Firm Learning and Growth *

Costas Arkolakis Theodore Papageorgiou Olga A. Timoshenko †

Yale University and NBER McGill University George Washington University

September 5, 2015

Abstract

We study the implications of introducing learning (Jovanovic, 1982) in a standard monopolistically competitive environment with firm productivity heterogeneity, á la Melitz (2003). The model predicts that firm growth rates decrease with age, holding size constant, and decrease with size, holding age constant, a fact that models focusing on idiosyncratic productivity shocks have difficulty matching. We calibrate the model using Colombian plant-level data and find that it matches growth and survival patterns well. Unlike the canonical Melitz (2003) model or the Jovanovic (1982) model our economy is not efficient. Subsidies to the fixed costs of young firms can be welfare enhancing: they allow young firms to avoid early exit and thus, benefit consumers through access to a larger number of varieties.

*We are grateful to Manolis Galenianos, Illenin Kondo, Logan Lewis, Ryan Monarch, Giuseppe Moscarini, Emanuel Ornelas, Andres Rodriguez-Clare, Kim Ruhl, Erick Sager, James Tybout as well as various seminar participants. This research is supported by NSF grant SES-0921673. This work was supported in part by the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center. All remaining errors are ours.

†Contact: Arkolakis: costas.arkolakis@yale.edu, Papageorgiou: theodore.papageorgiou@mcgill.ca, Timoshenko: timoshenko@gwu.edu

1 Introduction

Traditionally, empirical economists, driven by the availability of various measures of firm size in commonly used datasets, have stressed the role of firm size in firm growth. While some early work by Evans (1987a) and Dunne et al. (1989) has suggested that age might be an important determinant of growth, theoretical work has long ignored the role of age in firm growth. A recent wave of research, led by Haltiwanger et al. (2013), has returned the focus of empirical work on age, revealing its important role for firm growth. Yet the models of firm dynamics remain disconnected from these empirical results. Widely-used models that carefully model firm productivity heterogeneity, such as Hopenhayn (1992) and Melitz (2003), abstract from the role of age on firm growth. Learning models in the tradition of Jovanovic (1982), even though they provide a natural way to generate age dependent growth, have found little practical use.

In this paper we suitably adapt the standard learning mechanism of Jovanovic (1982) into the workhorse monopolistic competition framework of Melitz (2003). We characterize the properties of our setup and obtain three key results: First, we show that the model with learning is able to replicate the documented dependence of growth rates on the age of the firm, even conditional on size. Second, we find that smaller firms, conditional on age grow faster than larger firms. Third, we show that, unlike the canonical Melitz (2003) model or the Jovanovic (1982) model, our economy is no longer efficient. In particular age-dependent policies, such as subsidies provided to young firms (see for instance the recommendations of Commision (2010) and the analysis and recommendations in chapter 5 in the OECD (2013)), can lead to welfare gains.

In our model firms enter the market small and they are uncertain about the demand their product faces. Over time, as firms observe sales realizations, they may grow large if they have a very successful product or, if not, they shrink and may eventually exit the market. More specifically the demand for a firm's product is uncertain and demand realizations in each period are determined by an unobserved idiosyncratic firm demand component, namely the firm's product appeal in a market plus a noise component.

Each firm decides the quantity produced in the beginning of each period depending on its prior for its product appeal. Once demand is realized, the price adjusts to clear the firm's product market and, using the realized price, the firm revises its posterior for its product appeal.

A firm that experiences higher demand than initially expected revises upward its belief, expands production and grows in size. A firm that experiences lower demand than expected cuts back on production, and it may find it optimal to exit the market.

In our setup two different channels lead to faster growth of younger firms: beliefs updating and selection. First, younger firms face large uncertainty about their product appeal, they revise their beliefs relatively more, compared to firms that are older and better informed. In equilibrium, they are expected to grow faster compared to older firms. This holds even conditional on a firm's size, thus yielding the conditional age dependence of firms' growth rates that we observe in the data. Second, the endogenous selection of firms strengthens the above result: Since younger firms face larger uncertainty, they are much more likely to exit the market if they face a low demand realization. This selection implies that the measured growth rate of surviving young firms is even higher than the respective growth rate of older firms.

We investigate the quantitative implications of our setup, by calibrating the model to match moments from a panel of Colombian plant-level data, assuming that in our model each firm owns a unique plant. Guided by our theoretical results, we identify the importance of learning by targeting the impact of age on growth rates, conditional on size. In addition, the model correctly predicts moments that are not targeted in the calibration, such as the annual survival rate and the impact of size on growth rates, conditional on age. Given that learning also affects the exit behavior of firms, it is reassuring that our model is also able to match the survival rate.

Armed with the calibrated model, we evaluate the welfare impact of policies that subsidize young firms. While the entry and exit decisions are always optimal from the individual firm's point of view, they may not be optimal from an aggregate welfare perspective. For instance, when a firm exits (enters) it does not internalize the loss (addition) of a variety for the consumer. In addition, when a firm enters or exits it does not consider the negative impact of its decision on the profits of incumbent firms through the monopolistic distortion. In the canonical Melitz (2003) model these two effects exactly cancel out leaving no role for welfare improving policies. In our calibrated setup with learning however, in equilibrium, there is an inefficiently low mass of firms and varieties available to consumers.

We show that a subsidy to the fixed operating costs increases both the equilibrium mass of firms and welfare. We experiment with a range of different subsidies and find that, for instance, a 40% subsidy to the fixed cost of production applied to young firms leads to up to a 20% increase in the mass of operating firms in equilibrium. The subsidy is financed through lump-sum taxation. The welfare gains are of similar order of magnitude to the cost of the subsidy: a subsidy equal to 0.36% of GDP leads to a 0.23% increase in aggregate consumption (consumption multiplier of 0.64). A lower subsidy leads to an even higher consumption multiplier: for instance, a subsidy equal to 0.10% of GDP leads to a 0.08% increase in aggregate consumption (consumption multiplier of approximately 0.8).

Our paper contributes to a new, but growing literature on the quantitative implications of learning on firm growth. Eaton et al. (2012) consider a model where firms actively learn their product appeal by forming new matches with buyers. Abbring and Campbell (2003) also develop a structural model of firm learning, which they estimate with data on Texan bars. Both these frameworks are much richer in the specification of the learning mechanism, but they use a partial equilibrium model and as such, they cannot be used for macro policy evaluation, a key focus of our analysis. Ruhl and Willis (2014) modify a standard model with idiosyncratic productivity shocks by specifying a foreign demand function that increases with firm age and Albornoz et al. (2012) consider a model where a firm is uncertain about its demand and uncertainty is resolved by incurring a fixed cost. Finally, closer to our approach, Timoshenko (2015) develops a general equilibrium model of learning in the context of multi-product firms and shows that such a framework can predict well the age dependence of product switching among exporters.

Models that focus on idiosyncratic productivity shocks (Hopenhayn (1992), Luttmer (2007), Arkolakis (2015)) have difficulty explaining the dependence of growth rates on age, conditional on size.[1] The reason for this failure is that growth is based on an underlying Markov process. This assumption implies that all firms of the same size have the same growth profile, which is independent of their age.[2] Financial constraints together with idiosyncratic productivity shocks, as in Cooley and Quadrini (2001), can explain age dependence (even conditional on size) if the

[1] Hopenhayn (1992) notes that "since size is a sufficient statistic for [productivity], age has no extra predictive role" (page 1141).

[2] A similar reasoning applies to the Klette and Kortum (2004) approach. However, even if the productivity process is not a simple Markov process, but depends on a finite number of past realizations as well, the two approaches give distinct predictions regarding the variance of sales as a function of age. The learning explanation proposed here implies that the variance of sales declines with age, as large firms become better informed regarding the demand they face. However, productivity models have no such implication. In general, in a learning model, the dependence of firm actions on past realizations does not erode away as the firm ages. This implication is used in Ericson and Pakes (1998) to distinguish learning from a model of productivity shocks, in which the state dependence has ergodic characteristics.

entry of new firms is at high productivity levels.[3] D'Erasmo (2011) and Clementi and Palazzo (2013) show that adding convex and non-convex adjustment costs to Hopenhayn (1992) can generate age-dependent growth. Adjustment costs imply that there is no longer a one-to-one mapping between productivity and capital stock. As a result, firms of the same size can have different growth rates which depend on both productivity and the capital stock. In equilibrium younger firms are more productive, but have less capital, so by controlling for age in addition to size, one effectively controls for both productivity and the underlying level of the capital stock. Our approach does not require idiosyncratic productivity shocks to generate age dependence.[4]

Finally, our work follows the line of research that studies the impact of government subsidies on growth. Acemoglu et al. (2013) extend the Klette and Kortum (2004) model to examine the optimal level of subsidies to innovating firms, whereas Atkeson and Burstein (2010) study the impact of policy on innovation and aggregate output in a wider class of models. Itskhoki and Moll (2015) study policy in a growth setup with financial frictions. Unlike these papers, we do not consider a setup with aggregate growth but instead a setup with firm dynamics in a stationary steady state. In our context, some government policies may be welfare improving because of the interaction of learning and monopolistic competition with love for variety. There are two key mechanisms through which this interaction affects efficiency. First, in our setup with demand uncertainty, equilibrium entry is lower than the efficient level, as noted above. Second, by introducing a subsidy, some firms that would have otherwise exited the market, receive additional signals, learn that in fact their underlying demand is higher than they originally thought and grow large.

