IDB WORKING PAPER SERIES No. IDB-WP-341

(In)Formal and (Un)Productive:

The Productivity Costs of Excessive Informality in Mexico

Matías Busso
María Victoria Fazio
Santiago Levy

August 2012

Inter-American Development Bank
Department of Research and Chief Economist

(In)Formal and (Un)Productive:

The Productivity Costs of Excessive Informality in Mexico

Matías Busso
María Victoria Fazio
Santiago Levy

Inter-American Development Bank

Inter-American Development Bank
2012

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Busso, Matías.
(In)formal and (un)productive : the productivity costs of excessive informality in Mexico / Matías Busso, María Victoria Fazio, Santiago Levy.
p. cm. (IDB working paper series ; 341)
Includes bibliographical references.
1. Informal sector (Economics)—Mexico. 2. Industrial productivity—Mexico. I. Fazio, María Victoria. II. Levy Algazi, Santiago. III. Inter-American Development Bank. Research Dept. IV. Title. V. Series.
IDB-WP-341

http://www.iadb.org

Documents published in the IDB working paper series are of the highest academic and editorial quality. All have been peer reviewed by recognized experts in their field and professionally edited. The information and opinions presented in these publications are entirely those of the author(s), and no endorsement by the Inter-American Development Bank, its Board of Executive Directors, or the countries they represent is expressed or implied.

This paper may be reproduced with prior written consent of the author.

Corresponding authors: Santiago Levy (email: slevy@iadb.org) and Matías Busso (email: mbusso@iadb.org)

Abstract[1]

The laws that regulate relations between firms and workers in Mexico distinguish sharply between salaried and non-salaried workers, and they are at the root of the existence of informality. This paper provides a clear definition of informality, distinguishing it from illegality. Using Mexico's Economic Census, the paper shows that the majority of firms are informal but legal, that there are more small formal firms than large ones, and that some large firms are informal. It also shows that informality and illegality increased in the period 1998-2008. Using a simple model of monopolistic competition to measure the productivity losses due to distortions that misallocate resources, the paper finds that one peso of capital and labor allocated to formal and legal firms is worth 28 percent more than if allocated to illegal and informal firms, and 50 percent more than if allocated to legal and informal firms. The paper concludes arguing that the distortions in the labor market created by informality reduce total factor productivity.

JEL classifications: D24, O47, L25
Keywords: Total factor productivity, Informality, Distortions, Misallocation costs, Mexico, Latin America

[1] The three authors are affiliated with the Inter-American Development Bank, but their views do not necessarily reflect those of the institution. We thank Pedro Martínez for excellent research assistance, and Mexico's National Institute of Statistics, Geography and Informatics (INEGI) for giving us generous access to their data. We thank Chang-Tai Hsieh and Andres Neumeyer for careful comments. The usual disclaimer applies. Corresponding authors: slevy@iadb.org and mbusso@iadb.org .

1. Introduction

Productivity growth, typically estimated as the portion of GDP growth that cannot be explained by the accumulation of physical and human capital or the growth of employment, is critical for Mexico. Over the period 1960-2008 factor accumulation was faster than in the United States (US). If total factor productivity (TFP) had kept pace, relative income per capita would be 24 per cent higher in 2008 vs. 1960. However, the sharp fall in Mexico's TFP relative to the US since 1980 more than offset the gains from factor accumulation, with the result that in 2008 Mexico's relative income per capita was 14 per cent lower.[2] Many factors contribute to this phenomenon and we do not attempt a precise measurement of each (Hanson, 2010). Our purpose here is to explore the extent to which informality is one of them.

There has been an extensive debate on the nature of the informal sector and its implications for development (e.g., Guha-Khasnobis, Kanbur and Ostrom, 2006; de Soto, 2000); Fields, 2007; Levy, 2008; Maloney, 2004; and Fajnzylber et al., 2007). To some, the informal sector is a source of dynamism and innovation, where entrepreneurs escape burdensome regulations and create the needed jobs that the formal sector fails to create; to others, it is the refuge where low-productivity firms enjoy an unfair advantage over their formal peers by evading tax and other regulations (La Porta and Schleifer, 2008). We take an agnostic view and let the data speak. Our results are based on Mexico's Economic Censuses for 1998, 2003 and 2008, an unusually rich data set, which provide information on 3.6 million firms of all sizes in all sectors of the economy.

This paper makes several contributions. First, we provide a sharp definition of informality. We focus on the institutions and laws that regulate relations between firms and workers in Mexico, and argue that the asymmetry in the regulation of salaried and non-salaried labor is the root cause of informality. We make a distinction between informality and illegality, and in turn separate these two attributes from firm size. In fact, we show that the majority of firms in Mexico are informal but legal, that there are more small formal firms than large ones,

[2] For this comparison we measure TFP as the standard Solow residual derived from a Cobb-Douglas production function using data from Fernández-Arias and Daude (2010). To control for the effects of macroeconomic cycles, the series for both countries have been smoothed with a Hodrick-Prescott filter. With 1960 = 1, we find the ratio of TFP in Mexico to the US to be 0.69 in 2008. The declining evolution of productivity holds with other methods of estimation. In particular, following Jermanowski (2007), we use data envelope analysis to calculate an efficiency index of Mexico vs. the US which, with 1960 = 1, equals 0.87 in 2008. This technique requires only an assumption about constant returns to scale.

and that some large firms are informal. Second, we develop a framework where the issues of firm size, formal status and legal behavior can be separated, and where informal firms are not necessarily unproductive firms. Thus, we do not pre-judge the relationship between informality and productivity. Third, we describe the allocation of labor and capital across firms in 2008 and show that small informal firms capture a large share of resources, thus highlighting the fact that their performance matters greatly for aggregate TFP. Fourth, using the methodology developed by Hsieh and Klenow (2009), we show that there is huge heterogeneity in firms' productivities even when firms are grouped within sectors at the 6-digit level, and that the distortions affecting them are not only large compared to other countries, but increased between 1998 and 2008. The TFP gain if these distortions were eliminated would be close to 200 percent.

Finally, we study the extent to which informality and illegality account for this misallocation of resources. Our main empirical result is that the productivity losses associated with excessive informality are large. We find that in 2008, controlling for size and legal status, among legal firms, formal ones were on average 84 percent more productive than informal ones; that within the set of informal firms, illegal ones were on average 59 percent more productive than legal ones; and that within the set of firms with salaried contracts, legal ones were on average 44 percent more productive than illegal ones. We also find that large firms face more distortions than smaller ones, implying that more productive firms are taxed and less productive ones subsidized. Altogether, our results indicate that from the point of view of productivity there is an excess of informal firms and a scarcity of formal ones. The relevance of our results is enhanced by our finding that informality increased between 1998 and 2008. Thus, we find strong backing for the proposition that reforming the public policies that cause excessive informality could raise TFP considerably; and, conversely, that deepening those policies—as has occurred in recent years—is very costly to Mexico. This is a critical result given the country's poor growth record.

The remainder of the paper is structured as follows: section 2 summarizes the theory of distortions, resource misallocation and TFP losses developed by Hsieh and Klenow (2009). Section 3 defines firm informality and illegality based on the differences in the regulation of salaried and non-salaried labor. Section 4 discusses the data. Section 5 presents a set of stylized facts. It first describes resource allocation across firm-types and its evolution over time; second, it shows that informality is different from illegality and different from firm size; third, it presents

the relationships between productivity, informality and illegality. Section 6 contains our main results. It first provides measures of heterogeneity in productivities and distortions across firms, and then tackles the question of whether informality and illegality are hurting Mexico's productivity. Section 7 concludes.

2. A Theory of Distortions, Resource Allocation and Total Factor Productivity

We begin by briefly outlining Hsieh and Klenow's (2009) model. Consider a standard model of monopolistic competition with heterogeneous firms that face distortions in the prices they observe. These distortions introduce wedges between the marginal revenue products of capital and labor across firms, misallocating resources and lowering aggregate total factor productivity. There is a single final good Q produced by a representative firm out of a set of goods Q_s in a perfectly competitive final output market with a constant returns to scale ($\sum \theta_s = 1$) Cobb-Douglas production technology:

$$Q = \prod_{s=1}^{S} Q_s^{\theta_s} \tag{1}$$

Each sector output Q_s is produced by combining M_s differentiated goods Q_{si} produced by individual firms using a CES technology, where the elasticity of substitution is assumed the same for all industries:

$$Q_s = \left[\sum_{i=1}^{M_s} Q_{si}^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma-1}{\sigma}} \tag{2}$$

In turn, each good Q_{si} is produced with a Cobb-Douglas technology with capital share α_s and productivity A_{si}:

$$Q_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s} \tag{3}$$

where L_{si}, K_{si}, denote labor and capital services.[3] Letting w and R denote wages and the rental cost of capital, respectively, individual firm profits are given by:

$$\pi_{si} = \left(1 - \tau_{Q_{si}}\right) P_{si} Q_{si} - \left(1 + \tau_{L_{si}}\right) w L_{si} - R K_{si} \tag{4}$$

[3] The parameter α_s is the capital share, which is assumed to be constant for all firms within a given industry. As the elasticity of substitution between plant value-added σ increases, intermediate inputs become closer to perfect substitutes. At the limit, only the highest-productivity good is produced.

Two types of distortions affect the firm's decisions. Output distortions (τ_{Qsi}) distort the output price observed by the firm and affect both capital and labor. Examples of these are high transportation costs, bribes/costs that have to be paid in order to operate or government-issued size restrictions. There also exist relative factor price distortions (τ_{Lsi}) that change the marginal product of labor relative to capital observed in equilibrium. Examples of these are credit constraints and labor regulations that differ across firms. In the presence of distortions the marginal revenue products are given by:

$$\text{(5)} \qquad \text{MRPK}_{si} = \text{R}/\left(1-\tau_{Q_{si}}\right) \qquad \text{and} \qquad MRPL_{si} = w\left(1+\tau_{L_{si}}\right)/\left(1-\tau_{Q_{si}}\right)$$

At this point, it is important to distinguish between physical total factor productivity ($TFPQ_{si}$), measured by A_{si}, and total factor revenue productivity ($TFPR_{si}$), measured by $P_{si}A_{si}$. It can be shown that:

$$\text{(6)} \qquad \text{TFPR}_{si} \propto (\,\text{MRPK}_{\text{si}})^{\alpha_s}(\,\text{MRPL}_{\text{si}})^{1-\alpha_s} \propto \left(1+\tau_{L_{si}}\right)^{1-\alpha_s}/\left(1-\tau_{Q_{si}}\right)$$

In the absence of distortions, more resources would be allocated to firms with the highest physical productivities, A_{si}, until $TPPR_{si}$ is equated across firms within a sector. In this case, there should be no dispersion in the distribution of sector *TFPR*. More productive firms would attract more capital and labor, which would expand output and lower the price faced by those firms, until the revenue productivity of capital and labor equaled that of less productive firms. Departures from this benchmark determine the magnitude of distortions, which are measured through the dispersion of *TFPR*.

It is useful to relate aggregate *TFPQ* to firms' productivities and firm-level distortions. Industry *TFPQ* can be expressed as a weighted geometric average of firms A_{si}. Firms with *TFPR* smaller than the sector average, that is firms that face fewer distortions and therefore use more inputs than they would in an undistorted economy, receive a higher weight. Given the assumed aggregate production function, aggregate *TFPQ* can be expressed as:

$$\text{(7)} \qquad TFPQ = \prod_{s=1}^{S}[TFPQ_s]^{\theta_s} = \prod_{s=1}^{S}\left[\sum_{i=1}^{M_s}\left\{A_{si}\frac{\overline{TFPR_s}}{TFPR_{si}}\right\}^{\sigma-1}\right]^{\frac{\theta_s}{\sigma-1}}$$

In the absence of distortions, aggregate *TFPQ* will be highest because resources are reallocated from less to more productive firms. There will, however, be some dispersion in the

distribution of firms' physical productivities. The efficient *TFPQ* becomes a geometric average of A_{si}:

$$TFPQ^* = \prod_{s=1}^{S}[TFPQ_s]^{\theta_s} = \prod_{s=1}^{S}\left[\sum_{i=1}^{M_s} {A_{si}}^{\sigma-1}\right]^{\frac{\theta_s}{\sigma-1}} \tag{8}$$

*TFPQ** can then be used as a benchmark to compute the output cost of deviations from the efficient allocation of resources caused by distortions. In particular, the gap between the efficient and the distortion-driven level of *TFPQ* can be shown to be: [4]

$$\frac{TFPQ}{TFPQ^*} = \prod_{s=1}^{S}\left[\sum_{i=1}^{M_s}\left\{\frac{A_{si}}{\overline{A_s}}\frac{\overline{TFPR_s}}{TFPR_{si}}\right\}^{\sigma-1}\right]^{\frac{\theta_s}{\sigma-1}} \tag{9}$$

Most establishment-level surveys do not record individual, plant or product level prices. However, within the model we can compute physical productivity by means of the following expression which can be observed in the data:[5]

$$A_{si} = \frac{Q_{si}}{K_{si}(wL_{si})^{1-\alpha_s}} = \frac{(P_{si}Q_{si})^{\frac{\sigma}{\sigma-1}}}{K_{si}(wL_{si})^{1-\alpha_s}} \tag{10}$$

In addition, since workers' human capital levels are not observable, the plant wage bill is used instead of labor input as a way of adjusting for differences in human capital across plants. In turn, α_s is measured as one minus the labor share in industry *s* in the United States.[6] This is a simple way to control for distortions that could affect the capital share differently in different countries while the United States is taken as a benchmark of an undistorted economy.

Finally, using the first-order conditions and assuming that value added does not include any taxes or subsidies that differentially affect firms within the same industry, we can compute a measure of the distortions faced by firms:

$$1+\tau_{L_{si}} = \frac{1-\alpha_s}{\alpha_s}\frac{RK_{si}}{wL_{si}} \quad ; \quad 1-\tau_{Q_{si}} = \frac{\sigma}{\sigma-1}\frac{RK_{si}}{\alpha_s P_{si}Q_{si}} \tag{11}$$

---

[4] $\overline{TFPR_s} = \left[\frac{1}{\alpha_s}\sum_{i=1}^{M_s}\left\{\left(\frac{1}{1-\tau_{Q_{si}}}\right)\left(\frac{P_{si}Q_{si}}{P_sQ_s}\right)\right\}\right]^{\alpha_s}\left[\frac{w}{1-\alpha_s}\sum_{i=1}^{M_s}\left\{\left(\frac{1+\tau_{L_{si}}}{1-\tau_{Q_{si}}}\right)\left(\frac{P_{si}Q_{si}}{P_sQ_s}\right)\right\}\right]^{1-\alpha_s}$.

[5] The key assumption used is that the product demand is given by ${P_{si}} = {Q_{si}}^{-1/\sigma}$.

[6] As reported by the Manufacturing Industry Database hosted by the NBER.

3. Unbundling Distortions in Factor Prices: Informality and Illegality

3.1 Salaried and Non-Salaried Labor

Several forces drive wedges between firms' marginal revenue products of labor and capital (i.e., several forces could make τ_Q and τ_L differ across firms in the same sector). These forces range from market failures that differentially affect firms within the same industry to policy distortions that treat similar firms differently. Here we focus on labor regulations.

In Mexico, and more generally in Latin America, labor regulations make a critical distinction between salaried and non-salaried workers. A *salaried* worker is a person who works for a firm as a subordinated employee, from whom she receives instructions in exchange for a wage, a fixed amount of money per unit of labor time. A written contract is not required. The only requirement is that at least two people be involved: a principal (boss/manager) who gives orders and an agent (worker) who executes them in exchange for a certain payment. All other workers in Mexico are *non-salaried*. Workers of the latter type may be associated with a firm, but formally not subordinated to it, and receive payments in the form of a piece-rate, as a share of profits, or in proportion to the product regardless of the time the task requires. In Mexico these workers are not considered by law to be the firm's employees. Common examples include workers who sell door-to-door, workers in a cooperative sharing the output, workers on a temporary contract performing a non-recurrent task, workers in a family enterprise who share output or benefits and, very importantly, self-employed workers.

