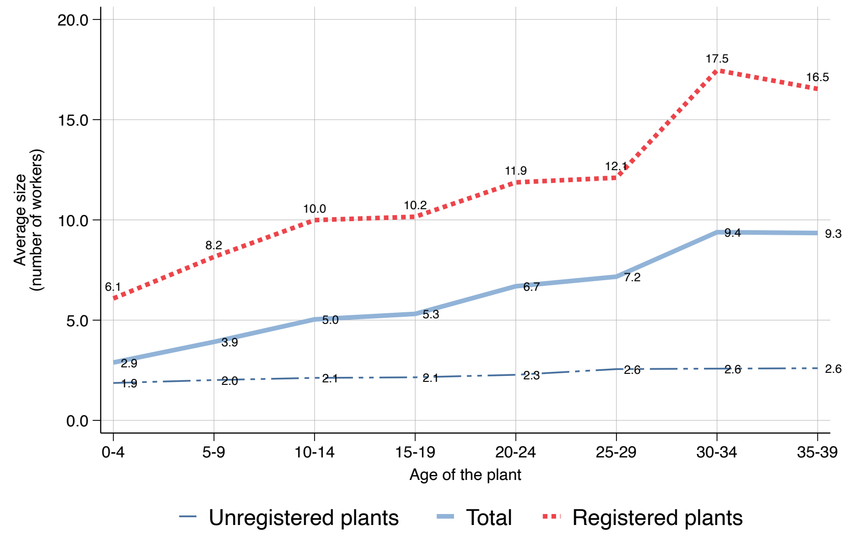
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Margins table

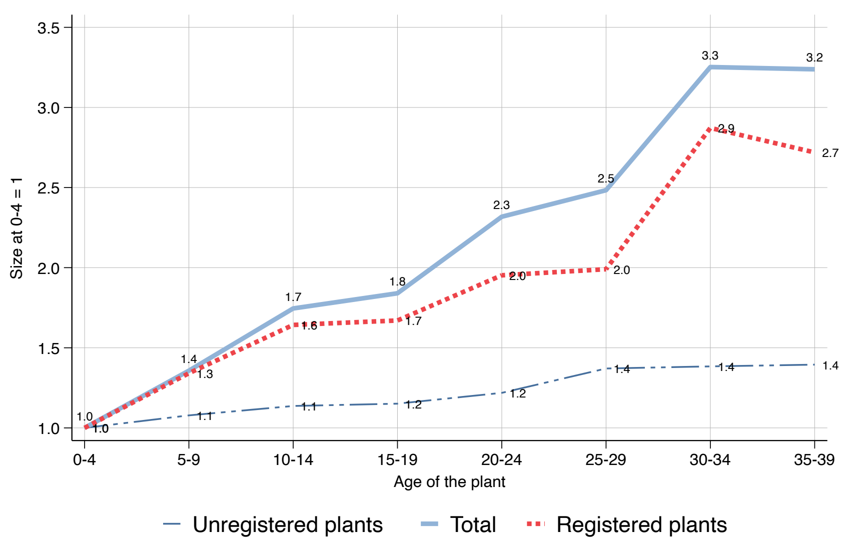
	(1) [0-4]	(2) [5-9]	(3) [10-14]	(4) [15-19]	(5) [20-24]
[0-4]xInformal	0.951				
[0-4]xFormal	0.928				
[5-9]xInformal	1.203	0.942			
[5-9]xFormal	1.445	0.922			
[10-14]xInformal	1.311	1.173	0.940		
[10-14]xFormal	1.670	1.304	0.925		
[15-19]xInformal	1.391	1.262	1.147	0.976	
[15-19]xFormal	1.678	1.414	1.261	0.954	
[20-24]xInformal	1.415	1.308	1.209	1.146	0.967
[20-24]xFormal	1.806	1.496	1.362	1.254	0.961
[25-29]xInformal		1.340	1.274	1.197	1.152
[25-29]xFormal		1.635	1.432	1.411	1.288
[30-35]xInformal			1.333	1.204	1.223
[30-35]xFormal			1.507	1.379	1.370
[35-39]xInformal				1.286	1.277
[35-39]xFormal				1.477	1.421

Figure A2: Cross-section estimates of the life cycle of plants. 2018.

