Establishment dynamics and the persistence of resource misallocation in Mexico: An analysis of longitudinal data for 1998-2013.*

Matías Busso, Santiago Levy, Jesica Torres†

July 2019

-Preliminary-

Abstract

We build on the work of Levy (2018) and examine the effects of the special provisions for non-salaried workers on the dynamics of establishments. We exploit the four waves of the Mexican establishment census available for 1998-2013 combined with the panel identifiers developed by Busso, Fentanes, and Levy (2018). We find that legally informal establishments are just as likely to survive as fully formal plants, but when an informal plant leaves, another informal plant enters. In other words, the net exit of legally informal establishments is zero. When we examine the balanced panel of surviving establishments, we find that when informal plants survive, they do not in general transition into formality nor do they add more workers over their life cycle. In contrast, the average 40-year-old fully formal plant employs four times more workers than its younger counterpart, but almost half of fully formal plants transition into informality (legal or illegal) when they survive. We also document that in Mexico in general the resource reallocation from the exit of high productivity establishments potentially offsets the reallocation from the exit of low productivity plants, while the reallocation from the entry of low productivity establishments likely offsets the resource reallocation from the entry of high productivity plants.

*Comments welcome. We thank the excellent research assistance of Óscar Fentanes and Edgar Ávalos. We also thank INEGI for granting us access to the establishment micro-data examined in this paper. All our work took place at INEGI's Micro-data Laboratory under strict surveillance to protect the information of individual establishments. The panel identifiers and the establishment data exploited here can only be examined with INEGI's approval. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Brookings Institution, the Inter-American Development Bank, its Board of Directors, or the member countries. All errors or omissions are our own.

†Levy: Brookings Institution. Busso and Torres: Inter-American Development Bank. Corresponding author: jesicatorrescoronado@gmail.com.

1. Introduction

Total Factor Productivity (TFP) in Mexico has remained stagnant since the mid-1990s. From 1996 to 2015, growth in GDP per capita averaged in Mexico only 1.2% per year, while growth in TFP averaged only 0.4% per year (and the growth rate is actually negative if we correct for the quality of the labor force; see Levy (2018)). That is, only growth in physical capital and growth in the labor force, with no improvements in efficiency, explain the improvement in Mexico's standards of living the past 25 years.

Policies and institutions in Mexico implicitly and explicitly favor small establishments, which turns them into a potential major culprit for the country's poor productivity performance. Roughly, aggregate productivity is the average of idiosyncratic productivities over all establishments in the economy weighted by their amount of capital and labor. Policies that distort relative prices (by favoring small plants, for example) misallocate in equilibrium too much resources to (relatively) unproductive establishments, and not enough resources to the most productive plants (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). Moreover, these same policies potentially affect as well both the selection of establishments operating in the economy (who enters and who exits) and their technology at a point in time (whether they want to become more productive and grow or not), which combined determines the distribution of physical productivity, and may result in less (physically) productive firms on average. In other words, distortions potentially affect both the distribution of (physical) productivity (by distorting the selection of plants alive at a point in time and their technology) and the allocation of resources across this distribution.

Diego Restuccia, in his remarks for the conference on the Second Generation Productivity Analysis and Policy, identifies these three channels that account for productivity differences across countries: (1) technology, (2) selection, and (3) misallocation.[1] Hsieh and Klenow (2014) also develop a similar argument. Similarly, Levy (2018) examines the effects of the special provisions for non-salaried workers on the dynamics of establishments between 2008 and 2013. He finds that with Mexico's large misallocation, establishment exit, survival, and entry do not correspond with their productivity. He argues that without resource misallocation, the most productive establishments would grow, the relatively bad ones would exit, and aggregate productivity would increase over time.

In this paper we build on the work of Levy (2018) and exploit the four waves of the Mexican establishment census available for 1998-2013 combined with the panel identifiers developed by Busso, Fentanes, and Levy (2018). The Mexican establishment census lists information on both formal and informal establishments, regardless of their size, and is an exceptionally rich database by Latin American standards. In addition, with the panel identifiers we can now follow establishments in the fifteen-year period between 1998 and 2013, which opens a remarkably broad window into the dynamics of formal and informal plants.

We argue that the institutional and regulatory framework in Mexico not only misallocates capital and labor across a fixed productivity distribution (static effect), but also affects the productivity distribution because a large fraction of surviving establishments do not grow over their life cycle, and many low productivity establishments enter while many high productivity establishments exit (dynamic effects).

We describe the establishment census and our data in section 2. In section 3 we argue that policies and institutions in Mexico implicitly tax salaried labor and subsidize non-salaried arrangements, which favors small plants. We show that the majority of establishments in Mexico (almost three fourths) are legally informal, whereas less than 3% are fully formal. Legally informal plants are small and relatively unproductive, and employ more than optimal amounts of capital and labor; fully formal plants are large and productive, but seem to face implied obstacles that prevent them from attracting additional resources. In section 4 we document the dynamics of formal and informal establishments and in section 5 we show that in general, establishment

[1] The conference took place at the World Bank in Wasington, DC on October of 2016. The conference presentations and remarks are available in this link.

entry and exit in Mexico are not productivity-increasing. We conclude in section 6.

2. Data

Our data opens a remarkably broad window into the dynamics of informal and formal establishments, which is unusual for a developing country. We exploit the Mexican establishment census, which surveys all establishments providing goods and services in localities with 2,500 or more inhabitants, and is administered by Mexico's statistics bureau INEGI. The census covers both formal and informal establishments, regardless of the size, but located within fixed premises, which means that activities carried out by street vendors or in open street markets are excluded. Still, establishments in the census account for between 42 and 44% of total employment in Mexico (Levy, 2018).

The census offers an exceptionally rich database. The questionnaire collects data on the location and the age of the establishment, the stock of capital, sales, value added, number of workers (blue collar, white collar, co-business owners and family members that work as employees), aggregate remunerations to blue and white collar employees, and contributions to social security, among other variables. In addition, the sorting of establishments into sectors in the census follows the North American Industrial Classification System (NAICS), a 6-digit classification system which offers a significant level of granularity: establishments within a 6-digit NAICS sector potentially produce very similar goods.

The census is administered every five years, and we exploit the last four waves available: 1998, 2003, 2008, and 2013. We only consider establishments in Manufacturing, Wholesale and retail, and Services. In addition, we exclude around 2% of observations due to measurement error or for consistency purposes (see Appendix B for details). Our final dataset contains 2.7 million plants in 1998, 2.9 million in 2003, 3.6 million in 2008, and 4.1 million in 2013.

The 2008 and 2013 waves of the Mexican establishment census include identifiers that allow researchers to connect observations across this five-year period. These identifiers are not available for the 1998 and 2003 waves of the census, but we exploit the synthetic identifiers developed by Busso, Fentanes, and Levy (2018) (BFL) to follow establishments in the fifteen-year period between 1998 and 2013. More precisely, we use the synthetic identifiers to follow establishments between 1998 and 2008 and then switch to the actual identifiers from 2008 through 2013.

BFL use information on the location, the name, and the 6-digit NAICS sector to link establishments between two consecutive waves. Roughly, if two units report the same name and location (e.g. the same block within the same statistical area within the same locality), the same location and sector, or the same name and sector across two consecutive waves, they are considered to be the same plant. If the algorithm does not find a match in *t+5* for a plant listed in *t*, BFL assume that the plant exited during the 5-year period; similarly, if the algorithm does not find a match in *t-5* for a plant listed in *t*, BFL assume that the plant entered during the 5-year period. Their measure of similarity of names, sectors, and locations across waves allows some room for error. When BFL use the actual identifiers for 08-13 to validate their results, they find that the algorithm results in around 10% of false positives and around 10% of false negatives.

We compute the physical and the revenue productivity (TFPQ and TFPR, respectively) for each plant within each census wave in our dataset following the methodology in Hsieh and Klenow (2009). In Appendix B we list in detail the variables that we extract from the census, and we also describe our methodology to impute the earnings of co-owners and family employees (which are not reported in the questionnaire) and to adjust the value of the stock of capital of the plant.

3. Institutional and regulatory framework and the legality of informality in Mexico

In this section we follow Levy (2018) and argue that policies and institutions in Mexico implicitly tax salaried labor and subsidize non-salaried arrangements. This regulatory distinction between salaried and non-salaried labor distorts the optimal combination of salaried and non-salaried labor in the plant, favors small establishments (since they optimally rely more heavily on non-salaried labor), and distorts occupational choices (and ultimately, the selection into entrepreneurship). We show that the effective contribution to social security reported in the data increases with the size of the plant, which suggests that regulations in general are particularly enforced on larger plants. This size-dependent enforcement, combined with other policies such as simplified tax regimes, likely aggravate the distortionary effects of the special provisions for non-salaried labor. We then show that the majority of establishments in Mexico (almost three fourths) are legally informal, whereas less than 3% are fully formal. Legally informal plants are small and relatively unproductive, and employ more than optimal amounts of capital and labor; fully formal plants are large and productive, but seem to face implied obstacles that prevent them from attracting additional resources.

1. Policies and institutions in Mexico implicitly tax salaried contracts and subsidize non-salaried labor.

Federal, state, and local institutions and regulations in Mexico sort workers into salaried and non-salaried. Salaried workers are employees who work in exchange for a fixed regular payment (a salary) and are officially considered subordinates to the employer. Non-salaried workers include the self-employed and employees not subordinated to the employer or who do not work for a wage: contractuals, workers officially employed by a third party (outsourced), and co-owners and employees in rent-sharing arrangements with their employer.[2] In general, regulations apply only to employer-subordinate relationships and explicitly exempt non-salaried arrangements (including the self-employed) from complying.[3] Employers of non-salaried workers then do not face the same statutory burden that employers of salaried workers do.

Salaried employees are entitled to a minimum number of days of paid leave, a yearly bonus irrespective of their performance, and a severance upon dismissal, which combined amount to a tax on salaried labor of 18% (Alaimo et al., 2017). They are also entitled to a minimum wage and to form a labor union with other salaried employees. Non-salaried workers are not entitled to any of these benefits. Employers of salaried workers additionally face a state payroll tax—with rates that vary across states between 2 and 3%—and the administrative burden of withholding their personal income tax. Employers of non-salaried workers are not required to withhold their income tax and are exempt from the state payroll tax.

Employers of salaried workers are also required to contribute around 23% of the average wage to social security, but this requirement does not extend to non-salaried workers whose social security is implicitly subsidized by non-contributory programs.[4] Contributions to social security in Mexico cover pension benefits, the provision of medical services, a work injury compensation insurance, and additional benefits such as access to day-care centers. They are levied on both employers and employees, but the estimate of 23% totals only the statutory rates for the employer of an average employee (whose wage is three times the minimum wage), and includes the mandatory contribution of 5% to the national housing fund (which offers salaried workers mortgages at relatively low rates). Non-contributory programs also cover medical insurance, retirement

[2] In rent-sharing arrangements earnings are not fixed but vary with the performance of the employee or the performance of the firm. These arrangements include family businesses, and workers on commission or paid per piece.

[3] Contractuals in the US, for example, are still required to contribute to their social security (which they pay when they file their income tax), even though they are not officially subordinates to their employer. Social security in Mexico, to the contrary, is linked to a salaried labor relationship. Articles 20 and 21 of the Federal Labor Law (*Ley Federal del Trabajo*), and Articles 12 and 13 of the Social Security Law (*Ley del Seguro Social*), for example, explicitly exempt "workers in family businesses or small retailers" (where rents are commonly shared) from contributing to social security.

[4] The labor reform of December of 2012 made explicit the right to the social security for workers officially employed by a third party (before that reform, legal disputes would commonly arise).

benefits, day care services, and even housing subsidies. Levy (2018) estimates that the contributory social insurance amounts to an implicit tax on salaried labor of 12% (as the services provided are usually of low quality and many workers will not qualify for benefits after retirement), whereas the non-contributory social insurance amounts to an implicit subsidy on non-salaried labor of 16%.

2. This regulatory distinction between salaried and non-salaried labor distorts several static margins.

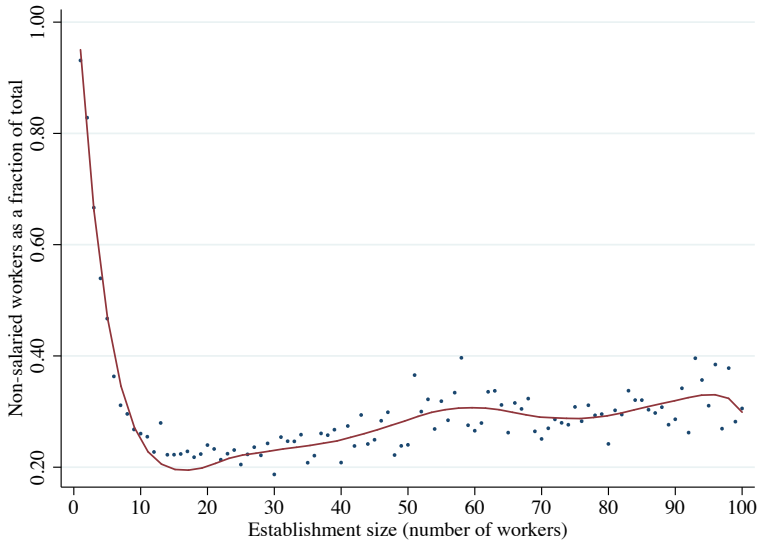
The regulatory distinction between salaried and non-salaried labor introduces then a wedge between the marginal product and the marginal cost of salaried labor, thus distorting the optimal combination of salaried and non-salaried workers: plants across the size distribution will employ more than optimal amounts of non-salaried labor. Moreover, if there are costs of compliance, the special provisions for non-salaried labor will implicitly favor small establishments as they will optimally rely on non-salaried arrangements more heavily than larger plants (all else equal).

Intuitively, in small establishments both the business owner and his or her employees can easily observe not only everyone else's effort but also the profits of the firm. Everyone can therefore disentangle the effects of shocks from the effects of effort on the performance of the firm. As the establishment grows, however, the cost of monitoring everyone else's effort will increase as well—some employees may optimally free ride on the effort of others—and employees will have a harder time verifying the profits of the firm—which increases the cost of bargaining with the owner. If complementarities in production across employees are strong, the owner will transition into salaried labor; if complementarities are not strong enough, the manager may still choose to share rents with his or her employees, or to contract out or outsource some tasks.[5]

In Figure 1 we show the relationship observed in the Mexican establishment census between the fraction of non-salaried workers and the size of the plant. More than 80% of all workers in micro-establishments (with 3 or fewer workers) are non-salaried (who are mainly in rent-sharing arrangements, as we show in Figure A1 in the Appendix), whereas the average fraction of non-salaried workers is around 20% in medium size-plants, and increases to fluctuate around 30% in establishments with 50 or more workers (and these non-salaried workers are mainly outsourced employees).

