Size-dependent policies, talent misallocation, and the return to skill*

José Joaquín López†
University of Memphis

Jesica Torres
Inter-American Development Bank

Abstract

We study the allocation of talent in knowledge-based hierarchies subject to a payroll tax that increases with firm size. The tax distorts the allocation of talent across occupations, as well as the sorting of all infra-marginal agents, thus attenuating the strength of the positive sorting throughout the entire economy. This talent misallocation results in lower output, smaller firms, higher self-employment, less wage employment, and lower returns to skill for the most able workers and managers. We quantify the effects of size-dependent policies in two sets of numerical exercises. First, we calibrate a distorted benchmark economy that matches the establishment-level evidence on size-dependent compliance of labor regulations in Mexico. Perfect enforcement of the average effective tax rate increases output by 2%, while the average return to skill for workers (managers) increases by 35% (1%). Eliminating labor regulations increases output by 5%, while the average return to skill for workers (managers) increases by 50% (2%). Second, when we introduce size-dependent policies like those observed in Mexico in an undistorted economy calibrated to the U.S., output decreases by as much as 12%, while the returns to skill for workers (managers) decrease by as much as 60% (7%).

Keywords: Organization of production, returns to skill, misallocation

---

*We thank Rafael Lopes de Melo, Chang-Tai Hsieh, Erik Hurst, and Casey Mulligan for their helpful feedback. We are especially grateful to Santiago Levy for many useful comments and constant encouragement. We have also benefited from useful discussions with Enghin Atalay, Matias Busso, Mariano Bosch, Luca Flabbi, Julio Leal, Bernabé López-Martín, Federico Mandelman, Andrés Neumeyer, Erwan Quintin, Paulina Restrepo, Ctirad Slavík, Jamin Speer, Cesar Tamayo, and John Van Reenen, as well as several participants at various conferences and seminars. All errors are our own.

†Corresponding author. Department of Economics, Fogelman College of Business & Economics, University of Memphis. 3675 Central Ave., Memphis, TN 38152. E-mail: jjlopez@memphis.edu.

1. Introduction

Workers in large firms are, on average, more skilled than workers in smaller firms. Further, larger firms tend to be run by more educated managers, relative to smaller firms. If more talented individuals form larger firms, which, in turn, tend to have a larger share of highly skilled workers, then policies that affect firms of different sizes differently will distort the allocation of talent, as well as the return to skill. In this paper we argue, in particular, that firm regulations which favor small businesses misallocate talent throughout the entire economy, and lower the return to skill.

The relationship between skill composition and firm size holds across countries, and for different measures of skill and size of the firm. Headd (2000) reports that in the U.S. small firms are more likely to employ workers with a high school diploma or less, whereas workers with at least some college are more likely to work in larger firms. Also for the U.S., Cardiff-Hicks et al. (2014) find that higher quality workers are sorted into large firms and large establishments in retailing. Fox (2009) documents evidence for Sweden consistent with hierarchical matching, while Busso et al. (2012) document a positive relation between cognitive skills and firm size for both employers and employees in OECD and Latin American countries.

In this paper, we study the allocation of talent in a knowledge economy when firms are subject to size-dependent regulations, which we model as a tax on labor that increases with firm size. To this end, we embed the hierarchical model of Garicano and Rossi-Hansberg (2004, 2006) into a production economy subject to decreasing returns to scale, as in Lucas (1978). This extension allows us to analyze how individuals with heterogeneous abilities are sorted into occupations and into firms of varying sizes, while delivering a realistic distribution of firm sizes. Under a specific parameterization detailed below, our environment generates closed-form solutions for all equilibrium objects, including a Pareto distribution of firm sizes.

In the model, the skill of an agent completely determines his occupation, the quality of his match, as well as the size of his firm, and the return to skill derives from matching with more talented workers in larger firms. The size-dependent tax encourages managers to constrain the size of their firm, which they achieve by matching with workers of lower skill. The tax distorts not only the allocation of the marginal individual in each occupation, but also the sorting of the infra-marginal workers, thus attenuating the strength of the positive sorting throughout the entire economy. Both talent misallocation and a lower return to skill originate from the same source: distortion on occupational choices, which reorganizes production by re-sorting everyone within occupations, generating a talent mismatch.

To understand the magnitude of these effects, we conduct two sets of numerical exercises. In the first part, we calibrate a size-dependent tax to match the establishment-level evidence on the size-dependent compliance of payroll taxes in Mexico, as documented in Busso et al. (2018). We use this policy distortion to calibrate a distorted benchmark economy that matches the observed allocation of talent in Mexico, as well as several moments of the distribution of firm sizes. We then consider two possible counterfactual policy scenarios: (1) perfect enforcement of the average effective tax rate under size-dependent enforcement, and (2) complete elimination of the payroll tax. Under perfect enforcement, self-employment decreases by 16%, while wage employment increases by 4.4%. Output and productivity in this case increase by 2%. Workers reap the highest benefits from this policy, as their average returns to skill increase by 35%. However, these gains are not evenly distributed across the skill distribution: high skill workers benefit the most, while low skill workers see a reduction in wages, brought about by the higher taxes paid by smaller firms. Not surprisingly, a complete elimination of the payroll tax has larger positive effects: self-employment is cut in half, and wage employment increase by 13%. Output and productivity both increase by 5%. Average returns to skill for workers increase by 50%, and, while all workers see an increase in their wages, high-skill workers experience the largest gains.

In the second part of our numerical experiments, we calibrate an undistorted version of our knowledge-based economy to match some features of the distribution of firm sizes and the allocation of workers in the U.S., and then conduct two counterfactual exercises. First, we introduce the size-dependent tax calibrated from the micro-evidence on payroll taxes in Mexico. We find that this tax generates output and productivity losses of 7%, while the returns to skill for workers (managers) decrease by 18% (4%). Second, we recalibrate the level of the size-dependent tax so as to match the high share of self-employment observed in Mexico (around 20%). The output and productivity losses implied by the observed allocation of talent in Mexico are 12%. The returns to skill drop by 60% for workers, and 7% for managers. The average tax rate under this policy is 61%, which we take as a measure of the efficiency wedge implied by the observed allocation of talent in Mexico, vis-à-vis the U.S.

We make three contributions. First, we contribute to the theory of Pareto distributions of firm sizes by providing a specific example of how a Pareto firm-size distribution can arise from positive sorting between heterogeneous workers and managers, with non-Pareto primitive distributions, and in a purely static environment. This stands in contrast with previous work where Pareto firm-size distributions arise from either assuming that some primitive distribution is itself Pareto—e.g. managerial talent, or firm-level productivity—as in Lucas (1978), or Perla and Tonetti (2014), or from a long sequence of large and persistent positive shocks, as in Luttmer (2007). This part of our work is closely related—yet independently and contemporaneously developed—to the more general theory of Pareto distributions by Geerolf (2017). Moreover, our resulting equilibrium distribution of firm sizes is a double-truncated Pareto distribution, which is a form commonly found in the literature on misallocation and TFP, such as Leal-Ordonez (2014). Unlike Leal-Ordonez (2014), we do not have to assume a double-truncated Pareto distribution of managerial talent to obtain this result.

Second, we show that the skill composition of the firm, which arises endogenously from positive sorting, is crucial to fully understand the effects of firm-specific distortions. In doing so, we fill an important gap in the vast literature on misallocation and TFP spurred by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), where the skill composition of firms has remained absent. Further, by fully considering the role of the skill distribution, our work also contributes to the more specific literature on size-dependent policies. For instance, the works of Braguinsky et al. (2011), Garicano et al. (2016), and Guner et al. (2008), consider heterogeneity in managerial skill, but this talent becomes useless if agents choose to work for a wage. The evidence on positive sorting in quality and quantity discussed above suggests potentially larger aggregate effects through talent misallocation. In a closely related paper, Alder (2016), shows that deviations from positive sorting between CEO and projects (firms) can have sizable aggregate effects, depending on the degree of complementarity between projects and managers, as well as the correlation between mismatch and project quality. Our paper differs from Alder's in that we focus on the effects of size-dependent distortions on talent misallocation across the entire skill distribution—including wage workers, the self-employed, and managers—as opposed to the allocation between heterogeneous managers and projects of varying quality.

Our third contribution is showing that size-dependent regulations could significantly lower the average return to skill in the economy. Our model features superstar effects, in the sense that the earnings schedule is convex in ability, as in Rosen (1981) and, more recently, Scheuer and Werning (2017). Therefore, because the best managers are matched with the best wage workers, distortions that increase with firm size act as an increasing marginal tax schedule that attenuates the strength of the positive sorting, disproportionately lowering the return to skill for the most able workers and managers.

Even though our model does not feature endogenous skill formation, as in Bobba et al. (2017), it still has important implications for the role of human capital and industrial policies in the process of development. Our model predicts misallocation of talent and a lower return to skill as consequences of size-dependent policies for a given distribution of skill. That means that even if individuals became more skilled for other reasons, they would still be misallocated as long as the size-dependent policies persisted. This prediction is in line with the evidence for Mexico presented in Levy and López-Calva (2016), who document a growing mismatch between the supply and demand for skilled labor: while the amount of available skilled workers has grown, the earnings of more educated workers have decreased. While Levy and López-Calva (2016) suggest a link between size-dependent policies and the return to skill, our paper is the first to formalize such link. Our results are also consistent with the long-lasting effects of development policies highlighted by Buera et al. (2013), and with the negative effects of costly formalization on the demand for skilled labor presented in D'Erasmo et al. (2014).

The rest of the paper is organized as follows. In the next section, we present an overview of our slightly modified version of the hierarchical model of Garicano and Rossi-Hansberg (2004, 2006), without any taxes. In section 3, we completely characterize a parametric example that serves as the basis of our analysis. In section 4 we study the effects of a size-dependent tax that attenuates, but does not break down positive sorting. Section 5 contains the quantitative analysis of the effects of the size-dependent tax described in section 4. We also show that such a tax policy fits well the establishment-level evidence on the size-dependent compliance of payroll taxes in Mexico, and present an overview of motivating facts about the Mexican economy that our model can address. In section 6, we study the qualitative effects of a size-dependent policy that breaks down positive sorting along a segment of the skill distribution. Section 7 concludes.

2. The allocation of talent: knowledge hierarchies and decreasing returns

The basic undistorted environment is the general equilibrium, continuous assignment model of Garicano and Rossi-Hansberg (2004, 2006), which we embedd into an economy subject to decreasing returns to scale, as in Lucas (1978). This basic extension is not necessary to derive our main analytical results, but we find it convenient for two reasons: 1) it allows for a more ample support for the distribution of firm sizes, and 2) serves as a useful—and familiar—general setup to study the effects of certain size-dependent policies that break positive sorting along a segment of the skill distribution. We consider an equilibrium where firms have two and one layers. The latter represent the self-employed in the economy—business owners without employees. Unlike Sattinger (1993), here the assignment is many-to-one, the densities in each sector are endogenous, and there is the option of not matching, that is, going into self-employment.

Individuals differ in a single trait—call it talent—and own one unit of time. They choose the use of their time and talent that maximizes their earnings. They can either produce alone or in a team (firm or business) with other workers. They work together to specialize either in managerial activities—running the business—or in production activities—working as one of the firm's employees—and thus exploit complementarities in production. Individuals have then three options: to produce alone, to run a business hiring others, or to work for a wage for someone else. The formation of firms in the model is also endogenous—the managers optimally select the size of their firm, as well as the quality of their employees, whereas the wage workers optimally select which of the continuum of teams to join.

To produce in this economy, workers solve problems which vary in their difficulty, z, according to some density $g(z)$. Skill is cumulative—a worker of skill z can solve all problems of difficulty less than or equal to z. Workers draw and attempt to solve one problem in their unit of time and produce only if they know the answer. The (expected) earnings of worker z are therefore the percentage of problems he is able to solve: $G(z)$. The skill endowment z varies continuously in the population according to the (given) skill distribution $F(z)$, with support $[L, H]$, and density $f(z)$.