The distinction between salaried and non-salaried labor is central for our purposes. This is because labor regulations establish certain rights for salaried workers as well as obligations to firms that hire them. These rights and obligations do not apply to non-salaried workers. First, firms are obligated to enroll salaried workers in the Mexican Social Security Institute (IMSS, for its Spanish acronym) and pay a social security tax whose revenue is used to fund a bundled set of social benefits. Second, a salaried worker has a right to a severance payment when fired (and the right to be re-instated in his job if the firm fired him for an unjust cause).[7] Non-salaried workers, on the other hand, are legally not the employees of the firm and cannot be fired by it; further, firms are not obliged to enroll these workers in IMSS. Third, salaried workers have the right to form a union, while non-salaried do not, as again they are not the firm's employees. Fourth,

[7] In which case the firm must pay all wages accumulated during the time the worker was without work.

salaried workers are subject to a minimum wage, while non-salaried workers are not since their earnings are not legally wages. On the other hand, finally, non-salaried workers are entitled to freely receive an unbundled set of social benefits financed from general government revenues, which are an imperfect substitute for the set of bundled benefits that salaried workers receive.

We argue that, depending on parameter values, these asymmetries in the regulation of labor may translate into distortions in factor prices, driving wedges between the marginal products of similar firms and distorting the allocation of resources away from the optimum. To capture this, let w_s and w_{ns} be the wage and wage equivalent remuneration of a salaried and non-salaried worker, respectively. Further, let T_s be the monetary cost of the regulations on salaried labor, which include the explicitly legislated social security taxes and the contingent costs associated with severance pay and other constraints on firing.[8] In parallel, let T_{ns} be the monetary costs of the regulations on non-salaried labor, which in this case include only the value of social benefits. Critically, note that T_s needs to be internalized in the contract between the firm and the worker, while T_{ns} is paid by the government. Finally, let $\beta_s, \beta_{ns} \in [0,1]$ be the value imputed by workers to the benefits derived from T_s and T_{ns}. As a result, the utility derived from each form of employment is:

(12) Utility of a salaried contract = $U_s = w_s + \beta_s T_s$

(13) Utility of a non-salaried contract = $U_{ns} = w_{ns} + \beta_{ns} T_{ns} + \alpha$

where $\alpha \geq 0$ is the non-pecuniary benefit of a non-salaried contract, i.e., the value to the worker of not having a boss (for ease of notation, henceforth assumed to equal zero).

On the other hand, the firm's labor costs depend on the type of contract, and on whether firms fully comply with their legal obligations with respect to the salaried workers that they hire. In other words, salaried workers can be hired legally (SL_{si}^{L}) if the firm complies with the regulations on salaried labor, or illegally (SL_{si}^{I}) if the firm evades them. From the point of view of the production process legally and illegally hired workers are perfect substitutes (i.e., $\mathrm{SL}_{si} = \mathrm{SL}_{\mathrm{si}}^{\mathrm{L}} + \mathrm{SL}_{\mathrm{si}}^{\mathrm{I}}$). Of course, if a firm hires a salaried worker illegally, it faces a probability λ_{is} of being detected and having to pay a fine $F > T_s$. This probability is an increasing function of the

[8] These include litigation costs, as the law makes a distinction between firing for a just or an unjust cause. T_s also includes taxes on salaried labor and the transaction costs of compliance with all these obligations; see Levy (2008) for a fuller description. Henceforth we use enrollment with IMSS and compliance with the regulations on salaried labor interchangeably.

number of illegally hired salaried workers, so that $\partial\lambda_{is} / \partial SL^{I}_{is} > 0$, and we take it to be firm-specific since there are other firm attributes that make it more or less likely to be detected violating the law.[9] Importantly, note that if a firm evades the regulations on salaried labor, the worker is not left without social benefits, as she can receive the same benefits offered to non-salaried workers.

Table 1 depicts the costs of each labor contract to firms, the benefits to workers, and the implicit taxes and subsidies to labor that, depending on parameter values, may be caused by the interaction between the asymmetries in the regulation of labor and their enforcement.

Table 1. Costs and Benefits of Labor Contracts

Labor contract	Cost to firms	Benefit to workers	Implicit tax/subsidy
Legal salaried	$w_s + T_s$	$w_s + \beta_s T_s$	$(1-\beta_s)T_s$
Illegal salaried	$w_{ns} + \lambda(.)F$	$w_{ns} + \beta_{ns} T_{ns}$	$\lambda(.)F - \beta_{ns} T_{ns}$
Non-salaried	w_{ns}	$w_{ns} + \beta_{ns} T_{ns}$	$-\beta_{ns} T_{ns}$

Clearly, when there are no regulations on either salaried or non-salaried labor, $T_s = T_{ns} = 0$, there are no taxes or subsidies, and no distortions; in that case no salaried workers would be hired illegally, as there are no regulations to evade. The same occurs when there are no social benefits for non-salaried and illegally hired salaried workers, $T_{ns} = 0$, and salaried workers fully value the monetary costs of the labor regulations that apply to them, $\beta_s = 1$. In these two cases, the allocation of capital and labor across firms with salaried and non-salaried contracts, and across firms of different sizes, will not be distorted by labor regulations (nor the level of self-employment).

As argued by Levy (2008), however, the relevant case for Mexico is the one where non-salaried workers and illegally hired salaried workers do receive social benefits, $T_{ns} > 0$, where salaried workers do not fully value the social benefits that pertain to them, $\beta_s < 1$, and where enforcement is imperfect with the probabilities of detection positively associated with firm size $\lambda(.) < 1, \lambda' > 0$. This case generates two distortions. First, there is a difference between the costs

[9] One such attribute is location, as for any firm size, firms in larger urban areas face greater probabilities since there is a greater density of inspectors. Another is client, as firms selling to the government are more likely to be caught. A third is activity, as manufacturing firms tend to have their workers in the same location, while services may have their workers more dispersed.

of salaried and a non-salaried worker that results purely from the asymmetry in labor regulations; second, the marginal cost of a salaried worker increases with firm size. As a result, large firms will hire salaried workers mostly legally and therefore are taxed, small firms hiring salaried workers will do so mostly illegally and are therefore partly subsidized, and firms with non-salaried contracts, regardless of size, are subsidized more than small firms hiring salaried workers (since they face no probability of being fined regardless of their number of workers).[10] The presumption is that these taxes and subsidies will be reflected in the corresponding *TFPR*s. Given the distribution of physical productivities across firms, more capital and labor will be allocated to firms with non-salaried contracts, and to smaller firms with salaried contracts, than would be warranted by these firms' underlying physical productivities (unless these distortions are offset by distortions in output prices, τ_{Qsi}).

To discuss the channels through which labor distortions lower productivity, it is convenient to recall that a firm's labor input (L_{si}) is a combination of salaried (SL_{si}) and non-salaried (NL_{si}) workers, which are not perfect substitutes. On the one hand, a firm will offer a worker a salaried contract when it needs to carefully control its labor: the worker is expected to show up and leave work at specified times and to perform the task instructed to him by the firm's manager. This allows the firm to organize the production process, monitor quality, control inventories, and to coordinate the worker's labor input with that of other salaried workers. On the other hand, a firm can offer workers non-salaried contracts when it wants to share risk, perform occasional tasks, elicit effort, or solve an information problem. When there are a few workers it is easy to monitor effort, coordinate activities, or reach agreements to distribute profits; this is the case of cooperatives or, more relevant for Mexico, family firms; that is why non-salaried contracts are more prevalent among smaller firms. It is more difficult to do this when there are a large number of workers in the firm, particularly if the production technology calls for a fixed place of work and close coordination between tasks performed by different workers, as well as for some minimum scale. In these cases salaried contracts are more appropriate than non-salaried

[10] Antón, Hernández and Levy (2012) use the same census data used in this paper to parameterize $\lambda(.)$ and reproduce the size distribution of firms and the composition of employment observed in 2008. They estimate that once wages adjust to these regulations, there is a 24 percent cost difference between a salaried and a non-salaried contract; that in the absence of labor regulations the cost of salaried labor would be about 10 percent lower and of non-salaried labor about 14 percent higher; and that firms hiring up to (more than) seven salaried workers are subsidized (taxed).

ones, because the relation of subordination and a wage payment become indispensable. As a result, on average, larger firms will have proportionately more salaried contracts than smaller ones.[11]

The tax-cum-subsidy on salaried and non-salaried labor will induce adjustments that will lower productivity, for two main reasons. First, the type distribution of firms will be biased in favor of firms intensive in non-salaried contracts; in parallel, to the extent that these firms tend to be small, the size distribution of firms will be biased in the direction of smallness. Differently put, there will be an excess of capital and labor in cooperatives and in family firms, and the level of self-employment will also be higher than optimal (as the one-man firm can only have a non-salaried contract).

Second, firms where salaried labor is indispensable will try to stay small to avoid regulations without being detected by the tax authority (La Porta and Shleifer, 2008; Levy, 2008). Indeed, it is very costly for IMSS to monitor small firms as they enter and exit the market at a much higher rate than larger firms. In addition, for very small firms the costs of collection will typically exceed the amount to be collected. Thus, smaller firms hiring salaried workers will cheat more and have lower expected average costs of labor than their larger peers (but more sharply increasing marginal costs), and thus attract more capital and labor than in the absence of these distortions. Production will tend to be dispersed among a large number of units, where economies of scale and scope may be under-exploited. Further, because these small firms are violating the law, they will have no access to formal credit. Third, more generally, an atmosphere of illegality and associated uncertainty will be less conducive to investments in labor training or technology adoption.[12]

3.2 Formality and Legality

The interaction between, on one hand, the asymmetries in the regulation of salaried and non-salaried labor and, on the other, their enforcement, gives rise to the definitions of (il)legality and (in)formality. Following Kanbur (2009), we define formality with reference to the observance of

[11] There can be large firms with non-salaried contracts: some firms in Mexico engaging in individual sales at home rather than at the store have thousands of workers. However, these cases are the exception, not the norm; historically, salaried labor has been associated with larger firm size, as the small family enterprise or cooperative with non-salaried contractual relations gives way to the larger firm with salaried relations.

[12] Fanjzylber, Maloney and Montes-Rojas (2009) study small Mexican firms and find that increases in formality translate into higher firm productivity; Fanjzylber, Maloney and Montes-Rojas (2011) find similar results for Brazil.

a particular regulation. Formal workers are those covered by regulations on salaried labor (materialized by enrollment in IMSS). Because these regulations are imperfectly enforced and cheating occurs, salaried workers can be hired formally or informally, in the latter case also illegally. On the other hand, non-salaried workers are informal, but legal. Note that if enforcement were perfect, there would be a one-to-one correspondence between (in)formal and (non)salaried workers. Note as well that legality is defined with reference to compliance with regulations on salaried labor; other dimensions of legality like complying with value added or income taxes that are equally relevant for firm behavior are not considered here.[13]

The simplicity of these classifications does not extend to firms, however, since they can both mix salaried and non-salaried workers, on one hand; and only enroll part of their salaried workers with IMSS, on the other. It is clear that a firm that only hires salaried workers and enrolls all of them with IMSS, is both formal and legal; and that a firm with only salaried workers that does not enroll any with IMSS, is informal and illegal; in contrast, a firm that only has non-salaried workers but does not enroll any with IMSS because it is not obligated to, is informal but legal. However, there are mixed cases as well. As a result, we develop indices of firm legality and firm formality. We measure firm legality by the ratio of total social security taxes to the wages of salaried workers. Legal firms pay all the social security taxes they owe on the salaried workers that they hire; illegal firms hire salaried workers but pay no social security taxes. Semi-legal firms hire salaried workers but pay fewer social security taxes than they should. On the other hand, we measure firm informality as the ratio of social security taxes to the remunerations of all workers (salaried and non-salaried). Informal firms have a formality index of zero, whether they are legal or not. Formal firms only hire salaried workers and comply with the law; semi-formal firms represent intermediate cases.

These definitions produce five types of firms, shown in Table 2.[14] Three points are worth emphasizing about these definitions. First, informality and illegality are not the same; as noted,

[13] In general one would expect that firms that cheat along one dimension (regulations on salaried labor) also cheat along others (value added or income taxes); this is so, but the association is not one-to-one because in Mexico the value added and income tax laws have various exemptions and special regimes by sector of activity and volume of firm's sales. This is an important observation to the extent that labor distortions may be correlated with output distortions, with some of the effects of informality being captured by the latter.

[14] Total taxes on salaried workers in Mexico are approximately 32 percent of the wage (excluding the contingent costs associated with firing and severance pay regulations). However, the census data excludes contributions to the housing fund, to sub-national taxes and contributions made by workers. As a result, a firm that fully complies with the components of salaried regulations included in the census would pay 18 percent of wages as social security taxes. To compute the indices we also imputed wages for unpaid workers using the average wage paid by firms with

firms and workers can be informal but legal. Second, informality is not defined by the size of firms; many firms can be large and informal or small and formal. Finally, informal firms are not necessarily equivalent to low productivity firms; many firms with non-salaried contracts may be highly productive.

Table 2. Classification of Firms by Formality and Legality Indices

	Index of Legality	Index of Formality
Legal and Formal	>18%	>18%
Legal and Informal	not defined	0%
Legal and Semi-formal	>18%	0% -18%
Semi-legal and Semi-formal	0% -18%	0% -18%
Illegal and Informal	0%	0%

To sum up: we take as given the underlying distribution of firms' physical productivities, A_{si}. In the absence of distortions, capital and labor are optimally allocated across firms. When distortions are present, misallocations occur and *TFPQ* is lowered: some firms are getting more capital and labor than what their underlying A_{si} justify, and others with higher A_{si} are getting less. We focus on distortions caused by labor regulations in a context of imperfect enforcement and argue that these distortions over-allocate capital and labor to firms with non-salaried contracts (including family firms and the self-employed), and to small firms with illegal salaried contracts, resulting in an excess of informal firms, some legal and some illegal.

4. Data

Our analysis relies upon detailed establishment level data from Mexico's Economic Census, conducted by the Mexican Statistical Office (INEGI). We have access to data for 1998, 2003 and 2008, although for reasons of space we concentrate mostly on 1998 and 2008.[15] The census measures economic activity taking place in private establishments with a fixed location in urban areas, and captures information on firm sales, value added, number of workers, types of contractual arrangements, labor remunerations, payments to IMSS, and value of fixed capital

up to 10 workers in the corresponding 6-digit sector in each year (firms in this category represent 70 percent of all unpaid workers). Results of the paper are basically the same if we do not perform this imputation.

[15] In the Appendix we present a complete set of results for all three years.

stock.[16] Although it covers all non-agricultural activity, we focus on manufacturing, services and retail commerce and exclude from our analysis the energy, mining and transportation sectors whose firms either operate in very concentrated markets or are dominated by state-owned enterprises. The census classifies activities with considerable detail, up to 6 digits of the North American Industrial Classification System. We excluded from the analysis industries with fewer than ten establishments, since few firms in an activity may be incompatible with our model of monopolistic competition.

Table 3 shows the total number of industries, firms and workers. In 2008 the three broad sectors covered more than 3.6 million establishments employing 17.7 million workers (97 percent of all firms and 87 percent of all workers captured in the census).

Table 3. Sample Sizes

		Industries	Establishments	Workers
1998	Manufacturing	297	344,118	4,512,595
	Retail & Wholesale	147	1,443,676	3,999,290
	Services	115	938,572	4,265,660
	Total	559	2,726,366	12,777,545
2008	Manufacturing	292	436,851	4,661,062
	Retail & Wholesale	262	1,858,550	6,134,758
	Services	153	1,348,581	6,859,659
	Total	707	3,643,982	17,655,479

Source: Authors' calculations based on Mexican Economic Census (INEGI).

