



WP/19/87


IMF Working Paper

Firm Size, Life Cycle Dynamics and Growth Constraints: Evidence from Mexico

by Florian Misch and Christian Saborowski

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

I N T E R N A T I O N A L M O N E T A R Y F U N D



WP/19/87

IMF Working Paper

Firm Size, Life Cycle Dynamics and Growth Constraints: Evidence from Mexico

by Florian Misch and Christian Saborowski

***IMF Working Papers* describe research in progress by the author(s) and are published to elicit comments and to encourage debate.** The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

I N T E R N A T I O N A L M O N E T A R Y F U N D

IMF Working Paper

Western Hemisphere Department

Firm Size, Life Cycle Dynamics and Growth Constraints: Evidence from Mexico*

Prepared by Florian Misch and Christian Saborowski

Authorized for distribution by Costas Christou

May 2019

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

Abstract

This paper examines the variation in life cycle growth across the universe of Mexican firms. We establish two stylized facts to motivate our analysis: first, we show that firm size matters for development by illustrating a close correlation with state-level per capita incomes. Second, we show that few firms grow as much as their U.S. peers while the majority stagnates at less than twice their initial size. To gain insights into the distinguishing characteristics of the two groups, we then econometrically decompose life cycle growth across firms. We find that firms that have financial access and multiple establishments and that are formal, part of diversified industries and located in population centers can grow at sizeable rates.

JEL Classification Numbers: D22, O12, O40

Keywords: Firm size, Firm growth, Life cycle dynamics, Distortions

Author's E-Mail Address: fmisch@imf.org; csaborowski@imf.org

* The views expressed in this paper are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management. We thank Costas Christou, Alfonso Cebreros, Gabriel Cuadra, Oscar Contreras, and Aldo Heffner for their comments and Juan Pablo Cuesta Aguirre for excellent research assistance. We are also grateful to Natalia Volkow and her team at INEGI for their support. The paper uses confidential data from Mexico's Economic Census, accessed through INEGI's Microdata Laboratory. All results and information shown were compiled by the authors and are not part of INEGI's official statistics.

Contents	Page
I. INTRODUCTION	3
II. CALCULATING FIRM LIFE CYCLE GROWTH	5
A. Previous Studies	5
B. This Paper	6
III. STYLIZED FACTS	7
A. Income Elasticity of Firm Size	7
B. Variation in Firm Life Cycle Dynamics	9
IV. LIFE CYCLE GROWTH DECOMPOSITION	10
A. Descriptive Decomposition	10
B. Econometric Decomposition—Baseline	12
C. Econometric Decomposition—Extensions	16
V. CONCLUSION	19
VI. REFERENCES	20
 Figures	
1. Firm Size and Subnational Income	8
2. Manufacturing Firm Size and Subnational Income	9
3. Average Firm Growth Over the Life Cycle	10
4. Cumulative Growth by 3-Digit Industry at age 15-19: Top and Bottom Five 3-digit	11
5. Top Performers: Firm Size and Birth Cohorts	12
6. Predicted Conditional Life Cycles	17
 Tables	
1. Baseline Regressions	15
2. Baseline Regressions	16
3. Manufacturing vs. Services	18
 Appendices	
1. Variable Definitions and Descriptive Statistics	22
2. Robustness Checks	23

I. INTRODUCTION

Mexico's low productivity growth in recent decades has puzzled many observers (Levy and Rodrik, 2017). Slow technology diffusion is unlikely to be the main explanation given that Mexico has benefited from large foreign investment inflows after opening up its economy as part of a series of sweeping structural reforms since the 1990s. A recent strand of work argues that Mexico's productivity has been held back in part because the allocation of labor and capital is distorted (e.g. Levy, 2018; Misch and Saborowski, 2018).¹

In this paper, we focus on within-firm growth as a complementary explanation for Mexico's low productivity growth.² We argue that individual firms do not invest enough, thus depressing firm growth and preventing firms from taking sufficient advantage of economies of scale. The paper calculates individual firm growth over the life cycle based on five waves of the Mexican Economic Census from 1993 to 2013, complemented by longitudinal firm identifiers constructed by Busso and others (2018) and expanded by INEGI staff. We then decompose our estimates of life cycle growth at the firm level to better understand the distinguishing features of firms that grow at sizeable rates compared to those that stagnate.

We establish two stylized facts to motivate our analysis. First, we show that firm size is positively correlated with state-level per capita incomes. The finding complements previous results at the cross-country level and highlights the close association between firm growth and economic development (Bento and Restuccia, 2018). Second, we show that there is a small minority of Mexican firms that does grow at rates similar to the average U.S. firm. While we confirm the finding of Hsieh and Klenow (2014) that the vast majority of Mexican firms stagnate at about 15-19 years of age, there are some firms that do grow continuously to several times their initial size by the age of 20-24.³ Given the close association between firm size and development, it is thus natural to ask what it is that allows this small minority of firms to perform well in an environment in which the majority does not.

We decompose firm life cycle growth into the contributions of individual firm characteristics using a simple regression framework in the second part of the analysis. Our findings suggest that Mexican firms that have financial access and multiple establishments and that are formal, part of diversified industries and located in population centers can grow at sizeable rates. The regressions predict that a firm that is both formal and has financial access would grow to some

¹ This approach is in line with the Schumpeterian view that emphasizes reallocation of output away from less to more productive firms as a source of aggregate productivity growth (van Reenen, 2018) and builds on a large literature on resource misallocation following Hsieh and Klenow (2009).

² Throughout the paper, we focus on the growth of individual establishments which we refer to as 'firm growth' for simplicity. We recognize however that establishment and firm and establishment growth could differ if firms grow through increasing the number rather than the size of establishments. In the context of Mexico, establishment and firm-level growth are identical in most cases as only very few firms have more than one establishment.

³ These findings are similar to Eslava and Haltiwanger (2017) who find evidence highlighting that only a small proportion of Colombian firms manages to grow at rates similar to the average firm in the U.S. economy.

2.4 times its initial size. Formal firms with financial access and multiple establishments, in turn, are predicted to grow to 3-4 times their initial size over their life cycles, an order of magnitude similar to that of the average U.S. firm (Hsieh and Klenow, 2014). We further illustrate that these ‘superstar’ firms tend to be very large firms in industries that form part of the North American supply chain.

We contribute to the existing literature in several ways. First, by using data that is perfectly comparable across Mexican states, our paper lends greater credibility to estimates in the literature that find evidence of a positive link between firm size and development. These include the findings by Bento and Restuccia (2017) and Bento and Restuccia (2018) who estimate an elasticity of establishment size with respect to GDP per capita of around 0.3.⁴ Second, we provide a more nuanced view on firm life cycle dynamics in Mexico than Hsieh and Klenow (2014). We document not only that the average Mexican firm grows less than the average U.S. firm, but also that firms that have managed to overcome a well identified set of distortions are able to grow at rates similar to the average U.S. firm. Third, our findings provide empirical evidence in support of the theoretical proposition that distortions that constrain productivity through misallocation of resources between more and less productive firms also depress investment by individual firms (Hsieh and Klenow, 2014, and Bento and Restuccia, 2017).⁵ We show that life cycle dynamics can indeed be decomposed into a very similar set of distortions to those identified in Misch and Saborowski (2018) as drivers of resource misallocation in Mexico.

From a policy perspective our findings highlight the importance of pushing ahead with structural reforms in priority areas. This includes efforts to reduce informality, combat the undue use of market power in concentrated industries and strengthen access to financial services. Furthermore, targeted infrastructure investments would help better connecting more remote regions to the major population centers in the country.

The paper is organized as follows. Section 2 discusses the data and methodology employed in calculating life cycle dynamics. Section 3 presents stylized facts. Section 4 presents the descriptive and econometric decompositions of firm life cycle growth, and Section 5 concludes.