Agents can also form teams, each consisting of identical production workers and one manager. In these teams, the manager attempts to solve the problem whenever his production workers do not know the answer, and production workers do not interact with each other. More precisely, a team with n employees draws n problems, and the (expected) output of the team is the percentage of tasks that a manager with skill z_m is able to solve in his n units of time,

$$y = G(z_m)\, n^{\alpha}$$

where $\alpha \in (0,1)$ denotes the degree of decreasing returns to scale in the use of time, as in Lucas (1978).

Communication in teams is costly: employees of skill z_p will ask the manager with probability $1 - G(z_p)$, and solving these problems costs a fraction $h(z_p)$ of the manager's time, per worker. We assume that these costs are bounded: $h(L) = \overline{h}$ and $h(H) = \underline{h}$, with $0 < \underline{h} < \overline{h}$. Thus, communication in teams is always costly: communicating with a worker who knows all the answers still takes a fraction of the manager's time. Communication costs then limit the entrepreneur's span of control, or firm size, n. Within his unit of time, a manager is able to coordinate at most $1/h(z_p)$ workers, and this way

$$nh(z_p) \leq 1.$$

In other words, if wage workers are of quality z_p, a manager can coordinate at most $n(z_p)$ employees. The problem of a manager with ability z_m is to choose the quality of his employees, z_p, and the size of his firm n, so as to solve

$$\begin{aligned} R(z_m) = \max_{z_p, n} \; & G(z_m)\, n^{\alpha} - w(z_p)\, n, \\ s.t. \; & n \leq n(z_p), \end{aligned} \tag{1}$$

where $w(z_p)$ denotes the equilibrium wage rate, and

$$n(z_p) = \frac{1}{h(z_p)}.$$

Without any distortion, the inequality constraint will always bind: from the manager's point of view, firm size l is feasible hiring wage workers of any type above $n^{-1}(l)$, but costs are minimal when employing workers of skill $n^{-1}(l)$, who earn the lowest wage, and that is true for all firm sizes. Firm sizes in the model are therefore simply a function of the knowledge in the bottom layer of the firm. Thus, we can rewrite output as

$$y\left(z_p, z_m\right) = G\left(z_m\right)\left[n\left(z_p\right)\right]^{\alpha}, \tag{2}$$

and managerial rents are thus

$$R\left(z_m\right) = \max_x G\left(z_m\right)\left[n\left(x\right)\right]^{\alpha} - w\left(x\right)n\left(x\right). \tag{3}$$

Individuals take wages and managerial rents as given, and choose the occupation that yields the highest earnings given their skill:

$$\max\left\{w\left(z\right), G\left(z\right), R\left(z\right)\right\}. \tag{4}$$

The equilibrium exhibits positive sorting in both quality and quantity, as in Garicano and Rossi-Hansberg (2004, 2006) and Garicano and Hubbard (2012). Specifically, the best managers match with the best wage workers to form the largest teams, the second-best managers match with the second-best wage workers to form the second-largest firms, and so on. The smallest firms in the market are thus the match of the least-skilled entrepreneurs to the least-skilled wage workers.[1]The equilibrium also exhibits perfect stratification of individuals into occupations based on their skill: the less skilled agents become production workers, those in the middle produce alone, while the most skilled ones work managing others. More formally, a competitive equilibrium is defined as follows.

Definition 1 *Given a distribution of talent $F(z)$ over $[L, H]$, a distribution of problems $G(\tilde{z})$, a communication technology $h(z)$, and a production technology $y = G(z_m)n(z_p)$, a competitive equilibrium consists of (i) an assignment function sorting individuals into occupations and into firms; (ii) a wage function $w(z)$; (iii) managerial rents $R(z)$; and (iv) a pair of cutoffs $\{z_1 < z_2\}$, where $[L, z_1]$ is the set of workers, $[z_1, z_2]$ is the set of owners producing alone, and $[z_2, H]$ is the set of firm managers, such that: (E1) no agent desires to switch to another occupation or firm; (E2) the supply of workers equals the demand for workers; (E3) the supply of managers equals the demand for managers.*

Since the assignment function matches workers to managers and firm sizes, conditions (E2) and (E3) are equivalent. Figure 1 shows the equilibrium allocation of workers into occupations according to their skill level.[2]

[1]The positive sorting of workers is guaranteed as long as the second-order sufficient condition of the entrepreneur's problem is satisfied.

[2]Unlike the equilibrium in the original framework, the equilibrium in this model could exhibit

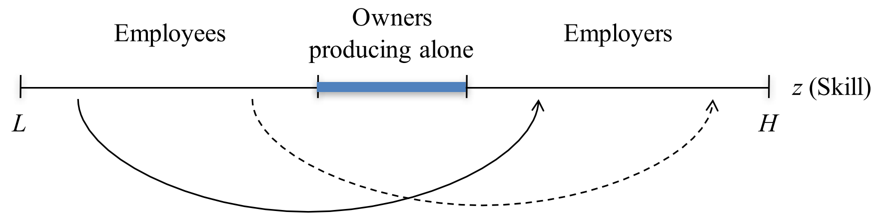

Figure 1
Allocation of Workers across Occupations

To find the equilibrium, we follow the algorithm in Sattinger (1993), as described by Garicano and Rossi-Hansberg (2004, 2006). We conjecture the existence of the two thresholds, z_1 and z_2, such that agents with skill $z \leq z_1$ optimally become wage workers, agents with skill $z_1 < z < z_2$ optimally choose to produce alone, and agents with skill $z \geq z_2$ optimally choose to manage others. The first step is to find the equilibrium assignment function $z_m = m\left(z_p\right)$, which denotes the manager z_m that corresponds to workers of skill z_p.

In equilibrium the market for managers clears: for all $s \in [L, z_1]$ it has to be the case that

$$\int_L^s \frac{f\left(x\right)}{n\left(x\right)} dx = F\left(z_m\left(s\right)\right) - F\left(z_2\right), \tag{5}$$

which implies

$$\frac{f\left(s\right)}{n\left(s\right)} = f\left(m\left(s\right)\right) \frac{dm}{ds}. \tag{6}$$

We solve for $m\left(z_p\right)$ using equation 6, with $m\left(L\right) = z_2$—which states that the worst wage workers are matched with the worst employers—as the boundary condition. The second step is to combine the first order condition (FOC) of the entrepreneur's problem with the equilibrium assignment function to find the equilibrium wage function. The firm's problem is to select z_p to maximize

$$R(z_m) = G\left(z_m\right)\left[n\left(z_p\right)\right]^{\alpha} - w\left(z_p\right) n\left(z_p\right).$$

segregation of workers, as in Kremer and Maskin (1996), in which the top agents match together whereas workers at the bottom produce alone. Whether the equilibrium entails segregation or not depends on the sensitivity of earnings to skill in the outside option, the skill distribution, and the strength of complementarities in production. In our analysis we focus on the equilibrium without segregation as the one displayed in Figure 1.

The FOC is

$$G\left(z_m\right)\alpha\left[n\left(z_p\right)\right]^{\alpha-1}\frac{dn}{dz_p}=\frac{dw}{dz_p}n\left(z_p\right)+w\left(z_p\right)\frac{dn}{dz_p}. \tag{7}$$

We then plug in $z_m = m\left(z_p\right)$ from step one above, and solve the differential equation for $w\left(z_p\right)$ using the equilibrium condition $w\left(z_1\right) = G\left(z_1\right)$, which states that the best wage workers are indifferent between working for a wage or producing alone.

The final step is to pin down the constants of integration in the two previous steps to completely characterize the assignment function $z_m = m\left(z_p\right)$, and the earnings functions $w\left(z\right)$ and $R\left(z\right)$. To do so, we solve $m\left(z_1\right) = H$ and $R\left(z_2\right) = G\left(z_2\right)$. The former condition states that the best managers match with the best workers, while the latter states that the worst managers are indifferent between running a team and producing alone.

3. Parameterization of the undistorted environment

We now fully characterize a specific parametric example, which we will rely upon in the calibration and counterfactual policy experiments that follow.

We assume that the distribution of problems, $G(z)$, is uniform in the interval $[L, H]$. The population skill distribution, $F(z)$, is assumed to be a double-truncated exponential distribution over the interval $[L, H]$, with parameter λ, that is,

$$F\left(z\right) = \frac{\exp\left[-\lambda L\right]-\exp\left[-\lambda z\right]}{\exp\left[-\lambda L\right]-\exp\left[-\lambda H\right]}$$

Let $q(x) = 1 - G(x)$ denote the probability that a worker of skill x asks a question to the manager. Communication costs take the form

$$\begin{aligned} h\left(x\right) &= a\exp\left[bq\left(x\right)\right], \\ &= a\exp\left[b\left(\frac{H-x}{H-L}\right)\right], \end{aligned}$$

where $a, b > 0$ are parameters.

Assumption 1 *Define*

$$\gamma \equiv \frac{a\lambda\left[H-L\right]}{\lambda\left(H-L\right)+b}.$$

Assume

$$\gamma exp\left[\lambda\left(z_2-L\right)+b\right] = 1$$

In this specification and under **Assumption 1**, closed form solutions exist for all the equilibrium objects. All proofs are provided in the Appendix.

Proposition 1 *Under Assumption 1, the equilibrium assignment is a linear function given by*

$$m(z)=z\left[1+\left(\frac{1}{\lambda}\right)b\left(\frac{1}{H-L}\right)\right]-\left(\frac{1}{\lambda}\right)ln\gamma-\left(\frac{1}{\lambda}\right)b\left(\frac{H}{H-L}\right).$$

The equilibrium wage function is

$$w\left(z\right) = \left[\frac{1}{n\left(z\right)}\right]\left\{exp\left[k_3+k_4z\right]\left[\frac{k_2}{k_4}-\frac{k_1}{k_4^2}+\frac{k_1z}{k_4}\right]+G\left(z_1\right)n\left(z_1\right)-C(z_1)\right\}.$$

Where $C(z_1) = exp\left[k_3+k_4z_1\right]\left[\frac{k_2}{k_4}-\frac{k_1}{k_4^2}+\frac{k_1z_1}{k_4}\right]$*. The constants* k_i*,* $i \in \{1,2,3,4\}$ *are defined in the Appendix.*

The cutoffs z_1 and z_2 are then

$$z_1 = H+\frac{(H-L)\ln\gamma}{b+(H-L)\lambda},$$

$$z_2 = L+\frac{w\left(L\right)n\left(L\right)}{n\left(L\right)^{\alpha}-1}\left[H-L\right].$$

The condition needed by Assumption 1 then becomes

$$\left(-\frac{1}{\lambda}\right)\left[b+\ln\gamma\right] = \frac{w\left(L\right)n\left(L\right)}{n\left(L\right)^{\alpha}-1}\left[H-L\right].$$

The specific functional forms in this example also deliver a closed form for the distribution of firm sizes.

Proposition 2 *The firm-size distribution (FSD) is a double-truncated Pareto distribution with power* $p \equiv \lambda [H - L] \left(\frac{1}{b}\right) + 1$*. The density thus writes*

$$\tilde{f}(n) = \frac{[n(L)]^p p n^{-p-1}}{1 - \left(\frac{n(L)}{n(z_1)}\right)^p}.$$

An obvious implication of Proposition 2 is that the power of the Pareto distribution is always strictly greater than one, which stands at odds with the evidence for the U.S. presented in Axtell (2001). The power parameter will be close to one when H is very close to L, and/or when the communication cost parameter b is very high.

Figure 2 displays the earnings profiles and occupational choices for this economy, whereas Figure 3 shows the equilibrium assignment function. As conjectured in the algorithm, in this example it is the case that $w(z) \geq G(z)$, and $w(z) \geq R(z)$, for all $z \in [L, z_1]$, $G(z) \geq w(z)$ and $G(z) \geq R(z)$ for all $z \in [z_1, z_2]$, and $R(z) \geq G(z)$ and $R(z) \geq G(z)$ for all $z \in [z_2, H]$. Therefore, the conjecture is correct and the allocation is an equilibrium.