Approximately 40 countries in the world report conducting economic censuses on a regular basis (UN, 2010). Nine of these are in Latin America: Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, El Salvador, Mexico and Uruguay. Mexico is the only one that carries out a census every five years covering all sectors of the economy and all firm sizes. This is unusual, as most small establishments are normally not captured in censuses typically available in other countries; in addition, it is one of the few censuses that collects information on services and commerce. Having detailed establishment level records for non-manufacture firms is rare. It is, by far, the largest and most comprehensive economic census in Latin America, and one of the largest available in the world.

[16] The census treats each plant as a firm. In the 2008 census only 2.5 percent of all plants belong to multi-plant firms (i.e., 97.5 percent of all firms have only one plant).

That said, Mexico's census also has some important limitations. It excludes economic activity in rural areas and in government offices, as well as in urban areas that takes place in mobile units (street vendors and the like). These are important shortcomings, particularly for the study of informality. Table 4 combines data from the 2008 census and the National Survey of Employment and Occupation (ENOE) to understand the implications of these exclusions.

Table 4. Distribution of Mexico's Employment

	Total Number of Workers	% of Col. [1] that is Informal
	[1]	[2]
Private urban employment captured in Census (All Sectors)	**19,629,896**	**55.5**
0 – 5	8,770,687	93.2
6 – 10	1,714,678	57.2
11 – 50	3,791,630	28.0
50+	5,352,895	12.8
Private urban employment not captured in Census	**13,223,008**	**86.8**
Self-employment	4,073,747	99.8
2 – 5	6,228,533	96.6
6 +	2,920,728	48.0
Public sector employment not captured in Census	**4,645,104**	**0.0**
Agriculture and other rural employment not captured in Census	**5,638,429**	**95.0**
Total	**43,136,437**	**64.3**

Source: Authors' calculations based on data from INEGI.

The census underreports informal employment, the number of active firms, the number of small firms, and (probably) the number of illegal firms. First, total urban employment equals almost 33 million workers, of which 40 percent are employed in firms not captured by the census. The majority of these excluded workers are informal (although not necessarily illegal). Second, there are more firms than the 3.6 million captured in the census. The ENOE asks workers the size of the firm they work for. Using this information we estimate that, in addition to the 3.6 million firms captured in the census, there are at least 1.2 million more firms, all of them informal.[17] Further, there are 4 million self-employed workers in urban areas which, again, are all informal. To this we need to add 5.6 million workers in rural areas, of which 95 percent are informal. Third, if account were taken of all firms regardless of whether they have a fixed establishment or not, the share of small firms would increase considerably and average firm size

[17] We say at least assuming that the 6.2 million workers in firms with 2 to 5 workers are all employed in firms with 5 workers; if we used the mean number of workers (3.5), there would be an additional 1.8 million firms (50 percent of what is captured in the census), not 1.2 million.

would fall (even if the self-employed are not counted as a firm). Unfortunately, for firms without a fixed location there are no equivalent data on sales, composition of employment, value added and capital as there are for those included in the census. Fourth, although we have no evidence as to whether informal firms not captured in the census offer their workers salaried or non-salaried contracts, the presumption is that it is mostly the latter, as their activities are undertaken outside of any fixed establishment, where effort is difficult to monitor, and where payments are more naturally structured as commissions (or profit-sharing arrangements in the case of family firms).

In sum, even though we have access to unusually rich and detailed data, it is still the case that they only account for approximately 50 percent of all non-public sector workers in Mexico (19.6 out of 38.5 million), for an undetermined share of the capital stock, and for an equally undetermined number of mostly small firms. This is an important observation to bear in mind when interpreting the results that follow.

5. Stylized Facts

5.1 Resource Allocation in the (Census Captured) Mexican Economy

We begin by summarizing in the first top three panels of Table 5 the shares of establishments, workers, capital and value added by sector, firm size and formality and legality status. We show these numbers for 2008.[18] Three stylized facts are worth noting. First, while the manufacturing sector only includes 12 percent of the establishments, it hires 26 percent of the labor force, has 48 percent of capital and produces half the value added of the economy.

Second, there is a very large number of small firms that employ a non-trivial proportion of workers (even ignoring firms excluded from the census). Almost 90 percent of firms employ less than five workers, and 6 percent of them employ between six and 10 workers. Together they account for 46 percent of employment, 18 percent of capital and only 15 percent of value added. Large firms (with over 50 workers), on the other hand, are only 1 percent of all firms and employ fewer workers than small firms (39 percent vs. 46 percent) but account for 73 percent of value added.[19] Note that if we were to truncate the sample so that it resembles the data typically available in other countries, including only firms with 10 or more employees, we would account for only 4 percent of the establishments and we would miss almost half of the workers.

[18] These numbers are very similar for 2003 and for 1998, and can be found in A1-A3 in the Appendix.
[19] In the Appendix we show that these patterns are present for all three broad sectors. See Appendix Tables A1-A3.

Third, the majority of firms are informal (legal and illegal) and employ a significant amount of resources (again, even within the share of the economy captured by the census). Two-thirds of all firms are informal but legal. These firms account for 35 percent of workers and 32 percent of capital, and they produce 21 percent of value added. On the other hand, legal and formal firms are only 3 percent of all firms, hire 20 percent of workers, account for 30 percent of capital and produce 31 percent of value added. Notably, almost 23 percent of all firms hire their salaried workers illegally. In fact, there are over seven times more firms that hire salaried workers illegally than legally, prima facie evidence that the incentives to evade are very large indeed. These illegal firms hire 20 percent of workers but produce only 7 percent of value added.

Table 5. Resource Allocation and Output

	Establishments	Workers	Capital	Value Added
Shares in 2008				
By Sector				
Manufacturing	12.0	26.4	48.3	49.5
Retail & Wholesale	51.0	34.7	17.3	21.0
Services	37.0	38.9	34.5	29.5
By Firm Size				
[0-5]	89.7	37.8	13.2	10.3
[6-10]	5.8	8.8	4.5	4.6
[11-50]	3.6	14.9	10.2	12.5
[+50]	0.9	38.5	72.1	72.5
By Formality/Legality Status				
Legal & formal	3.1	20.4	29.1	30.7
Legal & Informal	67.5	35.4	31.8	20.6
Legal & Semi-formal	1.7	3.1	2.4	1.9
Semi-legal & Semi-formal	5.1	22.1	30.4	39.9
Illegal and informal	22.7	19.1	6.3	6.8
Growth 1998-2008				
By Sector				
Manufacturing	26.9	3.3	78.8	154.8
Retail & Wholesale	28.7	53.4	190.7	78.9
Services	43.7	60.8	151.8	213.6
By Firm Size				
[0-5]	31.9	53.3	69.6	68.7
[6-10]	72.1	68.8	111.2	89.6
[11-50]	33.8	30.0	105.3	98.0
[+50]	20.7	24.0	127.2	182.2
By Formality/Legality status				
Legal & formal	-11.5	36.2	202.5	190.8
Legal & Informal	33.2	97.8	616.1	322.2
Legal & Semi-formal	-10.5	24.1	159.5	92.5
Semi-legal & Semi-formal	-29.6	-22.3	63.9	95.6
Illegal and informal	96.4	121.0	86.4	82.8

Source: Authors' calculations based on Mexican Economic Census (INEGI).

The bottom three panels of Table 5 show the change in the allocation of resources and value added over the 1998-2008 period. Changes in capital and value added are nominal, so one should interpret the numbers relative to a base category which, for simplicity, will be the first row of each panel. We highlight three trends.

First, resources moved away from manufacturing; the allocation of resources tilted towards commerce and services. Whereas the number of workers employed in these two sectors increased by 50 to 60 percent, manufacturing did not hire more workers; further, these two sectors also captured twice as much new capital as manufacturing. The number of firms grew faster in services than in the other two sectors.

Second, there is a significant jump in the number of small firms over the 10-year period. More than 780,000 new firms with five or less employees and 115,000 with six to 10 employees were created.[20] Employment in firms with less than five workers grew by 53 percent and in firms with six to 10 by 69 percent; in contrast, employment in firms with 11 to 50 workers grew by 30 percent, and in firms with more than 50 workers by 24 percent. In parallel, however, value added in larger firms grew more than in smaller ones (183 percent in firms with more than 50 workers vs. 69 percent in those with less than five). Differently put, workers moved towards smaller firms at the same time that these same firms lost relative importance in terms of value added in the economy.

Third, the number of informal firms increased dramatically over the 1998-2008 period. Among those, illegal ones increased three times more than the legal ones. Moreover, illegal-informal firms captured relatively more workers than any other type of firms. The number of legal and formal establishments decreased and the number of workers employed by them increased at a much slower rate than at any other type of establishment. Finally, legal-informal firms hired relatively more capital and increased their share in value added vis-à-vis all other firms. In sum, the census data shows an unambiguous trend towards more informality and illegality.

[20] This phenomenon is not a spurious result from a change in data collection methods or broader regional coverage; this was verified with INEGI.

5.2 Differences between Informality, Illegality and Size

Table 6 shows the share of establishments, resources and value added that are allocated to each type of firm disaggregating by formality and legality status and firm size. Each panel adds up to 100 percent. Note first that 66 percent of firms are small, informal and legal. At the same time, many small firms are formal; in fact, in absolute numbers there are more small formal firms than large ones (44,246 vs. 10,408).

Table 6. Shares by Firm Size and Informality Status

(Shares for 2008)

	Legal & formal	Legal & Informal	Legal & Semi-formal	Semi-legal & Semi-formal	Illegal and informal
Establishments					
[0-5]	1.21	65.60	1.18	2.49	19.25
[6-10]	0.69	1.11	0.33	1.23	2.46
[11-50]	0.87	0.56	0.15	1.10	0.90
[+50]	0.29	0.19	0.03	0.30	0.05
Workers					
[0-5]	0.77	23.60	0.79	1.69	10.94
[6-10]	1.09	1.63	0.50	1.92	3.66
[11-50]	3.84	2.39	0.57	4.72	3.36
[+50]	14.67	7.74	1.19	13.75	1.18
Capital					
[0-5]	0.73	7.17	0.64	1.33	3.35
[6-10]	0.85	0.78	0.29	1.23	1.30
[11-50]	3.18	2.15	0.35	3.29	1.28
[+50]	24.31	21.75	1.10	24.59	0.34
Value Added					
[0-5]	1.24	4.71	0.26	1.07	3.03
[6-10]	0.95	0.78	0.20	1.37	1.35
[11-50]	3.69	2.17	0.33	4.67	1.66
[+50]	24.85	12.98	1.15	32.83	0.73

Source: Authors' calculations based on Mexican Economic Census (INEGI).

Second, as expected, firm formality increases with size. Among large firms, there are more formal than informal and more legal than illegal ones. That said, observe that less than half of all large firms are formal and that a high proportion of large firms are informal but legal, offering their workers non-salaried contracts. We were expecting no illegal large firms, given that in their case the probability of being audited is high. There are however, about 1,800 establishments that are large, illegal and informal. We think these are establishments that are part of a multi-plant firm. Unfortunately, the census did not capture which plants of multi-plant firms

included their social security taxes in another plant of the same firm. As a result, it is very likely that the number of illegal plants of this size is overestimated, although we cannot measure this with precision; potentially there could be no illegal plants in this size.

Third, illegality is proportionally more important among firms with six to 10 workers than in those with up to five workers; this follows from the fact that most firms with up to five workers have mostly non-salaried workers (perhaps because they are by-and-large family firms), while firms with six to 10 workers have more salaried ones, albeit illegally hired.

The allocation of resources and output also yields interesting stylized facts. First, almost one quarter of workers are in small, legal and informal firms and 11 percent are in small, illegal and informal firms. While salaried workers hired illegally are concentrated in firms with up to 10 workers, even medium and large firms hire some workers illegally (although for the reasons stated above the number of illegal workers in large firms is probably overestimated). Employment in formal firms is almost the same as in illegal and informal firms (3.6 vs. 3.3 million workers, respectively), highlighting the importance of illegal salaried employment. On the other hand, employment in informal legal firms is almost double employment in informal illegal firms (6.2 vs. 3.3 million, respectively), highlighting in turn the importance of distinguishing between legal and illegal informal employment.[21]

5.3 Productivity, Informality and Illegality

We now introduce back into the analysis the notions of total factor productivity (*TFPQ*), revenue productivity (*TFPR*), and idiosyncratic taxes/subsidies faced by the firm (τ_L and τ_Q). In this section we analyze how these measures correlate with the indices of informality and illegality. Table 7 shows the results.[22]

[21] See Appendix Table A4-A7 for further results.
[22] In Appendix Tables A8-A10 for more detailed results.

Table 7: Correlations

	$Log\left[\frac{TFPQ_{si}}{\overline{TFPQ_s}}\right]$	$Log\left[\frac{TFPR_{si}}{\overline{TFPR_s}}\right]$	Formality Index	Legality Index	Firm Size (# Workers)	$Log\left[\frac{1-\tau_{Qsi}}{\overline{1-\tau_{Qs}}}\right]$
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.000					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.754	1.000				
Formality Index	0.316	0.136	1.000			
Legality Index	0.254	0.097	0.942	1.000		
Firm Size (# Workers)	0.362	0.067	0.159	0.127	1.000	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.414	-0.781	0.030	0.041	0.101	1.000
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.676	0.499	0.227	0.176	0.145	-0.339

Source: Author's calculations based on Mexican Economic Census (INEGI)

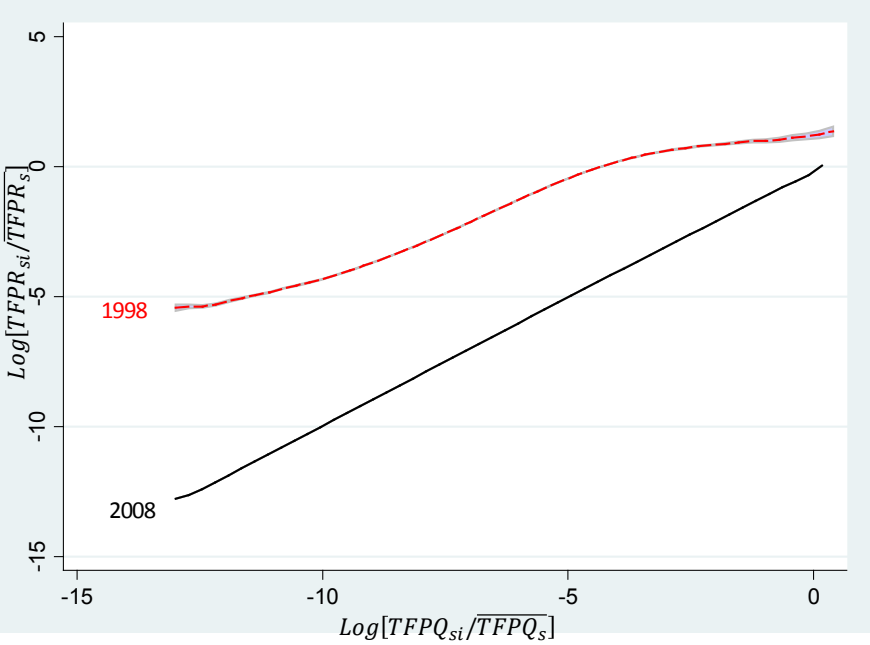
Various observations are relevant. First, it is often argued that firm size and firm productivity are interchangeable concepts. More productive firms will survive longer periods in the market, will attract more resources, and will grow over time. Therefore measuring firm size is the same as measuring firm productivity. In fact, there is a positive correlation between firm size, measured by the number of workers, and firm relative productivity; however, the correlation, at 0.362, is far from one. This suggests that in Mexico productive firms are not attracting as much resources as they should.

Second, there is a strong correlation between productivity (*TFPQ*) and distortions (*TFPR*). More productive firms face larger distortions. Figure 1 captures this showing a non-parametric regression of these two measures. It can be observed that the relation is almost linear and that the correlation (the slope of the curves) increased between 1998 and 2008, suggesting that more productive firms face higher distortions in 2008 than 10 years earlier (a trend documented in the next section).