⁴ Costa-Scottini (2018) and Bento and Restuccia (2017) review this literature. Poschke (2018) finds evidence of a relatively large income elasticity of around 0.75-0.8 using two separate data sources for small and medium firms and for large firms. Schoellman and others (2017) use data from labor force surveys to show that the share of medium sized (small) firms increases (falls) with increases in per capita income across countries. Earlier papers including Alfaro et al. (2009) find a negative relationship. This could be due to the fact that they use firm-level data which does not cover firms of all sizes. However, using Enterprise Survey data which are not representative of the universe of firms, García-Santana and Ramos (2015) still demonstrate a close correlation between the share of labor allocated to small plants and per capita incomes and examine the distortions that are associated with the share of labor allocated to small plants.

⁵ In their model, Bento and Restuccia (2017) demonstrate that the extent to which distortions that effectively penalize more productive relative to less productive establishments discourage productivity investment by all firms.

II. CALCULATING FIRM LIFE CYCLE GROWTH

A. Previous Studies

Data constraints often make it difficult to compute life cycle growth at the firm level. In an ideal setting, longitudinal information would be available for all firms in a country over a long time span. Firm growth over the lifecycle would then be computed as the ratio of firm size at a given age divided by the size of the firm at the time of its birth. However, most firm-level data sets available in developing economies do not fulfill these conditions. Typically, information at the firm level is either not available over an extended period, lacks firm identifiers or does not cover a representative sample of firms.

Two of the more well-known papers in the life cycle literature include Hsieh and Klenow (2014) and Eslava and Haltiwanger (2017). Hsieh and Klenow (2014) make use of the Mexican Economic Census as we do in this paper. The main advantage of the Mexican Economic Census is that it comprises information on the full universe of non-agricultural firms with a fixed location in Mexico, including micro and informal firms. The survey has been compiled every five years for several decades and thus provides a sufficiently extensive period of observation to study firm life cycle growth. The main downside is that the raw data does not comprise firm-level identifiers for more than the last two (2008 and 2013) waves of the Census.

Hsieh and Klenow (2014) use the Mexican Economic Census to calculate life cycles. In the absence of firm-level longitudinal information, the authors focus on growth rates in average firm sizes within highly disaggregated industries and age groups from the 2003 wave of the Economic Census to the 2008 wave.⁶ By combining average industry-wide firm growth rates across age groups into one life cycle, the approach has several weaknesses. First, while each growth rate from one age group to the next in the life cycle is based on data for a specific cohort of firms, the calculation does not follow the same cohort of firms throughout the life cycle. Second, the computation of the life cycle includes firms that cannot be observed since the time of their birth. Third, the calculation suffers from a likely attrition bias in that lower age-groups of firms may include a larger number of small and unproductive firms in younger age groups that do not survive to older ones.

Eslava and Haltiwanger (2017) avoid the problems in Hsieh and Klenow (2014) by taking advantage of rich microdata on Colombian establishments. The authors are able to make use of annual data on a sample of firms over a long time horizon which includes information on prices and quantities. The key disadvantage of their data from the perspective of estimating average life cycles is that it is not as representative as the Mexican Economic Census since

⁶ In particular, they use the 2003 and 2008 waves of the Census to calculate the growth rate of the average size of firms in a given industry and age group in the 2008 wave relative to firms with five years less of age in the same industry in the 2003 wave. The ratio is then used as the growth rate from one age group to the next for firms in that industry.

small and informal firms are underrepresented (due to exclusion criteria based on size that exclude micro establishments).

B. This Paper

In this paper, we make use of data from five waves of the Mexican Economic Census from 1993 to 2013.⁷ We complement this data with the recent work of Busso and others (2018) – expanded in coverage by INEGI staff – who use an innovative methodology to construct firm identifiers for the Mexican Economic Census, thereby obtaining longitudinal firm data covering the 1993-2013 period. In particular, they develop a sophisticated algorithm that links establishments based on location, legal entity, industry and other characteristics that are plausibly time-invariant. They then validate their procedure using the panel identifier that is already built into the data of the last two waves.

We calculate firm life cycle growth as the cumulative growth rate in a firm's number of employees from one wave of the Economic Census to the next. We associate each firm at each point in time with a 5-year age group based on information about its birth year (0-4 years, 5-9 years, 10-14 years, 15-19 years and 20-24 years). We then restrict the sample of firms in three ways. First, we only include firms that we can observe since their birth to ensure that we do not miss important dynamics (e.g. we exclude any firms that were older than 0-4 years in 1993). Second, we only include firms that we can observe continuously since their birth (e.g. for all 5 waves in the case of firms born in 1993, for 4 waves in the case of firms born in 1998, etc.) to ensure a balanced sample when aggregating life cycles across firms for a given birth cohort. In other words, we drop firms that cease to exist, or that cannot be linked through the identifier or only exist for one period in our sample (those born around 2013). Finally, we drop firms that report birth years that differ by more than 4 years across waves as an additional check on the firm identifier methodology. This leaves us with 25,653 firms born in the 1993 wave, 56,188 firms born in the 1998 wave, 139,475 firms born in the 2003 wave and 524,854 firms born in the 2008 wave.

By calculating life cycles at the firm level rather than at the industry level, our approach avoids the problems associated with the methodology of Hsieh and Klenow (2014). Compared to Eslava and Haltiwanger (2017), our approach has the advantage that the Mexican Economic Census reflects the full universe of surviving firms.

⁷ We exclude sectors in which estimates of firm growth could be misleading or difficult to compare to the remaining sectors, for instance due to the presence of natural monopolies, including financial services, utilities, real estate, professional / technical services sectors as well as the management of shell companies. In addition, we omit health and education as well as arts and culture, and thus sectors in which an important share of the entities involved are unlikely to pursue profit objectives. This leaves manufacturing, retail and wholesale trade, transportation and warehousing, accommodation and food services, information and other services in our sample. As is standard in the literature, we also exclude all entities with negative or zero reported value added, capital, sales or labor input (including labor provided by the owner of the firm) and omit industries with fewer than 10 firms. In addition, we remove the 1 percent tails of the distribution of firm-specific output wedges, capital wedges, and total factor productivity.

III. STYLIZED FACTS

A. Income Elasticity of Firm Size

This subsection examines the association between firm size and per capita GDP across Mexican states. Causation is likely to run both ways: Firms that do not grow are likely to invest less and are not able to take advantage of economies of scale, thereby subduing aggregate productivity. At the same time, richer jurisdictions may provide a more conducive environment for firm growth. For instance, an association between firm size and per capita incomes could be the result of distortive size-dependent policies (Guner et al., 2008).

Existing estimates of the elasticity of firm size with respect to aggregate, country-level income range from around 0.3 to 0.8 in the literature using data from a broad range of sectors (Bento and Restuccia, 2017; Bento and Restuccia, 2018; Poschke, 2018). The cross-country approach taken by these studies suffers from data issues that arise from pooling firm-level data from several countries which often involves having data reported using different definitions of employment and production units or having data that disproportionately include larger firms (Bento and Restuccia, 2017). Using data from one country and examining subnational variation across different states or regions is an obvious remedy for such challenges. However, there are almost no papers to date that take this route.

We estimate the relationship between non-oil GDP per capita and firm size across Mexican states in 2013. We focus on non-oil GDP since significant natural resource revenues may well distort the relationship between firm size and per capita incomes, in particular if the oil sector is relatively isolated from the rest of the economy. In Mexico, oil is extracted primarily offshore and in two of its 32 states, but the revenue is largely transferred to the government. We omit Mexico City which is an outlier in part because of its significant public sector presence.