The rest of the paper is organized as follows: Section 2 introduces a model that suitably adapts the standard learning mechanism of Jovanovic (1982) into the monopolistic competition

[3]In a multi-market setup, however, one can distinguish between our setup and the models of Jovanovic (1982) and Cooley and Quadrini (2001). Since a firm can have different demand realizations across markets, our framework is consistent with differential behavior of the same firm across the various markets in which it operates. For instance, it is possible that a firm expands production in a given market, while choosing to exit another one. If the financial constraints are important or if the firm is learning about its underlying cost structure, then its behavior is restricted to be the same across markets. Furthermore, the evidence in Eaton et al. (2008) and Albornoz et al. (2012) suggest that the age of a firm in a market is inversely related to firm growth and probability of exit in that market, as implied by our model.

[4]It is worth pointing out that while age dependence always takes place in a Jovanovic learning model the sign of that dependence and its quantitative magnitude depend on the functional form assumptions. In our context with CES demand and monopolistic competition we find that growth rates always decline with firm age, conditional on size, as discussed earlier.

framework of Melitz (2003). Section 3 calibrates the model using a panel of Colombian plant-level data, and investigates the growth and welfare implications. Section 4 concludes.

2 The Model

This section describes a model which introduces a demand-learning mechanism into a heterogeneous-firms model. The learning mechanism is similar to that of Jovanovic (1982) and firms must learn about their unobserved demand level which is subject to transient preference shocks. Firms operate in a monopolistically competitive environment as in Melitz (2003).

2.1 Environment

Time is discrete and denoted by t. The economy is populated by a continuum of consumers of mass L. Each consumer derives utility from the consumption of a composite good, C_t, according to the utility function

$$U = E \sum_{t=0}^{+\infty} \beta^t \ln \left(C_t \right),$$

where β is the discount factor. The composite good consists of the consumption of a continuum of differentiated varieties $c_t(\omega)$, aggregated using a constant elasticity of substitution (CES) aggregator with elasticity of substitution $\sigma > 1$

$$C_t = \left(\int_{\omega \in \Omega_t} \left(e^{a_t(\omega)} \right)^{\frac{1}{\sigma}} c_t \left(\omega \right)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where Ω_t is the continuum of differentiated varieties available at time t; $a_t \left(\omega \right)$ is the demand shock at time t for variety $\omega \in \Omega_t$. Consumers are endowed with one unit of labor that they inelastically supply to the market to receives a wage, w_t, in return. They also own an equal share of domestic firms.

Each good is supplied by a monopolistically competitive firm. These firms maximize the expected present discounted value of profits and have the same discount factor as consumers, β. The realization of the demand variable $a_t \left(\omega \right)$ for each firm is determined by a time invariant component, $\theta(\omega)$, and a shock, $\varepsilon_t(\omega)$, and is given every period by

$$a_t(\omega) = \theta(\omega) + \varepsilon_t(\omega), \; \varepsilon_t(\omega) \sim N\left(0, \sigma_\varepsilon^2\right) \text{ i.i.d.}$$

The time-invariant demand parameter, $\theta(\omega)$, can be interpreted as the true underlying demand for the firm's product, its product appeal, and it is unobserved by the firm. This is subject to transient preference shocks, $\varepsilon_t(\omega)$, which are independent over time and across firms.

Following Chaney (2008), we assume that every period there is an exogenous mass of potential entrants J. Entrants draw their unobserved demand parameter, $\theta(\omega)$, from a normal distribution with mean $\bar{\theta}$ and variance σ_θ^2. There is no sunk cost of entry.

Firms exit the market either exogenously with probability δ or endogenously if the sum of their discounted expected profits is less than zero. If the firm stays in the market, it decides the quantity to be produced, $q_t(\omega)$. Output is linear in labor, $l_t(\omega)$, which is the only factor of production, and there is a fixed cost of production, f, measured in the units of labor. Firms are heterogeneous in their productivity level, $e^{z(\omega)}$, which is drawn upon entry and is time-invariant (in Appendix A we relax this assumption and allow productivity to change over time). Unlike the demand parameter, $\theta(\omega)$, each firm's productivity level is observed by the firms.

The sequence of actions is the following: At the beginning of each period, each incumbent firm decides whether to stay in the market and what quantity to produce or to endogenously exit. Next, potential entrants draw their productivity, $e^{z(\omega)}$, and their (unobserved) demand parameter, $\theta(\omega)$, and decide whether to sell or exit. Those that decide to sell pay the fixed cost of production. Subsequently, the firm decides the quantity produced and production takes place. Next, the demand shock, $a_t(\omega)$, is realized and the price of each good, $p_t(\omega)$, adjusts so that the good's market clears. The firm observes the price, infers the underlying demand realization and updates its belief for its product appeal. Finally, firms exit the market exogenously with probability δ.

2.2 Consumer Demand

Each consumer chooses the consumption levels, $c_t(\omega)$, to minimize the cost of acquiring C_t taking into account the prices of varieties, $p_t(\omega)$, as well as his income. The resulting demand equation for each variety is given by

$$q_t(\omega) = e^{a_t(\omega)} \frac{(p_t(\omega))^{-\sigma}}{P_t^{1-\sigma}} Y_t, \tag{2}$$

where Y_t is the aggregate spending level and is given by $Y_t = w_t L + \Pi_t$, while Π_t is total profits of firms. P_t is the aggregate price index associated with the consumption of the composite good C_t and is given by

$$P_t^{1-\sigma} = \int_{\omega \in \Omega_t} e^{a_t(\omega)} (p_t(\omega))^{1-\sigma} d\omega.$$

To economize on notation, in what follows we drop the variety index ω.

Belief Updating

At the start of each period t a firm makes a quantity decision q_t before observing the current demand shock a_t. Once the quantities are set the goods market clears and the firm observes the market clearing price p_t and, using the inverse demand function (2), infers the value of the demand shock a_t. Hence, at the end of each period in which the firm operates, it receives and learns one observation of the demand shock. Taken together over time, these demand shocks are informative about the unobserved demand parameter θ. Denote by $\bar{a}$ the mean of the observed shocks and by n the number of the observed shocks. Since the firm receives one shock per period, n also corresponds to the firm's age. Using Bayes' rule, the firm updates its beliefs regarding θ: the posterior belief of a firm that has observed n signals with mean $\bar{a}$, is given by a normal distribution with mean

$$\mu_n = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2} \bar{\theta} + \frac{\sigma_\theta^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2} n\bar{a} \ , \tag{3}$$

and variance

$$\nu_n^2 = \frac{\sigma_\theta^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2}.$$

Thus, the belief of the firm regarding the realization of its demand shock, a_t, follows $N(\mu_n, \nu_n^2 + \sigma_\varepsilon^2)$. An entrant knows the distribution from which the demand parameter is drawn, and that distribution serves as the prior. Notice that in the long run, upon observing infinitely many signals, the posterior belief converges to a degenerate distribution centered at $\bar{a}$ and $\lim_{n\to\infty} \bar{a} = \lim_{n\to\infty} \frac{\sum_{i=0}^n a_i}{n} = E(a_t) = \theta$. Thus, in the long run, firms learn their product appeal level, while in the short run their knowledge can be summarized by the two state variables: the number of observed shocks n, and the mean of the observed shocks $\bar{a}$.

2.3 Firm Optimization

In describing the firm's problem, we focus on the stationary equilibrium.[5] Since there is no aggregate uncertainty, the aggregate expenditure, Y_t, the price index P_t, and the wage rate, w_t, are constant in the equilibrium. Hence, we suppress the time notation on the aggregate variables. Each firm has to make two decisions every period: whether to stay in a market or exit and how much to produce, if it chooses to stay. When the firm decides whether to produce or exit, it takes into account not only that period's profits, but also the value of learning by observing demand realizations. More specifically, learning allows the firm to make more informed decisions in the future about how much to produce and whether to exit or not. When deciding how much to produce however, learning considerations are not a factor, since the demand realization is independent of how much the firm produces. Therefore its quantity choice is a static one.[6] We begin by analyzing the static quantity decision and then consider the entry and exit decision.

Quantity Decision

The static per-period profit maximization problem is given by

$$\max_{q_t} E_{a_t|n,\bar{a}}\left[p_t\left(a_t\right)q_t\right] - w\frac{q_t}{e^z} - wf, \tag{4}$$

subject to the consumer inverse demand

$$p_t\left(a_t; q_t\right) = \left(\frac{e^{a_t}Y}{q_t}\right)^{\frac{1}{\sigma}} P^{\frac{\sigma-1}{\sigma}}. \tag{5}$$

Substituting the inverse demand constraint (5) into the static maximization problem (4) and taking the first order conditions leads to the optimal quantity choice given by

$$q_t\left(z, b(\bar{a}, n)\right) = \left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\left(\frac{b(\bar{a}, n)e^z}{w}\right)^{\sigma}\frac{Y}{P^{1-\sigma}}\ , \tag{6}$$

[5] The equilibrium exists as long as $\delta > 0$.