Where do these distortions come from? Table 7 shows that more productive firms face both larger scale distortions (smaller $1-\tau_Q$) and factor price distortions (larger $1+\tau_L$). Figure 2 shows a kernel density estimate of the distribution of *TFPQ* for each type of firm. We rescaled the abscissa so that we can interpret it as relative productivity. That is, firms near 1 are as productive as their (6-digit) industry average. Firms close to 1/4 have a productivity that equals 25 percent of its industry average, and so forth. In all three sectors, there are huge differences in productivity between the least and the more productive firms. Moreover, there is also a very large mass of firms that have productivities less than their sector average; the majority of those

firms are informal.[23] Lastly, observe that legal formal firms are the minority, and their density almost overlaps with the horizontal axis. As a result, it is not easy to see that on average these firms are more productive than the rest, although this is confirmed in the regressions presented in Table 9.

Figure 1: Non-Parametric Regression of $Log\left[\frac{TFPQ_{si}}{\overline{TFPQ_s}}\right]$ on $Log\left[\frac{TFPR_{si}}{\overline{TFPR_s}}\right]$

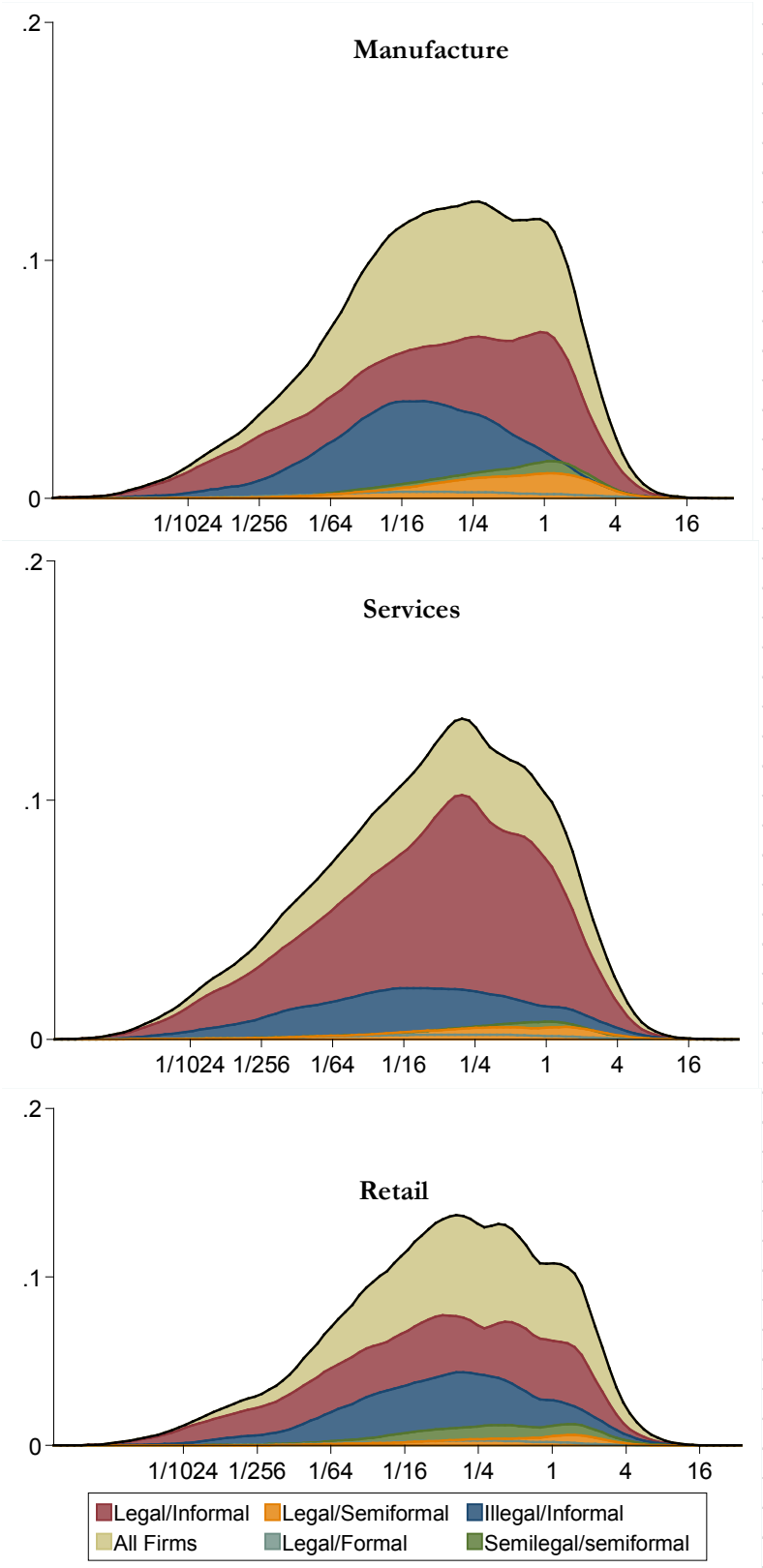


Source: Author's calculations based on Mexican Economic Census (INEGI)

Finally, Figure 3 shows a non-parametric regression of TFPQ and TFPR on the legality and formality index for 1998 and 2008. The correlations shown in Table 7 hide important information. First, there is a discontinuous jump both in *TFPQ* and *TPFR* between illegal and informal firms and all other firms. Second, conditional on having some salaried workers, the higher the formality index the more productive and more distorted against the firms are. These highly non-linear relations justify using categories rather than continuous indices in the regression analysis of the next section.

[23] IDB (2010) compares the distribution of TFPQ between Mexican and US firms in manufacturing at the four digit level, and finds that the dispersion in TFPQ is much larger in Mexico, with most of the variance explained by the presence of a thick left tail populated by many small firms (see Fig 4.3(a), p. 76). In parallel, Leal (2010) compares the distribution of employment by firm size in Mexico and the US. When only employment in formal firms is considered, the two distributions overlap considerably; when employment in informal firms is added, however, the two distributions differ importantly, again as a result of a large left tail of employment in mostly small and informal firms.

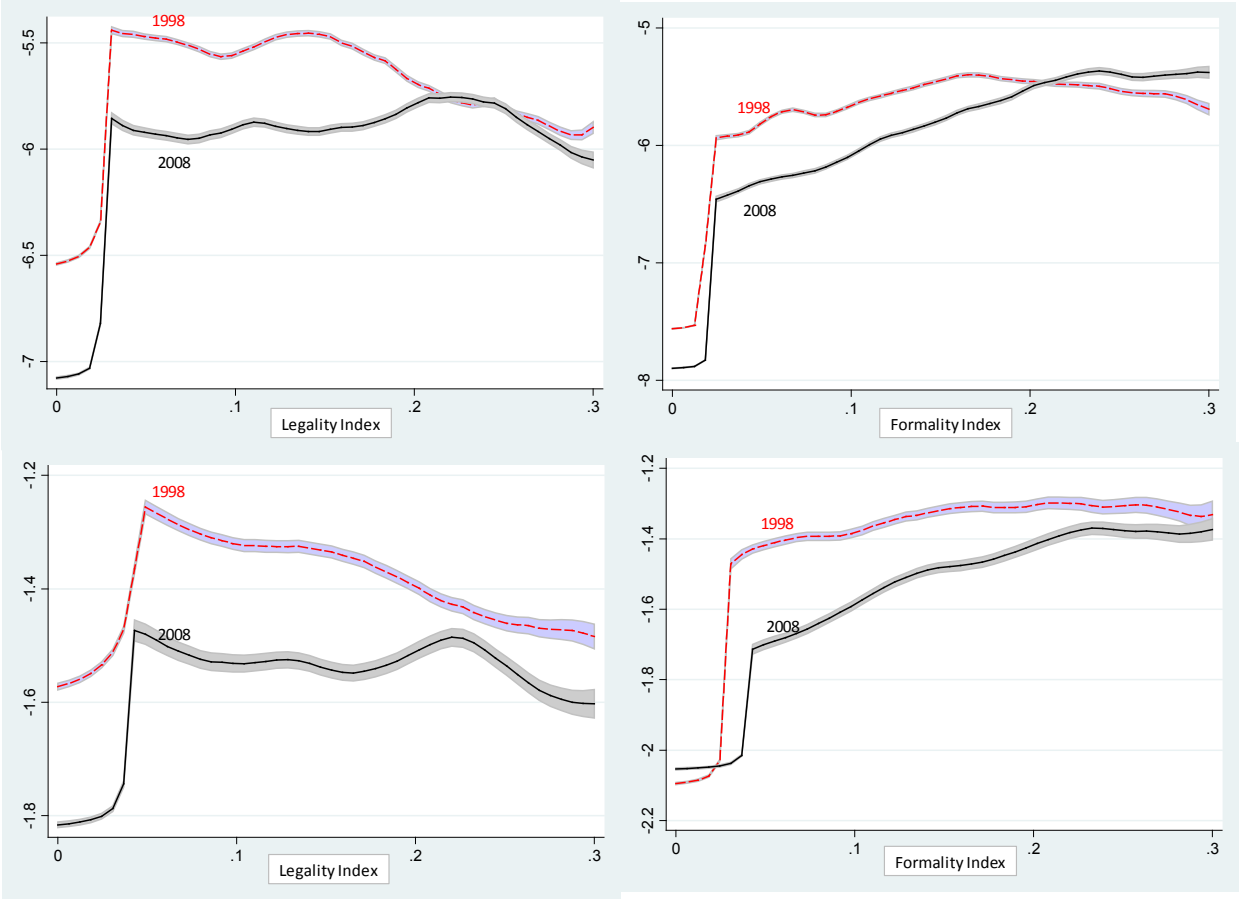
Figure 2: Distribution of $Log\left[\frac{TFPQ_{si}}{TFPQ_{s}}\right]$



Note: Kernel estimation is based on value-added weighted relative productivity.
Each subgroup is rescaled by its corresponding number of firms.

Source: Author's calculations based on Mexican Economic Census (INEGI)

Figure 3: Non-Parametric Regressions of TFPQ/TFPR on Legality/Formality



Source: Author's calculations based on Mexican Economic Census (INEGI)

6. Results

We now present two sets of results. First, we quantify the degree of heterogeneity in productivity and distortions and show that they are much larger than other results found in the literature for other countries, and that they have increased between 1998 and 2008. Second, we quantify the differences in productivity and distortions across firm types, and argue that the misallocations behind the productivity losses are at least in part related to labor regulations.

6.1 Distortions and Heterogeneity in Firm Productivities

To measure the heterogeneity in firms' productivities and distortions, we show statistics of the dispersion of *TFPQ*, *TFPR*, and the two components of *TFPR*, τ_L and τ_Q. Table 8 presents three

measures of dispersion for 1998 and 2008: the inter-quartile range, the difference between the 90th and 10th percentile, and the standard deviation.

We can observe that firms in the 90th percentile of productivity are between 400 to 500 percent more productive than firms in the 10th percentile implying that within sectors defined at the 6-digit level, certain firms are able to produce much more output out of the same amount of inputs than others. To put these numbers in perspective, Hsieh and Klenow (2009) report that in the US firms in the 90th percentile are around 85 percent more productive than firms in the 10th percentile. The productivity dispersion is larger in the retail sector than in services than in manufacturing. Productivity dispersion slightly increased between 1998 and 2008.

In the Appendix we show that if we truncate the sample to only manufacturing establishments with 10 or more workers (the typical dataset available for most countries), the dispersion of *TFPQ* is lower.[24] Nonetheless, even when we make the sample comparable, it is still the case that the dispersion of *TFPR* in manufacturing in Mexico is higher than in the US (Hsieh and Klenow, 2009) and in any other Latin American country for which similar computations have been performed (Busso, Madrigal and Pagés, 2010).

The second line in each panel shows the dispersion of the distribution of *TFPR*. In the absence of distortions, all firms should have the same *TFPR* and therefore dispersion should be zero. As discussed, higher dispersion indicates more distortions and a greater misallocation of resources across firms. We observe that according to these metrics Mexico suffers from a substantial degree of misallocation. Dispersion is higher in retail and services. Note again that distortions increase slightly between 1998 and 2008. A heuristic interpretation of the differences in *TFPR*, focusing on 2008 and the aggregate economy, is that one peso of capital and labor allocated to firms in the 90th percentile is worth three times more than the same peso allocated to firms in the 10th percentile. This relation falls from 3 to 1.5 comparing firms between the 75th and 25th percentile; a smaller but still substantial difference. Finally, observe that these differences are larger in 2008 vis-à-vis 1998.

[24] See Appendix Tables A11-A14.

Table 8: Dispersion of Productivity and Distorsions

Within 6-digits industries (all firm sizes)

	1998			2008		
	Std. Dev.	p75-p25	p90-p10	Std. Dev.	p75-p25	p90-p10
Agg.Economy						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.78	2.50	4.64	1.89	2.64	4.90
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.09	1.46	2.78	1.18	1.51	2.98
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.33	1.79	3.39	1.49	1.96	3.80
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.29	1.81	3.38	1.47	2.08	3.88
Manufacturing						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.81	2.54	4.65	1.89	2.66	4.89
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.00	1.31	2.52	1.11	1.43	2.78
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.23	1.62	3.09	1.34	1.74	3.36
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.27	1.79	3.28	1.41	1.91	3.68
Retail & Wholesale						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.78	2.49	4.67	1.93	2.64	5.03
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.22	1.68	3.17	1.33	1.77	3.43
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.42	1.95	3.68	1.69	2.21	4.44
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.27	1.77	3.41	1.53	2.10	4.12
Services						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.74	2.44	4.56	1.86	2.63	4.80
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.10	1.49	2.83	1.18	1.46	3.00
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.42	1.96	3.67	1.61	2.14	4.10
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.34	1.93	3.55	1.53	2.33	4.03

Source: Author's calculations based on Mexican Economic Census (INEGI)

The high level of dispersion in distortions suggest the existence of large potential gains in productivity that could be achieved by moving factors from firms with low to those with high marginal revenue products. Using the model we can calculate "efficient" output in order to compare it with actual output levels. In other words, we compute how much output an economy loses by allocating resources inefficiently. The aggregate TFP cost of misallocation is defined as C*=(TFPQ*/TFPQ-1)×100. If we eliminate all distortions and allow reallocation of existing capital and labor across firms, aggregate productivity in Mexico could increase by approximately 200 percent, a number that dwarfs the 43 percent found by Hsieh and Klenow (2009) for the US case and 86 percent for China. If we only eliminate factor price distortions, aggregate TFP would increase by 123 percent.[25,26]

[25] These gains consider reallocation only within six-digit industries. There could be further gains from reallocating across industries. Obviously, this exercise assumes no measurement error and a correct model specification. In

6.2 Are Informality and Illegality Hurting Mexico's Productivity?

We begin by recapitulating the main hypothesis relating informality and productivity; we then present regressions that are consistent with these hypothesis; we conclude making reference to other empirical work that show how the asymmetry in labor regulations is affecting the allocation of resources in Mexico in the direction of informality.

In general, there would be no *ex-ante* expectation that firms with non-salaried contracts are, on average, more or less productive than those with salaried contracts. In the case of Mexico, however, we argue that: i) there is an asymmetry in the regulation of salaried and non-salaried labor; ii) the relevant parameter values associated with these regulations result in a tax on salaried labor and a subsidy to non-salaried labor and, iii) the regulations on salaried labor are imperfectly enforced, allowing some firms to cheat. As a result, there will be a ranking in the average costs of labor: firms intensive in salaried labor that comply with the law will have higher costs than those that do not comply; and, in turn, these latter firms will have higher costs than those intensive in non-salaried labor. In this particular context, therefore, we expect legal and formal firms to face more distortions than illegal informal firms, which in turn face more distortions than informal legal firms.

Table 9 present OLS regressions of *TFPQ* and *TFPR* on legality and formality status using different samples, specified in the second row of the table, and controlling for size and age (except in columns [1] and [5]).