[5]The latter is the case of insurance companies or of firms who choose to outsource tasks that are not the core of the business (such as cleaning and maintenance, accounting, and legal services).

Figure 1: Fraction of non-salaried workers by establishment size.



Data for 2013 from the Mexican establishment census. Line smoothed using `lpoly` (kernel-weighted local polynomial regression) in STATA. We report only firms with 100 workers or fewer for illustration purposes. Non-salaried workers includes co-owners and family employees, piece workers, and workers officially employed by a third party (outsourced).

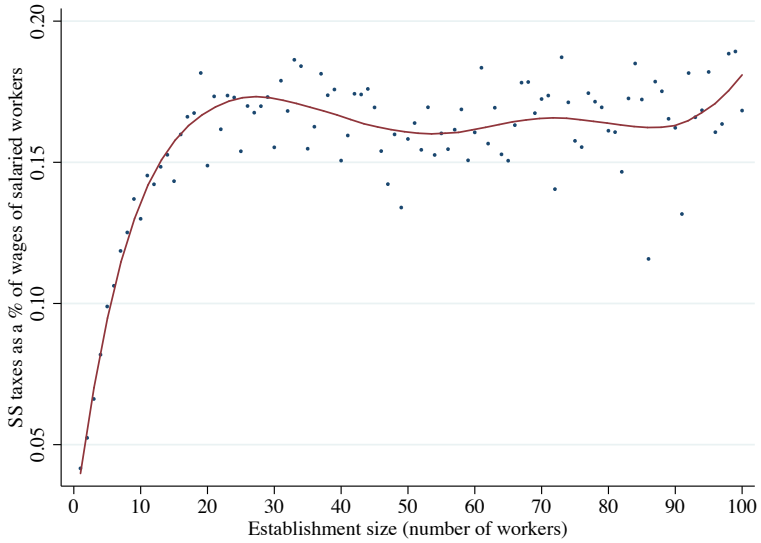
The special provisions for non-salaried labor also distort occupational choices because the regulatory burden on entrepreneurship is lower relative to salaried employment. Unlike salaried employees, the self-employed are exempt from contributing to social security and their social insurance is fully covered by the government with non-contributory programs. In addition, employers withhold the personal income tax on behalf of their salaried employees—which makes it harder for salaried employees to evade—whereas the self-employed file their own income tax and face preferential rates if they qualify for the simplified regime for the corporate tax (only 2% or less of their sales instead of the general rate for the corporate tax).

3. Size-dependent enforcement and additional features of the regulatory framework aggravate the distortionary effects of the distinction between salaried and non-salaried labor.

Regulations in Mexico are particularly enforced on large firms and as a result, effective tax rates increase with the size of the plant. Consider contributions to social security. The contribution of the employer of an average employee (who earns three times the minimum wage) and who fully complies with his or her statutory rates—without considering contributions to the housing fund which are omitted from the questionnaire in early waves of the establishment census—is approximately 18%.[6] Figure 2 shows that in the Mexican establishment census, establishments with 1-9 workers report an average contribution of only 8.8% (and establishments with 3 or fewer workers report an average contribution of around 5%), whereas the reported average contribution fluctuates around 16% among establishments with 30 or more workers.

[6]Statutory contributions have some flat components, and therefore the average contributions if the plant fully complies will in general vary with the wage composition of the salaried workforce. The census only surveys the plant on the number of salaried and non-salaried workers and the total wage bill on the salaried employees, and thus we cannot compute a more precise measure of compliance with social security contributions.

Figure 2: Average social security contribution on salaried employees by establishment size.



Data for 2013 from the Mexican establishment census. Line smoothed using `lpoly` (kernel-weighted local polynomial regression) in STATA. We report only firms with 100 workers or fewer for illustration purposes.

The wedge between the marginal product and the marginal cost of salaried labor is then larger for larger plants: note in Figure 1 that as social security regulations likely come into full effect (somewhere between 20 and 30 workers), the fraction of non-salaried workers actually increases instead of decreasing with the size of the plant. The enforcement of other provisions in Mexico likely follows this pattern as well, which results in higher compliance costs for larger plants—for example, salaried employees in large establishments presumably face higher effective tax rates on their income relative to salaried employees in smaller plants (which in addition, aggravates the distortion to occupational choices).

Other features of the regulatory framework also interact with the special provisions for non-salaried workers and explicitly or implicitly deincentivize both the employment of salaried labor and the growth of plants. The simplified regime for the corporate tax grants establishments with sales below a threshold both lower rates and lower compliance costs. Special regimes for the value-added tax exempt 42% of the tax base. Significant market concentration in the financial sector and incomplete enforcement of legal contracts limit the access to credit. Levy (2018) reviews in more detail how these features combined aggravate the distortionary effects of the distinction between salaried and non-salaried labor.

4. Most establishments in Mexico employ only non-salaried workers. These plants are small and relatively unproductive, and employ more than optimal amounts of capital and labor.

In the census we observe the number of salaried and non-salaried workers in each establishment, and both the wage bill and the contributions to social security (as reported by the establishment) that the plant paid on their salaried workers.[7] We compute the average contribution as the ratio of total social security contributions to the wage bill of salaried workers. The statutory social security contributions levied on employers approximate 18% for an average salaried worker (contributions to the national housing fund are

[7]The labor expenses on non-salaried workers are observed only for contractuals and outsourced employees.

omitted in some waves of the questionnaire), and we follow Levy (2018) and use this threshold to proxy for compliance.

Legally informal establishments employ only non-salaried workers: their workers are not covered by social security but the owners are not required to observe the provisions in the labor code. Full evaders or **illegally informal** plants report an average contribution on their salaried employees of 0%, and they may or may not also combine non-salaried workers. **Fully formal** plants do not employ non-salaried workers and report an average contribution of at least 18% on their salaried employees. The rest—**mixed establishments**–are not fully formal (they may employ non-salaried workers) nor are they fully legal (their average contribution is not necessarily above 18%). Table 1 shows the distribution of plants across these categories in 2013 with their corresponding employment and capital share.

Table 1: Fraction of plants and their workers by labor arrangement and compliance with Social Security contributions.

Compliance with contributions	Plants (%)	Workers (%)	Capital (%)
Informal but legal	72.76	39.92	31.69
Illegally informal	17.35	16.16	7.47
Mixed	7.28	30.07	41.59
Fully formal	2.60	13.85	19.25
Total	100	100	100

Data for 2013 from the Mexican establishment census.

Close to 73% of establishments employ only non-salaried workers and are thus exempt from complying. 40% belong to own account workers without employees, and the remaining 33% employ others, but not for a salary. They employ 40% of workers and 32% of the capital, and generate 28% of the value in these sectors (see Table A4 in the Appendix). Their average size is 2.3 workers and their stock of capital is 44% the capital of the average plant in the economy.

Only 2.6% of establishments are fully formal (they employ exclusively salaried workers and pay an average contribution above 18%). These plants employ 14% of workers and 19% of the capital in these sectors, and generate 19% of the value added. These plants employ on average ten times as much workers and seventeen times as much capital as informal but legal plants: their size is 22.4 workers on average and their stock of capital is 7.4 times the capital of the average plant in Mexico. Note that the fractions of legally informal and fully formal establishments and their allocation of labor and capital have remained relatively stable since 1998 (see Table A3 in the Appendix).

We compute the physical and the revenue productivity (TFPQ and TFPR respectively) for each plant within each census wave in our dataset following the methodology in Hsieh and Klenow (2009). Recall that in this monopolistic competition model TFPQ may vary across establishments but TFPR does not if there are no distortions: more physically productive plants receive more resources—which lowers their marginal revenue products of labor and capital—and sell more at a lower price, whereas less productive plants produce less and sell at higher prices. Idiosyncratic distortions act as barriers that prevent resources from flowing from low revenue productivity plants to establishments with higher TFPR; they misallocate labor and capital and, as a result, revenue productivity is not equated across establishments. A high TFPR would indicate that the establishment has less than optimal amounts of capital and labor (the marginal revenue products are too high). Conversely, a low TFPR would suggest that the plant has more than enough resources (the marginal

revenue products of capital and labor are too low). Measured TFPR in the model then captures the magnitude of the idiosyncratic distortions faced by the plant.

Both physical and revenue productivity in the data markedly vary across levels of compliance with social security. Tables A5 and A6 in the Appendix show the average percentage difference in TFPQ and TFPR (relative to industry means) for fully formal plants relative to informal but legal plants. Establishments that fully comply (regardless of whether they also employ non-salaried workers) are at least three times as (physically) productive as informal but legal plants. Moreover, relative to informal but legal establishments, additional resources are worth 76.2% more if allocated to fully formal plants (plants that only employ salaried employees and fully comply). The estimates in Table A6 show that additional resources are actually worth the least if assigned to informal but legal plants, and the most if assigned to fully formal establishments. The revenue productivity differences between legally informal and fully formal establishments suggest that fully formal establishments face implied obstacles that prevent them from attracting additional resources, whereas informal but legal plants enjoy implied advantages that secure them more than optimal amounts of capital and labor. In other words, policies and institutions in Mexico seem to channel too much resources to establishments that only employ non-salaried labor.

In addition to sorting establishments into levels of reported compliance, we follow Levy (2018) and split the revenue productivity distribution into three segments: establishments in the bottom quartile in the distribution of TFPR are considered **low productivity establishments** (the value of additional capital and labor flowing into these establishments is relatively low); establishments in the second and third quartiles are considered **medium productivity plants**; finally, establishments in the top quartile are considered **high productivity plants** (the value of additional resources is relatively high). Table A7 in the Appendix shows that plants in the bottom 25% of the revenue productivity distribution employ 17% of workers and around 18% of capital, and 87% of them are informal but legal whereas less than 1% are fully formal. In contrast, plants in the top 25% of the revenue productivity distribution employ around 30% of workers and 20% of capital, and around 60% of them are legally informal whereas close to 5% are fully formal. That is, low productivity establishments are smaller and less likely to fully comply relative to high productivity plants. Indeed, almost a third of legally informal establishments are in the low productivity segment and only around 21% are in the top of the productivity distribution, whereas 6% of fully formal plants are in the low productivity segment and around 44% are in the right-hand tail of the distribution (see Table A8 in the Appendix).

4. The dynamics of formal and informal establishments in Mexico

In this section we document the dynamics of establishments in Mexico. We find that since 1998, around 2 million establishments enter and around 1.5 million establishments exit within a 5 year period. Moreover, three quarters of all establishments in a given year will have exited 15 years later. These flows amount to aggregate rates of entry and exit that do not markedly differ from the US, but in Mexico small establishments enter and exit at significantly lower rates, whereas larger plants enter and exit at a faster pace. We also find that informal-but-legal plants and fully formal plants exhibit the same survival rates, but in general, when an informal-but-legal establishment leaves, another legally informal establishment enters. Finally, we document that when informal-but-legal establishments survive, they do not grow nor do they transition into formality, and actually their revenue productivity increases over their life cycle, which suggest that conditional on survival, these establishments seem too small relative to an allocation where revenue productivity is constant across plants.

4.1 Aggregate establishment dynamics

In Figure 3 we show the flows of establishment entry and exit between 1998 and 2013. More accurately, we show in the top panel the number of establishments listed in 1998 and the number of those that later exited or survived through 2013. Next we show the number of establishments that entered or survived between 1998 and 2003 and the number of those that exited or survived between 2003 and 2013. Similarly, we show in the third panel the number of establishments that entered or survived between 2003 and 2008 and the number of those that later survived or exited between 2008 and 2013. Finally, in the fourth panel we show the number of establishments that entered or survived between 2008 and 2013.

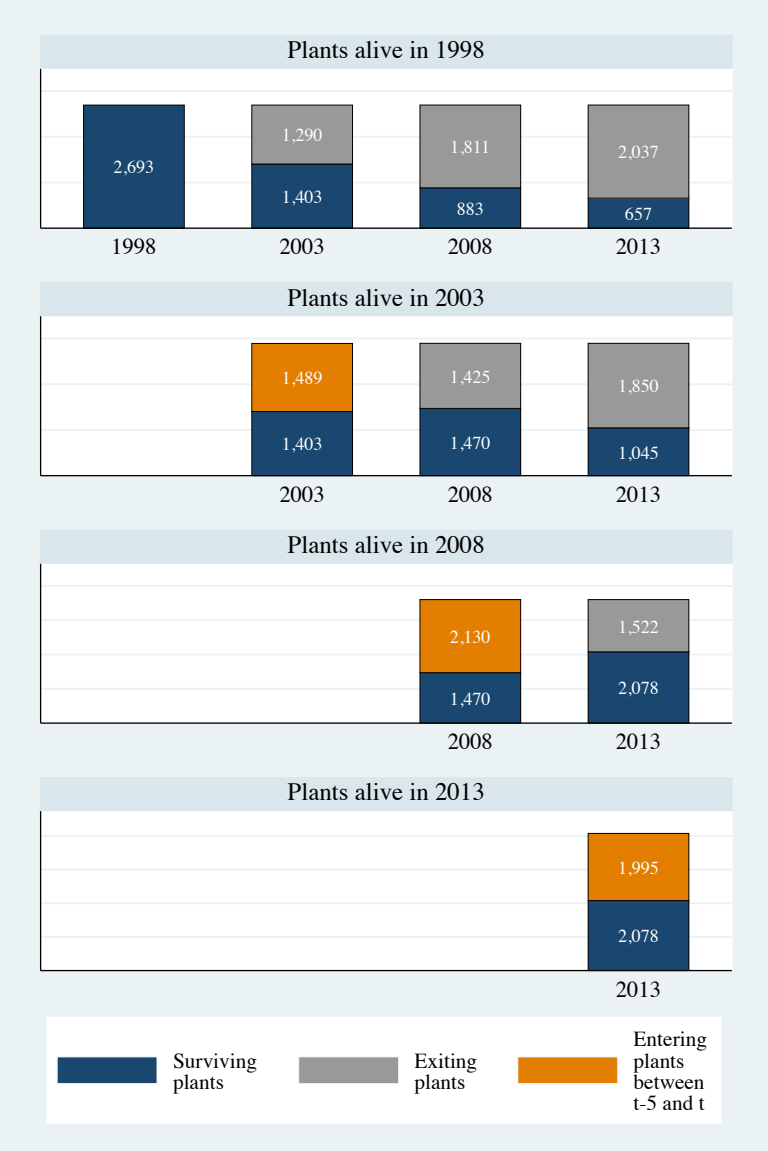
Note first that around 2 million establishments enter and around 1.5 million establishments exit within a 5 year period. These flows amount to average annualized rates of entry and exit between 2008 and 2013 of 8.8% and 9.6%, respectively.[8] Note in addition that on average, close to half of all establishments in a given year will survive through *t+5*, a third will survive through *t+10*, and only around one quarter will survive through *t+15*. In other words, close to 75% of all establishments in a given year will have exited 15 years later. Figure A2 in the Appendix breaks down the flows of entering, surviving, and exiting plants into levels of compliance with social security, but the results closely mimic the cross-sectional distribution: regardless of the span of time considered, around 70% of entering, surviving, and exiting plants are legally informal and only between 2 and 3% are fully formal.