(a) Average number of workers



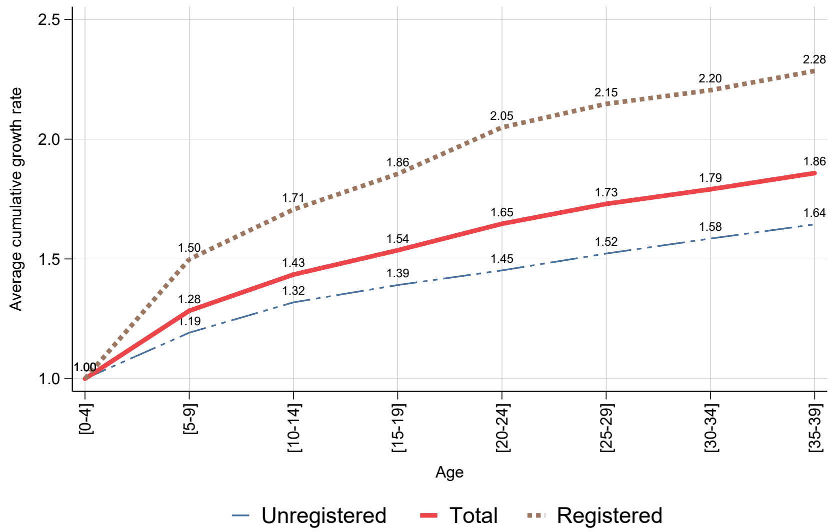
(b) Growth rates



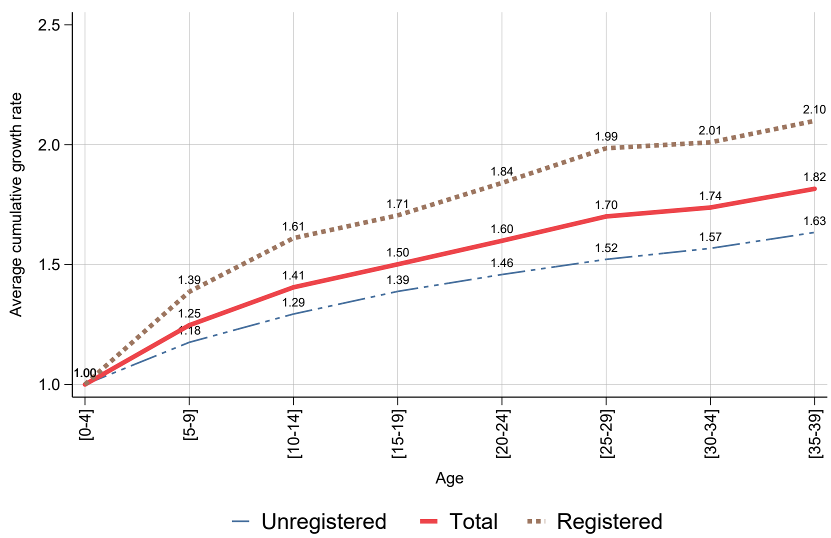
Notes: The top panel shows simple averages of the size of formal and informal plants in each age bin in the 2018 cross-section. Panel b computes growth rates using the data from panel a.

Figure A3: Cumulative growth rates in manufacturing and services.