Consider *TFPQ* first. The first column, corresponding to the sample of all firms, shows that all firms are less productive than formal ones which is the excluded category (specified for reference on the third row of the table). This is in accordance with our broad expectations. In the second column we show that results are basically the same once we control for size and age. We condition on firms size and age to control for other factors that might affect productivity and distortions and that are also correlated with firm size (e.g., enforcement of value added taxes, access to credit). The disadvantage is that these controls will wash the variation that might be causing the differences in the type of labor contracts. In general, the results are qualitatively the

---

particular, we measure labor input by the number of workers adjusting by the wage. Unobserved differences in worker quality will produce spurious differences in the marginal product of labor and capital. However, if measurement and modeling errors explain these results, they have to be much bigger in Mexico than in any other country reported by the literature (despite the fact that the Mexican census is arguably of higher quality). And within Mexico, measurement and modeling error have to be larger in manufacturing than in the other sectors.

[26] In Appendix Table 15 we present more detailed results.

same with and without size and age controls. If anything, controlling for these variables is more conservative in the sense that it makes the results weaker. Column 3 shows that, within the set of legal firms and even after controlling for size and age, formal firms are more productive than informal or semi-formal ones. This result is critical because it removes the effect of illegality and focuses only in the differences between salaried and non-salaried contracts. This result lends support to the proposition that it is the status of informality and not illegality that is associated with lower *TFPQ*. Finally, column 4 shows that, within the subset of informal firms, illegal ones with salaried workers are more productive than legal ones with non-salaried workers. These results are, again, consistent with our expectations.

Table 9: Correlation with Formality/Legality Status

Dependent Variable:	$Log\left[\frac{TFPQ_{si}}{\overline{TFPQ_s}}\right]$				$Log\left[\frac{TFPR_{si}}{\overline{TFPR_s}}\right]$			
Sample:	All Firms	All Firms	Only Legal Firms	Only Informal Firms	All Firms	All Firms	Only Legal Firms	Only Informal Firms
Excluded Category:	Legal and Formal	Legal and Formal	Formal	Legal	Legal and Formal	Legal and Formal	Formal	Legal
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Legal-semiformal	-1.616	-0.627	-0.569		-0.433	-0.368	-0.363	
	[0.0081]	[0.0071]	[0.0066]		[0.0045]	[0.0046]	[0.0043]	
Legal-informal	-2.426	-1.019	-0.842		-0.506	-0.372	-0.324	
	[0.0043]	[0.0041]	[0.0041]		[0.0024]	[0.0027]	[0.0027]	
Semilegal-semiformal	-0.211	0.053			0.020	0.034		
	[0.0046]	[0.0039]			[0.0025]	[0.0025]		
Illegal-informal	-2.089	-0.440		0.592	-0.284	-0.151		0.237
	[0.0045]	[0.0043]		[0.0031]	[0.0025]	[0.0028]		[0.0021]
Size Dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Age Control	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Obs	2,231,762	2,231,762	1,463,785	1,956,266	2,231,762	2,231,762	1,463,785	1,956,266
R-squared	0.188	0.399	0.440	0.255	0.032	0.041	0.053	0.018

Source: Author's calculations based on Mexican Economic Census (INEGI). Note: s.e. between brackets

Now consider *TFPR*. Columns 5 and 6 show the main results for the sample of all firms. Formal firms face more distortions than all other firms except semi-formal and semi-legal ones (although the difference is very small). Following on the heuristic interpretation of differences in *TFPR* offered before, the results show that one peso of capital and labor allocated to formal firms is worth 28 percent more than in illegal informal firms, and 50 percent more than in legal informal firms. Differently put, firms with salaried contracts that comply with the law face more distortions than those that do not comply; which in turn face more distortions than those with non-salaried contracts. Column 7 shows that, among legal firms, formal establishments have higher *TFPR* than semi-formal or informal ones, indicating that formal firms face higher distortions than all others. Finally, column 8 shows that, within the sub-set of informal firms,

legal ones with non-salaried workers face fewer distortions than illegal ones with salaried workers. Again, these results are consistent with our hypothesis.

Table 10 is a robustness check. It basically shows that all the relationships found in Table 9 survive even within firms in the same 6-digit industry and in the same firm size bin. Panel 1 shows regression results for the sample of all firms, panel two show regressions for the sample of legal firms, and panel three for the sample of informal firms. In each case, we estimate the regressions separately for every size bin. The results shown in Table 9 hold in all cases.[27]

Table 10: Correlation with Formality/Legality Status, by Firm Size

	Dependent Variable:	$Log\left[\frac{TFPQ_{si}}{\overline{TFPQ_s}}\right]$				$Log\left[\frac{TFPR_{si}}{\overline{TFPR_s}}\right]$			
Sample	Firm Size:	**Size [0-5]**	**Size [6-10]**	**Size [11-50]**	**Size [+50]**	**Size [0-5]**	**Size [6-10]**	**Size [11-50]**	**Size [+50]**
		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
All Firms	Legal-semiformal	-0.959	-0.910	-0.746	-0.574	-0.543	-0.569	-0.438	-0.099
		[0.0101]	[0.0215]	[0.0299]	[0.0711]	[0.0068]	[0.0144]	[0.0191]	[0.0405]
	Legal-informal	-1.759	-0.688	-0.395	-0.491	-0.606	-0.324	-0.317	-0.164
		[0.0068]	[0.0194]	[0.0195]	[0.0309]	[0.0046]	[0.013]	[0.0125]	[0.0176]
	Semilegal-semiformal	-0.214	-0.085	-0.092	0.159	-0.122	-0.048	-0.053	0.134
		[0.0081]	[0.0146]	[0.0149]	[0.0267]	[0.0055]	[0.0098]	[0.0095]	[0.0152]
	Illegal-informal	-0.881	-0.632	-0.701	-1.039	-0.300	-0.306	-0.260	-0.127
		[0.007]	[0.0142]	[0.0177]	[0.0826]	[0.0047]	[0.0095]	[0.0113]	[0.047]
Only Legal Firms	Legal-informal	-1.743	-0.684	-0.395	-0.495	-0.592	-0.326	-0.315	-0.166
		[0.0064]	[0.0205]	[0.0218]	[0.0329]	[0.0043]	[0.0137]	[0.0139]	[0.0189]
	Legal-semiformal	-0.961	-0.908	-0.748	-0.585	-0.545	-0.572	-0.435	-0.114
		[0.0094]	[0.0228]	[0.0335]	[0.0754]	[0.0063]	[0.0152]	[0.0213]	[0.0432]
Only Informal Firms	Illegal-informal	0.878	0.072	-0.315	-0.607	0.307	0.028	0.056	0.004
		[0.0031]	[0.0164]	[0.0219]	[0.091]	[0.0021]	[0.0111]	[0.014]	[0.0523]

Source: Author's calculations based on Mexican Economic Census (INEGI). Note: s.e. between brackets.

We end discussing the casual relation between the asymmetry in labor regulations and resource misallocation. To do this, we note that between 1998 and 2008 there have been no changes in the regulations affecting salaried labor, T_s. On the other hand, there have been notable changes in the regulations affecting non-salaried labor, T_{ns}. In particular, for the period 1998-2007 T_{ns} increased by 110 percent in real terms, from 1.1 percent of GDP in 1998 to 1.8 percent in 2007.[28] Assuming no changes in the enforcement of regulations on salaried workers, and on

[27] Appendix Tables 16-19 present more detailed results.

[28] Programs to provide social benefits to informal workers have been in place since the creation of IMSS in 1943. For a long time the resources channeled to them were small relative to those channeled for formal workers. As of 1996, however, there has been an important change, with a significant expansion of health, day care, pension and housing programs for workers not covered by IMSS; see Levy (2008) for discussion.

workers' valuation of these regulations, a higher T_{ns} increases the subsidy to non-salaried and illegal salaried contracts, shifting resource allocation in the direction of informality (see Table 1).

The increase in T_{ns} reflects the expansion or introduction of new health, day care, housing and pension programs for informal workers, by both the federal and state governments. Importantly for our purposes, some of these programs have been subject to impact evaluations. Four results are of interest. First, Juárez (2008) evaluates the impact of a program introduced in Mexico City in 2002 to provide free health benefits to women conditional upon informal employment. Taking advantage of the fact that the program was not introduced in Guadalajara and Monterrey, the second and third largest cities after Mexico City, she finds that the share of women with at most high school education and with formal employment in Mexico City fell from 0.55 to 0.45; she also finds a compensating decrease in informal wages. Second, Bosch, Cobacho and Pagés (2012) carry out a meta-evaluation of various evaluations of the impact of another health program for informal workers started in 2003, this one at the federal. They show that depending on the data used and the method of estimation, this program (known as *Seguro Popular*), reduced formal employment by between 160,000 and 400,000 workers, equivalent to between 8 to 20 percent of the formal jobs created between 2003 and 2008. Third, Bosch and Campos-Vázquez (2010) study the impact of Seguro Popular on firms and find that it induced about 36,000 firms with up to 50 workers to change from formal to informal status (about 5 percent of the stock of those firms in 2002). Finally, Galiani and Gertler (2009) evaluate the impact of the non-contributory federal pension program (known as Adults over Seventy). They find that workers close to but below the retirement age (65 years) switched from formal to informal status (while others dropped out of the labor force altogether).

Altogether, these papers indicate that changes in the regulations to non-salaried labor are affecting the allocation of capital and labor in Mexico in the direction of informality.[29] These results are consistent with the trends in resource allocation documented in Section 5. Of course,

[29] Antón, Hernández and Levy (2012), on the other hand, develop a simple general equilibrium model to evaluate the joint effects of the tax on salaried labor and the subsidy to non-salaried (the latter derived from all social programs for informal workers as opposed to the individual programs evaluated above). They find that in the absence of these taxes-cum-subsidies, there would be 26 million salaried workers in Mexico (all legal) and 13 non-salaried; that the tax on salaried labor reduces salaried employment to 21.6 million workers (and changes its composition, with 7.9 salaried workers now hired illegally), while non-salaried employment increases to 17.4 million; and that the subsidy to non-salaried labor further reduces salaried employment to 20 million workers and increases non-salaried employment to 18.9 million. These results indicate that the tax associated with T_s is more important than the subsidy associated with T_{ns}, suggesting that eliminating this tax would bring the largest productivity gains.

there may be other forces that are also inducing a re-allocation of resources towards the informal sector.[30] But that said, we interpret the evidence from this and the papers just discussed as pointing from distortions in the labor market created by asymmetric labor regulations towards informality and illegality and losses of TFP.

7. Conclusion

We study the relation between firm informality, illegality and productivity. In particular, we focus on asymmetries in the regulation of salaried and non-salaried contracts as a key cause of the existence of informality and illegality. The informal sector results from the interactions of firms' demands for salaried and non-salaried contracts, workers implicit valuations of the utility derived from these contracts, and the government's efforts to enforce them. The asymmetry in the regulation of labor is critical and allows us to disentangle informality from illegality and firm size.

Using Mexico's economic census we document two broad sets of stylized facts. The first one is that informality and illegality matter greatly for resource allocation. Two-thirds of all firms captured in the census are informal but legal; they account for 35 percent of all workers and 32 percent of the capital stock; in contrast, formal and legal firms are only 3 percent of all firms, hire 20 percent of workers and account for 30 percent of capital. Illegal behavior is also highly relevant: 23 percent of all firms hire workers illegally and 20 percent of all workers are hired illegally. To this we need to add that at least 1.2 million firms and almost 50 percent of all private sector workers are left out of the census, the vast majority of whom are informal. The second broad stylized fact is that there is a clear trend towards more informality and illegality. Between 1998 and 2008 the number of legal and formal establishments fell by 11.5 percent and the number of informal and legal ones grew by 33 percent, while informal and illegal ones grew by 96 percent. In turn, employment in legal and formal firms grew by 36 percent, in legal and informal firms by 97 percent, and in informal and illegal firms by 121 percent. Employment grew more in smaller than in larger firms, while the opposite happened to value added.

We use the framework developed by Hsieh and Klenow (2009) to measure the distortions in output and factor markets faced by firms. We find that there is a strong correlation between

[30] The trend towards more illegality and informality could result from many factors, not necessarily mutually exclusive. Aside from growth in social programs for informal workers, one can speculate if more intense competition from China has driven firms and workers into informality.

distortions and productivity; that larger firms face more distortions than smaller ones; that even within sectors defined at the 6-digit level there are huge disparities in firms' productivities; that these disparities are larger than for other countries where similar computations are available; and that distortions increased between 1998 and 2008. As a result, we find that misallocation of capital and labor increased: in 1998 one peso of capital and labor allocated to firms in the 90^{th} (75^{th}) percentile was worth 2.78 (1.46) times more than the same peso allocated to firms in the 10^{th} (25^{th}) percentile; in 2008 a peso of resources to firms in the 90^{th} (75^{th}) percentile was worth 2.98 (1.51) times more than the same peso allocated to firms in the 10^{th} (25^{th}) percentile.

We show that formal legal firms face more distortions than informal and illegal firms, who in turn face more distortions than informal and legal firms; this holds regardless of whether we control for firm size and age or not. In turn, we separate the effects of illegality from informality, and show that it is the status of informality and not illegality that is associated with lower productivity. Indeed, we show that with or without controlling for size and age, within each 6-digit sector informal and illegal firms are more productive than informal and legal firms. In parallel, we find that one peso of capital and labor allocated to formal and legal firms is worth 28 percent more than if allocated to illegal and informal firms, and 50 percent more than if allocated to legal and informal firms. All this argues that shifting resources from the informal to the formal sector would increase TFP.

We argue that changes to the regulations that apply to non-salaried labor, that *de facto* increase the subsidy to informality, are compounding the factors that misallocate resources towards the informal sector. Finally, we find support for the proposition that the distortions caused by informality are indeed lowering TFP in Mexico, and that the quantitative effects are large.

Our results are relevant to the debate on the causes and costs of informality in Mexico (and, more generally, Latin America), and on policies to deal with it. We make four observations in this regard. First, in any economy there will always be a need for non-salaried contracts (including self-employment). Clearly, if these contracts were subject to the same regulations as salaried ones, firms and workers would settle on the contract combination that maximized TFP. It follows that if we associate the informal sector with the segment of the economy where salaried contracts are not observed, there will always be an informal sector and, moreover, that its presence will contribute to maximize TFP. From the point of view of productivity the problem

arises when the differences in the regulations on salaried and non-salaried contracts generate idiosyncratic taxes-cum-subsidies; in particular, taxes on salaried contracts and subsidies to non-salaried. In that case, even under perfect enforcement, the informal sector will be larger than optimal; the type and size distribution of firms will be distorted and productivity will suffer. Thus, from the point of view of productivity, the problem is not that there are family firms or self-employed workers; the problem is that there are too many. And because enforcement of the regulations on salaried labor is imperfect and correlated with size, there are also too many mostly small illegal firms with salaried contracts. Excessive informality, not informality *per se*, reduces productivity.[31]

Second, small firm size and informality are sometimes conflated. This is understandable, since for the reasons explained earlier firms intensive in non-salaried contracts tend to be smaller than those intensive in salaried contracts; since the self-employed—the smallest possible firm—is necessarily non-salaried; and since evading firms with salaried contracts tend to be small. That said, our paper suggests that from the point of view of productivity, firm type is more important than firm size. This is relevant to the discussion of policies towards small and medium-size enterprises (SMEs). In an ideal world all firms would face the same policies, allowing the size distribution of firms to reflect the underlying distribution of firms' productivities. But if for whatever reason SMEs are to be promoted, it is key that formality and legality be a precondition for receiving government help. Using public resources to subsidize low-productivity informal firms is hardly a way to promote long-term growth; and while this might "create jobs" it will also ensure that workers get low-productivity jobs with low real wages. This observation is relevant as well with regards to policies towards micro-firms (like subsidizing micro-credits).