[8] We compute the 5-year rates of establishment entry and exit using the standard formula:

$$\frac{\textit{Entering (or exiting) establishments between 2008 and 2013}}{\left[\frac{\textit{Establishments alive in 2008}+\textit{Establishments alive in 2013}}{2}\right]}$$

We follow the corresponding formulas for the 5-year rates of job creation and job destruction. We then annualize these 5-year rates for comparison purposes.

Figure 3: Flows of establishment entry, survival, and exit (in thousands). 1998-2013.



In Table 2 we show the average size (in number of workers), the capital-labor ratio, and output per worker of establishments that entered between 1998 and 2003 and that survived or exited between 1998 and 2003, 1998 and 2008, and 1998 and 2013. Plants that survive after a 5-year period are on average larger and generate more output per worker relative to plants that exit: surviving plants employ on average 5.5 workers and 1.22 times the amount of capital per worker of the average plant in the economy, and generate 1.18 times the amount of output per worker of the average plant, whereas exiting plants employ on average 3.4 workers and only half the amount of capital per worker of surviving plants, and generate 68% the amount of output per worker of the average plant. Interestingly, plants that survive after 10 or 15 years are not markedly different from plants that survive through t+5.

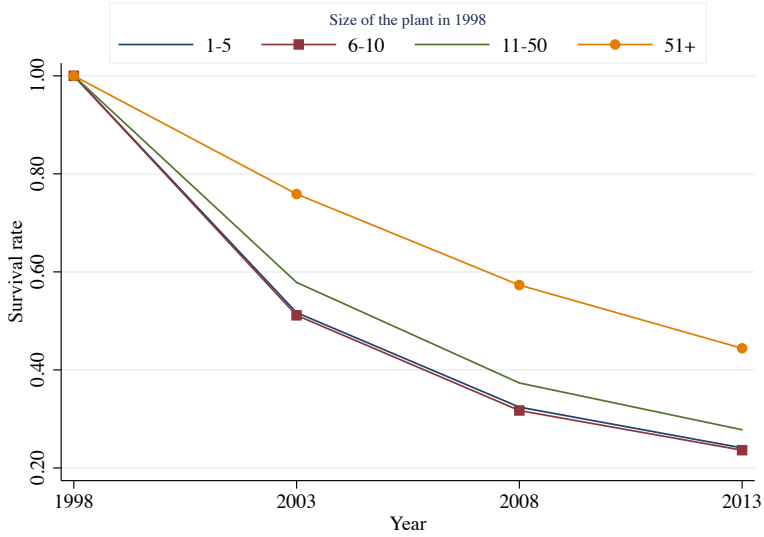
Table 2: Establishments that entered between 1998 and 2003 and that exited or survived between 1998 and 2003, 1998 and 2008, and 1998 and 2013.

	Plants (%)	Average size	Capital per worker	Output per worker
Entry				
t+5	51.82	3.24	0.58	0.63
Exit				
t+5	48.18	3.41	0.63	0.68
t+10	67.36	3.65	0.68	0.73
t+15	75.72	3.90	0.87	0.85
Survival				
t+5	51.82	5.53	1.22	1.18
t+10	32.64	6.27	1.39	1.32
t+15	24.28	6.40	1.25	1.28

The fraction of plants are relative to total in t+5 in the case of entry, and relative to total in t for exit and survival. Capital per worker and output per worker are relative to the capital and output per worker in the economy (in t+5 for entering plants, and t for exiting and surviving plants).

In Figure 4 we sort establishments by size in 1998 (measured using the number of workers) and then compute their average survival rate through 2013. Around half of establishments with 10 or fewer workers survived through 2003, and only 25% survived through 2013. In other words, 75% of small establishments will have exited within a 15-year period. Medium size establishments (with 11-50 workers) exhibit survival rates a few percentage points higher relative to smaller plants, whereas large plants (51 or more workers) are significantly more likely to survive. 76% of plants with 51 or more workers survived through 2003, 57% survived through 2008, and 44% survived through 2013, which is almost twice the 15-year survival rate of small plants. Still, 56% of large establishments will have exited within a 15-year period.

Figure 4: Survival rate (unconditional) by size of the plant. 1998-2013.



We sort establishments by number of workers in 1998. Survival rate = number of establishments alive through t+5, t+10, and t+15 as a fraction of establishments alive in that size category in t.

In Table 3 we show the flow of establishment entry between 2008 and 2013 by size. Note that half of all small establishments in 2013 entered between 2008 and 2013, whereas only 35% of large establishments (51 or more workers) were relatively young. In Table 4 we show rates of yearly entry and exit by size for both Mexico and the US. Note that the rates of entry and exit in Mexico are not as markedly dependent on size as in the US (see also Figures A3 and A4 in the Appendix), even if aggregate establishment dynamics does not significantly differ between the two countries. In particular, in Mexico small establishments enter and exit at significantly lower rates than in the US, whereas larger plants enter and exit at a faster pace.

Table 4 shows also the rates of job creation and job destruction by size. The rates of job destruction are very similar between Mexico and the US (except in establishments with 1-5 workers). To the contrary, job creation in Mexico does not markedly vary by size, and the rates are significantly lower than in the US (a third of the rate of job creation in the US in plants with 1-5 workers, and half in establishments with 6-10 workers). Table 5 shows the rates of exit, job creation, and job destruction by age, and also for the US. Young establishments (ages 1-5) in the US exit and create and destroy jobs at a faster pace than in Mexico, but more mature establishments behave similarly in both countries.

Table 3: Fraction of entering and surviving plants by size. 2008-2013

Size	Entering plants	Surviving plants	Total
1-5	49.82	50.18	100
6-10	40.64	59.36	100
11-50	42.28	57.72	100
51+	34.78	65.22	100

We sort establishments by number of workers in 2013 and then examine whether they are new entrants or survived from 2008. We show fractions of the number of establishments in that size category in 2013.

Table 4: Dynamics by establishment size in the US and Mexico.

Size of the	Entry		Exit		Job creation		Job destruction	
plant	US	MX	US	MX	US	MX	US	MX
1-5	17.3	8.9	15.3	9.7	28.3	10.1	25.4	14.4
6-10	4.1	7.4	3.3	9.6	18.3	9.3	16.2	17.3
11-20	3.1	7.6	2.3	8.7	15.8	9.8	12.6	14.7
21-50	2.4	7.2	1.9	7.3	14.1	9.8	10.9	12.2
51-100	1.9	5.8	1.5	5.2	12.8	9.4	10.0	9.9
101+	1.4	4.9	1.1	4.0	11.1	7.1	9.0	8.4
Total	10.1	8.8	8.8	9.6	13.9	9.1	11.4	12.2

US: Data for 2014 from the US Census Bureau with size bins 1-4, 5-9, 10-19, 20-49, 50-99, 100+. Mexico: Average annualized rates for 2008-2013.

Table 5: Dynamics by establishment age in the US and Mexico.

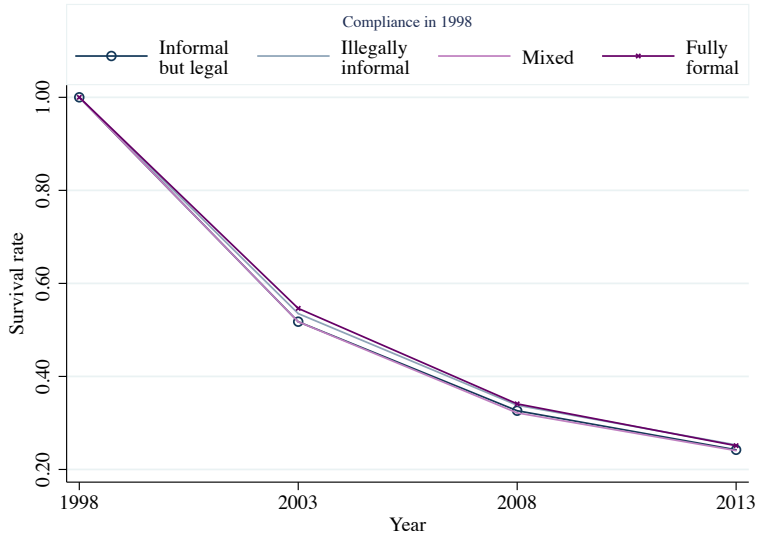
Age of the plant	Exit		Job creation		Job destruction	
	US	MX	US	MX	US	MX
1-5	16.7	11.32	17.1	12.71	19.5	15.5
6-10	9.1	7.52	11.2	7.28	13.1	11.72
11-15	7.1	5.42	9.5	6.75	11.2	10
16-20	6.1	6.79	8.5	6.37	9.8	9.73
21-25	5.8	4.51	7.9	5.69	9.1	8.8
26+	5.5	6.02	7.0	5.42	8.4	8.81

US: Data for 2015 from the US Census Bureau. Mexico: Average annualized rates for 2008-2013.

4.2 Exit, survival, and entry across levels of compliance

We sort establishments into levels of (reported) compliance with social security contributions in 1998, and compute their average survival rate through 2013. Note in Figure 5 that informal-but-legal plants exhibit the same survival rates that fully formal plants do. In other words, formal establishments are not significantly more likely to survive than informal plants, and both exhibit the same survival rates as the average plant: close to half will survive through *t+5*, a third will survive through *t+10*, and only around one quarter will survive through *t+15*. Formal plants are only slightly more likely than informal plants to survive through 2003, but the survival rates of both formal and informal establishments are very similar through *t+10* and *t+15*.

Figure 5: Survival rate by compliance with social security. 1998-2013.



We sort establishments into levels of compliance in 1998 and then follow their survival through 2013. Survival rate = number of establishments alive through t+5, t+10, and t+15 as a fraction of establishments alive in that category in t.

In Table 6 we compute survival probabilities controlling also for (revenue) productivity and the results are not markedly different. We run a series of pooled linear probability models to estimate the correlation between entry or exit and reported compliance with social security. The dependent variables in each model

are (1) a dummy variable which flags establishments that entered during the periods 1998-2003, 2003-2008, and 2008-2013; (2) a dummy which flags establishments that exited within each 5-year period; (3) a dummy which takes the value of 1 if the establishment exited between 2003 and 2008 or 2008 and 2013, conditional on surviving at least 5 years (that is, from 1998 to 2003 or 2003 to 2008); and (4) a dummy variable which flags establishments that exited between 2008 and 2013, conditional on surviving at least 10 years (from 1998 and 2008). We find that formal plants are 7 percentage points less likely to exit within 5 years, 6 percentage points less likely to exit within 10 years, and 5 percentage points less likely to exit within 15 years.

Table 6: Correlation between compliance and the probability of entry and exit.

	Entry between t and t+5		*Exit between t and t+5*		*Exit between t+5 and t+10*		*Exit between t+10 and t+15*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fully formal	-0.102***	-0.071***	-0.073***	-0.071***	-0.004**	-0.058***	0.006**	-0.050***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)
High productivity		-0.006***		-0.020***		-0.004***		-0.000***
		(0.000)		(0.000)		(0.001)		(0.001)
Low productivity		0.037***		0.046***		0.022***		0.021***
		(0.000)		(0.000)		(0.001)		(0.002)
Obs.	10,541,468	7,814,618	9,152,685	6,985,088	2,852,715	2,163,609	874,462	597,077
R-sq.	0.138	0.113	0.038	0.042	0.018	0.056	0.010	0.045

Results from a pooled linear probability model. Standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01. The dependent variable in columns (1) and (2) is a dummy variable which flags establishments that entered during the periods 1998-2003, 2003-2008, and 2008-2013. Similarly, the dependent variable in columns (3) and (4) is a dummy which flags establishments that exited within those 5-year periods. In columns (5) and (6) the dependent variable is a dummy which takes the value of 1 if the establishment exited within the 1998-2008 and 2003-2013 periods (conditional on surviving at least 5 years). Finally, in columns (7) and (8) the dummy variable flags establishments that exited between 1998 and 2013 (conditional on surviving at least 10 years). Fully formal establishments employ only salaried employees and fully comply with their social security contributions (their average contribution exceeds 18%). Odd-labeled columns use the sample before applying the methodology developed in Hsieh and Klenow (2009). The omitted category in these models is establishments that employ only non-salaried workers. Even-labeled columns use the sample with the TFPR estimates. *High productivity* are establishments in the top 25% of the revenue-productivity distribution and *Low productivity* are establishments in the bottom 25%. The omitted category in the even-labeled models is establishments that employ only non-salaried workers and are in the second and third quartiles of the revenue-productivity distribution. All models include controls for other levels of compliance, size, age, and 3-digit sector dummies. Full set of results is available upon request.

In Table 7 we examine whether formal or informal plants in 2013 survived or entered between 2008 and 2013. Half of all legally informal establishments in 2013 entered between 2008 and 2013, and the fraction of entering plants is very similar in other levels of compliance. In Table 8 we show yearly entry and exit rates of formal and informal plants. Note that the net exit of legally informal plants is almost 0: in a given year, 9 out of a 100 legally informal plants will have just entered and another 9.6 out of a 100 will exit. That is, on average, when an informal-but-legal establishment leaves, another legally informal establishment enters. Informal-but-legal and fully formal establishments enter and create jobs at very similar rates, though formal plants do leave and destroy jobs at lower rates relative to informal plants (see also Table A9 in the Appendix).

When fully formal establishments enter or exit, they are larger (they employ significantly more workers and more capital per worker) and produce significantly more output per worker than legally informal plants. Exiting informal establishments do not employ more workers and only slightly more capital per worker than their entering counterparts. Moreover, exiting informal plants produce half as much output per worker than entering informal plants. To the contrary, exiting formal plants are larger but only slightly more productive

(measured using output per worker) than entering formal plants.

Table 7: Fraction of entering and surviving plants by level of compliance with social security. 2008-2013.

Compliance	Surviving plants	Entering plants	Total
Informal but legal	48.68	51.32	100
Illegally informal	55.85	44.15	100
Mixed	60.44	39.56	100
Fully formal	46.92	53.08	100

We sort establishments into types in 2013 and then examine whether they are new entrants or survived from 2008. We show fractions of the number of establishments in that category in 2013.

Table 8: Size and productivity of establishments that enter and exit between t and t+5 by establishment type.