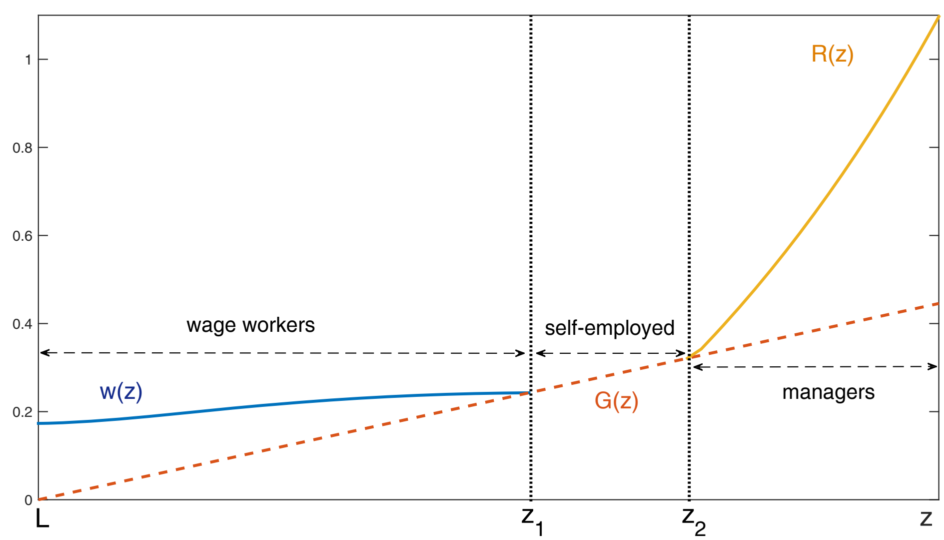

Figure 2
Earnings and Occupational Choices in Undistorted Environment
under Parametric Assumptions

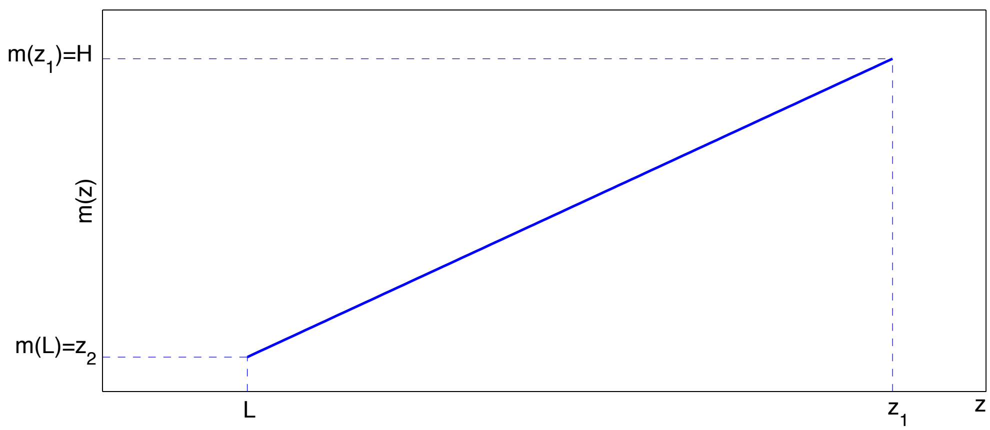

Figure 3
Equilibrium Assignment Function in Undistorted Environment under Parametric Assumptions

4. Effects of a size-dependent payroll tax

In this section we examine the effects of a general payroll tax whose enforcement increases with firm size. More precisely, we assume that the effective tax rate can be summarized by a function $\tau(n)$, with $\tau'(n) > 0$. Under this policy, the firm's optimization problem is now:

$$\max_{z_p, n} G(z_m) n^{\alpha} - [1 + \tau(n)] w(z_p) n, \text{ s.t. } n \leq n(z_p),$$

where $n(z_p) \equiv 1/h(z_p)$, as before. The tax policy we consider takes the following form

$$\tau(n) = \tau_s [1 - \exp(-\kappa n)]. \tag{8}$$

Where $\tau_s \in [0, 1)$ is the statutory tax rate, and $\kappa \geq 0$ is an enforcement parameter. Figure 4 plots this tax function for different parameter values. Note that the parameter τ_s is a level shifter, whereas the enforcement parameter κ affects the steepness of the tax policy.[3]

[3] This type of negative exponential tax functions can be used to model progressive taxation, as in López and Torres (2018a), and tax evasion, as in López (2017).

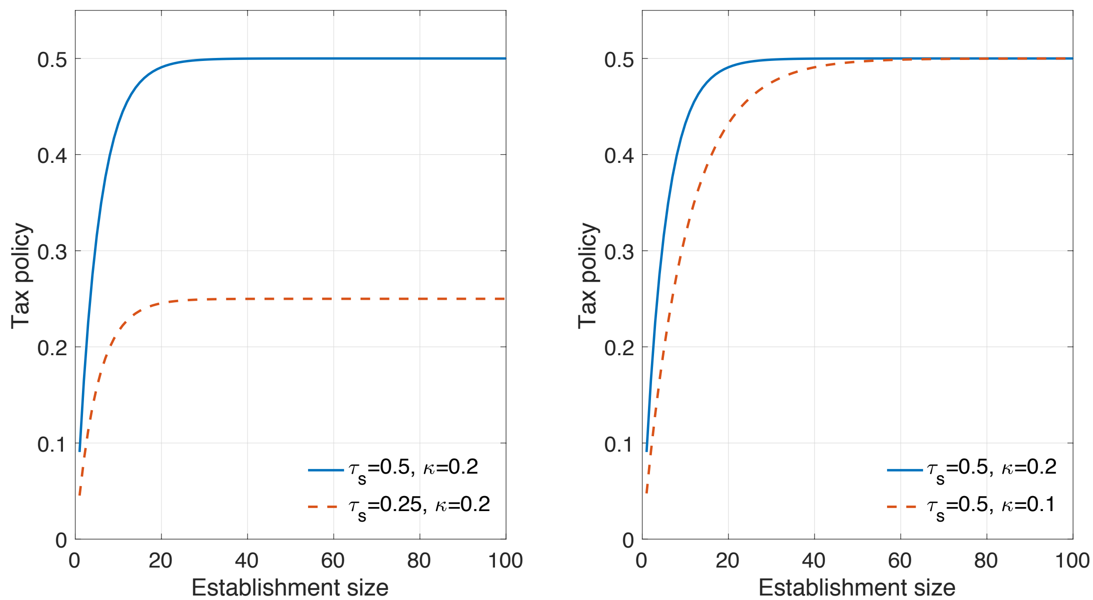

Figure 4
Size-dependent Tax Policy.

Because $\tau''(n) < 0$ and $n''(z_p) > 0$, the size constraint in the firm's problem always binds. Then, we can write the first order condition to the manager's problem as

$$G(z_m)\alpha[n(z_p)]^{\alpha-1}n'(z_p) = \left[w'(z_p)\,n(z_p) + w(z_p)\,n'(z_p)\right]\left[1+\tau(n(z_p))\right] + w(z_p)\,n(z_p)\,\tau'(n(z_p))\,n'(z_p)\,. \quad (9)$$

The algorithm to find the equilibrium under this type of size-dependent policy is similar to the one described above for the undistorted case. The labor market clearing condition in equation 6 used to derive the assignment function $z_m = m(z_p)$ remains unchanged, including the boundary condition $m(L) = z_2$. Once we solve for $m(z_p)$, we plug it into our first order condition in equation 9, and solve the differential equation for $w(z_p)$, using the indifference condition $w(z_1) = G(z_1)$ as the boundary condition. The final two equilibrium conditions used to pin down the cutoffs $\{z_1, z_2\}$ are $m(z_1) = H$—which states that the best manager matches with the best workers—and $R(z_2) = G(z_2)$—which states that the worst manager is indifferent between running a firm and being self-employed without employees.[4]

Even though managers still use the communication technology to determine the size of their firm, the size-dependent tax creates an incentive to form smaller firms, which can be achieved by matching with workers of lower quality compared to the undistorted equilibrium. Figure 5 shows the effects of two size-dependent taxes with the same level of enforcement, but different statutory tax rates. Notice that as the tax increases, the quality of the best worker, and therefore the size of the largest team, decreases, and workers are as a result matched with managers of increasingly more talent relative to the undistorted environment. Thus, the size-dependent tax not only changes the marginal individuals in each occupation, but also resorts everyone within occupations, weakening the positive sorting in this economy.

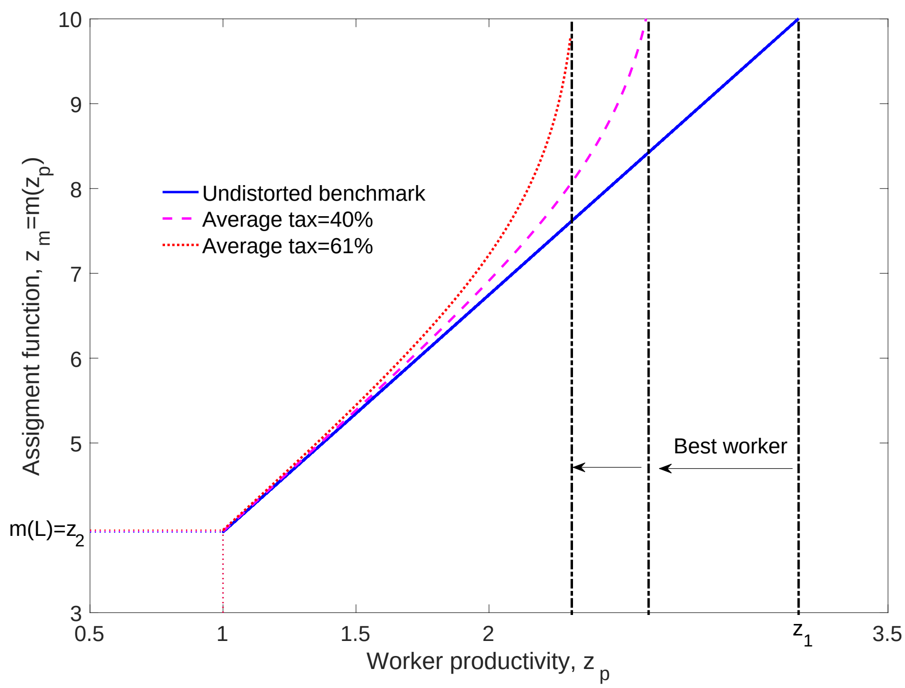

Figure 5

Effect of Size-dependent Policies on Equilibrium Assignments.

[4]The fact that the algorithm remains the same means that, starting from a distorted benchmark, we can derive closed form solutions for all equilibrium objects using a distorted version of the parametric example described in Section 3 (details are provided in the Appendix), which is particularly useful in the numerical exercises in the next section.

The tax reduces the demand for skill, which depresses wages and pushes the most talented workers into self-employment. This misallocation of talent results in lower output and productivity, and in lower earnings and returns to skill for the best workers and managers. Figure 6 shows the effects on earnings and the allocation of talent of the two size-dependent taxes used in Figure 5. Whereas the lower wages experienced by workers are the result of a decrease in the demand for skill, the lower earnings of managers come from two sources: a worse match quality, and the payroll tax.