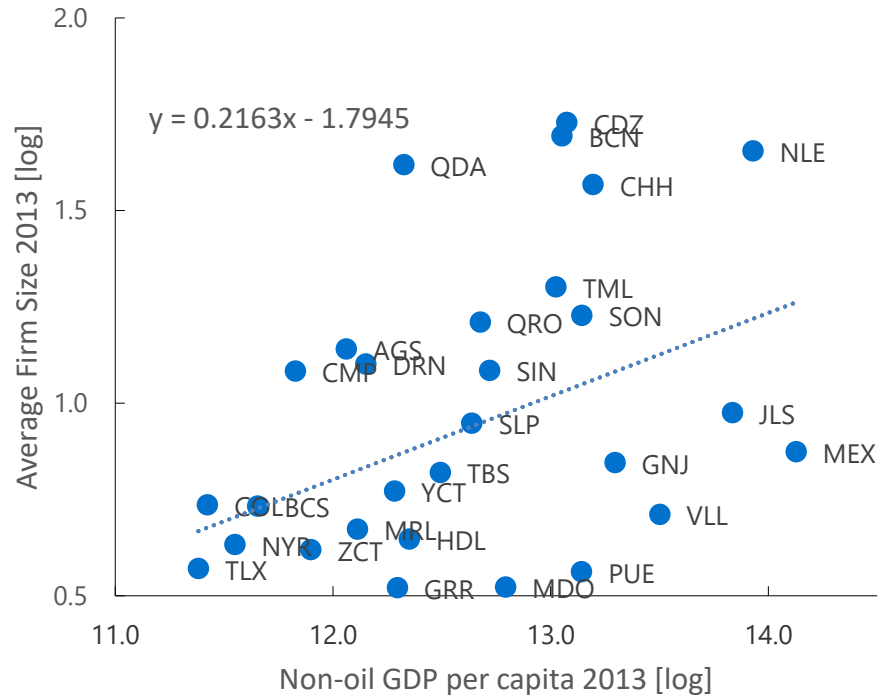
Our data allows controlling for the sectoral composition of firms across states in a very granular way. This is likely to be important given that firm size could differ across states solely because different production technologies imply differences in the optimal firm size. To control for industry composition, we regress firm size across states and industries on 4-digit industry fixed effects and then average the residual across industries in each state to establish our measure of state-level firm size.⁸

We find evidence of a strong positive relationship between state level incomes and average firm sizes. Figure 1 plots the relationship between non-oil GDP per capita and firm size when using all firms and controlling for industry composition. The estimated elasticity is 0.21 which is lower than the cross-country elasticities found in the literature, but higher than the elasticity of 0.16 we find when we do not control for industry composition (not shown in the chart). The correlation coefficient is slightly below 0.4. In Figure 2, we only focus on manufacturing firms but again control for differences in industry composition across states. The elasticity is much

⁸ As the residual is at times negative and we take logs in order to compute the income elasticity of firm size, we rescale the average of the residual so that the minimum across states is 1 before taking logs.

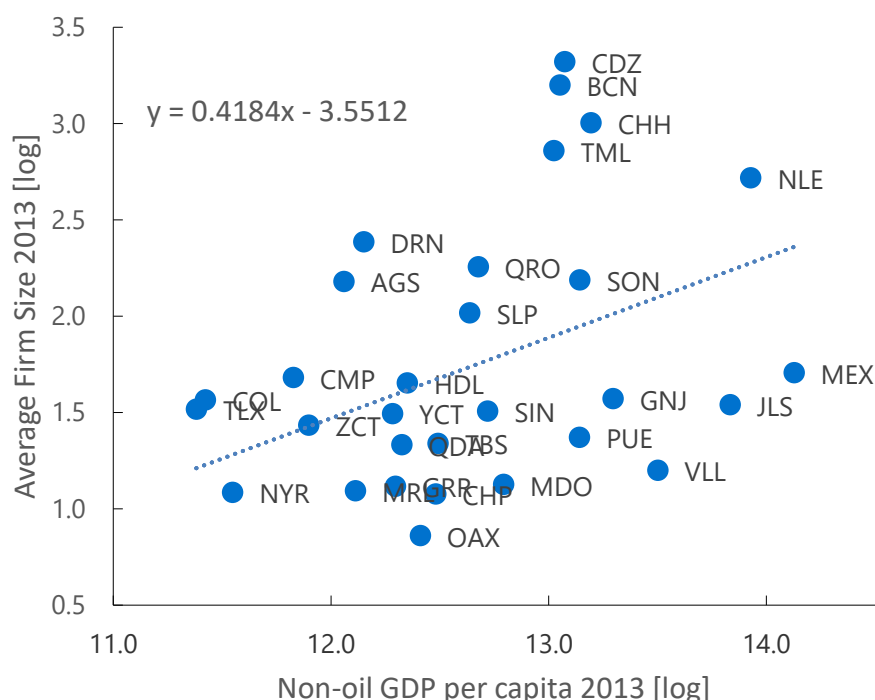
higher, at 0.41. The correlation coefficient is slightly above 0.4. However, if we do not control for differences in industry composition, the elasticity is 0.54, implying that not controlling for differences in industry composition can bias the elasticity estimates in either direction.⁹

Figure 1: Firm Size and Subnational Income



The data includes all firms in our sample and all states other than Mexico City. The firm size variable is corrected for the effects of industry composition. The state-level averages of the residuals were transformed so that the minimum is 1 before taking logs.

⁹ Lozada Cortés (2017) also examines the correlation between firm size and GDP per capita at the level of Mexican states but does not control for industry effects, ignores the role of oil revenue, and the data covers only medium and large firms which represent a very small subsample of the population of Mexican firms. The author estimates a similar elasticity to that established in Bento and Restuccia (2017) of 0.36 but fails to find evidence of a statistically significant link between firm size and income when including services sectors in addition to manufacturing firms.

Figure 2: Manufacturing Firm Size and Subnational Income

Notes: The data includes all manufacturing firms in our sample and all states other than Mexico City. The firm size variable is corrected for the effects of industry composition. The state-level averages of the residuals were transformed so that the minimum is 1 before taking logs.

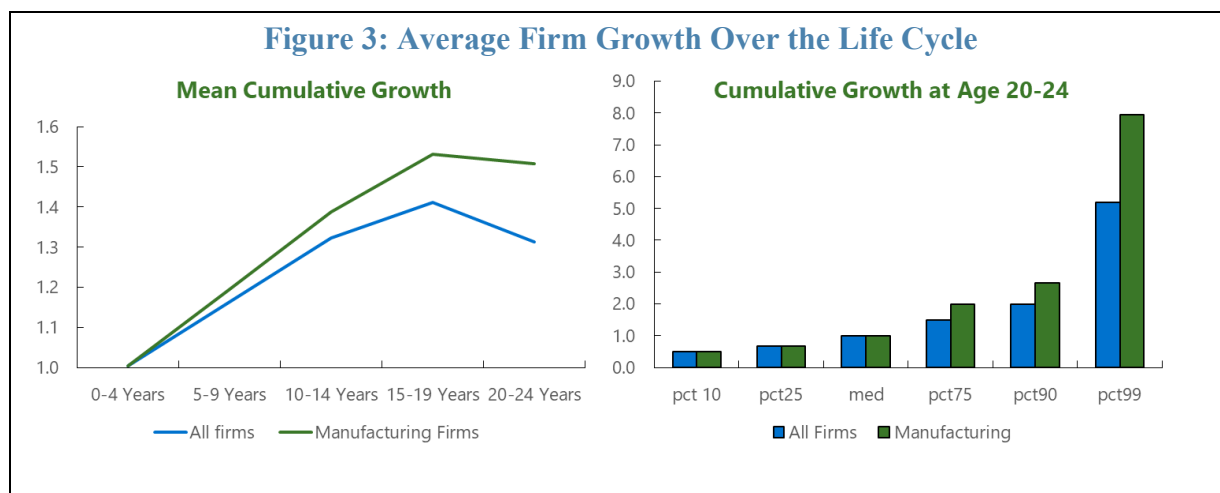
B. Variation in Firm Life Cycle Dynamics

In this subsection, we examine the variation in firm life cycle dynamics, using the data and methodology discussed in the previous section. For each firm, we normalize the number of employees with the number of employees it had when it was in the first age group (0-4 years), i.e., shortly after birth.

The average Mexican firm stagnates in size after some 15-19 years. The left panel of Figure 3 shows simple averages of life cycles across Mexican firms. It illustrates that manufacturing firms in Mexico grow to around 1.5 times their initial size on average until age 15-19 before stagnating below that level. This finding is even more discouraging than the results in Hsieh and Klenow (2014) who suggest that the average manufacturing firm roughly doubles in size over the same time span. For the economy as a whole, the growth rate is even smaller, with firms only expanding to about 1.4 times their initial size by age 15-19.