	Annualized rate	Average size at entry/exit	Capital per worker at entry/exit	Output per worker at entry/exit
Informal but legal				
Entry	9.08	2.04	0.55	0.62
Exit	9.56	2.09	0.69	0.33
Fully formal				
Entry	8.75	14.35	0.87	1.15
Exit	5.84	19.42	1.19	1.36

Informal but legal establishments employ only non-salaried workers. Fully formal establishments employ only salaried employees and fully comply with their social security contributions (their average contribution exceeds 18%). We examine establishments that entered or exited between 2008 and 2013 and show average annualized rates. Capital per worker and output per worker are relative to the capital and output per worker in the economy (in t+5 for entering plants, and in t for exiting plants).

Approximately 657 thousand establishments in the census survived from 1998 to 2013. We exploit this balanced panel, and examine first whether surviving establishments transitioned across levels of (reported) compliance with social security contributions during this 15-year period. Table 9 shows our results. 88% of informal but legal establishments in 1998 were still legally informal in 2013. That is, legally informal establishments do not in general transition into formality when they survive. Only 20% of fully formal establishments in 1998 were still fully formal in 2013—25% transitioned into legal informality, 22% into illegal informality, and 33% turned into mixed plants. In other words, almost half of fully formal plants transition into informality (legal or illegal) when they survive. Finally, note that between 1998 and 2013 around 40% of both illegally informal and mixed establishments transitioned into legal informality.

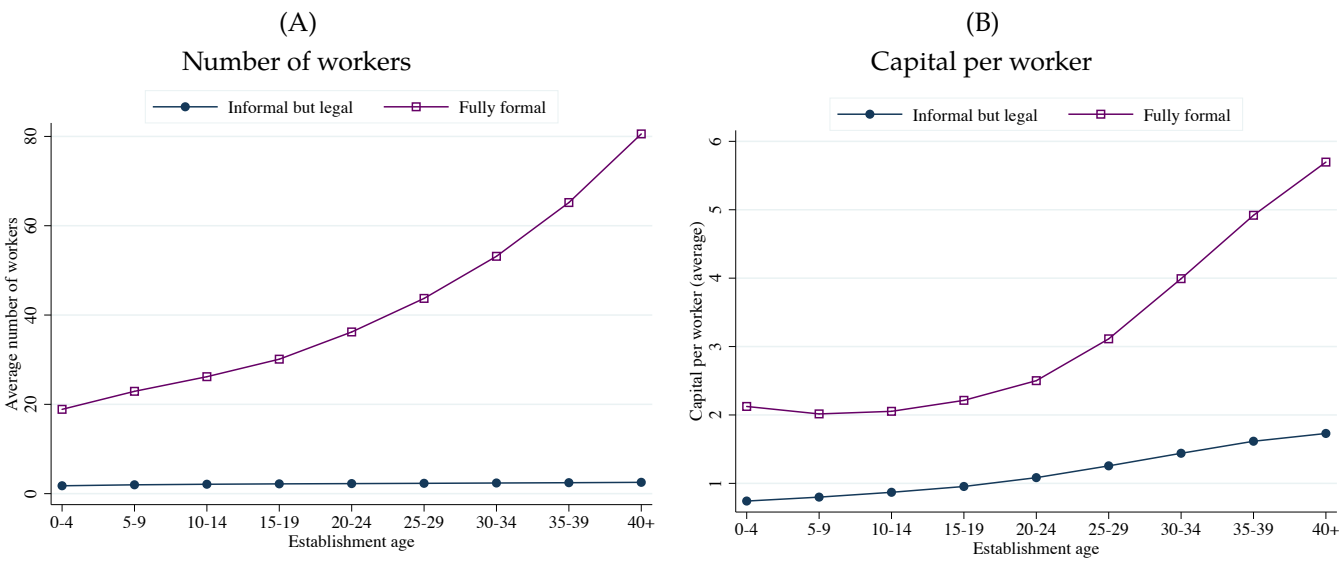
Table 9: Transition across levels of compliance between 1998 and 2013.

Compliance in 1998	Compliance in 2013				Total
	Informal but legal	Illegally informal	Mixed	Fully formal	
Informal but legal	**87.7**	6.9	4.9	0.4	100
Illegally informal	42.9	**23.4**	23.6	10.1	100
Mixed	39.2	21	**31.8**	8	100
Fully formal	25.2	22.1	32.9	**19.7**	100

Next we sort the balanced panel of surviving establishments into categories of reported compliance in 1998, and compute average size and average productivity over all plants of the same age within each category between 1998 and 2013. For example, 0-4 year old plants in 1998 were 5-9 in 2003, 10-14 in 2008, and 15-19 in 2013. Similarly, 5-9 year old plants in 1998 (born at some point between 1993 and 1998) were 10-14 in 2003, 15-19 in 2008, and 20-24 in 2013, and so on. We therefore have two groups of 5-9 year old plants which, even if they are from different cohorts, we can pool to compute their average size conditional on survival, without the noise of entering and exiting plants.

Size is measured in number of workers and capital per worker, and to measure productivity we use output per worker, TFPQ (physical productivity), and TFPR (revenue productivity), all in real terms. We deflate value added and capital to be expressed in 2013 prices using INEGI's publicly available implicit price deflators from the National Accounts and GDP time series, and then we run the methodology developed in Hsieh and Klenow (2009). Capital is deflated by the implicit price index for gross fixed capital formation, which is available for 3-digit sub-sectors in Manufacturing and the Service sector. Value added is deflated by the implicit price index for gross value added, which is available for 4-digit sub-sectors in Manufacturing, 2-digit sub-sectors in Wholesale & retail, and 3-digit sub-sectors in the Service sector. The wage bill is deflated using the consumer price index, publicly available also from INEGI.

Figure 6: The life cycle of establishments in Mexico: average size.

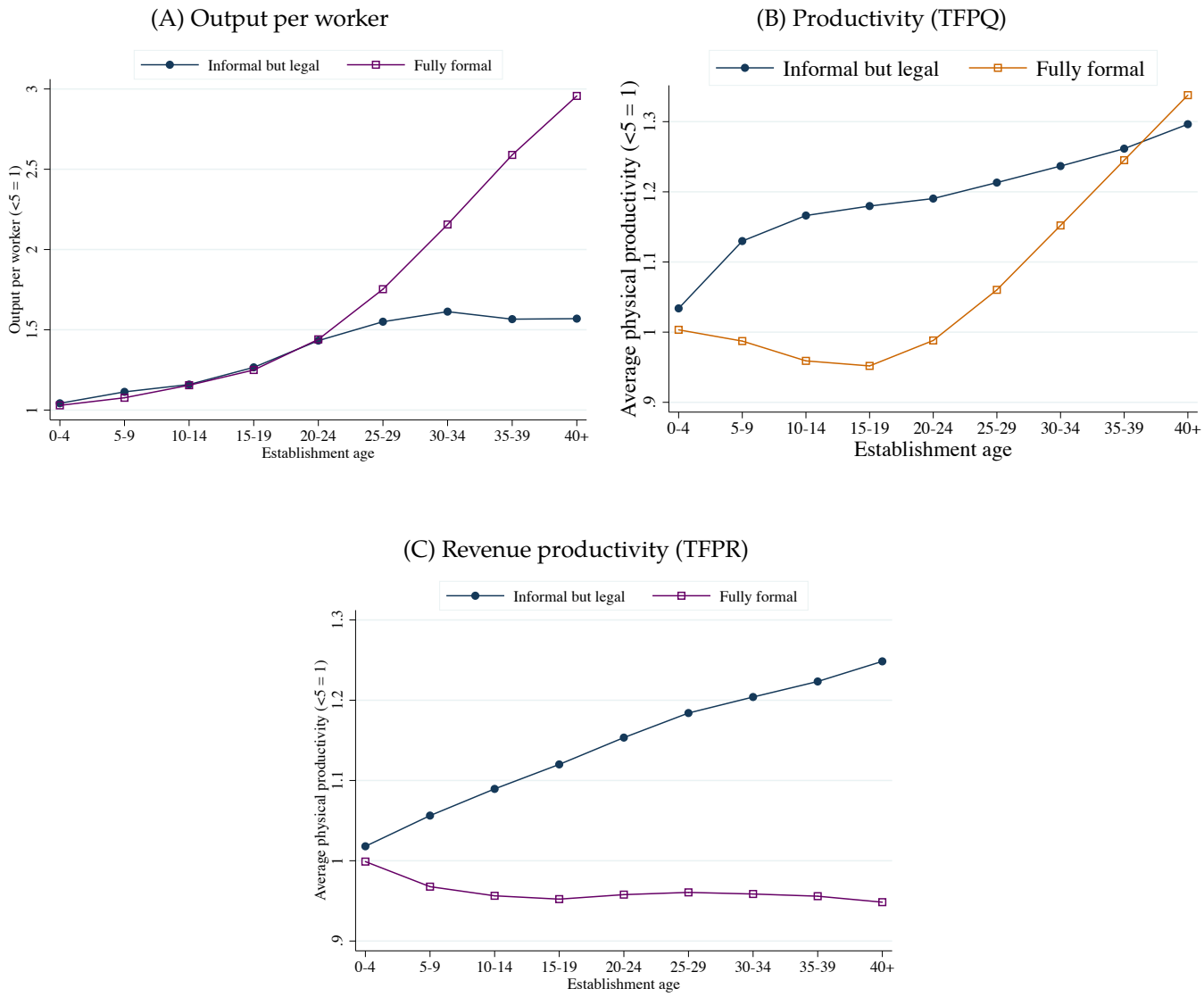


Lines smoothed using `lowess` (locally weighted regression) in STATA. Capital per worker is relative to the average in the economy in 2013.

Note in panel A of Figure 6 that the average 40-year-old fully formal plant employs four times more

workers than its younger counterpart, whereas the average informal-but-legal establishment does not add more workers over their life cycle (average size remains steady at 2 workers). Informal establishments do cumulate more capital per worker during their life cycle, but at slower rates than formal plants (see panel B). Moreover, conditional on surviving, informal establishments seem smaller than optimal: note in Figure 7 that the revenue productivity increases over the life cycle of informal-but-legal plants, which suggest that conditional on survival, these establishments seem to face implied obstacles that prevent them from attracting the amount of additional resources necessary for optimality (Hsieh and Klenow, 2009, 2014).

Figure 7: The life cycle of establishments in Mexico: productivity and revenue productivity.



Lines smoothed using `lowess` (locally weighted regression) in STATA. Output per worker is relative to the average in the economy in 2013. Both TFPQ and TFPR are computed relative to industry means.

5. Establishment dynamics, resource reallocation, and productivity

The productivity costs from resource misallocation in Mexico increased between 1998 and 2013. Table 10 shows, for each census wave between 1998 and 2013, the productivity gains from equalizing revenue productivity across plants and three alternative measures of the dispersion of revenue productivity. In 1998

the value of additional capital and labor flowing into a plant in the 90th percentile of the distribution was 2.5 times the value of these resources flowing into a plant in the 10th percentile. In 2013 this difference increased to 2.9. The increase in TFP from hypothetically reallocating labor and capital across establishments to eliminate this dispersion increased from 156% in 1998 to 211% in 2013.

Table 10: TFP gains and dispersion in revenue productivity. 1998-2013.

	1998	2003	2008	2013
TFP Gains	156.32	142.92	156.78	210.73
Dispersion in TFPR				
SD	0.99	1.03	1.18	1.15
p75-p25	1.30	1.33	1.58	1.49
p90-p10	2.52	2.61	3.06	2.93

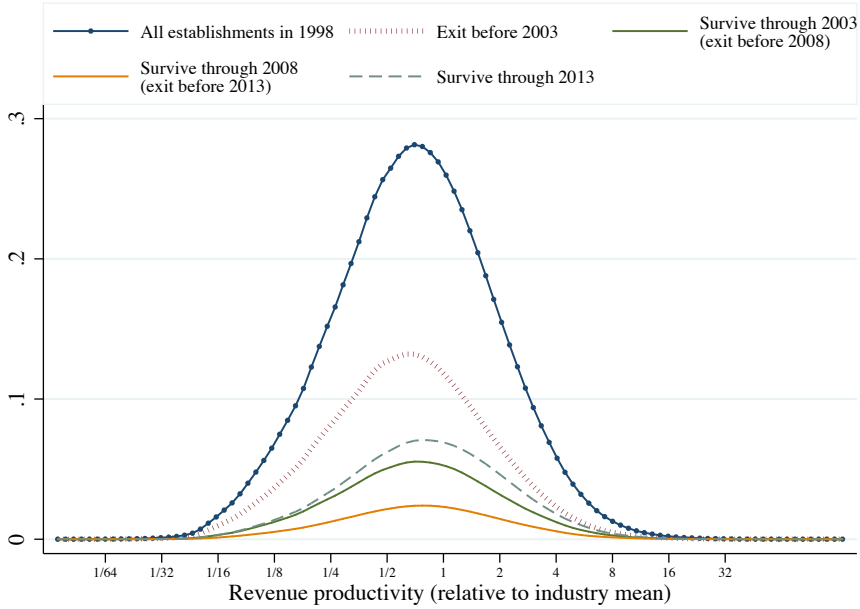
See Appendix for details on the computation of the revenue productivity for each plant following the methodology in Hsieh and Klenow (2009). TFP gains are the gains from equalizing revenue productivity across plants.

As Levy (2018) argues, if low revenue productivity establishments exited, releasing capital and labor, and these resources were absorbed by surviving or entering high revenue productivity establishments, aggregate productivity in Mexico would increase over time. Yet, the evidence in Table 10 indicates persistent resource misallocation, which in turn suggests a dysfunctional process of creative destruction in Mexico.

In this section we follow Levy (2018) and examine whether establishment entry and exit in Mexico are productivity-increasing or productivity-reducing. Exit and entry flows are productivity-increasing when low productivity plants exit or high productivity establishments enter; they are productivity-reducing when high productivity plants exit or low productivity establishments enter. We find that in Mexico in general the potential productivity losses from the exit of high productivity establishments likely offset the potential gains from the exit of low productivity plants, while the losses from the entry of low productivity establishments potentially offset the gains from the entry of high productivity plants. Our analysis suggest that the regulatory and institutional framework seems to account for a significant fraction of the productivity losses associated with establishment entry and exit in Mexico.