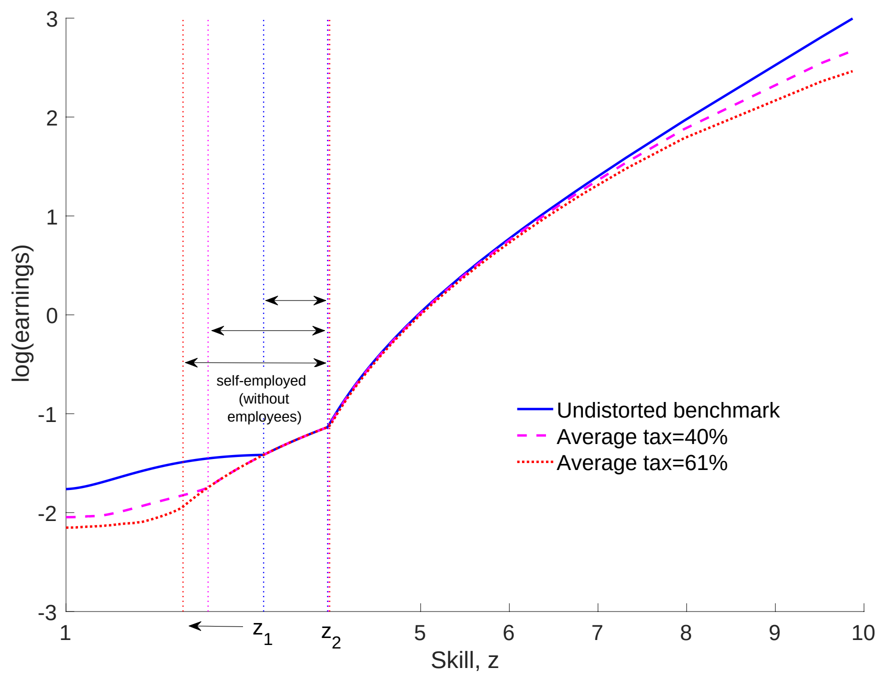

Figure 6
Effect of Size-dependent Policies on Equilibrium Earnings
and the Allocation of Talent.

In this model, the effects of size-dependent policies can be decomposed into the usual size effect, plus a talent mismatch effect not found elsewhere in the literature on misallocation. By distorting the selection of workers into occupations, size-dependent policies affect the rank of each individual within each occupation, and managerial talent is unambiguously wasted as the top managers now spend more time communicating with less-talented employees.

To sum up, the size dependent tax results in the following: i) a lower average firm size, ii) a talent mismatch between managers and workers, iii) worse average wage worker quality, iv) lower earnings and returns to skill, disproportionally affecting the best wage workers and the best managers, v) more self-employment, driven mainly by the best workers who choose to produce alone after the tax, and vi) less wage employment.

5. Data and quantitative exercises

Over the past 20 years, the labor market in Mexico has been characterized by persistently high levels of business ownership. Around 30% of workers run their own business—three times the entrepreneurship rate observed in the U.S. (U.S. Bureau of Labor Statistics). The fraction of non-employers, in particular, is one of the highest among OECD countries (OECD, 2017) and, as Figure 7 shows, has remained more or less stable since 1994.

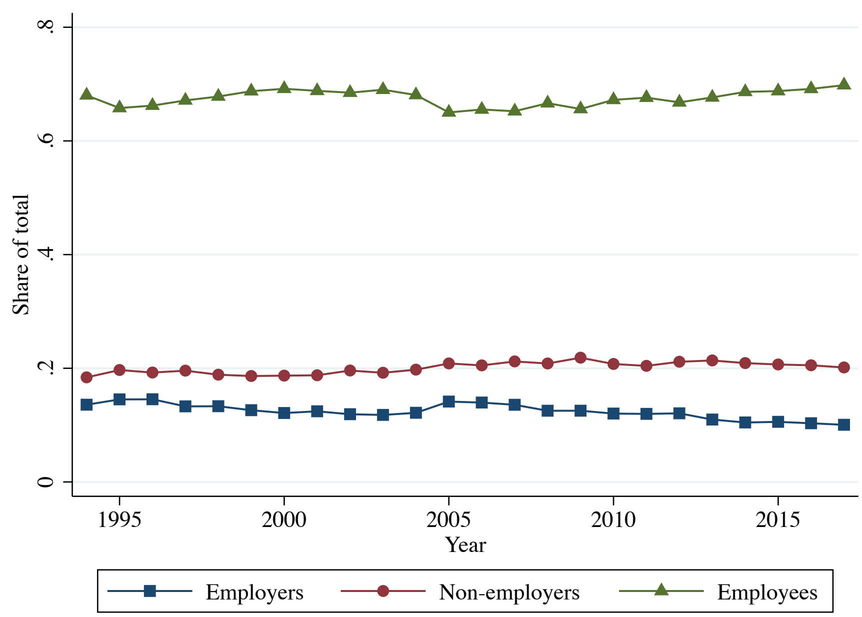

Figure 7

Distribution of Workers across Occupations in Mexico. 1994-2017.

Source: Author's calculations using data from Mexico's National Survey of Occupations and Employment (ENOE, by its Spanish acronym) and National Survey of Urban Employment. Data for workers ages 18-64 in urban areas.

The distributions of firm and establishment sizes in Mexico have remained stable as well.[5] Micro businesses make up for most of Mexican enterprises: the average firm size in Mexico is 5, 96% of firms have 10 or fewer employees, and roughly 1% of them employ more than 50 workers. In contrast, the average American firm employs 21 workers, 79% of U.S. firms employ 10 or fewer workers, whereas 4% of them employ more than 50 workers.

The last 20 years in Mexico have also seen a growing mismatch between the supply and demand for skilled labor: while the amount of available skilled workers has grown, the earnings of more educated workers have decreased (Levy and López-Calva, 2016), which suggests the presence of distortions that reduce the demand for, and hence, the returns to skill. This stands in sharp contrast with the evidence for advanced economies, such as the U.S., where the skill premium is often described as the result of supply not being able to keep up with the demand for skill.

In economies where workers sort positively in quality and quantity, size-dependent distortions disrupt the allocation of talent by increasing self-employment, resulting in smaller firms, lower returns to skill, and, ultimately, lower aggregate productivity. Busso et al. (2018) document evidence of two size-dependent distortions in Mexico: smaller establishments employ a higher fraction of non-salaried workers (which are exempt from complying with labor regulations) and pay on average a lower fraction of social security contributions for their salaried employees. We examine the effects of the latter using our hierarchical model with decreasing returns, and show that eliminating the size-dependent nature of the enforcement of these regulations accounts for a sizable share of the gains that would be realized if they were eliminated altogether.

Labor regulations in Mexico amount to 44% of the average wage of a formal employee: 23 percentage points from contributions to social security (pension benefits, medical services, and a social housing fund) and the rest from required days of paid leave, some dismissal restrictions, and a mandatory yearly bonus (Alaimo et al., 2017). Figure 8 shows a size-dependent tax policy that fits the compliance with social security contributions documented in Busso et al. (2018). We do not observe the relationship between compliance and firm size for the remaining labor regulations, but we assume that they share the same enforcement parameter calibrated from the

[5] Busso et al. (2018) document no major changes in the establishment size distribution since 1998. In Mexico the establishment size distribution does not markedly differ from the size distribution of firms (Mexico's National Institute of Statistics or INEGI, by its Spanish acronym).

data in Figure 8. In other words, we assume that all labor regulations in Mexico can be summarized by a negative exponential tax policy like the one in equation 8, with $\tau_s = 0.44$ and $\kappa = 0.18$. As seen in Figure 9, our assumption retains the calibrated slope, while changing the level of the tax.

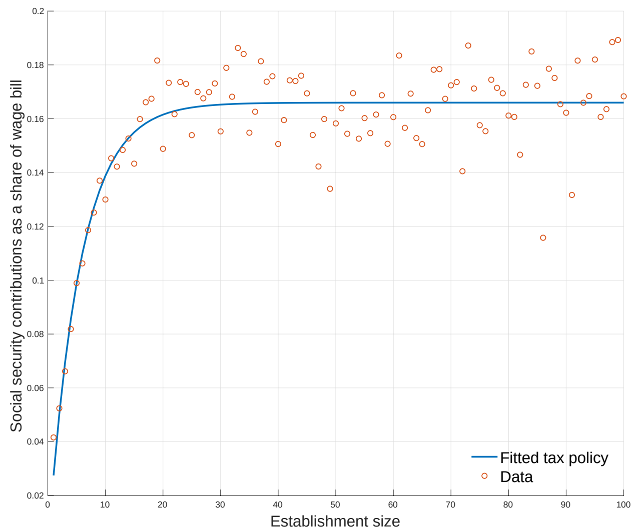

Figure 8

Average Social Security Contributions as a Fraction of Wages in Fixed Establishments in Mexico.

Source: Data from Busso et al. (2018). The fitted tax policy is $\tau(n) = \tau_s(1-\exp(-\kappa n))$, with $\tau_s = 0.1660$ and $\kappa = 0.18$.

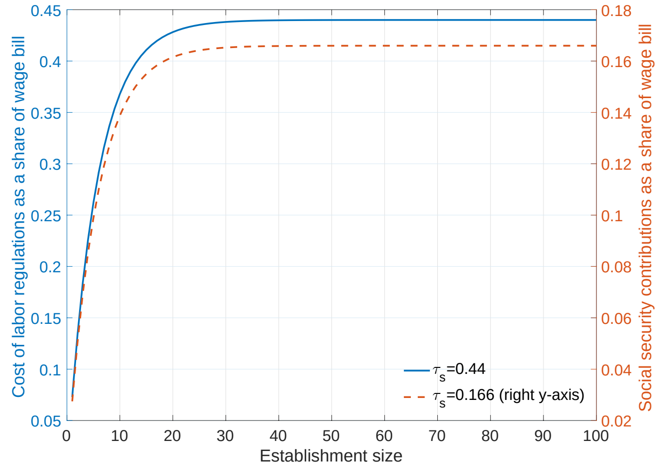

Figure 9

Calibrated Tax Policies for Social Security Contributions and the Total Labor Cost.

The fitted tax policy is $\tau(n) = \tau_s(1 - \exp(-\kappa n))$ with $\kappa = 0.18$ in both cases.

5.1. Calibration to Mexico

The first set of quantitative exercises we perform considers an economy with a size-dependent payroll tax like the one described in Section 4, calibrated to match the total cost of labor regulations in Mexico, as summarized in the previous section. Our benchmark economy with size-dependent taxes is then calibrated to match features of the allocation of workers and the distribution of establishment sizes in Mexico.[6] We then conduct two counterfactual policy experiments: i) perfect enforcement of the effective average payroll tax calculated under size-dependent enforcement, and ii) complete elimination of the payroll tax.

We set the lower bound of the distributions of problems and skill to $L = 1$, and then calibrate jointly the rest of the parameters $(\alpha, a, b, H, \lambda)$. The targets are: i) the parametric restriction in Assumption 1, necessary for a closed form solution to exist, ii) the average firm size, iii) the share of self-employed workers (business owners without employees), and iv) the power coefficient of a Pareto tail fitted using establishment-level data on the entire universe of Mexican establishments in manufacturing, services, and wholesale and retail (Busso et al., 2018).[7] All parameter

[6]We exploit the fact that closed-form solutions exist in a distorted benchmark version of the parametric example described in Section 3 to conduct our baseline calibration with distortions.

[7]We fit the parameters by minimizing the sum of squared errors between the targets and their

values are in Table 1.

Table 2 shows the results from the calibration of our benchmark distorted economy. The model matches very well all targeted moments, as well as the non-targeted shares of wage workers and employers. The model also generates a (non-targeted) distribution of firms that is similar to that in the data, but overstates the share of medium-sized firms.

Table 1
Parameter Values—Calibration to Mexico

Parameter	Definition	Value	Source
τ_s	Statutory payroll tax	0.44	Alaimo et al. (2017)
κ	Enforcement intensity	0.18	Calibrated
L	Lower bound for $G(\cdot)$ and $F(\cdot)$	1	Assigned
α	Returns to scale	0.7338	Joint calibration
a	Communication cost level	1.43×10^{-09}	Joint calibration
b	Communication cost slope	19.74	Joint calibration
H	Upper bound for $G(\cdot)$ and $F(\cdot)$	7.66	Joint calibration
λ	Exponential parameter (skills)	1.001	Joint calibration

model counterparts.