[6] See Bergemann and Välimäki (2000) and Eaton et al. (2012) for models of active learning and experimentation.

where we define

$$b(\bar{a}, n) \equiv E_{a_t|\bar{a},n}(e^{\frac{a_t}{\sigma}}) = \exp\left\{\frac{\mu_n}{\sigma} + \frac{1}{2}\left(\frac{\nu_n^2 + \sigma_\varepsilon^2}{\sigma^2}\right)\right\}, \tag{7}$$

i.e. the firm's belief regarding (e^{a_t}). Notice from equation (6) that firms which are more productive (have higher z) and firms which have higher beliefs regarding θ (have higher $\bar{a}$ and hence higher b) sell more.

Substituting the firm's quantity choice, equation (6), into the consumer inverse demand, equation (5), gives the market clearing price

$$p_t\left(a_t, z, b(\bar{a}, n)\right) = \frac{\sigma}{\sigma - 1}\frac{w}{e^z}\frac{e^{\frac{a_t}{\sigma}}}{b(\bar{a}, n)}.$$

The price depends on the firm's productivity, z and belief, b, as well as the realization of the demand shock, a. Taking expectation with respect to a, the expected market clearing price is given by

$$Ep_t = \frac{\sigma}{\sigma - 1}\frac{w}{e^z}\ .$$

Thus, in expectation, price is a constant mark-up over marginal cost.

Substituting the optimal quantity and price into equation (4), the expected per-period firm profits are given by

$$E\pi\left(z, b(\bar{a}, n)\right) = \frac{(\sigma - 1)^{\sigma-1}}{\sigma^\sigma} b(\bar{a}, n)^\sigma \left(\frac{e^z}{w}\right)^{\sigma-1} \frac{Y}{P^{1-\sigma}} - wf. \tag{8}$$

Notice that, since $\sigma > 1$, profits are a convex function of firm beliefs.

Exit and Entry Decision

As discussed above, a firm decides whether to continue paying the fixed cost and stay in the market or exit, taking the maximized expected profits, $E\pi(z, b)$, as given. Thus, the value of a firm of age $n \geq 0$, with productivity z and beliefs b, is given by

$$V(z, b, n) = \max\left\{E\pi(z, b) + \beta(1 - \delta)E_{b'|b,n}V(z, b', n + 1), 0\right\}, \tag{9}$$

where the value of exit is normalized to zero.

We now consider the entry decision. Potential entrants pay the fixed per-period cost and start to sell in the initial period if the value of entry is greater than zero, or otherwise immediately exit. Since all the potential entrants draw their demand shock from the same distribution they start with the same prior belief. From equation (7), the potential entrant's belief is given by

$$b^e = \exp\left\{\frac{\bar{\theta}}{\sigma} + \frac{1}{2}\left(\frac{\sigma_\theta^2 + \sigma_\varepsilon^2}{\sigma^2}\right)\right\}.$$

Notice that potential entrants do vary in terms of their productivity draw, e^z. Thus, the entry condition determines a productivity threshold value $\underline{z}$, such that an entrant that draws productivity $z < \underline{z}$, exits immediately. The condition

$$V(\underline{z}, b^e, 0) = 0. \tag{10}$$

pins down implicitly the productivity threshold $\underline{z}$.

2.4 Equilibrium

Denote the equilibrium mass of operating firms at every period by M. Each potential entrant draws its productivity e^z from the following Pareto distribution

$$f\left(e^z\right) = \frac{\xi}{\left(e^z\right)^{\xi+1}}, \tag{11}$$

where $\xi > 0$. They also draw their unobserved, demand parameter θ from $N\left(\bar{\theta}, \sigma_\theta^2\right)$. As discussed above, based on the productivity realization, each potential entrant decides whether to enter the market and begin selling or exit immediately.

To study the equilibrium, we normalize the wage rate, w, to 1. The stationary equilibrium is described by the entry productivity threshold $\underline{z}$, the aggregate price index P, the aggregate expenditure level Y, the mass of operating firms M, the probability mass function $m(z, b, n)$, firms' optimal quantity choice q, firms' optimal entry and exit decisions, consumers' optimal consumption choice c such that

1. Consumers are optimizing: given equilibrium values, c satisfies demand equation (2).

2. Firms are optimizing: given equilibrium values, q, $\underline{z}$ and the exit thresholds solve the firm's optimization problem in equations (4) and (9).

3. Goods market clears: $Y = L + \Pi$.

4. The probability mass function of active firms, $m(z, b, n)$, delivers the aggregate mass of firms, M, i.e. $M = \sum_n \int \int m(z, b, n)\, dz db$.

The computation of the stationary equilibrium is based on the method developed in Timoshenko (2015) and the setup is briefly outlined in Appendix C.

2.5 Implications

This section presents three key implications of introducing learning into a heterogeneous-firms model. The first two, relate to firm growth as a function of size and age: The expected growth rate of a firm declines with age, conditional on size, and it also declines with size, conditional on age. The third relates to the normative properties of the model: there exists some reallocation of resources in the economy which leads to an improvement in social welfare compared to the decentralized equilibrium.

Result 1: *In equilibrium, the expected growth rate of a firm declines in age, conditional on size.*

In the model, the learning mechanism lies at the heart of the conditional age dependence of growth rates, as we explain below. Let the expected growth rate of a firm be given by $E\left(q_{t+1}/q_t\right)$. As shown in Appendix B, using the expression for the optimal quantity choice in equation (6) and in the absence of endogenous exit, the expected growth rate is given by

$$E\left(\frac{q_{t+1}}{q_t}\right) = \exp\left(\frac{\lambda^2 \sigma_\varepsilon^2 \left(\sigma - 1\right)}{2\sigma\left(1 + n\lambda\right)\left(1 + \left(n+1\right)\lambda\right)}\right), \tag{12}$$

where $\lambda = \sigma_\theta^2/\sigma_\epsilon^2$. Notice that a firm's age n only appears in the denominator. Hence, for a given productivity level z and conditional on the firm's initial size, q_t, the expected growth rate of young firms is higher than the growth rate of older firms.

This result depends crucially on $\sigma > 1$. As seen from equation (6), a firm's optimal choice of output is a convex function of beliefs, b, through the term b^σ. The convexity of the output choice function implies that the increases in output when b increases are larger than the declines in output when b falls by the same amount. Since young firms face more uncertainty than older firms, their beliefs are expected to change by a larger amount. Higher volatility of beliefs, b, combined with the convexity of the output choice function in b, leads to a high expected growth rate. On the contrary, beliefs of older firms are less volatile leading to a lower expected growth rate. Note that with no endogenous exit, the level of firm size, q, has no impact on the expected growth rate, since it has no impact on the volatility of beliefs. Therefore the result holds even conditional on firm size.

Note that with the endogenous exit the observed growth rate is even larger for young firms near the exit threshold than what is implied by equation (12). The reason is that the firms that observe low sales realizations may choose to exit the market, especially if there is large uncertainty for their true demand, rather than stay in and record low or negative growth rates. Endogenous exit is the mechanism that is behind the next result.

Result 2: *In equilibrium, the expected growth rate of a firm declines in size, conditional on age.*

The conditional size dependence of growth rates arises solely due to endogenous exit as in the models of Hopenhayn (1992) and Melitz (2003). Appendix B provides the proof of this result, and here we provide the basic intuition. First, consider equation (12) which computes the expected growth rate of a firm in the absence of selection. Notice that, conditional on age, n, all firms are expected to grow at the same rate. More specifically, the expected growth rate is independent of a firm's productivity, z, and of the expected number of the observed demand shocks, $\bar{a}$. Thus, without controlling for selection, the growth rate is independent of a firm's initial size, q_t.

The size dependence, hence, can only arise due to the endogenous selection. This is demonstrated in Figure 1 which depicts, for a given productivity level z, the exit thresholds $\bar{a}^*(n, z)$ as a function of firm age, n. The depicted exits thresholds are solutions to a firm's entry and exit problem described in Section 2.3. A firm of age n continues to sell if $\bar{a} > \bar{a}^*(n, z)$ and exits otherwise. The exit threshold $\bar{a}^*(n, z)$ increases with n because as firms become more confident about their latent demand, i.e. n increases, the informational value of additional observations declines and firms remain in the market only if their (expected) profit flow is sufficiently high.

To gain intuition regarding Result 2, consider two firms, Firm A and Firm B, in the context of Figure 1. Both of these firms have the same age, n, and productivity, z, but vary in terms of their mean of the observed demand shocks $\bar{a}$. Since such mean is higher for Firm A, that firm is larger in its size, q_t, compared to Firm B. Notice that, due to its large size, Firm A is further away from the exit threshold compared to Firm B. Thus, Firm A can either grow or shrink and still survive in the market. Since Firm B is located much closer to the exit threshold, the only way it can survive in the next period is only if it grows by a sufficiently high amount. Hence, conditional on survival, Firm B is expected to grow faster compared to Firm A. As shown in Appendix B, the above result holds more generally in the case where the two firms have the same size, but different productivity levels, z.