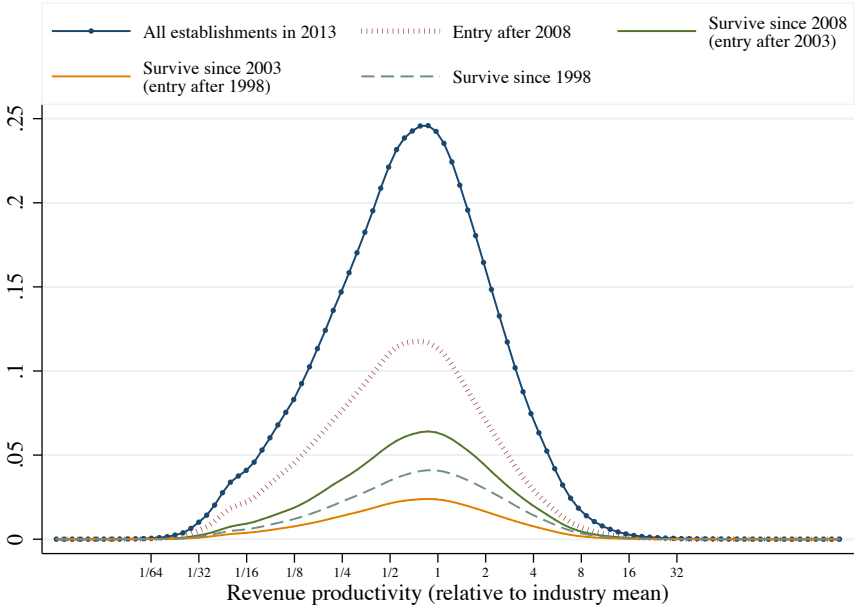
Figure 8 shows the revenue productivity distribution for all establishments in 1998 and separately for exiting and 5, 10, and 15-year surviving plants. The area within each sub-density totals the corresponding fraction of plants (thus, the largest mass corresponds to the density of exiting plants, as half of establishments exited within 5 years). Note that there are not significant differences in the shape of these distributions. For example, the fraction of establishments with TFPR below a quarter of the mean is 12% for plants that survived from 1998 through 2013, 13% for plants that survived from 1998 through 2008, 14% for plants that survived through 2003, and 17% for plants that exited before 2003. Similarly, the fraction of establishments with TFPR above four times the industry mean is 4.1% for plants that survived from 1998 through 2013, 3.9% for plants that survived from 1998 through 2008, 3.8% for plants that survived through 2003, and 3% for plants that exited before 2003. That is, the longer the survival, the lower the fraction of establishments with productivity less than or equal to 25% the industry mean and the higher the fraction of establishments with productivity higher than or equal to 4 times the industry mean, but the differences are not significant and still many low productivity establishments survive, and many high productivity establishments exit.

Figure 8: Revenue productivity distribution in 1998.



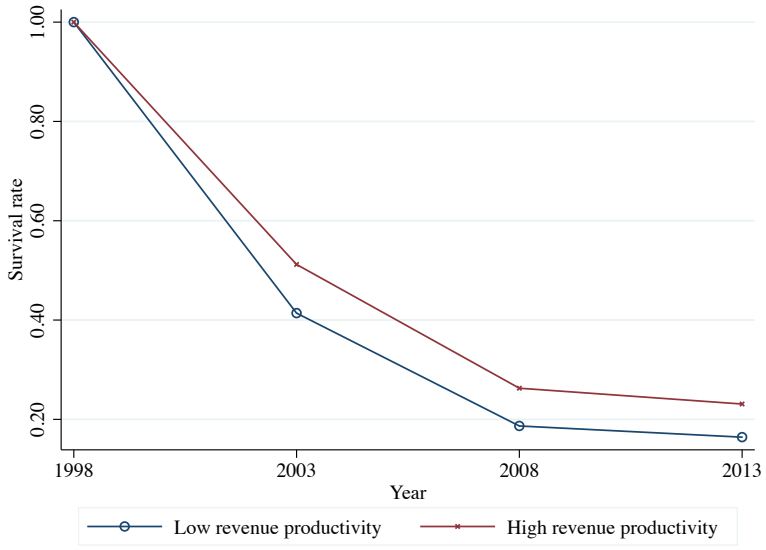
Similarly, Figure 9 shows the revenue productivity distribution for all establishments in 2013 and separately for entering and 5, 10, and 15-year surviving plants. As in the previous figure, there are not significant differences in the shape of these distributions. The fraction of establishments with TFPR below a quarter of the mean is 22% for plants that entered between 2008 and 2013 (which we observe for the first time in 2013), 17% for plants that entered between 2003 and 2008 and survived through 2013, 17% for plants that entered between 1998 and 2003 and survived through 2013, and 16% for plants that survived since 1998 (and entered before then). The fraction of establishments with TFPR above four times the industry mean is 4.2% for plants that entered between 2008 and 2013, 5.1% for plants that entered between 2003 and 2008 and survived through 2013, 4.9% for plants that entered between 1998 and 2003 and survived through 2013, and 5.8% for plants that survived since 1998. Again, the longer the survival, the lower the fraction of establishments with productivity less than or equal to 25% the industry mean and the higher the fraction of establishments with productivity higher than or equal to 4 times the industry mean. In short, not only high productivity establishments survive, but also many low productivity establishments enter.

Figure 9: Revenue productivity distribution in 2013.



We compute next the probability of survival across productivity segments. We sort establishments in 1998 into productivity segments and compute their average survival rates through 2013. Figure 10 shows these survival probabilities for low and high revenue productivity establishments only. Note that 59% of low productivity establishments exited before 2003, 81% exited before 2008, and 84% exited through 2013. High productivity establishments are more likely to survive, yet half exit before *t+5*, 74% exit before *t+10*, and 77% before *t+15*. That is, only 23% of high productivity establishments survived from 1998 through *t+15*. When we control for reported compliance, the results are slightly different. Recall that the analysis in Table 6 shows the estimates from a series of linear probability models controlling for both compliance with social security and productivity segment. The results in row 3 show that low productivity plants are 4.6 percentage points more likely to exit within 5 years (relative to establishments that employ only non-salaried workers and are in the second and third quartiles of the revenue-productivity distribution), 2.2 percentage points more likely to exit within 10 years, and 2.1 percentage points more likely to exit within 15 years. Interestingly, high productivity establishments are only 2 points less likely to exit within 5 years, but after that time threshold, there is no effect on their probability of survival.

Figure 10: Survival rate by segment of revenue productivity. 1998-2013.



In Table 11 we examine whether low and high productivity establishments in 2013 survived or entered between 2008 and 2013. Note 70% of all low productivity establishments in 2013 entered between 2008 and 2013, although that fraction is not markedly different in other productivity segments (62% for establishments in the second and third quartiles, and 58% among high productivity plants). Still, entering establishments take up a larger share in the low productivity segment.

Table 11: Fraction of entering and surviving establishments by productivity segment. 2008-2013.

Revenue productivity	Entering plants	Surviving plants	Total
Low Productivity	69.09	30.91	100
Medium Productivity	61.60	38.40	100
High Productivity	57.81	42.19	100

We sort establishments into productivity segments in 2013 and then examine whether they are new entrants or survived from 2008. We show fractions of the number of establishments in that category in 2013.

Next we examine transitions across productivity segments for surviving plants. Figure 11 shows the revenue productivity distribution of surviving plants between 1998 and 2013. Relative to 1998, the 2013 distribution exhibits a thicker left-hand tail and a lower fraction of establishments in the middle, with only a slightly higher fraction of establishments to the right of the mean. Interestingly, Table 12 shows that surviving low productivity establishments do not explain the higher mass in the left tail of the productivity distribution: only 26% of low productivity establishments in 1998 were still within the low productivity segment in 2013, whereas the remaining 74% moved into the medium or high productivity segment. In other words, the value of additional resources flowing into surviving low productivity establishments increased between 1998 and 2013. In contrast, only 34% of establishments in the high productivity segment in 1998 were still high productivity establishments in 2013.

Figure 11: Revenue productivity distribution of surviving plants. 1998-2013.

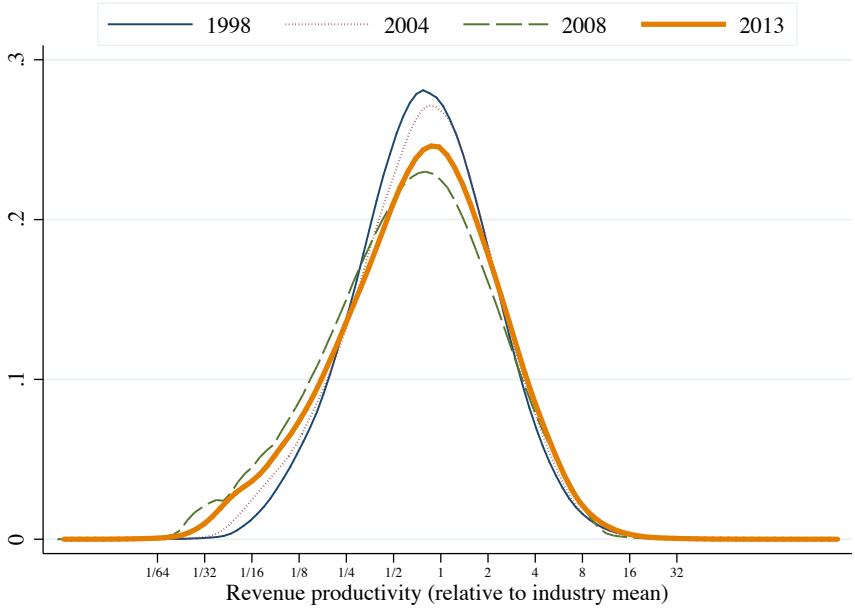


Table 12: Transition across productivity segments between 1998 and 2013.

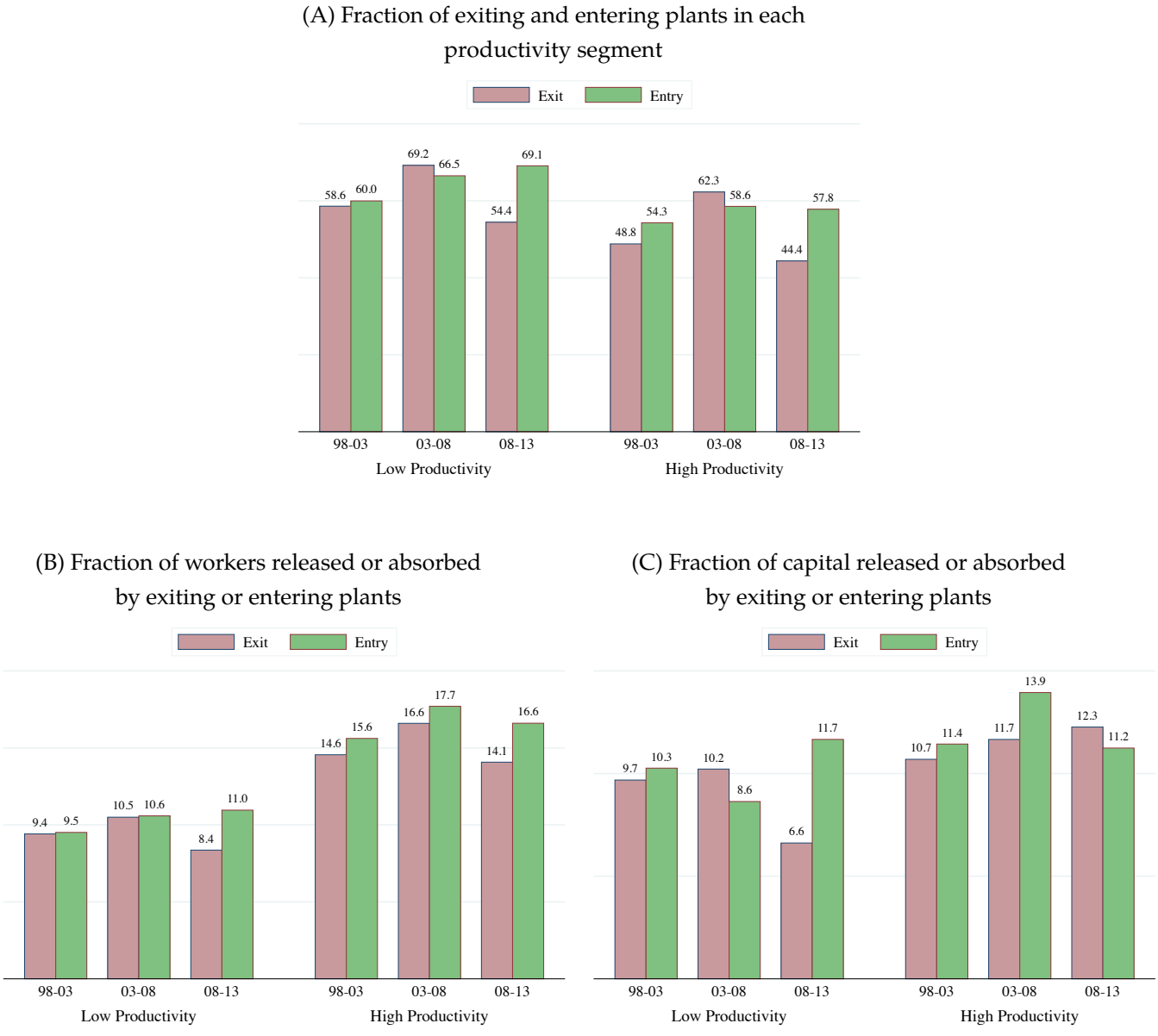
Productivity in 1998	Productivity in 2013			Total
	Low	Medium	High	
Low	**26.14**	49.47	24.39	100
Medium	20.26	**51.71**	28.04	100
High	16.98	49.20	**33.82**	100

Finally, we examine the resource reallocation that follows the entry and exit of low and high productivity establishments. In Panel A of Figure 12 we show the fraction of exiting and entering establishments in each productivity segment for the three consecutive census waves available, and Panels B and C show, respectively, the fraction of workers and capital released or absorbed by these exiting and entering plants. Consider first the low productivity segment of the revenue productivity distribution, where 87% of establishments are legally informal (see Table A7). 54.4% of low productivity establishments in 2008 exited before 2013, which is a productivity-increasing flow. When they left, these plants released 8.4% of workers and 6.6% of the stock of capital. During that same time period, however, many low productivity establishments entered, which is a productivity-decreasing flow: 69% of low productivity establishments in 2013 entered between 2008 and 2013, and these plants absorbed 11% of the stock of workers and 12% of the stock of capital available in 2013.

Consider next the right-hand tail of the distribution, which are establishments relatively more likely to comply. 44.4% of high productivity establishments in 2008 exited before 2013, releasing 14% of labor and 12% of capital, which is a productivity-reducing flow. Offsetting this aggregate productivity loss, high productivity establishments entered during the same time period: 57.8% of high productivity establishments in 2013 entered between 2008 and 2013, absorbing 16.6% of workers and 11.2% of the capital stock. Note in

Figure 12 that these productivity-increasing and productivity-reducing flows are not markedly different if we consider instead entry and exit between 1998 and 2003 or between 2003 and 2008.

Figure 12: Fraction of exiting and entering plants in each productivity segment and their corresponding labor and capital reallocation.



In Panel (A) we show the fraction of entering and exiting plants within each year and each productivity segment. The fraction is relative to plants in that productivity segment in t for exit and t+5 for entry. In Panels (B) and (C) we show the fraction of capital and labor released or absorbed by exiting or entering plants, respectively. The fraction is relative to the total amount of workers and capital in t for exit and t+5 for entry.