Table 2
Calibration Results: Mexico

Targeted	Data	Model
Average firm size	5	5
Share of non-employers	0.20	0.19
Share of firms of size $>= 50$	0.01	0.01
Pareto power	1.6	1.34
Non-targeted	Data	Model
Share of employers	0.10	0.13
Share of wage workers	0.70	0.68
Share of firms of size $<= 9$	0.96	0.90
Share of firms of size $10 - 19$	0.02	0.06
Share of firms of size $20 - 49$	0.01	0.03

Columns (1)-(3) of Table 3 display the results from the distorted benchmark economy, along with the results from two counterfactual experiments. In column (2) we ask what would be the effect of a perfectly-enforced payroll tax of the same level as the average effective tax rate under size-dependent enforcement (32%). Under perfect enforcement, wage employment increases by 4.4%, while self-employment decreases by 16%. Output and productivity increase by 2%. Returns to skill for wage workers increase by 35%, but the main winners from this policy change are high-skill workers. Figure 10 shows wages under the distorted benchmark, along with the two policy counterfactuals. Perfect enforcement increases the wages of the best workers, but lowers the wages of those at the bottom of the skill distribution. Notice that the share of managers remains unchanged, so all the gains in employment and wages come from the reallocation of the worst self-employed, who now become the best workers in the economy. The increase in the returns to skill for managers are much more modest, at 1%.

The fact managers do not seem to reap as high benefits from perfect enforcement as wage workers derives from several forces. First, there are two partially offsetting effects: the policy change hurts smaller firms, who now have to pay higher taxes. However, they pass through some of that burden to their workers via lower wages. Second, perfect enforcement benefits the more productive managers, as they are now able to form larger firms, matching with better workers, and face a lower effective tax rate. On the other hand, these large firms have to pay higher wages for their skilled workers. Thus, managers are better off on average, but only high-skill managers truly benefit from the policy change.

Table 3
The gains from removing size-dependent labor regulations

	(1) Distorted benchmark ($\tau_s = 0.44, \kappa = 0.18$)	(2) Perfect enforcement ($\tau_s = 0.32$)	(3) No distortion ($\tau_s = 0$)
Share of wage-workers	0.68	0.71	0.77
Share of self-employed	0.19	0.16	0.10
Share of managers	0.13	0.13	0.13
Output*	1	1.02	1.05
Productivity*	1	1.02	1.05
Return to skill (workers)*	1	1.35	1.5
Return to skill (managers)*	1	1.01	1.02
Average tax	32%	32%	0%

* Relative to distorted benchmark with ($\tau_s = 0.44, \kappa = 0.18$).

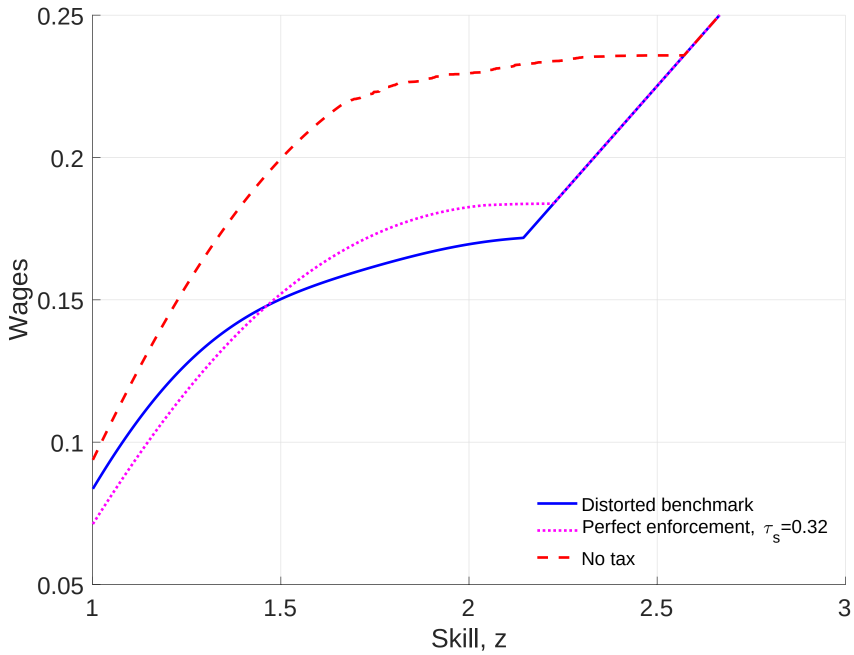

Figure 10
Effect of Removing Size-dependent enforcement on Equilibrium Wages.

Last, in column (3) of Table 3, we examine the effects from completely eliminating the payroll tax. Not surprisingly, the effects are larger under this policy change. Wage employment increases by 13%, while self-employment is nearly cut in half. Output and productivity both increase by 5%. Once again, wage workers reap the highest benefits from eliminating the payroll tax, as the return on their skill increases by 50%. In this case, wages increase across the board, but high-skill workers experience the largest gains, as seen in Figure 10. The mechanism that generates these gains is the same as in the case of perfect enforcement: reallocation from self-employment into wage work brought about from an increase in the demand for skill, as managers find it more profitable to match with better workers and form larger firms. The returns to skill for managers increase by 2%.

As before, there are offsetting forces that make managers in the bottom and the middle of the managerial skill distribution nearly indifferent to the policy change. On the one hand, those running smaller firms were not paying much taxes (if any) under size-dependent enforcement, so the elimination of the tax does not make a sizable difference in their earnings. The small gains from lower taxes for these managers are offset by the higher wages they have to pay. The managers at the top of the skill distribution, who were the most affected by the size-dependent tax, benefit from its elimination, but those gains are partially offset by the higher wages they have to pay as they match with better workers.

Even though perfect enforcement induces some distributional changes not present under a complete elimination of the tax, a large share of the gains from eliminating the tax can be obtained via enforcement. Gains from perfect enforcement amount to 40% of the productivity and output gains from removing the distortion. In the case of the returns to skill for workers (managers), the gains from perfect enforcement represent 70% (50%) of those from eliminating the tax.

Our results shed light on some of the driving forces behind the observed allocation of talent and the returns to skill in Mexico. First, size-dependent policies distort the allocation of the marginal individual in each occupation, and in particular, induce the best wage workers to enter into self-employment. In López and Torres (2018b) we find evidence of a misallocation of the most skilled workers into microbusiness ownership. In particular, we find that medium and high skill self-employed individuals would earn significantly more if they instead worked for a wage. Second, size-dependent policies also re-sort the infra-marginal workers and attenuate the strength of the positive sorting throughout the entire economy. Figure 11 shows the distribution of workers in the U.S. and Mexico across firm sizes, conditional on years of schooling. In both countries, workers with more years of schooling are more likely to work in larger firms, but in Mexico the relationship is not as strong as in the U.S. for those in the top of the skill distribution.

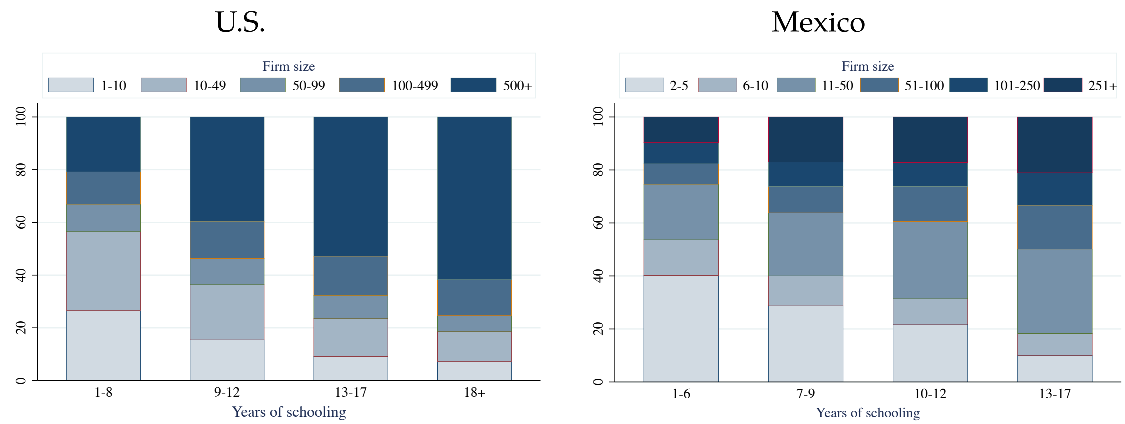

Figure 11
Sorting of Wage Workers Across
Years of Schooling and Firm Size in the U.S. and Mexico

Source: Author's calculations using the 2016 March Supplement of the Current Population Survey (CPS) and the National Survey of Occupations and Employment (ENOE, by its Spanish acronym), 2017-Q3.

5.2. Calibration to the U.S.

Next, we calibrate the undistorted parametric example described in Section 2 to match some features of the occupational shares and distribution of firm sizes in the U.S., and introduce two different size-dependent taxes. Specifically, we set the lower bound for the distribution of skill and problems to $L = 1$, and perform a joint calibration of the parameters a, b, λ, and $s \equiv {}^{L}/_{H}$. The targets are (i) the parametric restriction in *Assumption 1* necessary for the equilibrium to exist, (ii) the average firm size, (iii) the fraction of wage workers, and (iv) the share of firms with 50 or more employees. As in other works similar to ours, such as Garicano et al. (2016), we assume returns to scale equal to 0.8. Table 4 contains all parameter values. Table 5 shows the model's performance.[8]

[8] Data on the distribution of firm sizes for the U.S. come from publicly available statistics published by the U.S. Census Bureau, while data on occupational shares come from the U.S Bureau of Labor Statistics.

Table 4
Parameter Values—U.S. (undistorted)

Parameter	Value	Source
L	1	Assigned
α	0.8	Garicano et al. (2016)
a	1×10^{-8}	Joint calibration
b	16.5	Joint calibration
$s \equiv L/H$	0.0982	Joint calibration
λ	1	Joint calibration

Table 5
Calibration Results: U.S. (undistorted)

Targeted	Data	Model
Average firm size	21	17.09
Share of wage workers	0.90	0.89
Share of firms of size $>= 50$	0.039	0.042
Non-targeted	Data	Model
Share of firms of size $<= 19$	0.89	0.81
Share of firms of size $20 - 49$	0.07	0.14
Share of firms of size $50 - 99$	0.021	0.029
Share of firms of size $>= 100$	0.018	0.013

The model does an overall good job at matching the share of wage workers, and the share of firms with 50 or more employees. The calibrated average firm size is about 19% smaller than that in the data. Coincidentally, our calibrated average firm size is very close to the average establishment size in the data (17 in our model vs 16 in the data for 2014, according to the U.S. Census Bureau). The model also matches relatively well some non-targeted moments of the firm-size distribution.

The first counterfactual experiment is to introduce the size-dependent tax calibrated using the total labor cost and the observed compliance of social security contributions in Mexico. The second counterfactual experiment we conduct asks the question of what would be the tax necessary to induce the allocation of talent across occupations observed in Mexico. The strategy in this case is then to calibrate τ so as to match the rate of self-employed without employees (20%) in Mexico, while keeping the same enforcement parameter calibrated from the data on social security contributions. The calibrated parameter value is $\tau = 0.7325$, which generates a self-employment share of 21.6%.

Table 6 shows the results from these two experiments. Under the 44% imperfectly enforced labor cost, 9% of workers move from wage work into self-employment. Output and productivity decrease by 7% and the returns to skill for wage workers decrease by 18%. With a size-dependent tax of 73.25%, output and productivity losses are 12% . Again, all the loss in output is the result of productivity losses caused by the shifts in occupational choices. The main mechanism is the same as before: larger firms constrain their size as a result of the tax, which reduces the demand for skilled labor. The corresponding decline in wages for skilled labor induces a shift from wage work to self-employment. In this scenario, the returns to skill decline by 60% for workers, and 7% for managers. The average tax rate is 61%, which we interpret as a measure of the efficiency wedge implied by the observed allocation of talent in Mexico, compared to the U.S.