However, a small minority of firms grows continually at a high rate. The right panel of Figure 3 drills down on the variation of life cycle growth across firms in our sample. In particular, it presents the distribution of cumulative firm life cycle growth in the highest age group (at 20-24 years of age). The chart illustrates that it is indeed a very small share of firms that grow at rates significantly higher than the average. The median firm effectively stagnates at its initial size

while even the 90th percentile of firms grows by less than a factor of three. The 99th percentile of firms, in turn, grows to some five times its initial size (eight times in the case of manufacturing firms). Against the backdrop of the findings of Hsieh and Klenow (2014), namely that U.S. manufacturing firms grow to some 4-6 times their initial size by age 20-24, we may conclude that only a very small minority of Mexico's firms grows as much over the life cycle as the average U.S. firm.



IV. LIFE CYCLE GROWTH DECOMPOSITION

We have so far established two stylized facts: first, we showed that firm size is positively correlated with state-level per capita incomes, implying that firms in poorer states appear to expand less and do not take full advantage of economies of scale. Second, we showed that there is a small share of Mexican firms that significantly outperform the majority of their peers and grow at rates similar to the average U.S. firm. This section aims to better understand the characteristics of firms that grow at sizeable rates, both descriptively (Subsection A) and by means of simple regressions (Subsection B).

A. Descriptive Decomposition

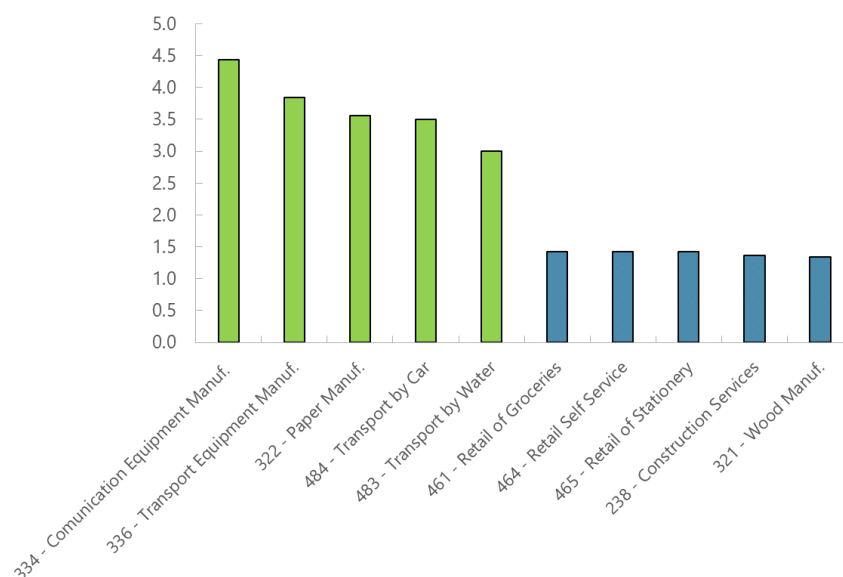
In this subsection, we present simple descriptive statistics to decompose firm life cycle growth across subsamples. The idea is to understand the distinguishing characteristics of those firms that do manage to grow at high rates and thus substantially outperform their peers in the Mexican economy.

Our first observation is that the majority of Mexico's fastest growing firms are those that are part of the export sector and the North American supply chain. Figure 4 shows simple averages of life cycle growth by 3-digit industries after 15-19 years of age.¹⁰ The chart includes the five industries with the highest growth rate and the five industries with the lowest growth rate to

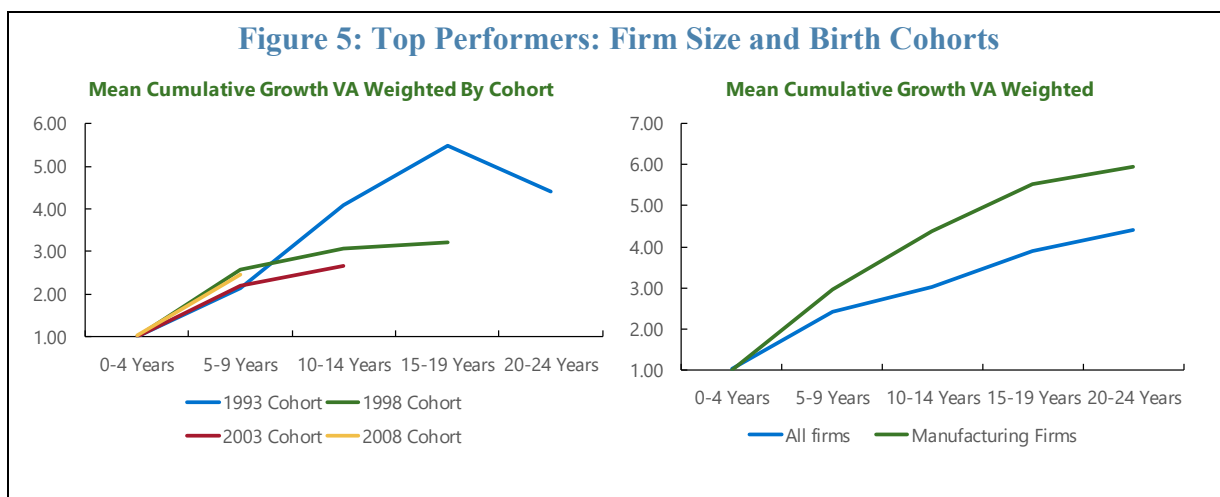
¹⁰ We chose 15-19 years of age rather than 20-24 years of age here in order to increase the number of firms available per industry. In particular, firms in the 1993 cohort are the only ones we can observe until 20-24 years of age while 15-19 years of age can be observed also for the 1998 cohort.

make the point. Among manufacturing industries, the top performers include the manufacture of communication and transport equipment (which includes the manufacturing of cars and car parts for instance), while among services industries they comprise largely transport services. Strikingly, the average cumulative growth rates of the top five industries are larger than those at the 90th percentile of the distribution of life cycle growth across all firms (Figure 2) and not much smaller than those at the 99th percentile. At the other end of the spectrum, firm life cycle growth is particularly weak in retail trade services which tend to include large numbers of micro and informal firms.

Figure 4: Cumulative Growth at age 15-19: Top and Bottom Five 3-digit Industries



In fact, it appears that firms that are part of the 1993 birth cohort have experienced somewhat stronger life cycle growth. Firms in the 1993 cohort are the only ones born prior to the signing of NAFTA. They might thus have had greater opportunities to take advantage of new export markets. The left panel of Figure 5 suggests that in the first years of existence, firms across the different birth cohorts grew at roughly similar rates. However, it does appear that firms born in or shortly before 1993 show somewhat stronger growth throughout the life cycle relative to the 1998 and the 2003 birth cohorts. Of course, it is important to bear in mind that the different cohorts were hit not just by one but several important shocks over their life spans. For example, the cohort of firms born in 1993 was initially hit not just by the enactment of NAFTA but also by the Tequila Crisis while the cohort of 1998 firms initially had to deal with increased competition emanating from China's WTO accession.



Second, very large firms grow faster. We make this point in the left panel of Figure 5 which illustrates average cumulative firm growth over the firms' life cycle. Contrary to Figure 3, we now compute averages weighted by firm-level value added. The average firm (using value added weights) shows a cumulative growth rate of more than four by age 20-24, and around six in the case of manufacturing firms. This is much higher than in Figure 3 which is based on simple averages and above the growth rate of firms at the 90th percentile (Figure 3, left chart). This implies that there is a small number of very large firms that grow significantly faster than their peers.

B. Econometric Decomposition—Baseline

In this subsection, we use simple regressions to further decompose average life cycle growth across firms with different characteristics. The objective is to identify potential distortions that are correlated with the significant heterogeneity in firm-level lifecycle growth that we observe.