The above two results provide a theoretical foundation for the empirical findings of Evans (1987b) and Haltiwanger et al. (2013). This literature establishes age as an additional determinant of firm growth, other than size. More precisely they find that a regression of firm growth rates on firm size and age yields negative and significant coefficients on both variables, which is consistent with results 1 and 2 in the model.[7]

It is worth pointing out that various authors have suggested models based on idiosyncratic productivity shocks combined with additional mechanisms in order to explain age dependence, conditional on size. Cooley and Quadrini (2001) generate conditional age dependence of growth rates assuming a persistent productivity process in a model with financial frictions. In their framework firms are heterogeneous in their productivity and new entrants start off with a high level of productivity. The interaction of that assumption with financial frictions imply that more productive young firms are able to borrow more than less productive older ones. Hence, the young grow faster, even conditional on size. Arkolakis (2015) can also generate conditional age dependence of growth rates by assuming that new firms have a higher long-run level of productivity and a mean-reverting productivity process. In his setup, comparing two firms with the same size, mean reversion implies the younger one is more likely to grow faster since its long-run steady-level size is higher. Finally, D'Erasmo (2011) and Clementi and Palazzo (2013) study a Hopenhayn (1992) model with convex and non-convex adjustment costs. The introduction of adjustment costs implies that young firms are below their target capital level and, as a result, are more likely to expand their capital and grow, even conditional on their initial size.[8]

[7] See also Eaton et al. (2012), Arkolakis (2015) for related evidence on exporters.

[8] See also Castro et al. (2004).

In our framework all that is required to generate conditional age dependence is a learning mechanism. While in our model we also assume heterogeneity in productivities, as can be seen from equation (12), age dependence holds for any productivity level. In Appendix A we calibrate a model with productivity evolution and find that productivity shocks account for less than ten percent of the conditional age dependence.

We next examine the normative predictions of the model. To do this we study the impact of a particular set of government policies on the equilibrium level of welfare.

Result 3: *There exists some reallocation of resources in the economy which leads to an improvement in social welfare compared to the decentralized equilibrium.*

To demonstrate the welfare properties of the model we introduce a rudimentary government and compare the resulting steady state with the one without a government. This government can influence entry and exit decisions of firms through a set of policy instruments. In particular, we endow the government with the ability to tax or subsidize the fixed per-period costs of young firms through a lump-sum transfer between consumers and producers.

Denote by τ a fixed cost subsidy or tax rate whereby a firm pays the share τ of its fixed costs. Hence, $0 < \tau < 1$ represents a fixed cost subsidy of $(1-\tau)$ percentage points. In this case the firms pays τf portion of the fixed costs and the government taxes the remaining amount lump-sum from consumers. In contrast, $\tau > 1$ represents a fixed cost tax of $(\tau - 1)$ percentage points. In this case a firm incurs fixed cost in the amount of τf, and the portion $(\tau - 1)f$ is lump-sum transferred to the consumers. Denote by $\bar{n}$ the maximum age of a firm which is subject to the described government's subsidy/tax rate. Thus, the set of policy instruments available to the government is given by a pair $(\tau, \bar{n})$.

With this set of policy instruments, the problem of the firm described in equation (9) becomes

$$V(z, b, n) = \max\{E\pi(z, b) - f + (1 - \tau)I(n \leq \bar{n})f + \beta(1 - \delta)E_{b'|b,n}V(z, b', n + 1); 0\},$$

where $I(n \leq \bar{n})$ is an indicator function which equals to 1 when $n \leq \bar{n}$ and zero otherwise.[9] Notice that the fixed cost subsidy or tax rate τ does not affect the intensive margin decisions of firms. The policy however affects entry and exit choices of firms as demonstrated in Figure

[9]The associated market clearing condition becomes $Y = L + \Pi + (\tau - 1)fM(\bar{n})$, where $M(\bar{n})$ is the mass of firms not older than $\bar{n}$ years.

2. In particular, Figure 2 depicts the behavior of the exit thresholds for a subsidy rate of 50 percent applied to firms of 4 years old or younger. Observe that in this case the exit thresholds decline for the duration of the subsidy. Hence, young firms would be more likely to survive.

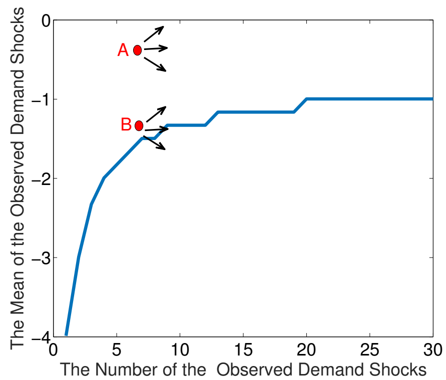
The higher survival of new firms has a direct impact on the equilibrium number of varieties. Panel A of Figure 3 depicts the equilibrium number of varieties as a function of the subsidy or tax rate as applied to firms of either age 1, age 3 or younger, or age 5 or younger. As can be seen in panel A of Figure 3, a fixed cost subsidy increases the equilibrium number of firms as the amount of the subsidy rises, and hence increase the equilibrium number of varieties available to consumers.

An increase in the number of varieties has a direct effect on the equilibrium welfare as measured by real consumption C given by equation (1). As can be seen from equation (1), holding all else constant, an increase in Ω increases the real consumption. Hence, one source of the welfare gains in this economy is associated with the increase in the equilibrium number of varieties generated by the fixed-cost subsidy. Grossman and Helpman (1991) refer to this phenomenon as the consumer-surplus effect. In addition to the consumer-surplus effect, the fixed-cost subsidy generates, what Grossman and Helpman (1991) refer to, as the profit-destruction effect which is illustrated in Panel B of Figure 3. Panel B depicts the equilibrium level of income, Y, as a function of the subsidy or tax rate as applied to firms of either age 1, age 3 or younger, or age 5 or younger. Since we normalize wage to 1, changes in the equilibrium level of income, Y, reflect change in aggregate profits. Notice, that since the mass of firms is increasing as a function of the subsidy, a greater amount of labor is used in paying the fixed costs (either by firms or by the consumer through lump-sum transfers). Thus, while the economy is populated by a greater amount of firms, less labor is used in the production of the final good, yielding less disposable income available to consumers, as shown in Panel B of Figure 3. We discuss the welfare implications of our calibrated setup in Section 3.3.

3 Quantitative Results

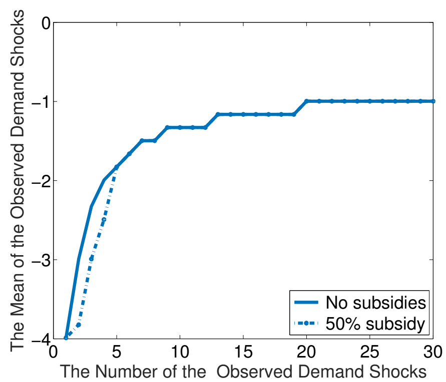
In this section we explore the quantitative implications of our model. In particular, we calibrate the model and examine how large are the potential welfare gains from a targeted policy intervention.

Figure 1: **Demand Shock and Market Participation Thresholds**



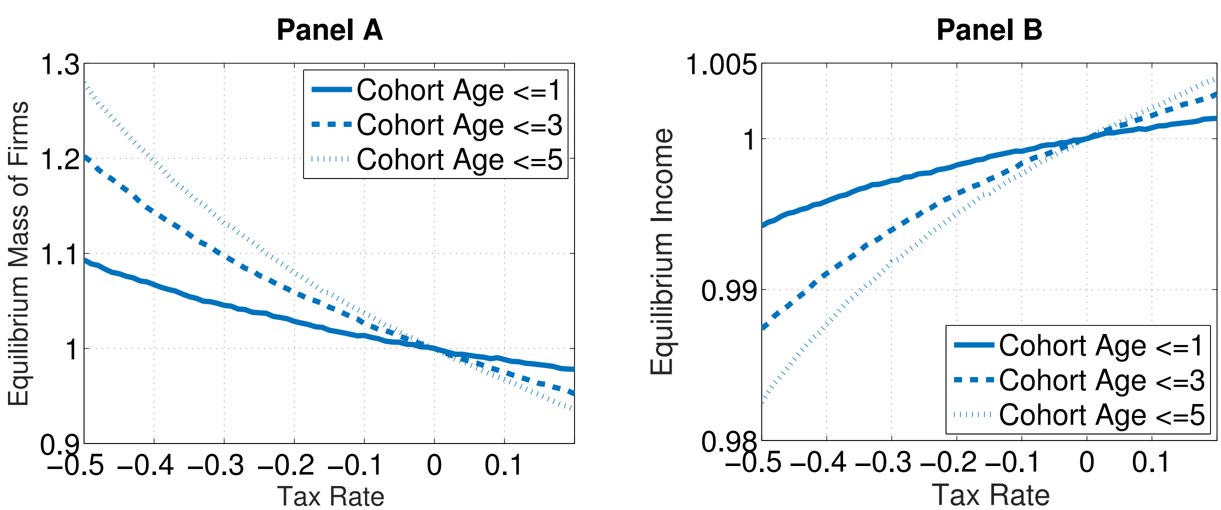
Source: Authors calculations on model simulations.

Figure 2: **Demand Shock and Market Participation Thresholds Under Different Subsidy Schemes**



Source: Authors calculations on model simulations. The fixed cost subsidy is applied for the first 4 years.

Figure 3: **External Benefits versus Costs of a Subsidy**



Source: Authors calculations on model simulations.

3.1 Calibration

For the calibration we assume that the data has been generated from the steady state of our model and match theoretical moments to their empirical counterparts. In the next section we use the calibrated model to investigate the welfare implications of a subsidy on young firms' per-period fixed operating costs.

We use the Colombian plant-level data collected in the DANE survey (see Roberts (1996)). In our calibration we treat a plant in the data as a firm in the model. The data cover all the manufacturing plants in Colombia with 10 or more employees for the period of 1983-1991.[10] For our purposes this database is particularly informative since it reports the real output of the plant, the age of the plant (i.e. the year of a plant's start-up which we use to infer the plant's age), and the decision of the plant to discontinue operations. In Appendix D we discuss additional details of the data and the construction of the data moments used in the calibration.