Table 6
The effects of size-dependent labor regulations ($\kappa = 0.18$).

	(1) Undistorted benchmark	(2) Total labor cost in Mexico ($\tau_s = 0.44$)	(3) Match self-employment in Mexico ($\tau_s = 0.7325$)
Share of wage-workers	0.89	0.80	0.73
Share of self-employed	0.06	0.15	0.22
Share of managers	0.05	0.05	0.05
Output*	1	0.93	0.88
Productivity*	1	0.93	0.88
Return to skill (workers)*	1	0.82	0.40
Return to skill (managers)*	1	0.96	0.93
Average tax	0%	40%	61%

* Relative to undistorted benchmark

6. Size-dependent regulations and bunching of firms

The size-dependent distortion considered until now attenuates, but does not break down, the equilibrium positive sorting. Therefore, it preserves the qualitative properties of the undistorted equilibrium. Further, such a policy does not produce a "bunching" of firms around a particular threshold in the equilibrium firm size distribution, as the size-dependent distortion studied by Garicano et al. (2016) for France.

In this section we describe the qualitative effects of a policy that can potentially break down positive sorting, and generates bunching in the distribution of firm sizes. Specifically, we consider the effects of the following Sigmoid tax policy:

$$\tau(n) = \frac{\tau_s}{1 + \exp\left[-\kappa(n - N)\right]},$$

where τ_s is the general, statutory tax rate, N is the Sigmoid's midpoint, which represents the threshold around which marginal enforcement significantly increases with firm size. The enforcement parameter κ reflects the steepness of the Sigmoid function around this threshold. Figure 12 shows an example with $\tau = 0.3$, $\kappa = 0.5$, and $N = 20$.

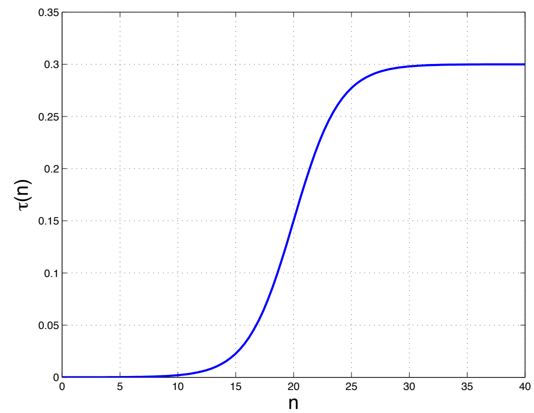

Figure 12
Sigmoid Tax Function ($\tau = 0.3$, $\kappa = 0.5$, $N = 20$)

Note that as κ becomes arbitrarily large, this continuous tax approaches a step tax policy of the following form: firms which hire more than N workers pay a tax $\tau \in (0, 1)$ on their wage bill, and firms which employ fewer than N workers pay no tax. Garicano et al. (2016) develop and estimate the effects of this policy in a simpler version of our model, in which wage workers are homogeneous in their abilities.[9]

Consider the FOC of the entrepreneur's optimization problem:

$$\begin{aligned} G(z_m)\,\alpha n^{\alpha-1} - w(z_p)\left[1 + \tau(n) + \tau'(n)\,n\right] - \mu &\geq 0, \\ -w'(z_p)\left[1 + \tau(n)\right] n + \mu n'(z_p) &\geq 0, \\ \mu\left[n(z_p) - n\right] &= 0, \\ \mu &\geq 0. \end{aligned}$$

Where μ is the Lagrange multiplier associated with the optimal size constraint. Note

[9] In our model, as in the simpler model in Garicano et al. (2016), a step-tax policy generates a mass point at the enforcement size threshold N. Because in their data on French firms they observe bunching across a range of sizes, and not a mass point at a single size value, Garicano et al. (2016) work around this complication by assuming measurement error in their data. The Sigmoid tax function we consider is the theoretical counterpart to their empirical methodology to fit the observed distribution of firm sizes under size-dependent policies.

that if the size constraint is binding, prices and assignments in equilibrium are solved for as in the undistorted case. In this case, managers constrain the size of their firm by matching with workers of lower skill, but still according to the communication technology $n(z_p)$. On the other hand, if the size constraint is not binding, then $\mu = 0$ and the optimal firm size solves:

$$\frac{G\left(z_m\right)\alpha}{w\left(z_p\right)} = n^{1-\alpha}\left[1 + \tau\left(n\right) + \tau'\left(n\right)n\right], \tag{10}$$

and from the second FOC it must be true that:

$$\begin{aligned} -w'\left(z_p\right)\left[1 + \tau\left(n\right)\right]n &= 0, \\ \Longleftrightarrow w'\left(z_p\right) &= 0. \end{aligned}$$

That is, the solution requires wages to be independent of skill if the size constraint is not binding: if managers constrain the size of their firm to avoid the tax, then the wages paid to their workers must be independent of their skill. Notice, however, that in this case sizes will depend on managerial skill, as dictated by equation 10—just as in a Lucas (1978) span-of-control model.

Figure 13 shows the assignment function for our undistorted benchmark economy, as well as for an economy subject to a Sigmoid tax policy.[10] The equilibrium assignment under this size-dependent tax can be characterized by six thresholds, $\{z_i\}_{i=1}^{6}$, such that : (i) $[L, z_1]$ is the set of low-skilled workers in small firms who pay low taxes and setup their size according to the communication technology $n(\cdot) \equiv 1/h(\cdot)$; (ii) $[z_1, z_2]$ is the set of medium-skilled workers in constrained firms that choose their size based on the skill of the manager, according to equation 10; (iii) $[z_2, z_3]$ is the set of high-skilled workers in constrained firms who choose their size according to the communication technology $n(\cdot) \equiv 1/h(\cdot)$; (iv) $[z_3, z_4]$ is the set of self-employed (without employees); (v) $[z_4, z_5]$ is the set of managers of small firms matched with workers in $[L, z_1]$; (vi) $[z_5, z_6]$ is the set of managers of medium-sized firms who set their size according to equation 10 and hire workers in $[z_1, z_2]$; (vii) $[z_6, H]$ is the set of managers of large firms matched with workers in $[z_2, z_3]$.

[10] The Sigmoid parameters for this example are $\tau_s = 0.3$, $\kappa = 0.25$, and $N = 20$.

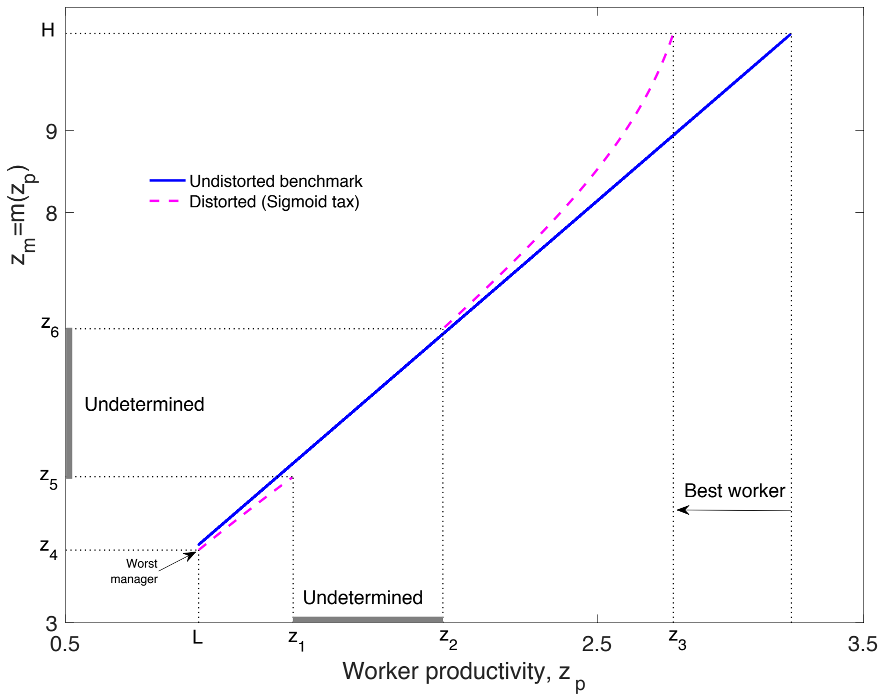

Figure 13
Equilibrium assignment before and after the tax.

The tax then encourages employers within a range of the size threshold N to constrain the size of their firm to avoid the tax, and discourages their wage workers from fully exploiting their talent. Positive sorting in the constrained firms who choose their size according to equation 10 is not guaranteed. More specifically, let $l^*(z_m)$ denote the size that solves the FOC in equation 10. Then, the minimum worker skill required to support l^* is given by $z_p^* = n^{-1}(l^*)$, where $n(\cdot) \equiv 1/h(\cdot)$. Notice, however, that in this segment of the skill distribution, a manager can setup a firm of size $l^*(z_m)$ using any worker with ability $z \in [z_p^*, z_2]$, because they all earn the same wage. Therefore, positive sorting is one of infinitely many possible outcomes among constrained firms and workers. This segment of the assignment function is therefore indeterminate. Managers could potentially mix employees from different types in a single firm, all of them earning the same wage.

Figure 14 displays the equilibrium firm-size distribution before and after the tax. In equilibrium, the size constraint for managers at the bottom and at the top of the managerial talent distribution always binds, and they maximize at an interior solution. To the contrary, managers in a segment around N constrain the size of their firm to sizes $[l^*(z_4), l^*(z_5)]$ following the FOC in equation 10, and no manager selects sizes $(l^*(z_5), n(z_2))$. In other words, the firm size distribution exhibits a bunching of firms in $[l^*(z_4), l^*(z_5)]$, with no mass in the segment $(l^*(z_5), n(z_2))$. When κ becomes arbitrarily large, this bunching of firms occurs at exactly N, as in the model of Garicano et al. (2016).

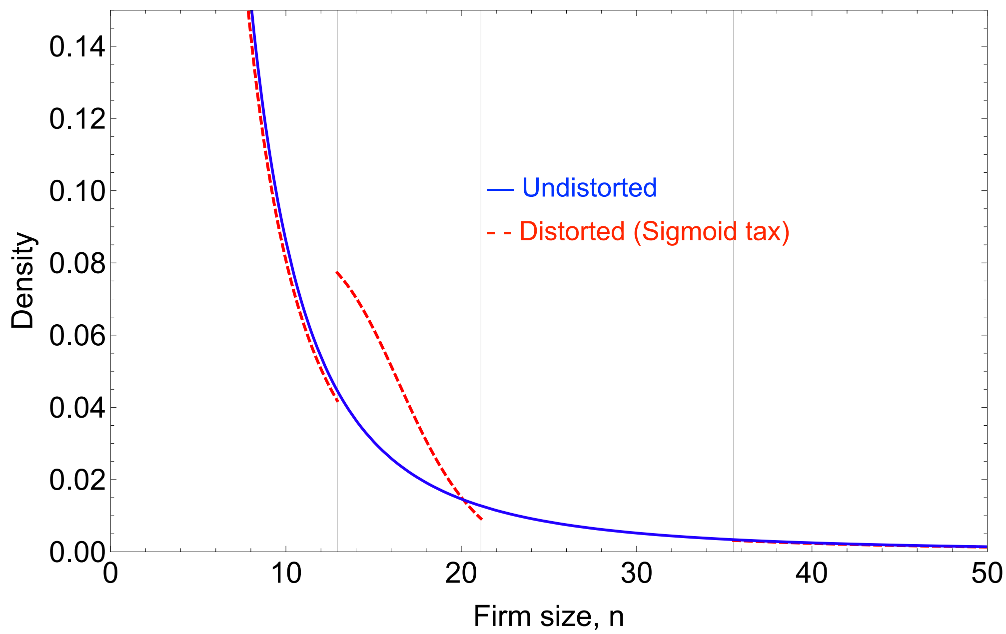

Figure 14
Equilibrium distribution of firm sizes before and after the tax.