We begin by estimating simple average life cycle dynamics as a benchmark. The specification we estimate is shown in Equation 1 where $Q_{f,t}$ is the cumulative growth rate in the number of employees of firm f between birth and wave t and D_{AG} is a set of age group dummies. Note that we do not include a constant in this regression. Moreover, the dummy for age group 0-4 is the one that is excluded, and the dependent variable takes the value 1 for all firms in age group 0-4 by construction. In the absence of a constant in the regression, the average life cycle across firms is thus given by the coefficients on the four age group dummies. As shown in the first column of Table 1, estimating Equation 1 suggests that the average life cycle increases until age group 15-19, before weakening somewhat in age group 20-24. The cumulative growth rate in the final age group reaches between 1.3 and 1.4, in line with the descriptive results in the previous section.

$$Q_{f,t} = \sum_{AG=2}^5 (\beta_{AG} D_{AG}) + \varepsilon_{f,t} \quad (1)$$

We move on to decomposing life cycles using a set of explanatory variables that capture several structural challenges facing the Mexican economy. The baseline specification is

$$Q_{f,t} = \alpha_{s,i,t} + \sum_{AG=2}^5 (\beta_{AG} D_{AG}) + \gamma C_{f,t} + \theta X_{f,t} + \varepsilon_{f,t} \quad (2)$$

where, compared to Equation 1, we include unobserved effects at the state-industry-year level (where industry i , state s and wave t reflect the 4-digit-industry classification, the state where the establishment is located and the wave of the survey, respectively) to control for possible confounding factors that could bias the coefficients on firm-level variables.¹¹ Moreover, we include other firm-level control variables $C_{f,t}$ and our variables of interest $X_{f,t}$. Throughout, we cluster the standard errors at the industry-state level. Appendix Tables A and B include definitions and summary statistics for all variables. In the Appendix, we test the robustness of our baseline results by using lagged explanatory variables and alternative fixed effects as part of an admittedly imperfect effort to address endogeneity-related concerns.

We include two control variables $C_{f,t}$ in the regressions. The first is the number of employees a firm had in the initial age group; the variable is meant to control for the possibility that firms that start with, say, 10 employees are less likely to double in size than firms that start with one employee (since moving from one to two employees may be less challenging than moving from 10 to 20). The regressions show that this concern is indeed valid in that the variable is significant in all regressions with a large negative and stable coefficient. The second control variable is a dummy indicating whether or not an establishment is part of a firm with multiple establishments. Such a firm might be expected to have better access to different markets and may more easily take advantage of economies of scale, thereby growing faster. The variable is significant, and the coefficient estimate of around 0.6-0.7 is relatively robust, implying that the average cumulative growth rate of firms with multiple establishments across age groups is some 60 to 70 percentage points higher than that of other firms.¹²

We now turn to the variables of interest $X_{f,t}$. Regressions 2 to 8 in Table 1 include them individually while Regressions 1 to 4 in Table 2 include them jointly. In Regressions 2 and 4 in Table 2 (in contrast to the remaining regressions), we drop observations with the 2.5 percent highest and lowest values of the dependent variable to control for the impact of outliers.

First, we consider two alternative measures of labor informality that have been used in the literature. These include the share of social security contributions a firm pays as a percentage of its salary payments to its employees (Busso et al, 2012) and the share of salaried employees in a firm's workforce (Dabla-Norris et al., 2018). Regressions 2 and 3 in Table 1 show that both variables are significant. The coefficients are also broadly similar in magnitude: a fully formal firm has a cumulative growth rate that is on average across age groups about 47 percent higher than a fully informal firm according to Regression 2 and some 38 percent higher according to Regression 3 (note that a fully formal firm in Mexico should pay around 16 percent of its workers' salaries in social security contributions). The variables remain highly

¹¹ We do not include firm fixed effects or more disaggregated industry or locational classifications in the baseline specification at this stage because some of our explanatory variables of interest would drop out of the regression. However, we do discuss the robustness of our results to the inclusion of such dummies later on.

¹² Note also that what is captured here is not growth through adding establishments but within-establishment growth as explained above.

significant with similar coefficients when included alongside other regressors in Table 2, although the coefficients decrease in magnitude in some specifications. The finding appears plausible, given that informal firms face various disincentives and obstacles to growth, such as the desire to remain small in order not to be detected by the authorities. It is also consistent with the link between informality and resource misallocation discussed in Misch and Saborowski (2018).

Second, in Regression 4 of Table 1, we include the Herfindahl index as a measure of market concentration. The index is calculated based on the number of employees across the 20 largest firms in each 6-digit industry i and state s in wave t . Our hypothesis is that market concentration should disincentivize investment and firm growth to the extent that it correlates with the undue use of market power. Surprisingly, Regression 4 in Table 1 suggests a positive association between the concentration variable and firm growth. In fact, the variable becomes significant with a negative coefficient once we exclude firms whose life cycle growth is above the 97.5th percentile or below the 2.5th percentile to remove the effects of outliers (Regressions 2 and 4 in Table 2). Taking these results together, it appears that, in general, there is a statistically significant negative link between market concentration and life cycle growth although some of the firms with the highest life cycle growth are located in highly concentrated industries. Finally, it is important to point out that the magnitude of the estimated coefficients is not economically significant - given that the concentration measure takes values between zero and 0.15 for most industries - potentially reflecting the fact that the fixed effects we include soak up much of the variation in the market concentration indicator.

We also include measures of geographic remoteness. Besides the fact that Misch and Saborowski (2018) show that geographic remoteness is associated with resource misallocation, there may also be a direct link with firm life cycle growth simply because firms that are far away from their customers, suppliers and competitors are likely to be less inclined to invest and may be less likely to have access to qualified workers. The three measures of geographic remoteness we include are the log of the firm's distance in km from the next major city (defined as a city with a population of at least 500,000), and the distance from the US border as well as from Mexico City (in logs). Columns 3-5 show that all three variables have negative coefficients, suggesting that firm growth is weaker the more geographically remotely a firm is located. The three variables generally remain significant in the regressions shown in Table 2. Nevertheless, similar to the market concentration indicator, the magnitude of the coefficient does not appear to suggest an economically significant relationship (the three variables take values generally between 0 and 10). However, note that our set of fixed effects includes state level dummies, implying that the coefficients can only capture intra-state differences in distances. It is conceivable that the impact would grow if variation in remoteness across states were exploited.

Table 1: Baseline Regressions

	1	2	3	4	5	6	7	8
Age Group 5 - 9	1.163*** (0.001)	0.344*** (0.005)	0.358*** (0.006)	0.353*** (0.006)	0.359*** (0.006)	0.358*** (0.006)	0.358*** (0.006)	0.381*** (0.008)
Age Group 10 - 14	1.323*** (0.003)	0.456*** (0.007)	0.468*** (0.008)	0.470*** (0.008)	0.480*** (0.008)	0.477*** (0.008)	0.477*** (0.008)	0.501*** (0.009)
Age Group 15 - 19	1.411*** (0.004)	0.543*** (0.010)	0.559*** (0.011)	0.557*** (0.011)	0.571*** (0.012)	0.570*** (0.012)	0.570*** (0.012)	0.575*** (0.013)
Age Group 20 - 24	1.314*** (0.007)	0.556*** (0.015)	0.583*** (0.015)	0.571*** (0.014)	0.594*** (0.016)	0.590*** (0.016)	0.589*** (0.016)	0.594*** (0.016)
Initial Size, log		-0.380*** (0.010)	-0.427*** (0.011)	-0.350*** (0.010)	-0.346*** (0.010)	-0.347*** (0.010)	-0.347*** (0.010)	-0.391*** (0.013)
Multi Establishments		0.580*** (0.031)	0.669*** (0.034)	0.698*** (0.033)	0.697*** (0.033)	0.698*** (0.033)	0.698*** (0.033)	0.679*** (0.039)
Social Security Share		0.024*** (0.001)						
Labor Paid Ratio			0.469*** (0.011)					
Concentration				0.150** (0.071)				
Distance Large City, log					-0.004*** (0.001)			
Distance US Border, log						-0.006** (0.002)		
Distance Mexico City, log							-0.011** (0.005)	
Financial Access								0.426*** (0.012)
Observations	1,654,117	1,651,110	1,527,530	1,620,017	1,590,701	1,478,496	1,478,496	1,252,742
State-Industry-Year FEs	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.401	0.145	0.157	0.121	0.140	0.136	0.136	0.145
Standard errors clustered at the state-industry level								
*** p<0.01, ** p<0.05, * p<0.1								

Finally, we also include a variable proxying for a firm's financial access in Regression 8 of Table 1, namely a dummy indicating whether or not a firm had a bank account in a given wave.¹³ Intuitively, one would expect financial access to be associated with higher life cycle growth as it alleviates investment constraints. This is indeed what the regression results suggest as the variable is significant despite the fact that its inclusion significantly reduces the number of observations as information on the variable is only available in the latest two waves of the Economic Census. The coefficient remains highly significant when we include it alongside other controls (Table 2) although its magnitude drops by about half when we drop outliers in the dependent variable.