Alternatively, note that the productivity losses from the resource reallocation that follows the exit of high productivity establishments potentially offset the productivity gains from the reallocation that follows the exit of low productivity plants—and indeed, when they exit, high productivity establishments release more labor and capital than low productivity plants. In parallel, the productivity losses from the reallocation that

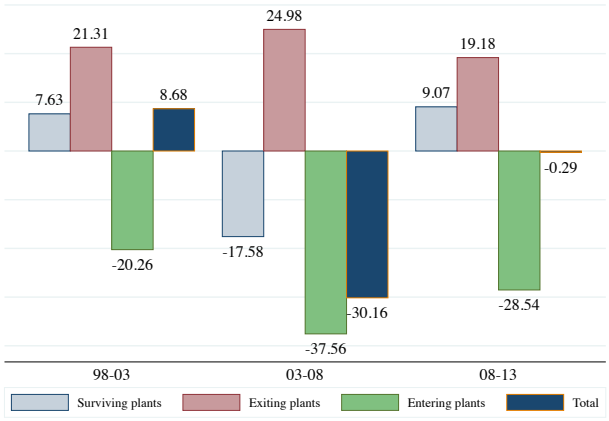
follows the entry of low productivity establishments likely offset the productivity gains from reallocation that follows the entry of high productivity plants: when they enter, low productivity establishments absorb the same fraction of capital (11%) and only 5 percentage points less labor than high productivity plants. In other words, in general establishment entry and exit in Mexico do not optimally reallocate capital and labor, but rather resources are wasted with the exit of high productivity establishments, and the entry and survival of low productivity plants (Levy, 2018).

We decompose the changes in aggregate productivity between 1998 and 2013 to quantify the contributions of surviving, entering, and exiting plants. More precisely, we follow Melitz and Polanec (2015) and decompose changes in aggregate productivity into three easily measured components: (1) the contribution of entry, which is positive (negative) if the productivity of entering establishments is higher (lower) than the productivity of surviving plants; (2) the contribution of surviving establishments, which is positive (negative) if surviving establishments become more (less) productive over time; and (3) the contribution of exit, which is positive (negative) if the productivity of exiting establishments is lower (higher) than the contribution of surviving plants. As productivity measures, we employ value added per worker in real terms, using employment shares as weights, and the Solow residual, with nominal value added shares as weights, and which we obtain using value added, capital, and employment in real terms.

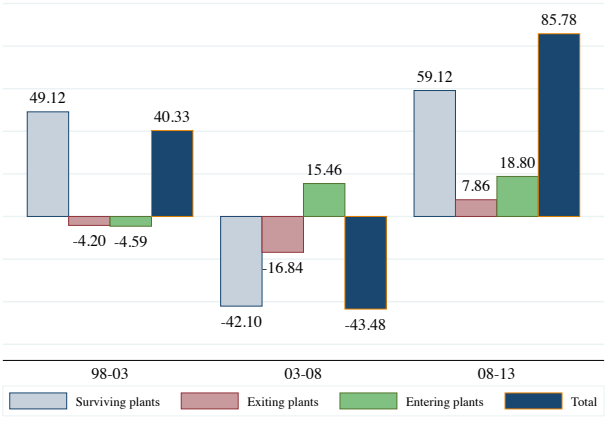
Note in Panel A of Figure 13 that for every pair of consecutive waves, the average value added per worker of new establishments is lower than the value added per worker of surviving plants: the contribution of entering establishments is negative. In other words, process of entry actually decreases value added per worker. Between 2008 and 2013, for example, the contribution of entering plants to the growth in value added per worker was -28.5 percentage points. The process of exit does increase average value added per worker (which indicates that the productivity of exiting plants is lower than the productivity of surviving plants), but these gains are not enough to offset the losses from entry. Surviving plants therefore mainly drove the growth in value added per worker. That is also the case in Panel B, where we measure productivity using the Solow residual. Between 2008 and 2013, entering plants contributed 19 percentage points while exiting plants contributed 8 percentage points, but between 2003 and 2008 their contributions nearly offset each other, and between 1998 and 2003 both contributions were negative. In other words, establishment entry and exit in Mexico are not in general productivity increasing.

Figure 13: Contribution of surviving, exiting, and entering plants to changes in aggregate productivity (percentage points).

A. Value added per worker



B. Solow residual



The regulatory and institutional framework in Mexico potentially accounts for a significant fraction of the

productivity losses associated with establishment entry and exit: too much resources seem to be wasted on the exit of high productivity establishments, and the survival and entry of low productivity plants. In Table 13 we show the fraction of fully formal and legally informal establishments among low and high productivity entering and exiting plants. For example, 68% of establishments in 2008 were legally informal. Yet, legally informal establishments accounted for 74% of low productivity surviving plants. Similarly, fully formal establishments accounted for less than 2% in the cross section, but were almost 4% of high productivity exiting plants. In short, legally informal plants are over-represented in low productivity survival, whereas fully formal plants are over-represented in high productivity exit. Informal but legal establishments are also over-represented in low productivity entry: informal but legal establishments accounted for 87% of low productivity entry, but only 73% of establishments in the cross section in 2013.

Table 13: Fraction of entering and exiting plants in each productivity segment across levels of compliance. 2008 and 2013.

Year	Reported compliance	Fraction of plants in the cross section	Low productivity (%)		High productivity (%)	
2008			Exiting	Surviving	Exiting	Surviving
	Informal but legal	67.64	78.91	74.36	52.15	51.37
	Illegally informal	22.70	17.60	20.42	30.01	29.48
	Mixed	7.90	3.01	4.62	13.94	15.19
	Fully formal	1.76	0.48	0.60	3.90	3.97
	Total	100	100	100	100	100
2013			Entering	Surviving	Entering	Surviving
	Informal but legal	72.76	87.37	84.54	60.60	56.52
	Illegally informal	17.35	10.07	10.42	24.68	24.33
	Mixed	7.28	2.02	4.10	10.51	13.84
	Fully formal	2.60	0.54	0.94	4.21	5.31
	Total	100	100	100	100	100

6. Concluding remarks

The exit, survival, and entry of establishments in Mexico seems fluid in some dimensions but dysfunctional at the same time:

- Aggregate establishment entry and exit in Mexico do not markedly differ from the US, but in Mexico small establishments, in general legally informal, enter and exit at significantly lower rates, whereas larger plants (which usually fully comply) enter and exit at a faster pace. Similarly, young establishments exhibit lower exit rates and lower rates of job creation and destruction. That is, small, young establishments in Mexico seem to behave very differently from their US counterparts.
- Half of both legally informal establishments and fully formal establishments in 2013 entered between 2008 and 2013. Indeed, both legally informal and fully formal establishments enter at the same yearly rate (9%), but when fully formal establishments enter, they are larger more productive than legally informal plants.
- Legally informal establishments (which tend to be small and unproductive) exhibit basically the same

survival rates than fully formal plants (which are larger and more productive). That is, informal but legal establishments are not significantly more likely to exit.

- Informal but legal establishments exit at a faster pace than fully formal plants, but the net exit of legally informal plants is 0: on average, when a legally informal establishment exits, another legally informal establishment enters.

- When they survive, legally informal establishments do not transition into formality nor do they add more workers over their life cycle. Yet, their revenue productivity increases over time, which suggests that, conditional on 15-year survival, they should attract the additional resources necessary for optimality.

- When low productivity establishments exit, a segment where legally informal establishments are over-represented, the gains from their release of capital and labor are offset by the exit of high productivity establishments, where fully formal establishments account for an above-average share. Similarly, when high productivity establishments enter, many low productivity establishments enter as well.

In an economy without distortions, small establishments will enter and exit at higher rates than larger plants, and surviving plants will eventually grow during their life cycle (Rossi-Hansberg and Wright, 2007). If productivity is realized at birth, those with bad draws will immediately leave, and thus the exit rate of small establishments will be disproportionately high. Establishments with successive positive shocks over their life cycle will increase their size, whereas unsuccessful plants will reduce their size and eventually die (and thus larger plants will exit at significantly lower rates). Distortions to relative prices affect these dynamics. They potentially affect the selection of firms: those with bad productivity draws may not ever leave. Moreover, if these distortions disproportionately increase costs for more productive plants (for example, due to higher compliance costs), then surviving plants will tend to invest less as they age relative to plants in undistorted economies. In other words, correlated distortions may disuade plants from investments that increase their productivity (Bento and Restuccia, 2017; Hsieh and Klenow, 2014).

In Mexico, the dynamics of formal and informal establishments are not symmetrical, and actually formal plants do not eventually prevail in the allocation of resources. Our evidence suggests that the special provisions for non-salaried labor in Mexico, combined with other features of the regulatory framework such as size-dependent enforcement and simplified tax regimes, distort both the selection of firms and the incentives for establishment growth. If small establishments do not comply, their probability of detection is low compared to larger plants; if they comply, they do so at preferential rates: they can join a special regime and pay, for example, a single tax on revenues instead of the general income and sales taxes. In addition, they can legally avoid contributing to the social security of their employees and observing other provisions in the labor code if they do not hire them for a salary and employ them instead as non-salaried employees. Increasing the size of the plant requires eventually transitioning into salaried employees and eventually complying with labor regulations (and likely with other provisions as well, such as the general corporate tax). If there are significant fixed costs of compliance, however, then only those plants that are productive enough will ultimately increase their size, while the rest will remain at their sub-optimal level of workers and their sub-optimal organization of tasks between salaried and non-salaried labor. As a result, establishments that only employ non-salaried workers but that do not grow over their life cycle will persistently coexist with productive plants intensive in salaried labor and complying with the provisions in the labor code.

References

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

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

Busso, M., M. V. Fazio, and S. Levy (2012). (In)Formal and (Un)Productive: The Productivity Costs of Excessive Informality in Mexico. *SSRN Electronic Journal 341*(August).

Busso, M., O. Fentanes, and S. Levy (2018). The longitudinal linkage of Mexico's economic census 1999-2014. *Manuscript, Inter-American Development Bank, Washington, DC*.

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

Hsieh, C. T. and P. J. Klenow (2014). The life cycle of plants in India and Mexico. *Quarterly Journal of Economics 129*(3), 1035–1084.

Levy, S. (2018). *Under-Rewarded Efforts: The Elusive Quest for Prosperity in Mexico*. Inter-American Development Bank.

Melitz, M. J. and S. Polanec (2015). Dynamic Olley-Pakes productivity decomposition with entry and exit. *The RAND Journal of Economics 46*(2), 362–375.

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

Rossi-Hansberg, E. and M. L. Wright (2007). Establishment size dynamics in the aggregate economy. *American Economic Review 97*(5), 1639–1666.

Appendix A: Additional tables and figures.

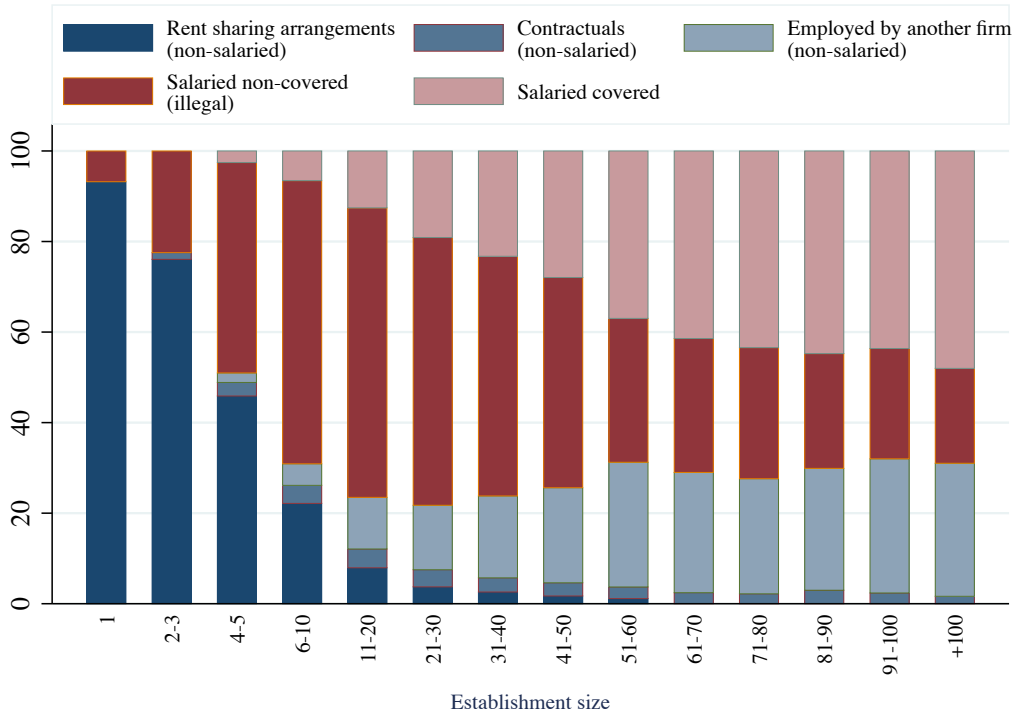
Table A1: Average establishment size.

	1998	2003	2008	2013
Aggregate	4.49	4.53	4.46	4.20
Manufacturing	12.65	12.46	10.43	10.14
Wholesale & retail	2.85	3.16	3.26	3.10
Services	4.04	4.12	4.19	3.78

Table A2: Distribution of workers in the establishment census across contractual arrangements.

	1998	2003	2008	2013
Manufacturing				
Salaried	84	81	70	68
Non-salaried	16	19	30	32
Rent-sharing arrangements	10	12	16	14
Employed by another firm	5	7	13	17
Contractuals	1	1	1	1
Wholesale & retail				
Salaried	51	49	38	40
Non-salaried	49	51	62	60
Rent-sharing arrangements	42	44	48	42
Employed by another firm	5	5	13	16
Contractuals	2	2	1	1
Services				
Salaried	58	57	51	51
Non-salaried	42	43	49	49
Rent-sharing arrangements	31	34	38	37
Employed by another firm	7	5	7	8
Contractuals	5	4	3	4
Aggregate				
Salaried	65	61	52	52
Non-salaried	35	39	48	48
Rent-sharing arrangements	27	31	36	32
Employed by another firm	6	6	11	14
Contractuals	3	2	2	2

Figure A1: Fraction of non-salaried workers and salaried workers covered and non-covered by social security in each establishment size.



Data for 2013 from the establishment census.

Table A3: Distribution of establishments and their resource allocation across sizes and levels of compliance with Social Security. 1998.