The tax unambiguously lowers the return to skill for workers in the constrained firms. Managers in these firms reduce the size of their firm to avoid the tax, but by doing so wage workers lose the ability to differentiate themselves, and therefore their returns to skill turn flat, whereas managers lose the reward for pairing with more skilled employees, which also lowers the return on their skill, as seen in Figure 15. Policies that produce a bunching of firms in the size distribution then only aggravate the misallocation of talent and its effects on both productivity and the return to skill.

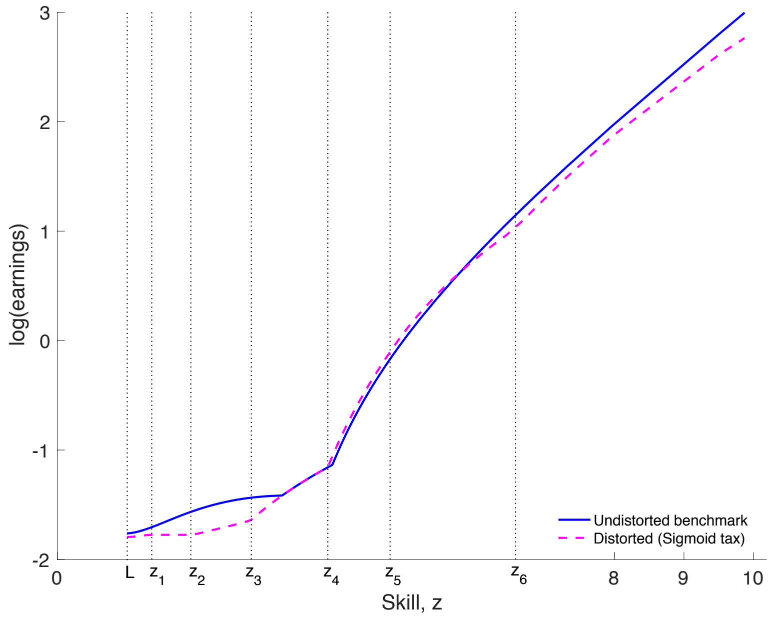

Figure 15
Equilibrium earnings before and after the tax.

7. Concluding remarks

We show that size-dependent regulations disproportionately lower the returns to skill for the most able workers and managers. Both the allocation of talent and the returns to skill are prominent concepts in the fields of growth and development—some might say, in the entire discipline of economics—yet remain largely unexplored in the vast literature on firm-specific distortions and misallocation. Our framework allows us to bridge this gap, and its analytical convenience opens many exciting lines of inquiry. Some of these include studying the effects of size-dependent policies on the acquisition of skills, as well as the endogenous evolution of occupational choices and the skill distribution throughout the process of development.

References

Alaimo, V., M. Bosch, M. Gualavisi, and J. M. Villa (2017). Measuring the cost of salaried labor in latin america and the caribbean. Technical report, Inter-American Development Bank.

Alder, S. D. (2016). In the Wrong Hands: Complementarities, Resource Allocation, and TFP. *American Economic Journal: Macroeconomics 1*(8), 199–241.

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

Bobba, M., L. Flabbi, and S. Levy (2017). Labor Market Search, Informality and Schooling Investments. IZA DP No. 11170.

Braguinsky, S., L. G. Branstetter, and A. Regateiro (2011). The incredible shrinking portuguese firm. Technical report, National Bureau of Economic Research.

Buera, F. J., B. Moll, and Y. Shin (2013). Well-intended policies. *Review of Economic Dynamics 16*(1), 216–230.

Busso, M., S. Levy, A. Neumeyer, and M. Spector (2012). Skills, informality and the size distribution of firms. *Buenos Aires, Argentina: Universidad Torcuato di Tella. Available at: https://dl. dropboxusercontent. com/u/499791/papers/Informality_December2012_2. pdf*.

Busso, M., S. Levy, and J. Torres (2018). Labor regulations and resource misallocation in Mexico. *Unpublished manuscript* (June).

Cardiff-Hicks, B., F. Lafontaine, and K. Shaw (2014). Do large modern retailers pay premium wages? Technical report, National Bureau of Economic Research.

D'Erasmo, P. N., H. J. M. Boedo, and A. Senkal (2014). Misallocation, informality, and human capital: Understanding the role of institutions. *Journal of Economic Dynamics and Control 42*, 122–142.

Fox, J. T. (2009). Firm-size wage gaps, job responsibility, and hierarchical matching. *Journal of Labor Economics 27*(1), 83–126.

Garicano, L. and T. N. Hubbard (2012). Learning About the Nature of Production from Equilibrium Assignment Patterns. *Journal of Economic Behavior and Organization 84*(1), 136–153.

Garicano, L., C. Lelarge, and J. Van Reenen (2016). Firm size distortions and the productivity distribution: Evidence from france. *The American Economic Review 106*(11), 3439–3479.

Garicano, L. and E. Rossi-Hansberg (2004). Inequality and the Organization of Knowledge. *American Economic Review, Papers and Proceedings 94*(2), 197–202.

Garicano, L. and E. Rossi-Hansberg (2006). Organization and Inequality in a Knowledge Economy. *The Quarterly Journal of Economics 121*(4), 1383–1435.

Geerolf, F. (2017). A theory of pareto distributions. *Unpublished Manuscript*.

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

Headd, B. (2000). The characteristics of small-business employees. *Monthly Lab. Rev. 123*, 13.

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

Kremer, M. and E. Maskin (1996). Wage Inequality and Segregation by Skill. National Bureau of Economic Research. Working Paper w5718.

Leal-Ordonez, J. C. (2014). Informal Sector, Productivity and Tax Collection. *Review of Economic Dynamics 17*(2), 262–286.

Levy, S. and L. F. López-Calva (2016). Labor Earnings, Misallocation, and the Returns to Education in Mexico. IDB Working Paper Series No IDB-WP-671.

López, J. J. (2017). A Quantitative Theory of Tax Evasion. *Journal of Macroeconomics 53*, 107–126.

López, J. J. and J. Torres (2018a). Modeling progressive income taxation. *Unpublished Manuscript*.

López, J. J. and J. Torres (2018b). The returns to entrepreneurship: selection, non-pecuniary benefits, and necessity in mexico. *Unpublished Manuscript*.

Lucas, R. E. (1978). On the Size Distribution of Business Firms. *The Bell Journal of Economics 9*(2), 508–523.

OECD (2017). *Entrepreneurship at a Glance 2017*. OECD.

Restuccia, D. and R. Rogerson (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics 11*(4), 707–720.

Rosen, S. (1981). The Economics of Superstars. *The American Economic Review 71*(5), 845–858.

Sattinger, M. (1993). Assignment Models of the Distribution of Earnings. *Journal of Economic Literature 31*(2), 831–880.

Scheuer, F. and I. Werning (2017). The taxation of superstars. *The Quarterly Journal of Economics 132*(1), 211–270.

Appendix

1) Proof of Proposition 1: Assignment and wages in equilibrium.

The assignment function in equilibrium follows the differential equation

$$\begin{aligned} \frac{f(x)}{n(x)} &= f(m(x))\, m'(x), \\ &= \frac{d}{dx} F(m(x)). \end{aligned}$$

Integrating both sides assuming a double truncated exponential distribution of skill yields

$$\begin{aligned} \int \frac{f(x)}{n(x)} dx &= F(m(x)), \\ \frac{-a\lambda [H-L]}{(\exp[-\lambda L] - \exp[-\lambda H])(\lambda(H-L)+b)} \left(\frac{\exp[-\lambda x]}{an(x)}\right) + c &= \frac{\exp[-\lambda L] - \exp[-\lambda m(x)]}{\exp[-\lambda L] - \exp[-\lambda H]}. \end{aligned}$$

Where c denotes the constant of integration. Using the boundary condition $m(L) = z_2$ to solve for c yields

$$c = \frac{\exp[-\lambda L] - \exp[-\lambda z_2]}{\exp[-\lambda L] - \exp[-\lambda H]} - \frac{-a\lambda [H-L]}{(\exp[-\lambda L] - \exp[-\lambda H])(\lambda(H-L)+b)} \left(\frac{\exp[-\lambda L]}{an(L)}\right),$$

and therefore

$$m(x) = -\frac{1}{\lambda} \ln \left[1 + \frac{a\lambda [H-L]}{\lambda(H-L)+b} \left(\frac{\exp[\lambda(z_2 - x)]}{an(x)}\right) - \frac{a\lambda [H-L]}{\lambda(H-L)+b} \left(\frac{\exp[\lambda(z_2 - L)]}{an(L)}\right)\right] + z_2.$$

Now define

$$\frac{a\lambda [H-L]}{\lambda(H-L)+b} \equiv \gamma,$$

and rewrite the assignment function using this new constant:

$$m(x) = \left(-\frac{1}{\lambda}\right) \ln \left[1 + \gamma \left(\frac{\exp[\lambda(z_2 - x)]}{an(x)}\right) - \gamma \left(\frac{\exp[\lambda(z_2 - L)]}{an(L)}\right)\right] + z_2.$$

Assume

$$\gamma \text{exp}\left[\lambda\left(z_2 - L\right) + b\right] = 1.$$

If this assumption holds, then we can simplify the assignment function even further to obtain

$$m\left(x\right) = x\left[1 + \left(\frac{1}{\lambda}\right) b\left(\frac{1}{H-L}\right)\right] + \left(-\frac{1}{\lambda}\right)\text{ln}\gamma + \left(-\frac{1}{\lambda}\right) b\left(\frac{H}{H-L}\right).$$

To obtain the wage function we must solve the following differential equation (see FOC in manager's problem):

$$G\left[m\left(x\right)\right]\left[\frac{d}{dx}\left[n\left(x\right)^{\alpha}\right]\right] = \frac{d}{dx}\left[w\left(x\right)n\left(x\right)\right].$$

The left-hand side can be rewritten as

$$G\left(m\left(x\right)\right)\left[\frac{d}{dx}\left[n\left(x\right)^{\alpha}\right]\right] = \left[k_1 x + k_2\right]\text{exp}\left[k_3 + k_4 x\right],$$

where
$k_1 = \left[1 + \left(\frac{1}{\lambda}\right)\left(\frac{b}{H-L}\right)\right]\alpha\left[\frac{b}{[H-L]^2}\right]\left(\frac{1}{a^{\alpha}}\right)$,
$k_2 = \left[\left(-\frac{1}{\lambda}\right)\text{ln}\gamma + \left(-\frac{1}{\lambda}\right)\left(\frac{b}{H-L}\right)H - L\right]\alpha\left[\frac{b}{[H-L]^2}\right]\left(\frac{1}{a^{\alpha}}\right)$,
$k_3 = -b\alpha\left(\frac{H}{H-L}\right)$,
$k_4 = b\alpha\left(\frac{1}{H-L}\right)$.
Then,

$$\begin{aligned}\int G\left[m\left(x\right)\right]\left[\frac{d}{dx}\left[n\left(x\right)^{\alpha}\right]\right]dx &= \int\left[k_1 x + k_2\right]\text{exp}\left[k_3 + k_4 x\right]dx \\ &= \text{exp}\left[k_3 + k_4 x\right]\left[\frac{k_2}{k_4} - \frac{k_1}{k_4^2} + \frac{k_1 x}{k_4}\right],\end{aligned}$$

and therefore (integrating also the right-hand side)

$$\exp\left[k_3 + k_4 x\right]\left[\frac{k_2}{k_4} - \frac{k_1}{k_4^2} + \frac{k_1}{k_4}x\right] + c = w\left(x\right)n\left(x\right).$$

Where c denotes the constant of integration. Using the boundary condition $w\left(z_1\right) = G\left(z_1\right)$ to solve for c yields

$$c = G\left(z_1\right)n\left(z_1\right) - \exp\left[k_3 + k_4 z_1\right]\left[\frac{k_2}{k_4} - \frac{k_1}{k_4^2} + \frac{k_1}{k_4}z_1\right].$$