¹³ Using access to bank credit as an alternative control variable does not change the results materially.

Table 2: Baseline Regressions

	1	2	3	4
Age Group 5 - 9	0.339*** (0.005)	0.238*** (0.004)	0.367*** (0.007)	0.264*** (0.004)
Age Group 10 - 14	0.447*** (0.007)	0.301*** (0.005)	0.477*** (0.008)	0.331*** (0.005)
Age Group 15 - 19	0.525*** (0.010)	0.336*** (0.005)	0.536*** (0.012)	0.353*** (0.006)
Age Group 20 - 24	0.531*** (0.013)	0.326*** (0.007)	0.543*** (0.014)	0.347*** (0.007)
Initial Size, log	-0.391*** (0.011)	-0.299*** (0.011)	-0.426*** (0.013)	-0.317*** (0.014)
Multi Establishments	0.569*** (0.035)	0.323*** (0.017)	0.566*** (0.042)	0.314*** (0.020)
Social Security Share	0.024*** (0.001)	0.014*** (0.000)	0.022*** (0.001)	0.012*** (0.000)
Concentration	0.110 (0.070)	-0.103*** (0.038)	0.126 (0.088)	-0.134*** (0.049)
Distance Large City, log	-0.001** (0.001)	-0.001*** (0.000)	-0.000 (0.001)	-0.001* (0.000)
Distance US Border, log	-0.007*** (0.002)	-0.005*** (0.001)	-0.005* (0.003)	-0.003* (0.002)
Distance Mexico City, log	-0.008 (0.005)	-0.002 (0.003)	-0.011* (0.006)	-0.003 (0.003)
Financial Access			0.345*** (0.012)	0.178*** (0.006)
Observations	1,401,970	1,331,871	1,071,044	1,017,492
State-Industry-Year FEs	Yes	Yes	Yes	Yes
R-squared	0.140	0.191	0.148	0.197

Standard errors clustered at the state-industry level

*** p<0.01, ** p<0.05, * p<0.1

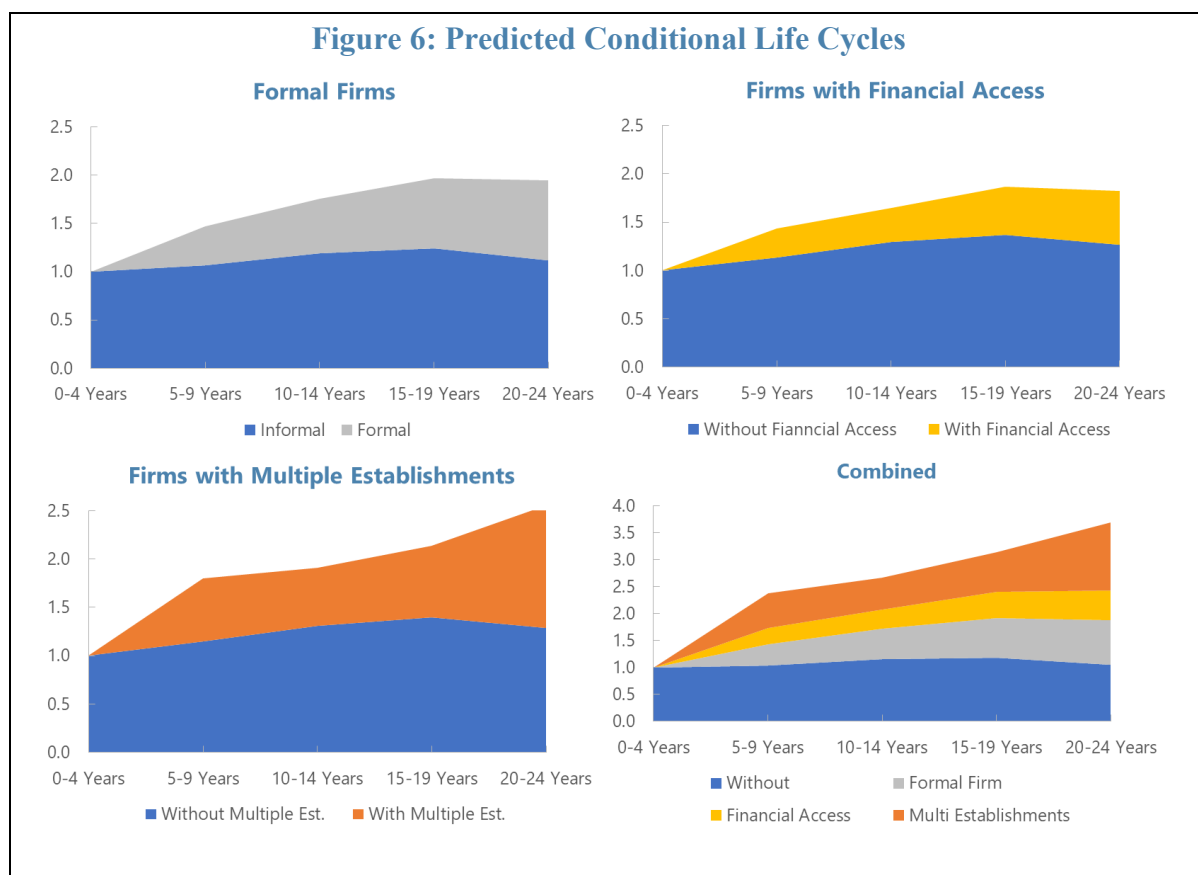
C. Econometric Decomposition—Extensions

In a first extension, we estimate life cycle growth conditional on our variables of interest $X_{f,t}$. The coefficient estimates presented in Tables 1 and 2 tell us the average impact of each of the regressors on firm life cycle growth rates across age groups. In this subsection, we analyze how our regressions would predict the cumulative firm growth rate to be affected differentially across age groups. We do so by estimating Equation 3 in which we add interaction terms $D_{AG}X_{f,t}$ between the age group dummies and our variables of interest $X_{f,t}$ to the specification in Equation 2. Holding all remaining variables constant at their means, we can then calculate life cycles conditional on different realizations of a given variable in $X_{f,t}$:

$$Q_{f,t} = \alpha_{s,i,t} + \sum_{AG=2}^5 (\beta_{AG} D_{AG} + \delta_{AG,X} D_{AG} X_{f,t}) + \gamma C_{f,t} + \theta X_{f,t} + \varepsilon_{f,t} \quad (3)$$

Figure 7 presents the predicted conditional life cycles. We only show the three regressors that appear to carry the most economically significant coefficients in the regressions in Tables 1 and 2, namely the indicators for multiple establishments, labor formality and financial access. The top left chart shows the predicted life cycle growth for the average firm with multiple establishments vs. the average firm without multiple establishments while controlling for all other factors. The top right chart shows the predicted conditional life cycle growth for formal and informal firms, the bottom left chart shows predicted conditional life cycles for firms with and without financial access, and the bottom right chart shows the predicted conditional life cycle growth for the average formal firm with multiple establishments and financial access.

It is clear at first glance that the results are economically significant. The findings suggest that the average formal firm would grow some 70 percentage more over the full life cycle than the average informal firm. A firm that is both formal and has financial access would see a cumulative growth rate of some 140 percentage points more than an informal firm without bank account. Finally, a formal firm with a bank account and multiple establishments would grow to 3-4 times its initial size over its full life cycle, and thus by an order of magnitude that is broadly similar to that of the average U.S. firm (Hsieh and Klenow, 2014).