		Plants (%)	Workers (%)	Capital (%)	Value added (%)	Relative capital	Relative value added	Capital-labor ratio	Capital-output ratio
Panel A: Size of the plant (number of workers)									
	1	47.08	10.60	5.46	3.78	0.12	0.08	0.51	1.57
	2-3	37.01	18.78	7.40	6.88	0.20	0.19	0.39	1.17
	4-5	6.84	6.70	3.65	4.82	0.53	0.70	0.54	0.83
	6-10	4.57	7.68	5.03	6.60	1.10	1.44	0.65	0.83
	11-20	2.25	7.32	4.69	6.77	2.08	3.01	0.64	0.75
	21-30	0.76	4.24	2.83	4.29	3.74	5.66	0.67	0.72
	31-50	0.60	5.23	4.34	5.23	7.25	8.75	0.83	0.90
	51-100	0.44	6.88	6.49	8.12	14.91	18.65	0.94	0.87
	101-250	0.29	10.30	13.51	13.64	45.82	46.27	1.31	1.08
	251-500	0.10	7.70	13.33	12.29	135.02	124.48	1.73	1.18
	501-1000	0.04	6.63	13.62	10.67	317.95	248.91	2.05	1.39
	+1,000	0.02	7.93	19.66	16.91	924.01	794.92	2.48	1.27
	Total	100	100	100	100	NA	NA	1	1.09
Panel B: Compliance with Social Security contributions									
Only NS		68.03	26.29	13.23	10.68	0.19	0.16	0.50	1.39
S and NS	Fully evades	8.81	6.98	2.51	2.58	0.29	0.29	0.36	1.09
	Partly complies	4.47	16.12	22.27	21.76	4.98	4.86	1.38	1.15
	Fully complies	3.02	7.30	8.16	6.49	2.70	2.15	1.12	1.41
Only S	Fully evades	6.67	5.19	7.88	6.97	1.18	1.05	1.52	1.27
	Partly complies	5.92	26.07	30.29	36.00	5.11	6.08	1.16	0.95
	Fully complies	3.08	12.04	15.65	15.50	5.08	5.03	1.30	1.14
	Total	100	100	100	100	NA	NA	1	1.09

Relative capital refers to the capital of the average plant in that category relative to the capital of the average plant in the economy. A similar definition applies to Relative value added. The capital-labor ratio shows the average amount of capital per worker in that category as a fraction of the capital per worker in the economy. The capital-output ratio is simply capital over value added. S = Salaried workers. NS = Non-salaried workers. Fully evades = Average contribution to social security on salaried workers is 0. Partly complies = Average contribution to social security on salaried workers is between 0 and 18%. Fully complies = Average contribution to social security on salaried workers is at least 18%.

Table A4: Distribution of establishments and their resource allocation across sizes and levels of compliance with Social Security. 2013.

		Plants (%)	Workers (%)	Capital (%)	Value added (%)	Relative capital	Relative value added	Capital-labor ratio	Capital-output ratio
Panel A: Size of the plant (size of the plant)									
	1	43.38	10.45	4.19	5.16	0.10	0.12	0.40	0.79
	2-3	40.91	22.30	7.05	6.17	0.17	0.15	0.32	1.11
	4-5	7.55	7.93	3.89	3.50	0.52	0.46	0.49	1.08
	6-10	4.35	7.78	5.15	6.56	1.18	1.51	0.66	0.76
	11-20	1.95	6.86	5.56	6.64	2.85	3.40	0.81	0.81
	21-30	0.61	3.69	3.73	3.42	6.12	5.62	1.01	1.06
	31-50	0.51	4.87	5.38	5.41	10.49	10.56	1.10	0.97
	51-100	0.35	6.01	9.91	7.56	27.96	21.34	1.65	1.27
	101-250	0.25	9.44	14.07	12.94	55.61	51.13	1.49	1.06
	251-500	0.08	6.38	11.85	10.57	155.01	138.27	1.86	1.09
	501-1000	0.03	5.79	10.17	11.11	292.72	319.75	1.76	0.89
	+1,000	0.02	8.48	19.07	20.96	980.49	1,077.79	2.25	0.88
	Total	100	100	100	100	NA	NA	1	0.97
Panel B: Compliance with Social Security contributions									
Only NS		72.76	39.92	31.69	28.24	0.44	0.39	0.79	1.12
S and NS	Fully evades	10.00	9.36	3.56	3.39	0.36	0.34	0.38	1.05
	Partly complies	2.31	8.98	13.18	12.61	5.72	5.47	1.47	1.05
	Fully complies	2.40	8.93	12.31	10.67	5.14	4.45	1.38	1.16
Only S	Fully evades	7.35	6.81	3.90	5.04	0.53	0.69	0.57	0.78
	Partly complies	2.58	12.16	16.10	20.88	6.24	8.09	1.32	0.77
	Fully complies	2.60	13.85	19.25	19.18	7.40	7.37	1.39	1.01
	Total	100	100	100	100	NA	NA	1	0.97

Relative capital refers to the capital of the average plant in that category relative to the capital of the average plant in the economy. A similar definition applies to Relative value added. The capital-labor ratio shows the average amount of capital per worker in that category as a fraction of the capital per worker in the economy. The capital-output ratio is simply capital over value added. S = Salaried workers. NS = Non-salaried workers. Fully evades = Average contribution to social security on salaried workers is 0. Partly complies = Average contribution to social security on salaried workers is between 0 and 18%. Fully complies = Average contribution to social security on salaried workers is at least 18%.

Table A5: Average difference of (log) TFPQ relative to establishments that employ only non-salaried workers.

	All plants		Sizes [0,5]		Sizes [6,10]		Sizes [11,50]		Sizes 50+	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Type 2	0.512***	0.570***	0.858***	0.867***	0.898***	0.899***	-0.111***	-0.089***	-0.268*	-0.249*
	[0.004]	[0.004]	[0.003]	[0.003]	[0.017]	[0.017]	[0.025]	[0.025]	[0.147]	[0.146]
Type 3	2.438***	1.746***	1.286***	1.249***	1.262***	1.247***	0.961***	0.960***	2.070***	1.871***
	[0.005]	[0.005]	[0.006]	[0.006]	[0.019]	[0.019]	[0.024]	[0.023]	[0.041]	[0.042]
Type 4	2.350***	1.952***	1.300***	1.260***	1.135***	1.113***	0.900***	0.880***	2.397***	2.312***
	[0.005]	[0.004]	[0.006]	[0.006]	[0.019]	[0.019]	[0.022]	[0.022]	[0.044]	[0.044]
Type 5	0.923***	0.974***	1.028***	1.048***	1.101***	1.103***	0.540***	0.580***	-0.000	0.013
	[0.004]	[0.004]	[0.003]	[0.003]	[0.018]	[0.018]	[0.021]	[0.020]	[0.064]	[0.063]
Type 6	2.165***	1.885***	1.553***	1.550***	1.453***	1.444***	0.837***	0.830***	1.808***	1.650***
	[0.004]	[0.004]	[0.005]	[0.005]	[0.018]	[0.018]	[0.019]	[0.019]	[0.040]	[0.041]
Type 7	2.463***	2.016***	1.586***	1.562***	1.350***	1.329***	1.011***	0.979***	1.970***	1.814***
	[0.004]	[0.004]	[0.005]	[0.005]	[0.017]	[0.017]	[0.019]	[0.019]	[0.036]	[0.037]
Constant	-7.032***	-7.310***	-7.636***	-7.779***	-6.821***	-6.904***	-6.076***	-6.265***	-5.116***	-5.319***
	[0.002]	[0.002]	[0.001]	[0.002]	[0.013]	[0.013]	[0.013]	[0.014]	[0.023]	[0.026]
Size		✓								
Age		✓		✓		✓		✓		✓
N	3,090,942	3,090,942	2,831,869	2,831,869	143,312	143,312	96,761	96,761	19,000	19,000
R^2	0.206	0.298	0.094	0.103	0.058	0.061	0.049	0.061	0.236	0.250

We compute the physical productivity (TFPQ) for each plant in the establishment census following the methodology in Hsieh and Klenow (2009). The dependent variable is the deviation of log TFPQ from the industry mean. The omitted category is establishments that employ only non-salaried workers (Type 1). Establishments of types 2-4 employ both salaried and non-salaried. Establishments of type 2 pay no social security contributions on their salaried employees, establishments of type 3 pay an average contribution between 0 and 18%, and establishments of type 4 pay on average a contribution of at least 18%. Establishments of types 5-7 employ only salaried employees. Establishments of type 5 pay no social security contributions on their salaried employees, establishments of type 6 pay an average contribution between 0 and 18%, and establishments of type 7 pay on average a contribution of at least 18%. Standard errors are reported in brackets. *p<0.1, **p<0.05, ***p<0.01.

Table A6: Average difference of (log) TFPR relative to establishments that employ only non-salaried workers.

	All plants		Sizes [0,5]		Sizes [6,10]		Sizes [11,50]		Sizes 50+	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Type 2	0.296***	0.300***	0.280***	0.285***	0.628***	0.628***	0.199***	0.203***	-0.117	-0.122
	[0.002]	[0.002]	[0.002]	[0.002]	[0.010]	[0.010]	[0.014]	[0.014]	[0.081]	[0.080]
Type 3	0.533***	0.492***	0.351***	0.328***	0.657***	0.650***	0.580***	0.580***	0.544***	0.593***
	[0.002]	[0.003]	[0.004]	[0.004]	[0.011]	[0.011]	[0.013]	[0.013]	[0.023]	[0.023]
Type 4	0.629***	0.604***	0.342***	0.319***	0.585***	0.576***	0.584***	0.580***	0.880***	0.901***
	[0.002]	[0.003]	[0.004]	[0.004]	[0.012]	[0.012]	[0.013]	[0.013]	[0.024]	[0.024]
Type 5	0.603***	0.608***	0.495***	0.507***	0.847***	0.848***	0.709***	0.717***	0.822***	0.818***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.011]	[0.011]	[0.012]	[0.012]	[0.035]	[0.035]
Type 6	0.677***	0.653***	0.599***	0.597***	0.909***	0.905***	0.656***	0.654***	0.669***	0.708***
	[0.002]	[0.002]	[0.003]	[0.003]	[0.011]	[0.011]	[0.011]	[0.011]	[0.022]	[0.023]
Type 7	0.762***	0.729***	0.601***	0.587***	0.804***	0.795***	0.757***	0.751***	0.822***	0.861***
	[0.002]	[0.002]	[0.003]	[0.003]	[0.010]	[0.010]	[0.011]	[0.011]	[0.020]	[0.020]
Constant	-0.578***	-0.616***	-0.586***	-0.670***	-0.677***	-0.714***	-0.598***	-0.634***	-0.515***	-0.464***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.008]	[0.008]	[0.007]	[0.008]	[0.013]	[0.014]
Size		✓								
Age		✓		✓		✓		✓		✓
N	3,090,942	3,090,942	2,831,869	2,831,869	143,312	143,312	96,761	96,761	19,000	19,000
R^2	0.074	0.076	0.037	0.044	0.061	0.063	0.069	0.071	0.119	0.122

We compute the revenue productivity (TFPR) for each plant in the establishment census following the methodology in Hsieh and Klenow (2009). The dependent variable is the deviation of log TFPR from the industry mean. The omitted category is establishments that employ only non-salaried workers (Type 1). Establishments of types 2-4 employ both salaried and non-salaried. Establishments of type 2 pay no social security contributions on their salaried employees, establishments of type 3 pay an average contribution between 0 and 18%, and establishments of type 4 pay on average a contribution of at least 18%. Establishments of types 5-7 employ only salaried employees. Establishments of type 5 pay no social security contributions on their salaried employees, establishments of type 6 pay an average contribution between 0 and 18%, and establishments of type 7 pay on average a contribution of at least 18%. Standard errors are reported in brackets. *p<0.1, **p<0.05, ***p<0.01.

Table A7: Resource allocation and fraction of plants of each level of compliance across revenue productivity segments.

Year	Productivity segment	Plants (%)	Workers (%)	Capital (%)	Fraction of informal but legal plants	Fraction of illegally informal plants	Fraction of mixed plants	Fraction of fully formal plants	Total
1998									
	Low	25.0	17.0	17.2	84.2	9.3	5.4	1.0	100
	Medium	50.0	52.1	60.3	65.3	17.8	13.9	3.0	100
	High	25.0	30.9	22.5	57.3	19.1	19.0	4.6	100
	Total	100	100	100	-	-	-	-	
2003									
	Low	25.0	16.6	17.9	86.7	9.0	3.6	0.7	100
	Medium	50.0	54.2	61.7	63.5	22.9	11.3	2.4	100
	High	25.0	29.2	20.4	56.1	25.6	14.6	3.7	100
	Total	100	100	100	-	-	-	-	
2008									
	Low	25.0	16.5	14.6	76.8	18.9	3.7	0.5	100
	Medium	50.0	52.0	60.2	56.0	31.3	10.8	2.0	100
	High	25.0	31.6	25.2	51.7	29.7	14.6	3.9	100
	Total	100	100	100	-	-	-	-	
2013									
	Low	25.0	16.9	18.2	86.5	10.2	2.7	0.7	100
	Medium	50.0	53.0	62.0	66.7	21.7	8.9	2.7	100
	High	25.0	30.1	19.8	58.9	24.5	11.9	4.7	100
	Total	100	100	100	-	-	-	-	

Low productivity plants are plants in the bottom 25% of the revenue productivity distribution; medium productivity plants are in the middle 50% of the revenue productivity distribution; high productivity plants are plants in the top 25% of the revenue productivity distribution.

Table A8: Distribution of formal and informal plants across productivity segments.

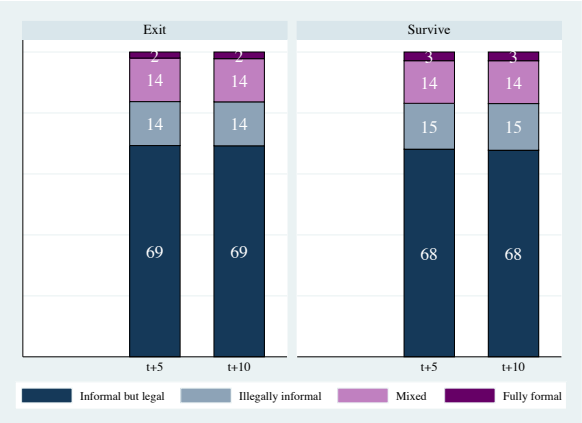
Year	Reported compliance	Low productivity	Medium productivity	High productivity	Total
1998					
	Informal but legal	31.0	48.0	21.0	100.0
	Illegally informal	14.5	55.6	29.9	100.0
	Mixed	10.4	53.3	36.3	100.0
	Fully formal	8.9	51.5	39.6	100.0
2003					
	Informal but legal	32.1	47.1	20.8	100.0
	Illegally informal	11.2	57.0	31.8	100.0
	Mixed	8.9	55.3	35.9	100.0
	Fully formal	7.9	51.5	40.6	100.0
2008					
	Informal but legal	31.9	46.5	21.5	100.0
	Illegally informal	17.0	56.3	26.7	100.0
	Mixed	9.4	53.9	36.7	100.0
	Fully formal	6.3	46.9	46.7	100.0
2013					
	Informal but legal	31.0	47.8	21.1	100.0
	Illegally informal	13.0	55.6	31.4	100.0
	Mixed	8.2	54.9	36.8	100.0
	Fully formal	6.2	50.3	43.5	100.0

Low productivity plants are plants in the bottom 25% of the revenue productivity distribution; medium productivity plants are in the middle 50% of the revenue productivity distribution; high productivity plants are plants in the top 25% of the revenue productivity distribution.