Third, our finding that legal informal firms are less productive than illegal informal firms suggests that the problem of informality cannot be solved only by stiffer enforcement of the regulations on salaried labor. As long as the taxes-cum-subsidies derived from asymmetries in the regulation of labor persist, stiffer enforcement might actually reduce productivity and increase informality; this depends on whether as a result of stiffer enforcement resources from evading firms shift towards firms with non-salaried contracts (and self-employment), or towards firms with salaried contracts that do comply with the law. This is all the more relevant given our

[31] In fact, if the asymmetry in labor regulations produced subsidies to salaried labor and taxes to non-salaried, excess formality would lower productivity.

finding that the majority of evading firms have six to 10 workers (as the majority of firms with one to five workers are informal but mostly legal, which we speculate are family firms). Stiffer enforcement may eliminate illegality; it will not end excessive informality.

Fourth, it is important to consider that excessive informality might affect productivity through other channels aside from static resource misallocation. The implicit tax on growth of firms intensive in salaried contracts may be part of the explanation of why over their life-cycle Mexican firms grow less than their peers in the US. Hsieh and Klenow (2012) find that in the manufacturing sector, whereas employment in the average 40-year-old firm in the US grows as much as eight times vis-à-vis the average plant five years or younger, in the Mexican counterpart it only grows two times. They find as well that this difference in life-cycle dynamics could lower aggregate manufacturing productivity in Mexico by 25 percent relative to the US. Further research needs to quantify the extent to which informality and illegality are affecting life-cycle dynamics and firms' investments in labor training, technology adoption and innovation.

To conclude, informality is usually thought of as a social problem, since when workers are informally employed they are not as protected against risks like ill health, unemployment, old-age poverty, death or disability as when they are formally employed; this is a serious concern. Informality is also thought of as a fiscal problem, as informal firms evade contributions and shrink the country's tax base while informal workers receive free social benefits; this is a serious concern as well. The central message of this paper, however, is that the distortions created by informality are also very costly from the point of view of productivity. Even if policymakers are willing to live with the social dimensions of informality and able to tolerate its burden on public finances, they should not treat Mexico's large informal sector with benign neglect unless they are willing as well to pay a high price in terms of productivity and, inevitably, medium-term growth and workers' living standards.

References

Antón, A., and F. Hernández. 2010. "VAT Collections and Social Security Contributions: Is There a Link?" Washington, DC, United States: Inter-American Development Bank.

Antón, A., F. Hernández and S. Levy. 2012. *The End of Informality in Mexico? Fiscal Reform for Universal Social Insurance*. Washington, DC, United States: Inter-American Development Bank.

Bell, L.A. 1997. "The Impact of Minimum Wages in Mexico and Colombia." *Journal of Labor Economics* 15: 103-35.

Bosch, M., and M.B. Cobacho. 2012. "Welfare Policies and Informalization: The Effect of Popular Insurance on Semi-Formal Firms in Mexico." Washington, DC, United States: Inter-American Development Bank. Mimeographed document.

Bosch, M., M.B. Cobacho and C. Pages. 2012. "Taking Stock of Eight Years of Implementation of Seguro Popular in Mexico." Washington, DC, United States: Inter-American Development Bank. Mimeographed document.

Bosch, M., and R. Campos-Vázquez. 2010. "The Trade-Offs of Social Assistance Programs in the Labor Market: The Case of the Seguro Popular Program in Mexico." Washington, DC, United States: Inter-American Development Bank. Mimeographed document.

Busso, M., L. Madrigal and C. Pagés. 2012. "Productivity and Resource Misallocation in Latin America." Working Paper IDB-WP-306. Washington, DC, United States: Inter-American Development Bank.

Calderón-Madrid, A. 2000. "Job Stability and Labor Mobility in Urban Mexico: A Study Based on Duration Models and Transition Analysis." Research Network Working Paper R-419. Washington, DC, United States: Inter-American Development Bank.

----. 2006. "Mobility of Workers between Formal and Informal Job Status: An Empirical Assessment of Earnings Variations and Exit Hazards in Mexico's Urban Labor Market." Mexico, Distrito Federal, Mexico: El Colegio de México. Mimeographed document.

----. 2010. "Employment Protection Legislation in Mexico: Is it Adversely Affecting Workers' Flows in the Manufacturing Sector?" Mexico, Distrito Federal, Mexico: El Colegio de México. Mimeographed document.

Daude, C., and E. Fernández-Arias. 2010. "On the Role of Productivity and Factor Accumulation in Economic Development in Latin America and the Caribbean." Working Paper IDB-WP-155. Washington, DC, United States: Inter-American Development Bank.

De Soto, H. 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. New York, United States; Basic Books.

Duarte, M., and D. Restuccia. 2009. "The Role of the Structural Transformation in Aggregate Productivity." Working Paper 348. Toronto, Canada: University of Toronto, Department of Economics.

Duval, R., and R. Smith. 2010. "Counterfactual Wages and Segmentation between Formal and Informal Sectors in Mexico." Washington, DC, United States: Inter-American Development Bank. Mimeographed document.

Fajnzylber, P., W. Maloney and G. Montes-Rojas. 2009. "Releasing Constraints to Growth or Pushing on a String? Policies and Performance of Mexican Micro-Firms." *Journal of Development Studies* 45(7): 1027-1047.

----. 2011. "Does Formality Improve Micro-firm Performance? Evidence from the Brazilian Simples Program." *Journal of Development Economics* 94: 262-276.

Fajnzylber, P. et al. 2007. Informality: Exit and Exclusion: Building Effective and Legitimate Institutions. Latin American and the Caribbean Studies. Washington, DC, United States: World Bank.

Fazio, M.V., and S. Levy. 2010. "Labor Mobility and Wage Rate Comparisons in Mexico: 2005-2008." Washington, DC, United States: Inter-American Development Bank. Mimeographed document.

Fields, G. 2007. "Labor Market Policy in Developing Countries: A Selective Review of the Literature and Needs for the Future." Policy Research Working Paper 4362. Washington, DC, United States: World Bank.

Galiani, S., and P. Gertler. 2009. "Informe Final sobre el Programa Setenta y Más." Mexico, Distrito Federal, Mexico: Secretaría de Desarollo Social.

Gong, X., A. Soest and E. Villagómez. 2002. "Mobility in the Urban Labour Market: A Panel Data Analysis for Mexico." *Labour Economics* 9(4): 513-529.

Guha-Khasnobis, B., R. Kanbur and E. Ostrom. 2006. "Beyond Formality and Informality." In: B. Guha-Khasnobis, R. Kanbur, and E. Ostrom, editors. *Linking the Formal and Informal Economy: Concepts and Policies*. Oxford, United Kingdom: Oxford University Press.

Hanson, G. 2010. "Why Isn't Mexico Rich?" *Journal of Economic Literature*, December 48(4): 987-1004.

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

Hsieh, C., and P. Klenow. 2012. "The Life-Cycle of Plants in India and Mexico." Chicago, United States: University of Chicago. Mimeographed document.

Inter-American Development Bank. 2010. *The Age of Productivity: Transforming Economies from the Bottom Up.* Development in the Americas Report. Washington, DC, United States: Inter-American Development Bank.

Instituto Nacional de Estadística, Geografía e Informática (INEGI). 1999. *Censos Económicos 1999*. Mexico, Distrito Federal, Mexico: INEGI.

----. 2004. *Censos Económicos 2004*. Mexico, Distrito Federal, Mexico: INEGI.

----. 2008. *Censos Económicos 2008*. Mexico, Distrito Federal, Mexico: INEGI.

Jermanowski, M. 2007. "Total Factor Productivity Differences: Appropriate Technology versus Efficiency." *European Economic Review* 51: 2080-2110.

Juárez, L. 2008. "Are Informal Workers Compensated for the Lack of Fringe Benefits? Free Health Care as an Instrument for Formality." Working Papers 0804. Mexico, Distrito Federal, Mexico: Instituto Tecnológico Autónomo de México.

Kanbur, R. 2009. "Conceptualizing Informality: Regulation and Enforcement." *Indian Journal of Labour Economics* 52(1): 33-42.

La Porta, R., and A. Shleifer. 2008. "The Unofficial Economy and Economic Development." NBER Working Paper 14520. Cambridge, United States: National Bureau of Economic Research.

Leal, J. 2010. "Informal Sector and Economic Development." Tempe, United States: Arizona State University. Ph.D. Dissertation.

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

Maloney, W. 1999. "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico." *World Bank Economic Review* 13: 275-302.

----. 2004. "Informality Revisited." *World Development* 20: 1-20.

Maloney, W., and J. Nuñez Mendez. 2004. "Measuring the Impact of Minimum Wages: Evidence from Latin America." In: J. Heckman and C. Pagés, editors. *Law and Employment: Lessons from Latin America and the Caribbean*. Chicago, United States: University of Chicago Press.

Navarro, S., and P. Schrimpf. 2004. "The Importance of Being Formal: Testing for Segmentation in the Mexican Labor Market." Chicago, United States: University of Chicago, Department of Economics. Mimeographed document.

Restuccia, D. 2008. "The Latin American Development Problem." Working Paper 432. Toronto, Canada: University of Toronto, Department of Economics.

United Nations (UN). 2010. "Economic census: challenges and good practices. A technical report." Mimeographed document.

Valero-Gil, J., and C. Sánchez-Vela. 2010. "Incidencia de la Ley del Impuesto sobre la Renta y la Ley Federal del Trabajo sobre la Economia y la Estructura Industrial de México." Washington, DC, United States: Inter-American Development Bank. Mimeographed document.

Appendix

Tables and Figures

Appendix Table 1. Resource Allocation and Output

1998

	Establishments		Workers		Capital*		Value Added*		Sales*	
	Number	%	Number	%	Number	%	Number	%	Number	%
Manufacturing										
[0-5]	282,861	82	590,263	13	32,392	4	15,971	3	47,290	3
[6-10]	25,709	7	191,380	4	11,040	1	9,489	2	26,109	1
[11-50]	23,617	7	518,208	11	41,546	5	35,758	6	121,151	7
[+50]	11,931	3	3,212,744	71	802,578	90	519,895	89	1,663,923	90
Total	344,118	100	4,512,595	100	887,556	100	581,113	100	1,858,473	100
Retail										
[0-5]	1,362,557	94	2,212,570	55	66,695	34	97,416	28	391,818	21
[6-10]	41,909	3	313,460	8	17,028	9	39,520	11	209,906	11
[11-50]	33,221	2	675,712	17	38,397	20	95,381	27	597,624	32
[+50]	5,989	0	797,548	20	73,015	37	118,941	34	696,940	37
Total	1,443,676	100	3,999,290	100	195,134	100	351,258	100	1,896,288	100
Services										
[0-5]	834,316	89	1,550,103	36	157,324	35	69,245	25	141,565	24
[6-10]	55,878	6	415,911	10	41,233	9	24,214	9	52,961	9
[11-50]	40,481	4	826,889	19	84,030	19	57,713	21	124,611	21
[+50]	7,897	1	1,472,757	35	167,533	37	129,782	46	262,326	45
Total	938,572	100	4,265,660	100	450,121	100	280,954	100	581,463	100
Aggregate										
[0-5]	2,479,734	91	4,352,936	34	256,411	17	182,632	15	580,673	13
[6-10]	123,496	5	920,751	7	69,301	5	73,224	6	288,976	7
[11-50]	97,319	4	2,020,809	16	163,973	11	188,852	16	843,387	19
[+50]	25,817	1	5,483,049	43	1,043,127	68	768,618	63	2,623,189	60
Total	2,726,366	100	12,777,545	100	1,532,811	100	1,213,325	100	4,336,224	100

Note: * Millions of nominal pesos..

Appendix Table 2. Resource Allocation and Output

2003

	Establishments		Workers		Capital*		Value Added*		Sales*	
	Number	%	Number	%	Number	%	Number	%	Number	%
Manufacturing										
[0-5]	274,222	83	581,613	14	32,311	3	23,911	3	62,316	2
[6-10]	24,456	7	180,490	4	13,703	1	12,239	1	30,376	1
[11-50]	19,754	6	431,768	10	58,614	5	50,797	5	156,507	6
[+50]	10,286	3	3,004,708	72	1,153,807	92	841,041	91	2,574,561	91
Total	328,718	100	4,198,579	100	1,258,435	100	927,987	100	2,823,761	100
Retail										
[0-5]	1,475,867	93	2,677,714	54	106,708	33	142,911	25	546,400	18
[6-10]	57,998	4	423,227	8	28,348	9	64,200	11	318,123	10
[11-50]	39,000	2	787,759	16	65,894	20	154,055	27	934,379	31
[+50]	7,722	0	1,108,666	22	125,231	38	219,027	38	1,238,033	41
Total	1,580,587	100	4,997,366	100	326,180	100	580,194	100	3,036,934	100
Services										
[0-5]	889,302	89	1,712,914	35	170,164	23	93,259	15	179,250	17
[6-10]	61,610	6	456,170	9	51,873	7	34,789	6	65,916	6
[11-50]	42,916	4	880,783	18	100,274	14	93,030	15	187,083	17
[+50]	9,498	1	1,890,111	38	414,634	56	398,799	64	651,406	60
Total	1,003,326	100	4,939,978	100	736,945	100	619,877	100	1,083,656	100
Aggregate										
[0-5]	2,639,391	91	4,972,241	35	309,183	13	260,081	12	787,966	11
[6-10]	144,064	5	1,059,887	7	93,924	4	111,228	5	414,415	6
[11-50]	101,670	3	2,100,310	15	224,782	10	297,881	14	1,277,969	18
[+50]	27,506	1	6,003,485	42	1,693,672	73	1,458,868	69	4,464,000	64
Total	2,912,631	100	14,135,923	100	2,321,560	100	2,128,058	100	6,944,350	100

Note: * Millions of nominal pesos.

Appendix Table 3. Resource Allocation and Output

2008

	Establishments		Workers		Capital*		Value Added*		Sales*	
	Number	%	Number	%	Number	%	Number	%	Number	%
Manufacturing										
[0-5]	367,744	84	814,332	17	43,812	3	26,391	2	71,739	1
[6-10]	36,412	8	266,381	6	21,100	1	15,762	1	42,520	1
[11-50]	22,349	5	467,197	10	69,840	4	63,571	4	200,865	4
[+50]	10,346	2	3,113,152	67	1,452,239	92	1,375,098	93	4,650,789	94
Total	436,851	100	4,661,062	100	1,586,991	100	1,480,821	100	4,965,912	100
Retail										
[0-5]	1,724,815	93	3,422,720	56	168,572	30	133,846	21	741,677	13
[6-10]	78,984	4	570,660	9	54,555	10	64,827	10	451,701	8
[11-50]	45,471	2	908,634	15	115,217	20	170,614	27	1,337,883	24
[+50]	9,280	0	1,232,744	20	228,868	40	259,017	41	3,003,826	54
Total	1,858,550	100	6,134,758	100	567,212	100	628,303	100	5,535,086	100
Services										
[0-5]	1,177,533	87	2,435,752	36	222,382	20	147,838	17	303,432	17
[6-10]	97,093	7	716,903	10	70,726	6	58,246	7	110,856	6
[11-50]	62,417	5	1,251,220	18	151,522	13	139,797	16	285,037	16
[+50]	11,538	1	2,455,784	36	688,721	61	535,192	61	1,047,234	60
Total	1,348,581	100	6,859,659	100	1,133,351	100	881,073	100	1,746,559	100
Aggregate										
[0-5]	3,270,092	90	6,672,804	38	434,766	13	308,075	10	1,116,847	9
[6-10]	212,489	6	1,553,944	9	146,380	4	138,834	5	605,076	5
[11-50]	130,237	4	2,627,051	15	336,579	10	373,982	13	1,823,785	15
[+50]	31,164	1	6,801,680	39	2,369,829	72	2,169,306	73	8,701,849	71
Total	3,643,982	100	17,655,479	100	3,287,554	100	2,990,197	100	12,247,557	100

Note: * Millions of nominal pesos.