In a second extension, we examine differences in life cycle dynamics between manufacturing and service sector firms. We estimate Equation 4 in which we interact all right-hand side variables with a dummy $M_{f,t}$ that takes the value one in the case of manufacturing firms:

$$Q_{f,t} = \alpha_{s,i,t} + \sum_{AG=2}^5 (\beta_{AG} D_{AG} + \rho_{AG} D_{AG} M_{f,t}) + \gamma C_{f,t} + \theta X_{f,t} + \vartheta C_{f,t} M_{f,t} + \mu X_{f,t} M_{f,t} + \varepsilon_{f,t} \quad (4)$$

Table 3 presents the results of estimating Equation 4. In line with the descriptive results in the previous section, the age group dummy interactions with the manufacturing indicator carry positive coefficients and are all highly significant, signaling that manufacturing firms on average have somewhat stronger life cycle growth. However, almost none of the interactions of our regressors of interest with the manufacturing dummy appear to matter for life cycle growth, suggesting that our variables matter for both manufacturing and services firms in a very similar way. The only exception is the multi establishment dummy whose interaction with the manufacturing dummy is highly significant and carries a positive coefficient in both regressions in Table 3. This suggests that the advantages of a firm structure with multiple establishments are particularly significant in the case of manufacturing firms.

Table 3: Manufacturing vs. Services

	Regression 1		Regression 2	
	Level Term	Interaction	Level Term	Interaction
Age Group 5 - 9	0.329*** (0.005)	0.071*** (0.014)	0.357*** (0.007)	0.067*** (0.017)
Age Group 10 - 14	0.432*** (0.008)	0.109*** (0.020)	0.463*** (0.008)	0.099*** (0.024)
Age Group 15 - 19	0.507*** (0.010)	0.109*** (0.020)	0.519*** (0.012)	0.116*** (0.029)
Age Group 20 - 24	0.512*** (0.014)	0.126*** (0.028)	0.526*** (0.014)	0.105*** (0.034)
Initial Size, log	-0.405*** (0.012)	0.128*** (0.034)	-0.438*** (0.016)	0.056** (0.022)
Multi Establishments	0.572*** (0.036)	0.667** (0.310)	0.568*** (0.043)	1.098** (0.527)
Social Security Share	0.023*** (0.001)	0.001 (0.002)	0.021*** (0.001)	0.002 (0.002)
Concentration	0.111* (0.067)	0.056 (0.216)	0.140* (0.079)	0.013 (0.260)
Distance Large City, log	-0.002** (0.001)	0.002 (0.002)	-0.000 (0.001)	0.003 (0.003)
Distance US Border, log	-0.007*** (0.003)	-0.006 (0.006)	-0.004 (0.003)	-0.006 (0.007)
Distance Mexico City, log	-0.007 (0.005)	-0.007 (0.012)	-0.009 (0.006)	-0.013 (0.015)
Financial Access			0.326*** (0.013)	0.132*** (0.038)
Observations	1,401,970		1,071,044	
State-Industry-Year FEs	Yes		Yes	
R-squared	0.142		0.15	

Standard errors clustered at the state-industry level

*** p<0.01, ** p<0.05, * p<0.1

V. CONCLUSION

This paper focuses on within-firm growth as an explanation for Mexico's low productivity growth. A recent strand of work argues that Mexico's productivity has been held back in part because the allocation of labor and capital is distorted (e.g. Levy, 2018; Misch and Saborowski, 2018). In this paper, we argue that individual firms do not invest enough, thus depressing firm growth and preventing firms from taking sufficient advantage of economies of scale.

We establish two stylized facts to motivate our analysis: First, we show that firm size is positively correlated with state-level per capita incomes. The finding mirrors previous results at the cross-country level and highlights the close association between firm growth and economic development (Bento and Restuccia, 2018). Second, we show that there is a small minority of Mexican firms that do not grow at rates similar to the average U.S. firm while most of their peers stagnate at less than two times their initial size.

The remainder of the paper analyzes the distinguishing factors of the small minority of firms that manage to grow in an environment in which the majority does not. We decompose life cycle growth into the contributions of individual distortions using a simple regressions framework. Our findings suggest that Mexican firms that are formal, have multiple establishments and financial access, are part of diversified industries as well as located in population centers can grow at sizable rates. The regressions predict that a firm that is both formal and has financial access would see a cumulative growth rate of some 140 percentage points more than an informal firm without financial access. Formal firms with financial access and multiple establishments, in turn, are predicted to grow to 3-4 times their initial size over their life cycles, and thus by an order of magnitude similar to that of the average U.S. firm.

Our findings support the theoretical proposition that distortions that constrain productivity through misallocation also depress investment by individual firms (Hsieh and Klenow, 2014, and Bento and Restuccia, 2017). We show that life cycle dynamics can indeed be decomposed into a very similar set of distortions to those that have been identified as drivers of resource misallocation in Misch and Saborowski (2018).

From a policy perspective, our findings are consistent with continued emphasis on structural reforms. This includes efforts to reduce tax evasion, limit the use of market power in concentrated industries and strengthen access to financial services. Furthermore, targeted infrastructure investments that help to better connect the more remote regions to the major population centers in the country also matter for firm growth. We leave several issues to future work. These include a more detailed study of differences in the distribution of life cycle growth across cohorts and age groups. Another avenue could be to more explicitly examine the role of market power in driving life cycle growth.

VI. REFERENCES

- Alfaro, L., A. Charlton and F. (2009), “Plant-Size Distribution and Cross-Country Income Differences,” in “NBER International Seminar on Macroeconomics 2008”, University of Chicago Press, pp. 243–272.
- Bento, P. and D. Restuccia (2018), “Misallocation, Establishment Size and Productivity”, *American Economic Journal: Macroeconomics*, 9(3), pp. 267-303.
- Bento, P. and D. Restuccia (2018), “On Average Establishment Size Across Sectors and Countries”, NBER Working Paper 24968.
- Busso, M., Fazio, M. and Levy, S. (2012), “(In)Formal and (Un)Productive: The Productivity Costs of Excessive Informality in Mexico,” IDB Working Paper 341, Inter-American Development Bank.
- Busso, M., F. Oscar and S. Levy (2018), “The Longitudinal Linkage of Mexico’s Economic Census 1999-2014”, IDB Technical Note 1477.
- Costa-Scottini, L. (2018), “Firm-Level Distortions, Trade, and International Productivity Differences,” Unpublished manuscript. Available at https://www.tuck.dartmouth.edu/uploads/content/Costa-Scottini_JMP_flyout.pdf.
- Dabla-Norris, E., L. Jaramillo, F. Lima and A. Sollaci (2018), “Size Dependent Policies, Informality and Misallocation”, IMF Working Paper, 18/179.
- Eslava, M. and J. Haltiwanger (2017), “The Drivers of Life-cycle Growth of Manufacturing Plants,” 2017 Meeting Papers 1540. Society for Economic Dynamics.
- García-Santana, M. and Ramos, R. (2015) “Distortions and the size distribution of plants: evidence from cross-country data,” *SERIEs*, p, pp. 279-312.
- Guner, N., G. Ventura and X. Yi, (2008), “Macroeconomic Implications of Size-Dependent Policies,” *Review of Economic Dynamics*, 11(4), pp 721-744.
- Hsieh, C. T., and Klenow, P. J. (2009), “Misallocation and manufacturing TFP in China and India,” *The Quarterly Journal of Economics*, 124(4), 1403–48.
- Hsieh, C.-T., and P. J. Klenow (2014), "The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics*, 129(3), pp. 1035–84.
- Levy, S. (2018), “Under-Rewarded Efforts: The Elusive Quest for Prosperity in Mexico,” Inter-American Development Bank, Washington, D.C.
- Levy, S., and D. Rodrik (2017), “The Mexican Paradox,” available at <http://www.projectsyndicate.org>.
- Lozada Cortés, E. A. (2017) “Mala asignación de recursos y distorsiones correlativas en México,” Unpublished manuscript. Available at <http://repositorio-digital.cide.edu/handle/11651/2465>.

- Misch, F., and C. Saborowski (2018), “Resource Misallocation and Productivity: Evidence from Mexico”, IMF Working Paper, 18/112.
- Poschke, M. (2018), “The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology,” *American Economic Journal: Macroeconomics*, 10(3), pp. 1-41.
- Schoellman, T., Lu, J., and Donovan, K. (2017) “Firm Size Distributions and Cross-Country Labor Market Outcomes,” 2017 Meeting Papers, 1571, Society for Economic Dynamics.
- Van Reenen, J. (2018), “Increasing Differences Between Firms: Market Power and the Macro-Economy,” CEP Discussion Papers, dp1576, Centre for Economic Performance, LSE.