(a) **Manufacturing**

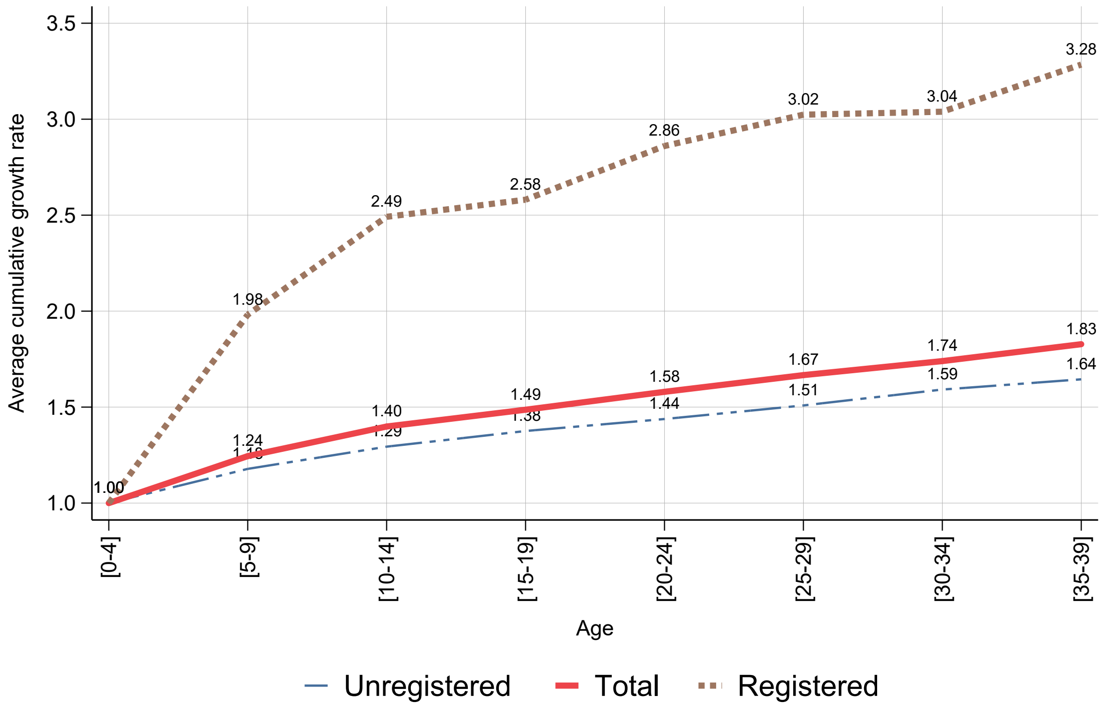


(b) **Services**



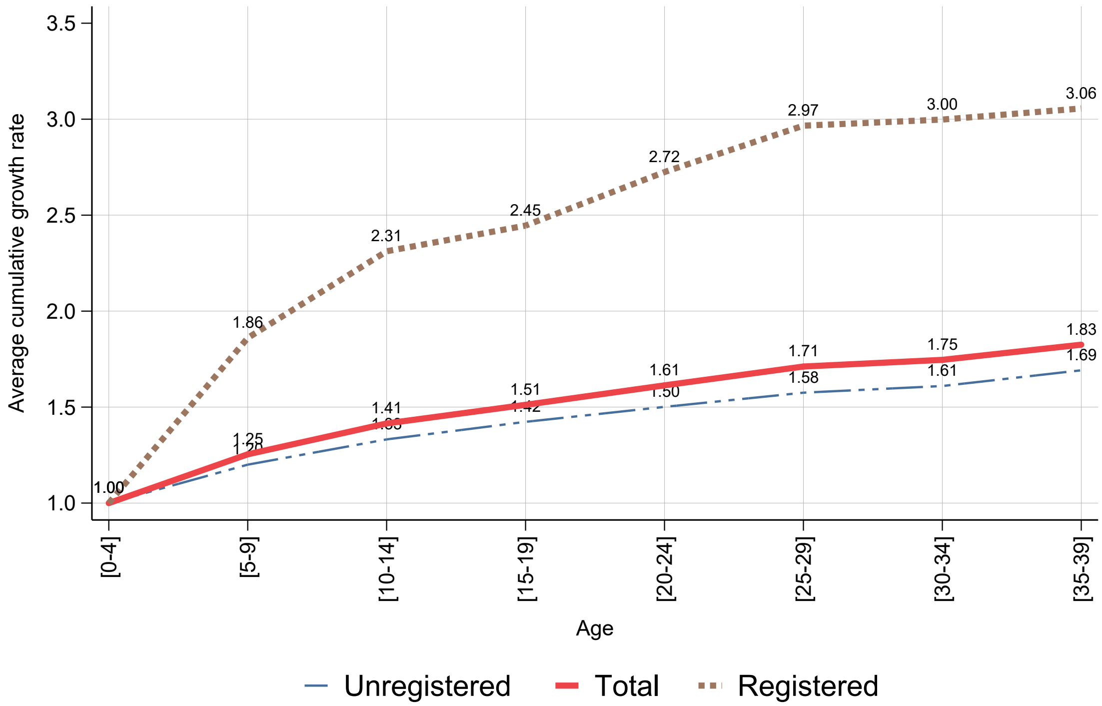
Notes: Formal establishments are registered with SAT, IMSS, or both. Establishments that are not formal are informal.

Figure A4: Estimates for the life cycle of formal and informal plants using a strict definition of informality.



Notes: Formal establishments are registered with both SAT and IMSS. Informal establishments are not registered neither with SAT nor with IMSS (they answered no to both questions).

Figure A5: Estimates for the life cycle of formal and informal plants using the definition of informality on the intensive margin from Busso et al. (2019).



Notes: Informal establishments either: (a) employ non-salaried workers only; (b) employ salaried workers only or a mixture of salaried and non-salaried, but in any case they fully evade their fiscal obligations in social security payments. Formal establishments employ at least one salaried worker and can also employ non-salaried workers; additionally they pay any positive amount in social security.