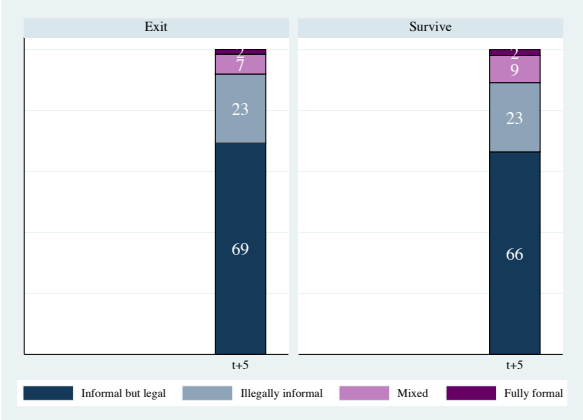
Figure A2: Distribution of entering, surviving, and exiting plants across levels of compliance with social security contributions.

Plants alive in 1998: exit and survival through 2013

Plants alive in 2003: exit and survival through 2013



Plants alive in 2008: exit and survival through 2013



Entry between t and t+5

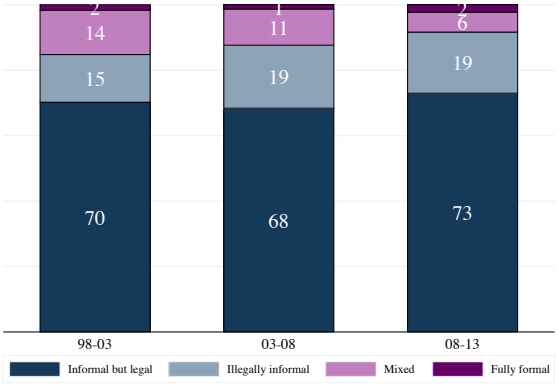
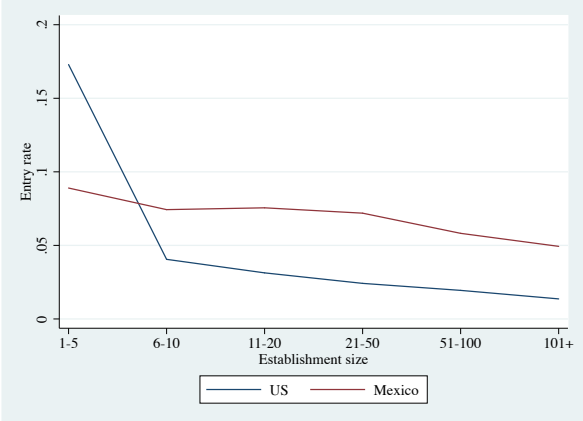


Figure A3: Entry and exit of establishments by size in the US and Mexico.

Establishment entry



Establishment exit

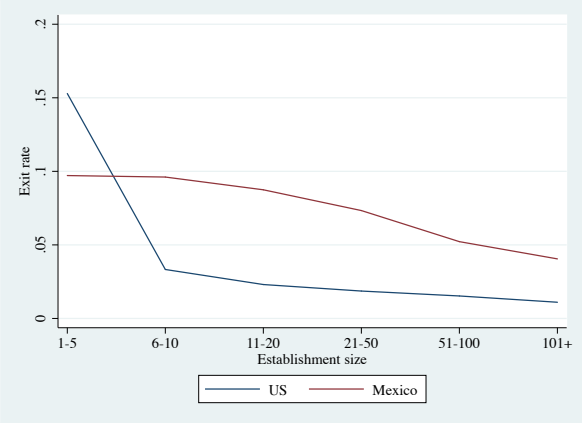


Figure A4: Rates of job creation and destruction by establishment size in the US and Mexico.

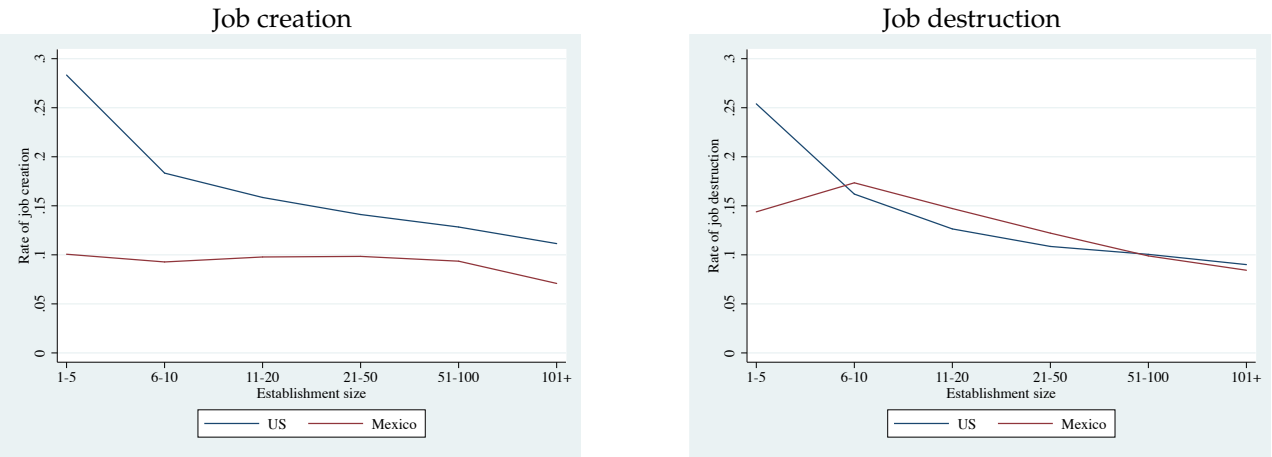


Table A9: Establishment dynamics by level of compliance with social security. Average annual rates for the period 2008-2013.

	Entry	**Exit**	**Job creation**	**Job destruction**
Informal but legal	9.08	9.56	9.75	12.5
Fully formal	8.75	5.84	8.27	8.01

Appendix B: Data Appendix.

We exploit data from the four waves the Mexican Economic Census for the years 1998, 2003, 2008, and 2013. The census is administered by the National Institute of Statistics and Geography of Mexico (INEGI) and covers all localities in Mexico with more than 2,500 inhabitants.[9] The unit of observation is a fixed or partly fixed establishment, or an establishment within a household.

The Economic Census covers establishments in Agriculture, Mining, Electricity, Construction, Manufacturing, Wholesale and retail, Transportation, Services, and Governmental Activities. The codification of industries follows the North American Industrial Classification System (NAICS). We focus only on Manufacturing, Wholesale and retail, and Services because the other sectors are either highly concentrated or mainly dominated by State-owned firms. Table B1 shows the number of establishments captured in these 3 sectors.

Table B1: Establishments by Year and Sector in the Economic Census.

Sector	1999	2004	2009	2014
Commerce	1,442,624	1,580,587	1,858,550	2,042,641
Manufacturing	342,659	328,718	436,851	489,530
Services	937,540	1,013,743	1,367,287	1,637,362
Total	2,722,823	2,923,048	3,662,688	4,169,533

In addition, we restrict our sample as follows:

1. We exclude the subsector *52 Financial Services* because there the unit of observation is the firm and not the establishment.

2. We exclude industries that presented a decrease in capital not consistent with the growth reported in National Accounts (see Table B4).

3. We exclude industry *5613 Employment Services* (or Outsourcing) because many of these establishments reported their employment at the firm level (not at the establishment level), which is inconsistent with the rest of the Census.

4. We exclude establishments belonging to firms that reported the same information for 10 or more of their subsidiaries. Note that less than 3% of all establishments belong to multi-establishment firms, and among them, less than a tenth reports the same information for all subsidiaries.

[9] A locality is usually equivalent to a municipality. 95% of all municipalities have only one locality. The remaining 5% are large municipalities with 2 to 10 localities.

Table B2: Industries with Capital Misreporting

Sector	Code	Industry
Manufacturing	312142	Distilleries
	312221	Cigarette Manufacturing
	321920	Wood Container and Pallet Manufacturing
	321999	Other Wood Product Manufacturing
	322122	Newsprint Mills
	324110	Petroleum Refineries
	327213	Glass Container Manufacturing
	327310	Cement Manufacturing
	334110	Computer Manufacturing
Services	512112	Motion Picture Production
	513120	Television Broadcasting
	531113	Lessors of Residential Buildings and Dwellings
	531114	Lessors of Residential Buildings and Dwellings
	532411	Commercial Transportation Equipment Rental
	561110	Office Administrative Services
	611311	Colleges, Universities, and Professional Schools
	622111	General Medical and Surgical Hospitals
	711311	Promoters of Arts, Sports with Facilities
	713941	Fitness and Recreational Sports Centers
	721111	Hotels (except Casino Hotels) and Motels

Table B3 shows the size of our final sample which consists of around 98% of all establishments reported in the census.

Table B3: Final Sample

Sector	1999	2004	2009	2014
Commerce	1,442,182	1,579,491	1,849,755	2,018,882
Manufacturing	338,947	324,701	432,506	485,319
Services	911,051	979,483	1,311,374	1,567,147
Total	2,692,180	2,883,675	3,593,635	4,071,348

Table B4: Percentage of Establishments dropped by criteria

Criteria	1999	2004	2009	2014
Final Sample	98.9	98.7	98.1	97.6
Financial Services	0.2	0.4	0.5	0.6
Capital Misreporting	0.8	0.9	1.1	1.1
Outsourcing	0.1	0.1	0.1	0.1
Multiplant Misreporting	0.0	0.1	0.2	0.6
Total	100.0	100.0	100.0	100.0

We list the variables we exploit in Table B5. Note that even if in a given year the code for a variable may change, the definition has remained constant across the four waves.

Table B5: Variable Codes and Description

1999	2004	2009	2014	Description
popto	h101a	h101a	h101a	Blue Collar Workers
emact	h200a	h200_a	h203a	White Collar Workers
pnrt	h300a	h300_a	h020a	Owners and Family Members
torst	i100a	i100a	i100a	Outsourced
pch	i200a	i200a	i200a	Contractual Workers
sa	j101a	j100a	j010a	Salaries (Blue Collar)
su	j200a	j200a	j203a	Wages (White Collar)
pagsub	k610a	k610a	k610a	Outsourcing Expenditures
pcys	k620a	k620a	k620a	Contractual Workers Expenditures
pbt	a111a	a111a	a111a	Production
it	a121a	a121a	a121a	Inputs
tida	m000a	m000a	m000a	Income
tgda	k000a	k000a	k000a	Expenditures
afts	q000a	q000a	q000a	Fixed Assets
cpss	j300a	j300a	j300a	Social Insurance Payments
N/A	N/A	k001a	k001	Value Added Taxes Payed
ogfis	l210	l210	l210	Income Taxes
iin	l220	l220	l220	Other Taxes
N/A	N/A	N/A	j301	Workers Covered by Social Insurance
vacb	a131a	a131a	a131	Value Added
N/A	N/A	o720e	o720e	Gains or Loses
abmi	k500a	k500a	k050a	Rented Capital Expenditures
ops	j400a	j400a	j400a	Other Social Benefits to Workers
ur	j500a	j500a	j500a	Profits Sharing
g111	g111	g111	g111	Year of Birth of the Establishment
e23	e23	e23	e23	Enterprise Identifier
f01	f01	f01	f01	Headquarter Identifier

The census questionnaire does not ask about the remunerations for co-owners and family members who work as employees. We follow Busso et al. (2012) to impute a remuneration for these workers, which is needed to apply Hsieh and Klenow's model. Even though the expenditures on outsourcing and contractual workers are reported, we impute these earnings too because these variables present significant misreporting. The imputation process consists of 3 steps:

1. We compute the average remuneration of paid workers (blue and white collar) from the same industry

and State and working in establishments with 10 or fewer workers. Then we impute this average to all unpaid workers from the same Industry and State.[10] This step covers almost 99% of establishments.

2. If an establishment was not covered by step 1, we repeat the process but omitting State in the imputation process.

3. If an establishment was not covered by steps 1 and 2, we repeat the process but omitting both State and Industry.

We concentrate on establishments from the same State so the opportunity costs are calculated taking as a reference the local labor market. Moreover, by concentrating on establishments of the same size and in the same six-digit sector, workers with similar abilities are considered.

The variable that captures Fixed Assets in the census considers only assets that belong to the business. That is, rented capital is not included in that variable even if it is also used in the production process. We therefore define a new measure of capital as follows:

$$Adjusted\,Capital \;=\; Own\,Capital \;+\; \frac{RentedCapital}{0.10}$$

This measure assumes the rented capital as owned by the establishment and is capitalized at a 10% interest rate. Since we treat rented capital as belonging to the business, we return these expenditures to the value added. Thus, our adjusted measure of value added is:

$$Adjusted\,Value\;Added \;=\; Value\;Added \;+\; Rented\,Capital$$

The Hsieh and Klenow model assumes monopolistic competition. We then drop all the establishments in industries with less than 10 establishments since presumably they would be incompatible with the model. To compute the measures of distortions and productivity, we use the imputations defined in the previous section.

- We measure labor as the wage bill, that is, the product of the average wage and the number of workers.

$$w_s L_{si} = Wage\;bill$$

- The average wage by industry is computed as:

$$w_s = \frac{\sum_i^s Wage\;bill_i}{\sum_i^s Workers_i}$$

- The wage bill of an establishment is computed as:

$$Wage\,bill_i = Salaries + Wages + (Imputed\,salary) * (Owners + Outsourcing + Contractual\,Workers)$$

- Total number of workers considers all arrangements:

$$Workers_i = Blue\,Collar + White\,Collar + Owners + Outsourcing + Contractual\,Workers$$

- Capital is measured as the sum of own capital plus the capitalization of rented capital at a rate of 10%.

- Value added is the adjusted version, which excludes expenditures in rented capital.

[10] Unpaid Workers = Owners and Family Members + Outsourced + Contractuals.

Hsieh and Klenow assume a Cobb-Douglas production function with capital share α_s differing across industries (same for labor share $1 - \alpha_s$). Our capital share is computed as:

$$\alpha_s = 1 - \frac{\sum_i^s w_s L_{si}}{\sum_i^s Value\ Added_i}$$

We drop all establishments whose α_s is non positive or greater than one. We trim the percentiles 1 and 99 of the distribution (within industry) of $TFPQ_{si}$, $\tau_{Y_s i}$ and $\tau_{K_s i}$. All our regressions are weighted by Value Added. Finally, when we use measures of $TFPQ$ and $TFPR$, we refer to deviations from industry means.