We first calibrate some of the parameters independently. In particular, we set the elasticity of substitution across goods, σ, to 7.49 following Broda and Weinstein (2006). Similarly we set the discount factor, β, to 0.9606, which implies a quarterly discount rate of 1%. In order to calibrate the exogenous exit probability, δ, we use the model's prediction that only the smallest firms exit endogenously, due to an adverse demand realization. Given this, we set δ to 0.025 to generate an exit rate of 2.56% among the largest 5% of plants in the Colombian data.

The remaining parameters are jointly calibrated. In particular, in order to identify the importance of learning we exploit implications of Results 1 in the previous section and run a regression of firm growth rates on age and size. If we compare two firms of the same size, one young and one old, in our setup the younger firm is more uncertain of its true demand. As a result, it is more likely to grow as new information arrives, compared to the older firm.

To be precise, we calibrate the following four parameters: the per period fixed cost, f, the standard deviation of demand shocks, σ_ε, the shape parameter of the distribution of the productivity draws, ξ, and the standard deviation of the distribution of the unobserved demand parameters, σ_θ.[11] In order to calibrate the parameters, we use the following four moments: the mean of log sales; the sales-weighted growth rate of entrants; the entrants' share of sales; and,

[10] Earlier data from 1977 to 1982 cover all manufacturing plants in Colombia but the data on real production are not as reliable for that time period. To avoid measurement bias caused by the change in the coverage we use only the latest part of the survey, 1983-1991.

[11] Note that $\bar{\theta}$ in this setup is not separately identified from f, so we set $\bar{\theta} = 0$.

Table 1: Parameter Values

Parameter	**Value**
ξ	13.09
σ_ϵ	1.46
σ_θ	0.997
f	253,181

Sources:Model Baseline Calibration

Table 2: Data and Simulation Moments

Targeted Moments:	**Data**	**Model**
Mean of log-sales	14.99	14.99
Share of sales from entrants	2.66%	2.66%
Sales-Weighted growth rate of entrants	23%	23%
Age Coefficient	-0.035	-0.035
Other Moments:		
Annual Survival Rate	91%	91%

Sources: DANE survey data (see text for details) and Model Baseline Calibration

following the discussion above, the age coefficient in a regression of firms growth rates on age and size.[12]

Although a rigorous identification argument is not possible due to the complexity of the setup, we give an informal argument of how each parameter is identified from the data. The per period fixed cost, f, is pinned down by the mean log sales, since increasing the fixed cost pushes less productive firms out of the market thereby increasing the mean sales. The shape of the distribution of the initial productivity draws, ξ, is pinned down by entrants' share of sales: A lower value for ξ, all else equal, leads to a higher initial productivity dispersion and therefore more entrants starting off relatively large. Finally, the parameters related to the learning process, σ_ε and σ_θ are pinned down by the sales-weighted growth rate of entrants and the age coefficient. The calibrated parameters for this baseline calibration are presented in Table 1 and the targeted moments are presented in Table 2.

As shown in Table 2, the fit for all four moments is close to exact. In addition, Table 2 reports the annual survival rate. This moment was not targeted in the calibration and, nevertheless, the predicted coefficient is very close to the empirical one. Moreover, the calibrated model does well in predicting the annual survival rate of firms, which was also not targeted in the calibration.

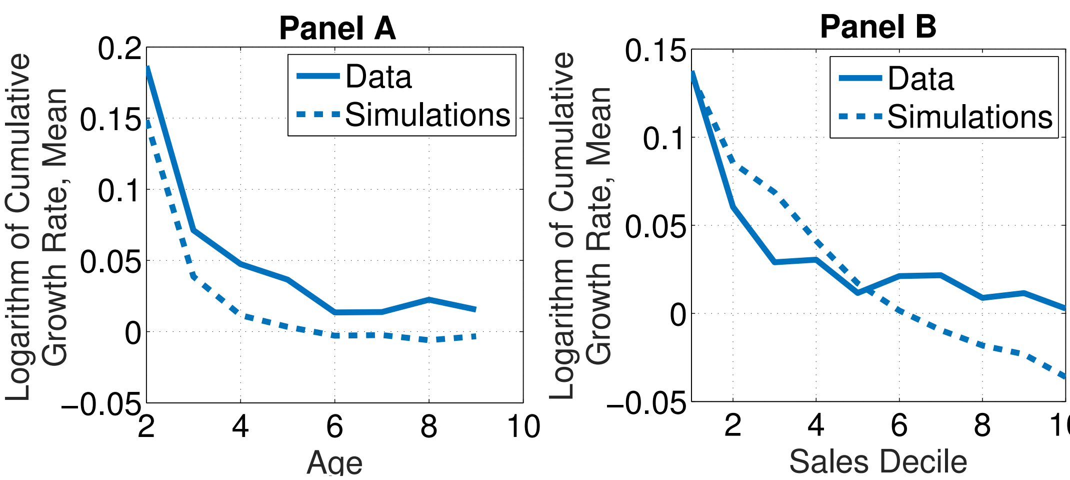
[12] The simplex search method is used to search over the parameter space $(\xi, f, \sigma_\epsilon, \sigma_\theta)$. To compute the simulated moments, 40,000 firms are simulated.

Given that learning affects both the growth and the exit behavior of firms, it's reassuring that our model is able to match this moment well.

3.2 Firm Growth Results

In this section we explore the quantitative implications of the calibrated model in terms of its ability to generate unconditional and conditional growth results. Figure 4 depicts the *unconditional* relationship between a firm's growth and age (Panel A), and a firm's growth and size (Panel B) for the simulated and actual data. A firm's size is measured as the sales in a given period; the growth rate measures the growth in a firm's sales between two consecutive periods. Observe from both panels of Figure 4 that Colombian data resembles similar patterns as found in the empirical literature: the growth rate declines in age and the growth rate declines in size.[13]

Figure 4: **Growth Rates with Age and Size**



Source: DANE survey data and authors calculations on model simulations.

The learning model is able to capture reasonably well both of these dependencies. As shown in Panel A, the model is able to replicate endogenously the non-linear relationship between growth and age: young firms have a high growth rate, which quickly declines as they learn their latent demand parameters. This non-linearity comes from the structure of the model and was not one of the calibration's targets, which used instead the age coefficient in a linear regression of growth rates on age and size. Notice from Panel A that in the simulated data, in contrast to the actual data, the growth rate gradually approaches zero as age rises. This is expected, since

[13]See Evans (1987a,b), Dunne et al. (1989), D'Erasmo (2011) and also Eaton et al. (2008) and Arkolakis (2015) for exporters.

Table 3: OLS Regression: the Age-Size Dependence of Growth Rates

	Data	Simulation
	(1)	(2)
$\log(RP_{i,t})$	-0.022***	-0.022***
	(0.002)	(0.001)
$\log(Age_{i,t})$	-0.035***	-0.035***
	(0.002)	(0.001)
Constant	0.428***	0.411***
	(0.022)	(0.012)
Sample	$Age_{i,t} \leq 20$	$Age_{i,t} \leq 20$
No. Obs.	36,882	205,882
R^2	0.02	0.03

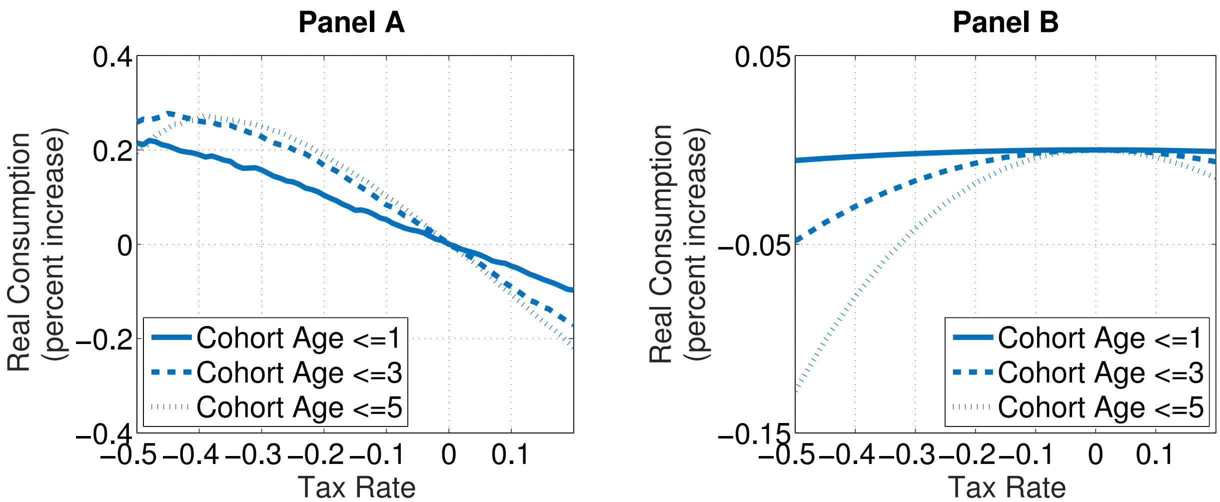
*** statistically significant at 1% level.

only source of firm growth in the model is learning. Once firms learn their true demand level, they experience no further growth. Further, as shown in Panel B of Figure 4, the relationship between growth and size is more linear in the model compared to the data. This is partially driven by large firms in the model exhibiting negative growth rates. In the model, sales depend on price realizations, which, in turn, depend on realizations of the preference shocks. Hence, an unusually large positive preference shock increases observed sales. However, since these shocks are transitory, there is mean reversion, so the following period realized sales will be lower on average yielding the negative growth rates observed in the figure.[14]

Moreover, the model is able to capture well the *conditional* size dependence of growth rates as presented in Table 3. The table present the results of a simple ordinary least squares regression of a firm's growth rate on its initial age and size. Column (1) presents results for the Colombian data and column (2) for the simulated data. Notice, that the age coefficient in column (1) was one of the target moments in the calibration, hence it is not surprising that the match is perfect. The size coefficient, however, was not targeted in the calibration. As can be seen from Table 3,

[14]Indeed, when the variance of the preference shock, σ_ϵ, shrinks, the growth rate of large firms converges to zero from below.