Appendix 1: Variable Definitions and Descriptive Statistics

Table A. Definitions

Variable	Definition
Age Group 0 - 4	Dummy taking value 1 when firms of age 0-4 years
Age Group 5 - 9	Dummy taking value 1 when firms of age 5-9 years
Age Group 10 - 14	Dummy taking value 1 when firms of age 10-14 years
Age Group 15 - 19	Dummy taking value 1 when firms of age 15-19 years
Age Group 20 - 24	Dummy taking value 1 when firms of age 20-24 years
Initial Size	Number of employees when firm was of age 0-4 years
Multi Establishments	Dummy taking value 1 if firm has multiple establishments
Social Security Share	Share of firm's social security payments in salaries
Labor Paid Ratio	Share of salaried employees in the total
Concentration	Herfindahl index at 6 digit industry level based on number of employees
Distance Large City	Distance in kilometers to next city with at least 500 thousand inhabitants
Distance US Border	Distance in kilometers to US border
Distance Mexico City	Distance to Mexico City
Financial Access	Dummy taking the value 1 when firm has a bank account

Table B. Descriptive Statistics

Variable	Mean	Median	Std. Dev.	5th pctl	95th pctl
Initial Size	3.6	2	28.9	1	7
Multi Establishments	0.02	0	0.14	0	0
Social Security Share	1.28	0	4.59	0	12.8
Labor Paid Ratio	0.19	0	0.34	0	1
Concentration	0.075	0.06	0.03	0.05	0.16
Distance Large City	77.4	48.5	85.6	110	250.6
Distance US Border	721.5	746.5	251.5	206.3	1097
Distance Mexico City	429.3	307.6	446.3	9.1	1298.2
Financial Access	0.03	0	0.16	0	1

Appendix 2: Robustness Checks

In our baseline specification, we cannot rule out omitted variable bias and reverse causality as firm growth could conceivably drive some of our explanatory variables. For example, the decision to be informal or the decision to have a bank account could either be driven by high growth rates the firm experienced over the period or could be driven by an unobserved factor that also boosted the firm's growth rate. To address this concern to the extent possible, we lag our variables of interest in a robustness check. Appendix Table C shows our basic regressions when all of our regressors are lagged by five years (by one wave). The table illustrates that our findings are remarkably robust to this change. In particular, the variables most likely to be subject to an endogeneity bias (the informality, multi establishment and financial access variables) remain highly significant with the expected coefficients.

Table C. Lagged Regressors						
	1	2	3	4	5	6
Age Group 5 - 9	0.356*** (0.006)	0.266*** (0.004)	0.370*** (0.008)	0.283*** (0.005)	-0.172*** (0.010)	-0.074*** (0.006)
Age Group 10 - 14	0.448*** (0.007)	0.318*** (0.005)	0.470*** (0.009)	0.343*** (0.005)	-0.090*** (0.008)	-0.022*** (0.005)
Age Group 15 - 19	0.515*** (0.011)	0.346*** (0.006)	0.534*** (0.013)	0.367*** (0.007)		0.024*** (0.005)
Age Group 20 - 24	0.534*** (0.014)	0.343*** (0.007)	0.562*** (0.015)	0.371*** (0.007)	-0.003 (0.011)	
Initial Size, log	-0.549*** (0.009)	-0.430*** (0.011)	-0.590*** (0.010)	-0.460*** (0.012)	-0.568*** (0.013)	-0.456*** (0.016)
Lag Multi Establishments	0.749*** (0.050)	0.499*** (0.025)	0.847*** (0.058)	0.564*** (0.030)	0.956*** (0.083)	0.641*** (0.037)
Lag Social Security Share	0.023*** (0.001)	0.014*** (0.001)			0.023*** (0.001)	0.015*** (0.001)
Lag Labor Paid Ratio			0.410*** (0.011)	0.258*** (0.007)		
Lag Concentration	0.091 (0.095)	-0.106** (0.045)	0.149 (0.099)	-0.068 (0.055)	0.120 (0.153)	-0.135* (0.074)
Lag Distance Large City, log	-0.003*** (0.001)	-0.002*** (0.001)	-0.003*** (0.001)	-0.002*** (0.001)	-0.005*** (0.001)	-0.004*** (0.001)
Lag Distance US Border, log	-0.009** (0.004)	-0.006*** (0.002)	-0.004 (0.004)	-0.002 (0.002)	-0.005 (0.005)	-0.004 (0.003)
Lag Distance Mexico City, log	-0.014 (0.009)	-0.003 (0.005)	-0.013 (0.009)	-0.002 (0.005)	-0.030** (0.013)	-0.012* (0.007)
Lag Financial Access					0.325*** (0.026)	0.169*** (0.011)
Observations	845,879	803,585	745,527	708,305	515,495	489,720
State-Industry-Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.159	0.164	0.164	0.165	0.144	0.155
Standard errors clustered at the state-industry level						
*** p<0.01, ** p<0.05, * p<0.1						

As an additional robustness check, we include alternative fixed effects. As discussed above, our baseline specifications controls for unobserved industry-state-year effects. Appendix Table D illustrates how our results are affected by expanding this set of fixed effects. In Regressions 1 and 2, we instead include state-sector-year fixed effects at the 6-digit level. In Specification 3 and 4, we also control for municipality fixed effects. In regressions 5 and 6, we additionally control for firm fixed effects. Note that more and more of our explanatory variables drop out as we add fixed effects at higher levels of disaggregation. For example, the concentration indicator cannot be included in any of the regressions since it is computed at the 6-digit industry level and thus captured by the new sets of fixed effects. The geographical variables, in turn, drop out as soon as we include fixed effects on a locality basis (because they are computed by locality) while the multi establishment and initial size indicators drop out when we include firm fixed effects.

Table 4 illustrates that our results are qualitatively unchanged by the inclusion of the additional fixed effects. Almost all variables remain statistically significant with only relatively small variations in coefficient size. The only case in which the magnitude of the coefficient changes notably is the informality indicator in the regressions that include firm fixed effects. The reason is likely related to the fact that there is high persistence in the informality indicator at the firm level.

Table D: Alternative Fixed Effects

	1	2	3	4	5	6
Age Group 5 - 9	0.340*** (0.005)	0.368*** (0.007)	0.341*** (0.005)	0.368*** (0.007)	0.433*** (0.006)	0.223*** (0.004)
Age Group 10 - 14	0.451*** (0.008)	0.484*** (0.009)	0.452*** (0.007)	0.484*** (0.008)	0.616*** (0.010)	0.179*** (0.006)
Age Group 15 - 19	0.530*** (0.010)	0.549*** (0.012)	0.533*** (0.010)	0.551*** (0.012)	0.769*** (0.013)	0.083*** (0.007)
Age Group 20 - 24	0.540*** (0.014)	0.567*** (0.015)	0.543*** (0.014)	0.569*** (0.015)	0.926*** (0.019)	
Initial Size, log	-0.416*** (0.011)	-0.435*** (0.013)	-0.412*** (0.010)	-0.433*** (0.013)		
Lag Multi Establishments	0.568*** (0.035)	0.613*** (0.045)	0.566*** (0.033)	0.614*** (0.043)		
Lag Social Security Share	0.024*** (0.001)	0.025*** (0.001)	0.024*** (0.001)	0.025*** (0.001)	0.014*** (0.000)	0.012*** (0.001)
Lag Distance Large City, log	-0.001 (0.002)	-0.001 (0.003)				
Lag Distance US Border, log	0.011 (0.008)	0.017* (0.010)				
Lag Distance Mexico City, log	-0.042*** (0.011)	-0.050*** (0.013)				
Lag Financial Access		0.398*** (0.031)		0.326*** (0.013)		0.298*** (0.012)
Observations	1,422,093	1,088,177	1,645,378	1,249,606	1,615,140	1,184,486
State-6DIndustry-Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FEs			Yes	Yes	Yes	Yes
Firm FEs					Yes	Yes
R-squared	0.197	0.203	0.184	0.190	0.560	0.697
Standard errors clustered at the state-industry level						
*** p<0.01, ** p<0.05, * p<0.1						