Appendix Table 4: Resource Allocation and Output

By Legality and Formality Status

	Establishments		Workers		Value Added		Capital	
	Number	%	Number	%	Number	%	Number	%
	125,940	5	2,641,450	21	315,959	26	407,367	27
	1,846,254	68	3,155,683	25	146,181	12	248,171	16
	69,059	3	434,237	3	29,998	2	42,751	3
nal	264,747	10	5,017,026	39	610,481	50	613,840	40
	420,366	15	1,529,149	12	110,707	9	220,682	14
	2,726,366	100	12,777,545	100	1,213,325	100	1,532,811	100
	118,575	4	4,216,830	30	854,259	40	844,997	36
	1,994,048	68	4,065,636	29	298,772	14	469,131	20
	68,501	2	543,559	4	61,319	3	72,374	3
nal	179,566	6	3,184,285	23	747,854	35	794,853	34
	551,941	19	2,125,613	15	165,854	8	140,206	6
	2,912,631	100	14,135,923	100	2,128,058	100	2,321,560	100
	111,503	3	3,597,936	20	918,783	31	955,836	29
	2,458,733	67	6,241,996	35	617,143	21	1,046,814	32
	61,786	2	539,033	3	57,743	2	77,840	2
nal	186,418	5	3,896,905	22	1,194,194	40	1,000,728	30
	825,542	23	3,379,609	19	202,335	7	206,336	6
	3,643,982	100	17,655,479	100	2,990,197	100	3,287,554	100

Appendix Table 5. Resource Allocation and Output

By Size and Legality/Formality Status, 1998

		Legal & formal	Legal & informal	Legal & semi-formal	semi-legal & semi-formal	Illegal and informal
Establishments	[0-5]	74,007	1,825,223	54,728	142,892	382,884
	[6-10]	22,022	13,826	9,291	52,561	25,796
	[11-50]	22,140	5,884	4,361	54,427	10,507
	[+50]	7,771	1,321	679	14,867	1,179
	Total	125,940	1,846,254	69,059	264,747	420,366
Workers	[0-5]	197,687	2,667,988	160,989	431,335	894,937
	[6-10]	166,916	101,066	67,965	397,926	186,878
	[11-50]	471,382	114,062	84,176	1,156,676	194,513
	[+50]	1,805,465	272,567	121,107	3,031,089	252,821
	Total	2,641,450	3,155,683	434,237	5,017,026	1,529,149
Capital	[0-5]	18,795	149,031	9,149	38,338	41,098
	[6-10]	11,372	9,711	4,670	33,772	9,775
	[11-50]	33,016	14,802	4,410	96,527	15,218
	[+50]	344,184	74,627	24,522	445,203	154,590
	Total	407,367	248,171	42,751	613,840	220,682
Value Added	[0-5]	17,223	81,235	6,994	40,041	37,139
	[6-10]	13,178	12,142	3,518	36,104	8,282
	[11-50]	37,298	12,174	4,904	121,302	13,175
	[+50]	248,260	40,631	14,582	413,034	52,111
	Total	315,959	146,181	29,998	610,481	110,707
Sales	[0-5]	60,140	248,488	23,467	136,584	111,993
	[6-10]	53,613	37,420	13,967	150,433	33,543
	[11-50]	164,666	68,165	24,475	536,501	49,578
	[+50]	801,677	223,040	53,339	1,304,178	240,955
	Total	1,080,097	577,113	115,249	2,127,696	436,069

Appendix Table 6. Resource Allocation and Output

By Size and Legality/Formality Status, 2003

		Legal & formal	Legal & informal	Legal & semi-formal	semi-legal & semi-formal	Illegal and informal
Establishments	[0-5]	50,234	1,959,846	48,786	96,237	484,288
	[6-10]	22,573	22,285	12,384	39,713	47,109
	[11-50]	31,803	9,317	6,300	35,260	18,990
	[+50]	13,965	2,600	1,031	8,356	1,554
	Total	118,575	1,994,048	68,501	179,566	551,941
Workers	[0-5]	143,598	3,158,314	153,374	302,883	1,214,072
	[6-10]	172,901	156,267	90,763	298,461	341,495
	[11-50]	709,741	198,507	120,065	724,412	347,585
	[+50]	3,190,590	552,548	179,357	1,858,529	222,461
	Total	4,216,830	4,065,636	543,559	3,184,285	2,125,613
Capital	[0-5]	16,828	178,179	13,202	30,073	70,901
	[6-10]	18,092	17,741	8,181	26,802	23,107
	[11-50]	73,609	46,122	11,789	70,406	22,855
	[+50]	736,468	227,089	39,201	667,572	23,343
	Total	844,997	469,131	72,374	794,853	140,206
Value Added	[0-5]	21,526	114,041	9,915	35,928	78,671
	[6-10]	24,924	13,936	7,467	37,626	27,275
	[11-50]	107,057	41,330	12,167	102,949	34,378
	[+50]	700,752	129,465	31,770	571,350	25,530
	Total	854,259	298,772	61,319	747,854	165,854
Sales	[0-5]	84,341	344,223	32,137	111,305	215,960
	[6-10]	109,295	69,759	28,945	125,800	80,616
	[11-50]	495,172	196,003	60,624	405,542	120,628
	[+50]	2,085,448	582,346	124,075	1,577,531	94,600
	Total	2,774,256	1,192,332	245,780	2,220,178	511,804

Appendix Table 7. Resource Allocation and Output

By Size and Legality/Formality Status, 2008

		Legal & formal	Legal & informal	Legal & semi-formal	semi-legal & semi-formal	Illegal and informal
Establishments	[0-5]	44,246	2,390,623	43,121	90,708	701,394
	[6-10]	25,299	40,607	12,201	44,909	89,473
	[11-50]	31,550	20,446	5,377	40,019	32,845
	[+50]	10,408	7,057	1,087	10,782	1,830
	Total	111,503	2,458,733	61,786	186,418	825,542
Workers	[0-5]	136,480	4,166,755	139,130	298,764	1,931,675
	[6-10]	193,202	287,076	89,146	338,119	646,401
	[11-50]	678,029	421,333	101,036	832,846	593,807
	[+50]	2,590,225	1,366,832	209,721	2,427,176	207,726
	Total	3,597,936	6,241,996	539,033	3,896,905	3,379,609
Capital	[0-5]	24,155	235,583	20,935	43,799	110,294
	[6-10]	27,859	25,641	9,497	40,504	42,879
	[11-50]	104,457	70,670	11,398	108,135	41,919
	[+50]	799,365	714,920	36,010	808,291	11,243
	Total	955,836	1,046,814	77,840	1,000,728	206,336
Value Added	[0-5]	37,096	140,769	7,650	31,903	90,656
	[6-10]	28,308	23,309	5,887	41,075	40,255
	[11-50]	110,381	64,805	9,766	139,541	49,489
	[+50]	742,997	388,259	34,440	981,674	21,935
	Total	918,783	617,143	57,743	1,194,194	202,335
Sales	[0-5]	125,717	524,401	32,198	117,186	317,345
	[6-10]	139,768	151,189	26,573	152,957	134,590
	[11-50]	555,998	441,511	56,111	606,732	163,433
	[+50]	2,578,193	2,138,196	146,297	3,789,320	49,842
	Total	3,399,675	3,255,297	261,179	4,666,195	665,211

Appendix Table 8: Correlation Matrix

1998

Manufactures	$Log\left[\frac{TFPQ_{si}}{\overline{TFPQ_s}}\right]$	$Log\left[\frac{TFPR_{si}}{\overline{TFPR_s}}\right]$	Formality Index	Legality Index	Firm Size (# Workers)	$Log\left[\frac{1-\tau_{Qsi}}{\overline{1-\tau_{Qs}}}\right]$
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.7394	1				
Formality Index	0.2669	0.1087	1			
Legality Index	0.1816	0.0616	0.9325	1		
Firm Size (# Workers)	0.2136	0.0078	0.0666	0.0536	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.4023	-0.7688	-0.0255	0.0016	0.0331	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.5855	0.4585	0.1488	0.0902	0.0592	-0.3811
Retail & Wholesale						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.8483	1				
Formality Index	0.1163	0.0418	1			
Legality Index	0.0579	0.0044	0.9052	1		
Firm Size (# Workers)	0.26	0.0307	0.0422	0.012	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.6673	-0.8901	-0.0146	0.0058	0.0684	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.6597	0.5622	0.0883	0.0404	0.0166	-0.4447
Services						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.7853	1				
Formality Index	0.1904	0.1098	1			
Legality Index	0.0811	0.0254	0.8924	1		
Firm Size (# Workers)	0.126	0.0363	0.0349	0.0192	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.3723	-0.7248	-0.0184	0.0155	-0.0367	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.6962	0.6084	0.0862	0.0058	0.0284	-0.3574

Appendix Table 9: Correlation Matrix

2003

Manufactures	$Log\left[\frac{TFPQ_{si}}{\overline{TFPQ_s}}\right]$	$Log\left[\frac{TFPR_{si}}{\overline{TFPR_s}}\right]$	Formality Index	Legality Index	Firm Size (# Workers)	$Log\left[\frac{1-\tau_{Qsi}}{\overline{1-\tau_{Qs}}}\right]$
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.697	1				
Formality Index	0.3208	0.1128	1			
Legality Index	0.2463	0.0747	0.9527	1		
Firm Size (# Workers)	0.397	0.0636	0.0784	0.043	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.2847	-0.7049	-0.0714	-0.0492	0.0184	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.5415	0.3559	0.2051	0.1417	0.0818	-0.3046
Retail & Wholesale						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.8261	1				
Formality Index	0.274	0.1005	1			
Legality Index	0.2035	0.0537	0.9293	1		
Firm Size (# Workers)	0.2878	0.0486	0.2892	0.2287	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.5977	-0.8793	0.0415	0.0593	0.1015	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.6686	0.5587	0.135	0.0788	0.0887	-0.3913
Services						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.7567	1				
Formality Index	0.1802	0.08	1			
Legality Index	0.0992	0.0189	0.9132	1		
Firm Size (# Workers)	0.1503	0.0238	-0.0032	-0.0146	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.5339	-0.8244	-0.0265	0.0138	-0.0313	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.5446	0.4257	0.0239	-0.042	0.0364	-0.2627

Appendix Table 10: Correlation Matrix

2008

Manufactures	$Log\left[\frac{TFPQ_{si}}{\overline{TFPQ_s}}\right]$	$Log\left[\frac{TFPR_{si}}{\overline{TFPR_s}}\right]$	Formality Index	Legality Index	Firm Size (# Workers)	$Log\left[\frac{1-\tau_{Qsi}}{\overline{1-\tau_{Qs}}}\right]$
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.6954	1				
Formality Index	0.3009	0.0862	1			
Legality Index	0.2284	0.0509	0.9438	1		
Firm Size (# Workers)	0.3889	0.0425	0.0922	0.0536	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.4039	-0.8165	0.0021	0.0207	0.0124	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.6599	0.4288	0.2708	0.2104	0.1739	-0.3159
Retail & Wholesale						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.8257	1				
Formality Index	0.2597	0.1477	1			
Legality Index	0.2084	0.1067	0.9442	1		
Firm Size (# Workers)	0.3172	0.0353	0.1303	0.1093	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.5192	-0.8438	-0.0093	0.0061	0.1882	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.7503	0.5992	0.241	0.1908	0.1579	-0.372
Services						
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1					
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.7415	1				
Formality Index	0.3864	0.1744	1			
Legality Index	0.3258	0.1331	0.9367	1		
Firm Size (# Workers)	0.3788	0.1224	0.2542	0.2177	1	
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	-0.3175	-0.6834	0.0959	0.0961	0.1014	1
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	0.618	0.4698	0.1681	0.1274	0.1043	-0.3278

Appendix Table A11: Dispersion of Productivity and Distorsions

6 dígits, all establishments

	1998			2003			2008		
	Std. Dev.	p75-p25	p90-p10	Std. Dev.	p75-p25	p90-p10	Std. Dev.	p75-p25	p90-p10
Agg.Economy									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.78	2.50	4.64	1.75	2.38	4.49	1.89	2.64	4.90
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.09	1.46	2.78	1.08	1.41	2.74	1.18	1.51	2.98
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.33	1.79	3.39	1.40	1.87	3.61	1.49	1.96	3.80
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.29	1.81	3.38	1.41	2.01	3.69	1.47	2.08	3.88
Manufacturing									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.81	2.54	4.65	1.74	2.37	4.41	1.89	2.66	4.89
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.00	1.31	2.52	0.95	1.23	2.40	1.11	1.43	2.78
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.23	1.62	3.09	1.22	1.60	3.05	1.34	1.74	3.36
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.27	1.79	3.28	1.34	1.88	3.47	1.41	1.91	3.68
Retail									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.78	2.49	4.67	1.83	2.47	4.76	1.93	2.64	5.03
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.22	1.68	3.17	1.25	1.69	3.23	1.33	1.77	3.43
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.42	1.95	3.68	1.56	2.13	4.06	1.69	2.21	4.44
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.27	1.77	3.41	1.42	2.01	3.80	1.53	2.10	4.12
Services									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.74	2.44	4.56	1.70	2.31	4.37	1.86	2.63	4.80
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.10	1.49	2.83	1.10	1.40	2.80	1.18	1.46	3.00
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.42	1.96	3.67	1.54	2.03	4.00	1.61	2.14	4.10
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.34	1.93	3.55	1.49	2.21	3.94	1.53	2.33	4.03

Appendix Table A12: Dispersion of Productivity and Distorsions

6 dígits, only establishments with more than 10 employees

	1998			2003			2008		
	Std. Dev.	p75-p25	p90-10	Std. Dev.	p75-p25	p90-10	Std. Dev.	p75-p25	p90-10
Agg.Economy									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.46	1.97	3.74	1.45	1.98	3.68	1.63	2.23	4.14
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.98	1.28	2.47	0.98	1.31	2.46	1.07	1.30	2.70
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.26	1.65	3.18	1.29	1.72	3.24	1.37	1.69	3.47
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.30	1.79	3.35	1.32	1.78	3.35	1.45	1.95	3.70
Manufacturing									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.40	1.96	3.60	1.32	1.78	3.34	1.57	2.15	4.01
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.90	1.18	2.27	0.85	1.12	2.11	0.98	1.22	2.40
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.18	1.56	2.98	1.11	1.48	2.79	1.22	1.55	3.00
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.34	1.89	3.43	1.28	1.77	3.30	1.35	1.74	3.35
Retail									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.53	1.98	3.90	1.54	2.02	3.93	1.63	2.08	4.08
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.11	1.45	2.80	1.13	1.43	2.84	1.18	1.41	2.96
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.29	1.68	3.23	1.35	1.71	3.38	1.46	1.69	3.71
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.14	1.46	2.90	1.22	1.53	3.06	1.35	1.60	3.45
Services									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.49	2.00	3.83	1.56	2.23	3.97	1.75	2.49	4.41
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.98	1.27	2.48	1.04	1.47	2.61	1.17	1.36	3.01
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.38	1.81	3.52	1.48	2.10	3.79	1.57	1.92	4.07
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.43	1.99	3.71	1.45	2.02	3.68	1.69	2.55	4.45

Appendix Table A13: Dispersion of Productivity and Distorsions

4 dígits, all establishments

	1998			2003			2008		
	Std. Dev.	p75-p25	p90-10	Std. Dev.	p75-p25	p90-10	Std. Dev.	p75-p25	p90-10
Agg.Economy									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.86	2.60	4.89	1.83	2.56	4.76	1.96	2.73	5.12
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.15	1.56	2.97	1.14	1.49	2.92	1.25	1.66	3.20
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.39	1.89	3.58	1.47	1.97	3.81	1.56	2.11	4.01
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.31	1.80	3.47	1.45	2.08	3.84	1.52	2.18	4.02
Manufacturing									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.89	2.63	4.96	1.84	2.59	4.76	1.96	2.74	5.09
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.08	1.42	2.75	1.03	1.34	2.62	1.18	1.55	3.01
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.30	1.70	3.28	1.31	1.72	3.32	1.41	1.86	3.61
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.29	1.74	3.35	1.40	1.93	3.68	1.48	2.09	3.90
Retail									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.93	2.76	5.11	1.91	2.65	5.04	2.09	2.93	5.56
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.31	1.85	3.44	1.30	1.77	3.40	1.42	1.93	3.66
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.52	2.13	3.96	1.62	2.21	4.25	1.77	2.37	4.56
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.31	1.82	3.58	1.47	2.09	3.96	1.62	2.37	4.37
Services									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.72	2.36	4.49	1.74	2.42	4.50	1.88	2.58	4.87
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.09	1.49	2.83	1.14	1.46	2.92	1.24	1.66	3.18
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.44	1.97	3.72	1.57	2.10	4.12	1.67	2.33	4.30
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.36	1.91	3.60	1.51	2.28	3.97	1.51	2.21	3.99