Figure 5: **Welfare Results**



Source: Authors calculations on model simulations.

the model is able to perfectly capture the conditional size dependence of growth rates.

3.3 Welfare Results

In this section we quantitatively explore the welfare implications of the learning model. As discussed in section 2.5 there are two opposing forces affecting welfare when a government introduces a fixed costs subsidy. On the one hand, there are more varieties in equilibrium which increases welfare (consumer-surplus effect). On the other hand, because of the increased number of firms, more resources are spent on the fixed costs and less on the production of the consumption good (profit-destruction effect).

Panel A of Figure 5 depicts the impact of a fixed-cost subsidy on real consumption. The y-axes displays the level of real consumption relative to its decentralized equilibrium level. In other words, 0.2 refers to a 0.2% increase in real consumption. Each curve in the figure corresponds to a subsidy applied to all firms below a certain age. For example, the curve corresponding to the label 'Cohort Age $\leq$ 5' refers to a subsidy rate applied to all firms that are 5 years old or younger.

Panel A of Figure 5 demonstrates that the equilibrium of our learning model is not efficient: welfare can be improved by subsidizing young firms. Following the discussion in Section 2.5 on the benefits versus the costs of the subsidy, we conclude that in our calibrated setup there are too few varieties in equilibrium. The introduction of a fixed-cost subsidy that is financed by taxing the consumers, leads to higher entry of newcomers and higher survival rates among the subsidized firms; hence, there is overall a larger number of varieties in equilibrium.

Furthermore, in an environment of demand uncertainty, firms which form beliefs that their demand conditions are not favorable, exit. As a result, some high-demand firms exit too early after observing a sequence of negative demand shocks which push firms beliefs sufficiently downwards to induce exit. The subsidy increases the survival rate of firms, and hence allows firms to receive more signals. Those high demand firms, which underestimated their demand due to a series of unlikely negative shocks, are able to grow and survive in the market as opposed to exiting early due to the lack of information. Indeed it is not only the mass of young firms that increases with the introduction of the subsidy: the mass of firms of all cohorts increases as a result of the subsidy (results available).

To underscore the difference between our setup and one with monopolistic competition and heterogeneous firms but no learning, Panel B of Figure 5 replicates an identical exercise as in Panel A in the context of the canonical Melitz (2003) model. Since the Melitz (2003) model without learning is Pareto optimal, any form of government intervention will not improve welfare, as illustrated by Panel B. This is not ex-ante obvious: with firm entry and exit, firms do not internalize the benefits to consumers of an additional variety. In addition, their entry has adverse effects on the profits of incumbent firms through the monopolistic distortion.[15] These effects exactly cancel out in the basic setup with CES demand and, as a result, efficiency holds even in the presence of firm heterogeneity, as pointed out by Dhingra and Morrow (2012). The key difference between our model and the Melitz (2003) model is that demand uncertainty implies that, when a firm is active, there is an option value from learning. The firm takes this option value into account when deciding whether to enter or exit. Put differently, the presence of learning alters firms' entry and exit decisions compared to the Melitz (2003) model.

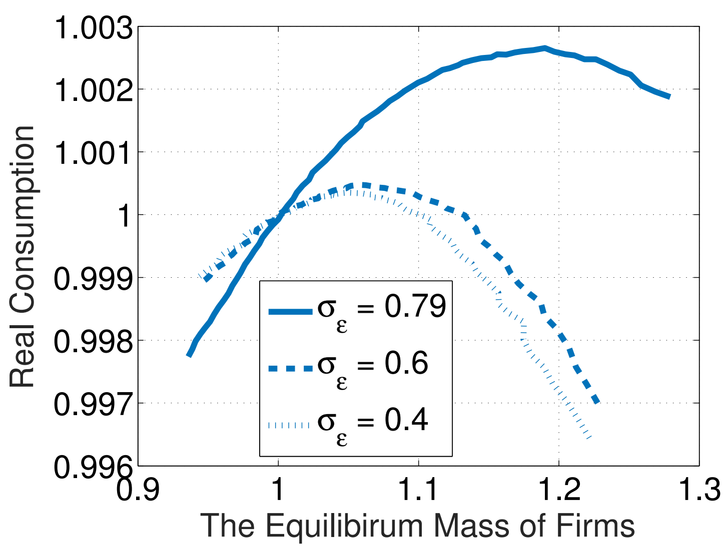
At an intuitive level, a key distinction behind the contrasting welfare results is that in a static Melitz model, a small variety enters and stays small despite any potential assistance. In a dynamic learning model, the same small variety enters and *may eventually become large*. Government assistance makes that outcome more probable and hence, enhances the consumer-surplus effect.

In addition, welfare gains are higher in more uncertain environments. Recall that the higher

[15]See also discussion in Grossman and Helpman (1991), page 82.

the variance of the preference shock, σ_ϵ, the more difficult the inference problem becomes for firms and the more unsure they are about their underlying demand. To examine how potential welfare gains depend on demand uncertainty, in Figure 6 each curve represents, for given level of the variance of the preference shock (noise), σ_ϵ, different combinations of real consumption (welfare) and equilibrium mass of firms relative to the decentralized equilibrium, for different subsidy levels. All the three curves intersect at point (1,1), at which point there is no subsidy or tax. As the subsidy level increases, so does the equilibrium mass of firms and therefore the varieties available to consumers and also welfare. In all three cases depicted, the curve is hump-shaped, suggesting that past a certain point there are no further welfare gains by further subsidizing firms. It is worth noting that the potential welfare gains are larger in more uncertain environments, whereas as the uncertainty approaches zero the gains are minimized. Following our earlier discussion, when uncertainty approaches zero, the consumer-surplus effect exactly offsets the profit-destruction effect as in Melitz (2003) and Chaney (2008) and there are no welfare gains from increasing the mass of firms/varieties. Conversely, as uncertainty increases and there is option value from learning that alters firms entry and exit decisions. Now the two effects no longer offset each other and there are potential welfare gains, as the consumer-surplus effect dominates the profit-destruction effect, at least initially.

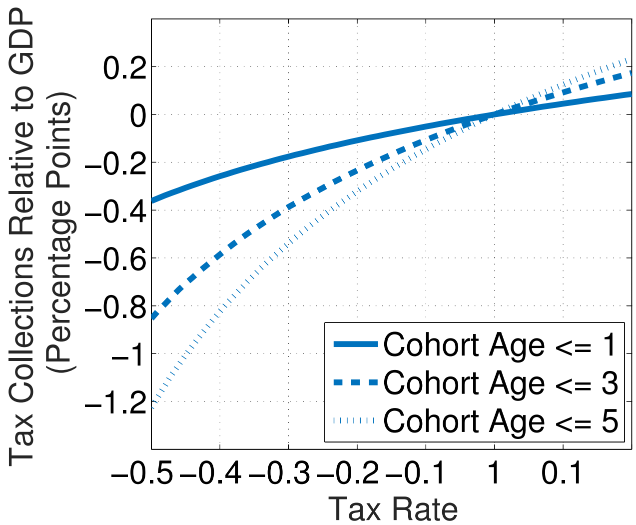
Figure 6: **Welfare and Learning**



Source: Authors calculations on model simulations. The subsidy is applied to all firms of 5 years or younger.

Finally, it is worth noting that despite a small magnitude of the potential welfare gains

Figure 7: **Tax Collections to GDP Ratio**



Source: Authors calculations on model baseline calibration.

depicted in Panel A of Figure 5, the gains are of a similar order of magnitude to the cost of the subsidy. For instance, a 35% subsidy on the fixed cost of all firms up to two years of age, costs 0.36% of GDP (see Figure 7) and leads to a 0.23% increase in aggregate consumption and therefore welfare implying a consumption multiplier of 0.64.[16] A lower subsidy leads to an even higher consumption multiplier: a subsidy for the first two years equal to 0.10% of GDP leads to a 0.08% increase in aggregate consumption (consumption multiplier of approximately 0.8). Notice that while the relative impact is large since the policy makes an impact only if applied to young firms the absolute increase of the GDP is small.[17]

4 Conclusion

In this paper we develop a framework to evaluate the importance of learning for firm growth by suitably adapting the standard learning mechanism of Jovanovic (1982) into the monopolistic competition framework of Melitz (2003). Our setup captures the dependence of growth rates on

[16]Subsidies in innovation has been shown to have a larger absolute effect as in, for example, Acemoglu et al. (2013). However, in that case the relative positive impact is not very large. For example the authors note that for a large innovation subsidy to new entrants amounting to 5% of GDP the effect of welfare is only about 0.5%

[17]In standard neoclassical models the consumption multiplier as a result of a government tax/policy is typically negative due to a strong wealth effect that decreases both consumption and leisure (see e.g. Baxter and King (1993) and Ramey (2011) and references therein). In our model there is no leisure and the negative impact of the lower wealth on consumption is compensated by the large increase in varieties as a result of the policy, resulting in an increase in real consumption.

age, even conditional on firm size. We calibrate our setup using this moment, as well as other moments from a panel of Colombian plant-level data and illustrate how targeted subsidies to the fixed costs of operations of young firms can lead to positive welfare gains. In our baseline setup we estimate gains from these targeted subsidies which are of similar order of magnitude to the cost of the subsidy.