To solve for the equilibrium we have two final conditions:

$$G\left(z_2\right)n\left(L\right)^{\alpha} - w\left(L\right)n\left(L\right) = G\left(z_2\right),$$

$$m\left(z_1\right) = H.$$

Take the second equation to solve for z_1:

$$\begin{aligned}
m\left(z_1\right) = H &= a^{\alpha}\left[H - L\right]\left[\frac{z_1 k_1 + k_2}{k_4}\right] + L, \\
H - L &= a^{\alpha}\left[H - L\right]\left[\frac{z_1 k_1 + k_2}{k_4}\right], \\
z_1 &= \left[\frac{\left(\frac{1}{a^{\alpha}}\right)k_4 - k_2}{k_1}\right], \\
&= H + \frac{(H - L)\ln\gamma}{b + (H - L)\lambda}.
\end{aligned}$$

Use the first equation to solve for z_2

$$
\begin{aligned}
G(z_2)\, n(L)^{\alpha} - w(L)\, n(L) &= G(z_2)\,, \\
G(z_2)\left[n(L)^{\alpha} - 1\right] &= w(L)\, n(L)\,, \\
G(z_2) = \frac{z_2 - L}{H - L} &= \frac{w(L)\, n(L)}{n(L)^{\alpha} - 1}, \\
z_2 - L &= \frac{w(L)\, n(L)}{n(L)^{\alpha} - 1}\left[H - L\right], \\
z_2 &= L + \frac{w(L)\, n(L)}{n(L)^{\alpha} - 1}\left[H - L\right].
\end{aligned}
$$

Now recall that from assumption we need:

$$
z_2 = L - \left(\frac{1}{\lambda}\right) b - \left(\frac{1}{\lambda}\right) \ln\gamma.
$$

Then, the parametric restriction becomes:

$$
\begin{aligned}
L - \left(\frac{1}{\lambda}\right) b - \left(\frac{1}{\lambda}\right) \ln\gamma &= L + \frac{w(L)\, n(L)}{n(L)^{\alpha} - 1}\left[H - L\right], \\
-\left(\frac{1}{\lambda}\right) b - \left(\frac{1}{\lambda}\right) \ln\gamma &= \frac{w(L)\, n(L)}{n(L)^{\alpha} - 1}\left[H - L\right], \\
\left(-\frac{1}{\lambda}\right)\left[b + \ln\gamma\right] &= \frac{w(L)\, n(L)}{n(L)^{\alpha} - 1}\left[H - L\right].
\end{aligned}
$$

2) Proof of Proposition 2: Firm size distribution.

First notice that using the communication technology, we can write the skill of worker z_p as a function of his firm size,

$$
z_p(n) = \frac{H - L}{b}\ln(n) + \frac{H - L}{b}\ln a + H.
$$

Then, using the equilibrium assignment function, we can write $m(n) = m(z_p(n))$, which represents the manager that corresponds to firm size n.

$$
m(n) = \left(\frac{H - L}{b}\ln(n) + \frac{H - L}{b}\ln a + H\right)\left[1 + \left(\frac{1}{\lambda}\right) b \left(\frac{1}{H - L}\right)\right] - \frac{1}{\lambda}\ln\gamma - \frac{b}{\lambda}\left(\frac{H}{H - L}\right).
$$

Then

$$\begin{aligned} \frac{dm}{dn} &= \left(\frac{H-L}{b} + \frac{1}{\lambda}\right)\left[\frac{1}{n}\right], \\ &= \lambda^{-1} p n^{-1}. \end{aligned}$$

Where $p \equiv \frac{\lambda(H-L)}{b} + 1$.

The managerial skill distribution follows $\frac{F(m)-F(z_2)}{1-F(z_2)}$. Using the change of variable technique yields

$$\begin{aligned} \Pr\left(n \leq \overline{n}\right) &= \Pr\left(n\left(m\right) \leq \overline{n}\right), \\ &= \Pr\left(m \leq n^{-1}\left(\overline{n}\right)\right), \\ &= \frac{F\left(n^{-1}\left(\overline{n}\right)\right) - F\left(z_2\right)}{1 - F\left(z_2\right)}. \end{aligned}$$

Then

$$\begin{aligned} \frac{f\left(m\left(n\right)\right)}{1 - F\left(z_2\right)}\left(\frac{dm}{dn}\right) &= \frac{\frac{\lambda \exp[-\lambda m(n)]}{\exp[-\lambda L] - \exp[-\lambda H]}}{\frac{\exp[-\lambda z_2] - \exp[-\lambda H]}{\exp[-\lambda L] - \exp[-\lambda H]}}\left(\frac{dm}{dn}\right) \\ &= \frac{\exp\left[-\lambda m\left(n\right)\right] p n^{-1}}{\exp\left[-\lambda z_2\right] - \exp\left[-\lambda H\right]}. \end{aligned}$$

We can then write the firm-size density as

$$\tilde{f}\left(n\right) = \frac{B p n^{-p-1}}{\exp\left[-\lambda z_2\right] - \exp\left[-\lambda H\right]}.$$

We want to show that

$$C \equiv \frac{B}{\exp\left[-\lambda z_2\right] - \exp\left[-\lambda H\right]} = \frac{[n(L)]^p}{1 - \left(\frac{n(L)}{n(z_1)}\right)^p}$$

Notice that

$$f(m(n)) = \exp\left[-\lambda\left(\frac{H-L}{b}\ln(n) + \frac{H-L}{b}\ln a + H\right)\left[1 + \left(\frac{1}{\lambda}\right)b\left(\frac{1}{H-L}\right)\right] + \ln\gamma + \frac{bH}{H-L}\right].$$

Then we can write the constant B as

$$B = \exp\left[-\lambda\left(\frac{H-L}{b}\ln a + H\right)\left[1 + \left(\frac{1}{\lambda}\right)b\left(\frac{1}{H-L}\right)\right] + \ln\gamma + \frac{bH}{H-L}\right].$$

Thus,

$$\begin{aligned}
C &= \frac{\exp\left[-\lambda\left(\frac{H-L}{b}\ln a + H\right)\left[1 + \left(\frac{1}{\lambda}\right)b\left(\frac{1}{H-L}\right)\right] + \ln\gamma + \frac{bH}{H-L}\right]}{\exp\left[-\lambda z_2\right] - \exp\left[-\lambda H\right]}, \\
&= \frac{\exp\left[-p\ln a + (-p+1)\frac{bH}{H-L} + \ln\gamma\right]}{\gamma\exp\left[-\lambda L + b\right] - \exp\left[-\lambda H\right]}, \\
&= \frac{a^{-p}\exp\left[(-p+1)\frac{bH}{H-L} + \lambda L - b\right]}{1 - \frac{1}{\gamma}\exp\left[-\lambda(H-L) - b\right]}, \\
&= \frac{a^{-p}\exp(-pb)}{1 - \frac{exp(-pb)}{\gamma}}, \\
&= \frac{[n(L)]^p}{1 - \left(\frac{n(L)}{n(z_1)}\right)^p}.
\end{aligned}$$

In the derivation above we have used **Assumption 1**, and the fact that $[n(L)]^p = a^{-p}\exp(-pb)$ and $[n(z_1)]^p = a^{-p}\gamma$. Then, the firm-size density writes

$$\begin{aligned}
\tilde{f}(n) &= Cpn^{-p-1}, \\
&= \frac{[n(L)]^p pn^{-p-1}}{1 - \left(\frac{n(L)}{n(z_1)}\right)^p}.
\end{aligned}$$

which is a double truncated Pareto density.

3) Conditions for a closed-form solution with distortions.

First notice that if the manager's problem has an internal solution (which is the case with the negative exponential tax function considered), then managers use the communication technology to setup the size of their firm. That implies that the market-clearing condition used to derive the assignment function remains the same under the size-dependent policy.

To obtain the wage function we must solve the differential equation from the FOC in the manager's problem:

$$G\left[m\left(x\right)\right]\left[\frac{d}{dx}\left[n\left(x\right)^{\alpha}\right]\right] = \frac{d}{dx}\left[(1+\tau(n(x)))w\left(x\right)n\left(x\right)\right].$$

The left-hand side can be rewritten as before,

$$G\left(m\left(x\right)\right)\left[\frac{d}{dx}\left[n\left(x\right)^{\alpha}\right]\right] = \left[k_1 x + k_2\right]\exp\left[k_3 + k_4 x\right],$$

where the constants k_i, $i \in \{1,2,3,4\}$ are the same as before. Then

$$\begin{aligned}\int G\left[m\left(x\right)\right]\left[\frac{d}{dx}\left[n\left(x\right)^{\alpha}\right]\right]dx &= \int\left[k_1 x + k_2\right]\exp\left[k_3 + k_4 x\right]dx \\ &= \exp\left[k_3 + k_4 x\right]\left[\frac{k_2}{k_4} - \frac{k_1}{k_4^2} + \frac{k_1 x}{k_4}\right],\end{aligned}$$

and therefore (integrating also the right-hand side)

$$\exp\left[k_3 + k_4 x\right]\left[\frac{k_2}{k_4} - \frac{k_1}{k_4^2} + \frac{k_1}{k_4}x\right] + c = (1+\tau(n(x)))w\left(x\right)n\left(x\right).$$

Where c denotes the constant of integration. Using the boundary condition $w\left(z_1\right) = G\left(z_1\right)$ to solve for c yields

$$c = (1+\tau(n\left(z_1\right)))G\left(z_1\right)n\left(z_1\right) - \exp\left[k_3 + k_4 z_1\right]\left[\frac{k_2}{k_4} - \frac{k_1}{k_4^2} + \frac{k_1}{k_4}z_1\right].$$

To solve for the equilibrium we have two final conditions. The first one is $m\left(z_1\right) =$

H, which is exactly the same as in the undistorted case. The second one is given by

$$\begin{aligned}
G\left(z_2\right) n\left(L\right)^{\alpha} - (1+\tau(n\left(L\right)))w\left(L\right) n\left(L\right) &= G\left(z_2\right), \\
G\left(z_2\right)\left[n\left(L\right)^{\alpha} - 1\right] &= (1+\tau(n\left(L\right)))w\left(L\right) n\left(L\right), \\
G\left(z_2\right) = \frac{z_2 - L}{H - L} &= \frac{(1+\tau(n\left(L\right)))w\left(L\right) n\left(L\right)}{n\left(L\right)^{\alpha} - 1}, \\
z_2 - L &= \frac{(1+\tau(n\left(L\right)))w\left(L\right) n\left(L\right)}{n\left(L\right)^{\alpha} - 1}\left[H - L\right], \\
z_2 &= L + \frac{(1+\tau(n\left(L\right)))w\left(L\right) n\left(L\right)}{n\left(L\right)^{\alpha} - 1}\left[H - L\right].
\end{aligned}$$

Then, the parametric restriction from Assumption 1 then becomes

$$\begin{aligned}
L - \left(\frac{1}{\lambda}\right) b - \left(\frac{1}{\lambda}\right) \ln\gamma &= L + \frac{(1+\tau(n\left(L\right)))w\left(L\right) n\left(L\right)}{n\left(L\right)^{\alpha} - 1}\left[H - L\right], \\
-\left(\frac{1}{\lambda}\right) b - \left(\frac{1}{\lambda}\right) \ln\gamma &= \frac{(1+\tau(n\left(L\right)))w\left(L\right) n\left(L\right)}{n\left(L\right)^{\alpha} - 1}\left[H - L\right], \\
\left(-\frac{1}{\lambda}\right)\left[b + \ln\gamma\right] &= \frac{(1+\tau(n\left(L\right)))w\left(L\right) n\left(L\right)}{n\left(L\right)^{\alpha} - 1}\left[H - L\right].
\end{aligned}$$