Appendix Table A14: Dispersion of Productivity and Distorsions
4 dígitos, only establishments with more than 10 employees

	1998			2003			2008		
	Std. Dev.	p75-p25	p90-10	Std. Dev.	p75-p25	p90-10	Std. Dev.	p75-p25	p90-10
Agg.Economy									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.49	2.02	3.84	1.51	2.05	3.89	1.68	2.35	4.38
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.00	1.30	2.51	1.05	1.37	2.65	1.14	1.47	2.93
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.28	1.67	3.24	1.36	1.82	3.42	1.47	1.94	3.81
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.31	1.72	3.38	1.38	1.89	3.55	1.52	2.16	3.96
Manufacturing									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.45	2.04	3.78	1.40	1.93	3.64	1.60	2.23	4.19
$Log[TFPR_{si}/\overline{TFPR_s}]$	0.92	1.21	2.29	0.89	1.15	2.25	1.03	1.34	2.63
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.20	1.57	3.04	1.17	1.55	2.97	1.29	1.71	3.29
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.33	1.75	3.45	1.34	1.82	3.49	1.44	2.00	3.72
Retail									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.53	2.03	3.91	1.57	2.10	4.06	1.76	2.36	4.52
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.11	1.46	2.82	1.15	1.46	2.91	1.28	1.60	3.30
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.29	1.68	3.27	1.38	1.74	3.47	1.55	1.90	4.06
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.15	1.45	2.96	1.24	1.58	3.05	1.41	1.73	3.64
Services									
$Log[TFPQ_{si}/\overline{TFPQ_s}]$	1.51	1.99	3.86	1.63	2.19	4.11	1.78	2.52	4.58
$Log[TFPR_{si}/\overline{TFPR_s}]$	1.01	1.30	2.55	1.20	1.63	3.02	1.21	1.58	3.19
$Log[1-\tau_{Qsi}/\overline{1-\tau_{Qs}}]$	1.43	1.87	3.63	1.63	2.29	4.05	1.71	2.35	4.49
$Log[1+\tau_{Lsi}/\overline{1+\tau_{Ls}}]$	1.46	1.97	3.77	1.56	2.28	4.10	1.73	2.74	4.58

Appendix Table 15. TFP Gains

	1998	**2003**	**2008**
6 dígits, All establishments			
Manufacturing	118.9	114.8	133.4
Retail commerce	338.6	294.3	335.2
Services	248.2	202.8	213.8
Aggregate Economy	212.5	189.4	199.5
6 dígits, establishments > 10			
Manufacturing	138.0	122.1	136.5
Retail commerce	407.5	377.0	356.4
Services	248.5	193.7	209.8
Aggregate Economy	241.6	212.5	204.3
4 dígits, All establishments			
Manufacturing	157.5	157.9	175.6
Retail commerce	374.5	273.4	375.0
Services	297.2	207.2	240.5
Aggregate Economy	252.7	203.7	236.6
4 dígits, establishments > 10			
Manufacturing	174.8	175.3	198.3
Retail commerce	400.1	426.1	487.5
Services	312.1	132.3	275.1
Aggregate Economy	271.8	231.2	281.7

Appendix Table A16. Correlates of TFPQ

	All firms				Size [0-5]	Size [6-10]	Size [11-50]	Size [+50]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Legal-semiformal	-1.616	-0.619	-1.551	-0.627	-0.959	-0.910	-0.746	-0.574
	[0.0081]***	[0.0071]***	[0.008]***	[0.0071]***	[0.0101]***	[0.0215]***	[0.0299]***	[0.0711]***
Legal-informal	-2.426	-1.056	-2.232	-1.019	-1.759	-0.688	-0.395	-0.491
	[0.0043]***	[0.0041]***	[0.0043]***	[0.0041]***	[0.0068]***	[0.0194]***	[0.0195]***	[0.0309]***
Semilegal-semiformal	-0.211	0.061	-0.204	0.053	-0.214	-0.085	-0.092	0.159
	[0.0046]***	[0.004]***	[0.0045]***	[0.0039]***	[0.0081]***	[0.0146]***	[0.0149]***	[0.0267]***
Illegal-informal	-2.089	-0.497	-1.843	-0.440	-0.881	-0.632	-0.701	-1.039
	[0.0045]***	[0.0043]***	[0.0045]***	[0.0043]***	[0.007]***	[0.0142]***	[0.0177]***	[0.0826]***
Size [6-10]		0.907		0.865				
		[0.0043]***		[0.0043]***				
Size [11-50]		1.700		1.626				
		[0.0038]***		[0.0038]***				
Size [+50]		3.257		3.152				
		[0.0038]***		[0.0038]***				
5-10 years			0.602	0.380	0.328	0.139	0.223	0.888
			[0.0041]***	[0.0036]***	[0.0039]***	[0.0133]***	[0.0174]***	[0.0385]***
+10 years			0.929	0.435	0.296	0.178	0.230	1.067
			[0.0034]***	[0.003]***	[0.0034]***	[0.0112]***	[0.0142]***	[0.03]***
Constant	-4.420	-6.534	-5.070	-6.778	-6.182	-5.656	-4.995	-4.260
	[0.0033]***	[0.004]***	[0.004]***	[0.0043]***	[0.0067]***	[0.0137]***	[0.015]***	[0.0293]***
Obs	2,231,762	2,231,762	2,231,762	2,231,762	1,964,980	154,394	92,401	19,987
R-squared	0.188	0.393	0.215	0.399	0.084	0.031	0.029	0.098

Appendix Table A17. Correlates of TFPR

	All firms				Size [0-5]	Size [6-10]	Size [11-50]	Size [+50]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Legal-semiformal	-0.433	-0.360	-0.431	-0.368	-0.543	-0.569	-0.438	-0.099
	[0.0045]***	[0.0046]***	[0.0045]***	[0.0046]***	[0.0068]***	[0.0144]***	[0.0191]***	[0.0405]***
Legal-informal	-0.506	-0.378	-0.480	-0.372	-0.606	-0.324	-0.317	-0.164
	[0.0024]***	[0.0027]***	[0.0024]***	[0.0027]***	[0.0046]***	[0.013]***	[0.0125]***	[0.0176]***
Semilegal-semiformal	0.020	0.039	0.018	0.034	-0.122	-0.048	-0.053	0.134
	[0.0025]***	[0.0025]***	[0.0025]***	[0.0025]***	[0.0055]***	[0.0098]***	[0.0095]***	[0.0152]***
Illegal-informal	-0.284	-0.161	-0.252	-0.151	-0.300	-0.306	-0.260	-0.127
	[0.0025]***	[0.0028]***	[0.0026]***	[0.0028]***	[0.0047]***	[0.0095]***	[0.0113]***	[0.047]***
Size [6-10]		0.100		0.087				
		[0.0028]***		[0.0028]***				
Size [11-50]		0.275		0.254				
		[0.0024]***		[0.0024]***				
Size [+50]		0.204		0.179				
		[0.0024]***		[0.0025]***				
5-10 years			0.210	0.189	0.160	0.068	0.119	0.452
			[0.0023]***	[0.0023]***	[0.0026]***	[0.0089]***	[0.0111]***	[0.0219]***
+10 years			0.145	0.111	0.050	0.020	0.069	0.345
			[0.0019]***	[0.002]***	[0.0023]***	[0.0075]***	[0.0091]***	[0.0171]***
Constant	0.090	-0.102	-0.030	-0.179	0.029	0.062	0.156	-0.283
	[0.0019]***	[0.0026]***	[0.0023]***	[0.0028]***	[0.0045]***	[0.0092]***	[0.0096]***	[0.0167]***
Obs	2,231,762	2,231,762	2,231,762	2,231,762	1,964,980	154,394	92,401	19,987
R-squared	0.032	0.038	0.036	0.041	0.022	0.018	0.016	0.042

Appendix Table A18. Correlates of TFPQ
(Only legal firms. Excluded category: formal)

	All firms	Size [0-5]	Size [6-10]	Size [11-50]	Size [+50]
	(1)	(2)	(3)	(4)	(5)
Informal	-0.842	-1.743	-0.684	-0.395	-0.495
	[0.0041]***	[0.0064]***	[0.0205]***	[0.0218]***	[0.0329]***
Semiformal	-0.569	-0.961	-0.908	-0.748	-0.585
	[0.0066]***	[0.0094]***	[0.0228]***	[0.0335]***	[0.0754]***
Size [6-10]	1.156				
	[0.0065]***				
Size [11-50]	1.970				
	[0.005]***				
Size [+50]	3.336				
	[0.0045]***				
5-10 years	0.517	0.419	0.203	0.255	0.987
	[0.0046]***	[0.0048]***	[0.0251]***	[0.0291]***	[0.0488]***
+10 years	0.559	0.418	0.207	0.232	1.137
	[0.0038]***	[0.0042]***	[0.0201]***	[0.0234]***	[0.038]***
Constant	-7.103	-6.252	-5.685	-5.002	-4.318
	[0.0048]***	[0.0065]***	[0.0185]***	[0.0212]***	[0.0347]***
Obs	1,463,785	1,365,881	48,858	37,165	11,881
R-squared	0.440	0.069	0.045	0.022	0.092

Appendix Table A19. Correlates of TFPR
(Only legal firms. Excluded category: formal)

	All firms	Size [0-5]	Size [6-10]	Size [11-50]	Size [+50]
	(1)	(2)	(3)	(4)	(5)
Informal	-0.324	-0.592	-0.326	-0.315	-0.166
	[0.0027]***	[0.0043]***	[0.0137]***	[0.0139]***	[0.0189]***
Semiformal	-0.363	-0.545	-0.572	-0.435	-0.114
	[0.0043]***	[0.0063]***	[0.0152]***	[0.0213]***	[0.0432]***
Size [6-10]	0.177				
	[0.0042]***				
Size [11-50]	0.334				
	[0.0032]***				
Size [+50]	0.190				
	[0.0029]***				
5-10 years	0.291	0.235	0.076	0.135	0.594
	[0.003]***	[0.0033]***	[0.0168]***	[0.0185]***	[0.028]***
+10 years	0.210	0.152	-0.010	0.120	0.462
	[0.0024]***	[0.0028]***	[0.0134]	[0.0149]***	[0.0218]***
Constant	-0.297	-0.029	0.076	0.123	-0.377
	[0.0031]***	[0.0043]***	[0.0124]***	[0.0135]***	[0.0199]***
Obs	1,463,785	1,365,881	48,858	37,165	11,881
R-squared	0.053	0.019	0.032	0.023	0.051

Appendix Table A20. Correlates of TFPQ
(Only illegal firms. Excluded category: formal)

	All firms	Size [0-5]	Size [6-10]	Size [11-50]	Size [+50]
	(1)	(2)	(3)	(4)	(5)
Illegal & Informal	0.592	0.878	0.072	-0.315	-0.607
	[0.0031]***	[0.0031]***	[0.0164]***	[0.0219]***	[0.091]***
Size [6-10]	0.970				
	[0.0049]***				
Size [11-50]	1.852				
	[0.0046]***				
Size [+50]	3.489				
	[0.005]***				
5-10 years	0.367	0.366	0.079	0.407	0.594
	[0.0039]***	[0.004]***	[0.0177]***	[0.0298]***	[0.0704]***
+10 years	0.351	0.336	0.275	0.198	0.535
	[0.0034]***	[0.0036]***	[0.0151]***	[0.0245]***	[0.0628]***
Constant	-7.847	-7.960	-6.379	-5.408	-4.408
	[0.0027]***	[0.0025]***	[0.0169]***	[0.0235]***	[0.0522]***
Obs	1,956,266	1,829,601	88,752	32,548	5,365
R-squared	0.2553	0.0474	0.0039	0.0131	0.026

Appendix Table A21. Correlates of TFPR
(Only illegal firms. Excluded category: formal)

	All firms	Size [0-5]	Size [6-10]	Size [11-50]	Size [+50]
	(1)	(2)	(3)	(4)	(5)
Illegal & Informal	0.237	0.307	0.028	0.056	0.004
	[0.0021]***	[0.0021]***	[0.0111]***	[0.014]***	[0.0523]
Size [6-10]	0.080				
	[0.0032]***				
Size [11-50]	0.278				
	[0.003]***				
Size [+50]	0.299				
	[0.0033]***				
5-10 years	0.187	0.176	0.053	0.201	0.342
	[0.0025]***	[0.0027]***	[0.012]***	[0.0191]***	[0.0405]***
+10 years	0.082	0.085	0.085	0.076	0.055
	[0.0022]***	[0.0024]***	[0.0102]***	[0.0157]***	[0.0361]*
Constant	-0.563	-0.590	-0.291	-0.180	-0.273
	[0.0017]***	[0.0017]***	[0.0114]***	[0.015]***	[0.03]***
Obs	1,956,266	1,829,601	88,752	32,548	5,365
R-squared	0.0183	0.0139	0.0008	0.0038	0.0165

Appendix Table 22. Multi-plant Firms and Firm Size

(All Sectors, 2008)

Firm size	Legal			Semi-Legal	Illegal	Sum
	Informal	Semi-formal	Formal	Semi-formal	Informal	
Firms with 1 establishment						
[0-2]	1,988,257	13,131	11,853	23,122	331,185	2,367,548
[3-5]	386,637	29,612	22,747	60,134	369,033	868,163
[6-10]	31,090	11,940	19,860	40,528	89,032	192,450
[11-20]	8,580	3,529	13,892	20,853	24,081	70,935
[21-30]	1,793	761	5,120	6,649	5,194	19,517
[31-50]	1,136	506	4,125	5,133	3,119	14,019
[51-100]	912	301	3,119	3,390	1,337	9,059
[101-250]	659	211	1,915	1,699	224	4,708
[251-500]	286	72	702	631	57	1,748
[501-1000]	132	20	341	295	18	806
[+1000]	58	13	150	209	7	437
Total	2,419,540	60,096	83,824	162,643	823,287	3,549,390
Firms with 2 to 5 establishments						
[0-2]	898	60	1,169	883	161	3,171
[3-5]	540	154	1,305	1,118	153	3,270
[6-10]	516	122	1,186	1,079	91	2,994
[11-20]	485	117	1,247	1,273	74	3,196
[21-30]	302	74	684	611	34	1,705
[31-50]	337	73	743	618	31	1,802
[51-100]	397	97	734	694	19	1,941
[101-250]	355	69	673	586	16	1,699
[251-500]	168	31	278	230	6	713
[501-1000]	98	19	170	171		458
[+1000]	41	6	109	101	1	258
Total	4,137	822	8,298	7,364	586	21,207
Firms with 5+ establishments						
[0-2]	6,858	41	2,823	1,828	382	11,932
[3-5]	7,433	123	4,349	3,623	480	16,008
[6-10]	9,001	139	4,253	3,302	350	17,045
[11-20]	4,142	162	3,308	2,647	156	10,415
[21-30]	1,773	86	1,443	1,010	78	4,390
[31-50]	1,898	69	988	1,225	78	4,258
[51-100]	2,020	109	866	1,182	76	4,253
[101-250]	1,467	81	786	1,182	47	3,563
[251-500]	355	49	322	253	18	997
[501-1000]	78	5	164	95	3	345
[+1000]	31	4	79	64	1	179
Total	35,056	868	19,381	16,411	1,669	73,385