References

Abbring, J. H. and J. R. Campbell (2003): "A Structural Empirical Model of Firm Growth, Learning, and Survival," *NBER working paper*, 9712.

Acemoglu, D., U. Akcigit, N. Bloom, and W. Kerr (2013): "Innovation, Reallocation and Growth," Tech. rep., National Bureau of Economic Research.

Albornoz, F., G. Corcos, E. Ornelas, and H. F. C. Pardo (2012): "Sequential Exporting," *Journal of International Economics*, 88, 17–31.

Arkolakis, C. (2015): "A Unified Theory of Firm Selection and Growth," *mimeo, Yale University.*

Atkeson, A. and A. Burstein (2010): "Aggregate Implications of Innovation Policy," *mimeo.*

Baxter, M. and R. G. King (1993): "Fiscal Policy in General Equilibrium," *American Economic Review*, 83, 315–334.

Bergemann, D. and J. Välimäki (2000): "Experimentation in markets," *The Review of Economic Studies*, 67, 213–234.

Broda, C. and D. Weinstein (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121, 541–585.

Castro, R., G. L. Clementi, and D. Corbae (2004): "Information and Firm Dynamics," Manuscript, University of Texas at Austin.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98, 1707–1721.

Clementi, G. L. and B. Palazzo (2013): "Entry, exit, firm dynamics, and aggregate fluctuations," *NBER working paper*, 19217.

Commision, E. (2010): "Communication from the Commision to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions," *European Commission.*

Cooley, T. F. and V. Quadrini (2001): "Financial Markets and Firm Dynamics," *American Economic Review*, 91, 91.

D'Erasmo, P. (2011): "Investment and firm dynamics," *Available at SSRN 994663.*

Dhingra, S. and J. Morrow (2012): "The Impact of Integration on Productivity and Welfare Distortions Under Monopolistic Competition," *mimeo, LSE.*

Dunne, T., M. J. Roberts, and L. Samuelson (1989): "The Growth and Failure of U.S. Manufacturing Plants," *Quarterly Journal of Economics*, 104, 671–698.

Eaton, J., M. Eslava, C. J. Krizan, M. Kugler, and J. Tybout (2012): "A Search and Learning Model of Export Dynamics," Manuscript, Pennsylvania State University.

Eaton, J., M. Eslava, M. Kugler, and J. Tybout (2008): "The Margins of Entry Into Export Markets: Evidence from Colombia," in *Globalization and the Organization of Firms and Markets*, ed. by E. Helpman, D. Marina, and T. Verdier, Massachusetts: Harvard University Press.

Ericson, R. and A. Pakes (1998): "Empirical Implications of Alternative Models of Firm Dynamics," *Journal of Economic Theory*, 79, 1–45.

Evans, D. S. (1987a): "The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries," *The Journal of Industrial Economics*, 35, 567–581.

——— (1987b): "Tests of Alternative Theories of Firm Growth," *The Journal of Political Economy*, 95, 657–674.

Grossman, G. and E. Helpman (1991): *Innovation and Growth in the Global Economy*, Cambridge, Massachussets: MIT press.

Haltiwanger, J., R. Jarmin, and J. Miranda (2013): "Who Creates Jobs? Small Vs. Large Vs. Young," *The Review of Economics and Statistics*, 95, 347–361.

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–1150.

Itskhoki, O. and B. Moll (2015): "Optimal Development Policies with Financial Frictions," *mimeo, Princeton University.*

Jovanovic, B. (1982): "Selection and the Evolution of Industry," *Econometrica*, 50, 649–670.

Klette, J. and S. Kortum (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112, 986–1018.

Luttmer, E. G. J. (2007): "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 122, 1103–1144.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725.

OECD (2013): "OECD Science, Technology and Industry Scoreboard 2013," *OECD Publishing.*

Ramey, V. A. (2011): "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature*, 49, 673–685.

Roberts, M. J. (1996): *Colombia, 1977-85: Producer Turnover, Margins, and Trade Exposure*, Oxford, UK: Oxford University Press, 227–259.

Ruhl, K. J. and J. Willis (2014): "New Exporter Dynamics," Manuscript New York University.

Timoshenko, O. A. (2015): "Product Switching in a Model of Learning," *Journal of International Economics*, 95, 233–249.

A Learning Model with Productivity Evolution

In this appendix we modify the learning model to allow for firm productivity, e^{z_t}, to change over time. We also calibrate this extended model with both learning and productivity evolution.

In particular, conditional on a firm's survival, productivity $z_t(\omega)$ changes over time according to

$$z_t(\omega) = \rho z_{t-1}(\omega) + u_t(\omega), \; u_t(\omega) \sim N\left(0, \sigma_u^2\right) \text{ i.i.d.}$$

Firms observe this period's productivity when making their quantity decision, but do not know next period's productivity realizations. In addition, today's quantity choice does not affect the evolution of productivity. Thus, the firm's optimal quantity decision is still given by equation (6). The firm's value function (dropping the ω and the t) is now given by

$$V(z,b,n) = \max\left\{E\pi(z,b) + \beta(1-\delta)E_{z'|z,b'|b,n}V(z',b',n+1), 0\right\}.$$

Since productivity now changes over time, firms take this into account both when making their entry decision, as well as when considering whether to remain in a market or exit. In particular, there is an option value of higher productivity draws in the future, especially given the convexity of e^z.

We now turn to the calibration. As in the calibration of the baseline model we calibrate some of the parameters independently. In particular, β and δ take the values of 0.9606 and 0.025 respectively, while the autocorrelation coefficient, ρ, is set to 0.999.

The remaining parameters are jointly calibrated. In particular, we calibrate the following five parameters: the per period fixed cost, f, the standard deviation of shocks to demands, σ_ε, the shape parameter of the distribution of the initial productivity draws, ξ, the standard deviation of the distribution of the unobserved demand parameter, σ_θ and the standard deviation of shocks to productivity, σ_u. We use the following five moments: mean log sales, the standard deviation of log sales, the sales-weighted growth rate of entrants, the entrants' share of sales and the annual survival rate.

As in the baseline model, it is impossible to provide a rigorous identification argument, but we offer an informal argument of how each parameter is identified from the data. The per period fixed cost, f, is pinned down by mean log sales, since increasing the fixed cost pushes less productive firms out of the market increasing mean sales. The standard deviation of log

Table 4: Parameter Values for the Learning Model with Productivity Evolution.

Parameter	Value
ξ	10.80
σ_ϵ	1.32
σ_θ	0.95
σ_u	0.02
f	248,572

Sources:Model Calibration with Productivity Evolution

Table 5: Data and Simulation Moments for the Learning Model with Productivity Evolution.

Targeted Moments:	**Data**	**Productiv. Model**	**No Productiv. Evolution**
Mean of Log-Sales	14.99	15.09	15.08
Share of Sales from Entrants	2.66%	2.65%	2.50%
Sales-Weighted Growth Rate of Entrants	23%	21%	19%
Age Coefficient	-0.035	-0.035	-0.032
Annual Survival Rate	91%	90%	92%
Other Moments:			
Size Coefficient	-0.022	-0.023	-0.017

Sources: DANE survey data (see text for details); Model calibration with productivity evolution; Counterfactual model simulation when $z_{t+1} = z_t$.

sales informs us about the standard deviation of the distribution of the unobserved demand, σ_θ, since higher dispersion in that distribution translates into a more disperse sales distribution. The shape of the distribution of the initial productivity draws, ξ, is pinned down by entrants' share of sales: A lower value for ξ, all else equal, leads to a higher initial productivity dispersion and therefore more entrants starting off relatively large. Finally, the sales-weighted growth rate of entrants and the annual survival rate identify the standard deviation of shocks to demand, σ_ε and the standard deviation of shocks to productivity, σ_u. Intuitively, a lower σ_ε implies faster learning and therefore, all else equal, low demand firms exit sooner, while surviving firms grow faster. Similarly, a higher dispersion of productivity shocks, σ_u, leads to both a higher exit rate, as more firms are hit by large negative shocks, but also higher growth rates. Therefore the growth and survival moments allow us to identify these two moments.

The calibrated parameters are presented in Table 4 and the targeted moments are presented in Table 5. Two comments are in order: First the fit of the model is quite good, with all moments quite close to their targets. In addition as shown in Table 5, the model also matches the size coefficient of the regression of growth of sales on age and size. This moment was not targeted in the calibration and nonetheless the fit is close to exact.

Second, the calibrated parameters indicate that most of the action is driven by learning rather than productivity. Indeed the standard deviation of productivity shocks, σ_u, necessary to match the data is very small, especially compared to the two parameters capturing the importance of learning, σ_ε and σ_θ.

This becomes even more apparent when we use the calibrated model with productivity growth and shut down the productivity evolution channel (set $z_t = z_0$ for all t). The resulting moments are presented in the last column of Table 5. Notice, that with the exception of the size coefficient, all other moments do not change much. In particular, the age coefficient changes by less than ten percent suggesting that learning is the key mechanism in accounting for the conditional age dependence of growth rates. The absolute value of the size coefficient however, declines by almost a third indicating that time-varying productivity plays a much larger role in accounting for the conditional size (rather than age) dependence of growth rates.

B Expected Growth Rate

Expected growth rate conditional on size

The expected growth rate of a firm with current size q_t is given by

$$E_{\bar{a}'}\left(\frac{q_{t+1}(z,\bar{a}',n+1)}{q_t(z,\bar{a},n)}\right) = \frac{E_{\bar{a}'}\left(q_{t+1}(z,\bar{a}',n+1)\right)}{q_t(z,\bar{a},n)}. \tag{13}$$

Using equation (6), we can substitute in for the firm's quantity choice each period to obtain

$$\frac{E_{\bar{a}'}\left(q_{t+1}(z,\bar{a}',n+1)\right)}{q_t(z,\bar{a},n)} = \frac{E_{\bar{a}'}\left(\left(\frac{\sigma-1}{\sigma}\right)^\sigma \left(\frac{b_{t+1}(\bar{a}',n+1)e^z}{w}\right)^\sigma \frac{Y}{P^{1-\sigma}}\right)}{\left(\frac{\sigma-1}{\sigma}\right)^\sigma \left(\frac{b_t(\bar{a},n)e^z}{w}\right)^\sigma \frac{Y}{P^{1-\sigma}}}$$

$$= \frac{\left(\frac{\sigma-1}{\sigma}\right)^\sigma \left(\frac{e^z}{w}\right)^\sigma E_{\bar{a}'}\left(b_{t+1}(\bar{a}',n+1)^\sigma\right)\frac{Y}{P^{1-\sigma}}}{\left(\frac{\sigma-1}{\sigma}\right)^\sigma \left(\frac{b_t(\bar{a},n)e^z}{w}\right)^\sigma \frac{Y}{P^{1-\sigma}}} = \frac{E_{\bar{a}'}\left(b_{t+1}(\bar{a}',n+1)^\sigma\right)}{b_t(\bar{a},n)^\sigma}.$$

Denote $b_{t+1}(\bar{a}',n+1)$ by b', and $b_t(\bar{a},n)$ by b. We know that $(b^\sigma)'$ is log normally distributed, with mean

$$m_n = \log\left(b^\sigma\right) - \frac{v_n^2 - v_{n+1}^2}{2\sigma}$$

and variance given by

$$s_n^2 = \frac{\lambda^2 \left(v_n^2 + \sigma_\varepsilon^2\right)}{\left(1 + (n+1)\,\lambda\right)^2}$$

where

$$\lambda = \frac{\sigma_\theta^2}{\sigma_\varepsilon^2}$$

and

$$v_n^2 = \frac{\sigma_\theta^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2} = \frac{\lambda \sigma_\varepsilon^2}{1 + n\lambda}$$

and n is the firm's age (number of observations).

Thus the expected growth rate is given by

$$E_{\bar{a}'}\left(\frac{q_{t+1}}{q_t}\right) = \frac{E_{\bar{a}}\left((b^\sigma)'\right)}{b^\sigma} = \frac{\exp\left(m_n + \frac{s_n^2}{2}\right)}{b^\sigma} = \frac{b^\sigma \exp\left(\frac{1}{2}\frac{\lambda^2\left(v_n^2+\sigma_\varepsilon^2\right)}{(1+(n+1)\lambda)^2} - \frac{v_n^2 - v_{n+1}^2}{2\sigma}\right)}{b^\sigma} =$$

$$\exp\left(\frac{1}{2}\frac{\lambda^2\left(v_n^2 + \sigma_\varepsilon^2\right)}{\left(1+(n+1)\,\lambda\right)^2} - \frac{v_n^2 - v_{n+1}^2}{2\sigma}\right).$$

Straightforward calculations show that we can rewrite the above growth rate as

$$\exp\left(\frac{\lambda^2 \sigma_\varepsilon^2 (\sigma - 1)}{2\sigma (1 + n\lambda)(1 + (n+1)\,\lambda)}\right).$$

The derivative of the above growth rate with age (n) gives

$$\frac{\partial \exp\left(\frac{\lambda^2\sigma_\varepsilon^2(\sigma-1)}{2\sigma(1+n\lambda)(1+(n+1)\lambda)}\right)}{\partial n}$$

$$\begin{aligned} &= \exp\left(\frac{\lambda^2 \sigma_\varepsilon^2 (\sigma - 1)}{2\sigma (1 + n\lambda)(1 + (n+1)\,\lambda)}\right) \frac{\partial\left(\frac{\lambda^2\sigma_\varepsilon^2(\sigma-1)}{2\sigma(1+n\lambda)(1+(n+1)\lambda)}\right)}{\partial n} \\ &= -\exp\left(\frac{\lambda^2 \sigma_\varepsilon^2 (\sigma - 1)}{2\sigma (1 + n\lambda)(1 + (n+1)\,\lambda)}\right) \frac{2\sigma\lambda^2\sigma_\varepsilon^2(\sigma-1)\left(2\lambda + 2n\lambda^2 + \lambda^2\right)}{\left(2\sigma(1+n\lambda)(1+(n+1)\,\lambda)\right)^2} < 0 \end{aligned}$$

since $\sigma > 1$ and all other parameters are positive.

Expected growth rate conditional on age and survival

As discussed in Section 2.3, a firm's state is given by the triplet $(z, \bar{a}, n)$. The solution to the firm's entry and exit problem described in (9) determines a set of market participation thresholds. Denote by $\bar{a}^*(n, z)$ such market participation thresholds, such that firm (z, n) stays in the market if $\bar{a} \geq \bar{a}^*(n, z)$. Hence, the expected growth rate conditional on survival can be written as

$$E_{\bar{a}'}\left(\frac{q_{t+1}(z,\bar{a}',n+1)}{q_t(z,\bar{a},n)}|\bar{a}'\geq \bar{a}^*(n+1,z),\bar{a},n\right)=$$

$$=E_{\bar{a}'}\left(\frac{\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\left(\frac{b(\bar{a}',n+1)e^z}{w}\right)^{\sigma}\frac{Y}{P^{1-\sigma}}}{\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\left(\frac{b(\bar{a},n)e^z}{w}\right)^{\sigma}\frac{Y}{P^{1-\sigma}}}|\bar{a}'\geq \bar{a}^*(n+1,z),\bar{a},n\right)=$$

$$=E_{\bar{a}'}\left(\frac{(b(\bar{a}',n+1))^{\sigma}}{(b(\bar{a},n))^{\sigma}}|\bar{a}'\geq \bar{a}^*(n+1,z),\bar{a},n\right)$$

From equations (3) and (7), $b(\bar{a},n)$ is monotonically increasing in $\bar{a}$. Hence, the conditional expected growth rate can be written as

$$E_{b'}\left(\frac{(b')^{\sigma}}{b^{\sigma}}|b'^{\sigma}\geq b^*(n+1,z)^{\sigma},\bar{a},n\right),$$

where b'^{σ} is log-Normally distributed with the mean and variance as specified before. Thus, the expected growth rate can be computed as

$$\frac{\int_{b^*(n+1,z)^{\sigma}}^{+\infty}\frac{b'^{\sigma}}{b^{\sigma}}pdf(b'^{\sigma})db'^{\sigma}}{Prob(b'^{\sigma}\geq b^*(n+1,z)^{\sigma})}=$$

$$=\frac{\frac{1}{b^{\sigma}}e^{m_n+\frac{1}{2}s_n^2}\Phi\left(\frac{m_n+s_n^2-\log(b^*(n+1,z)^{\sigma})}{s_n}\right)}{\Phi\left(\frac{m_n-\log(b^*(n+1,z)^{\sigma})}{s_n}\right)}=$$

$$=\exp\left(\frac{\lambda^2\sigma_{\varepsilon}^2(\sigma-1)}{2\sigma(1+n\lambda)(1+(n+1)\lambda)}\right)\frac{\Phi\left(\frac{m_n+s_n^2-\log(b^*(n+1,z)^{\sigma})}{s_n}\right)}{\Phi\left(\frac{m_n-\log(b^*(n+1,z)^{\sigma})}{s_n}\right)}$$

Denote by $K(b^{\sigma},z|n)=m_n-\log(b^*(n+1,z)^{\sigma})$. Variable K is monotonically increasing in a firm's size conditional on age. A firm's size is determined by the product $b^{\sigma}e^z$. Thus, an increase in firm size corresponds to either an increase in b^{σ}, or z, or both. As b^{σ} increases, m_n increases, and so does K. As z increases, b^* declines, and hence K rises. As both b^{σ} and z increase, m_n increase and $\log(b^*(n+1,z)^{\sigma})$ declines; hence, K rises.

The conditional on age expected growth rate can now be written as

$$E_{\bar{a}'}\left(\frac{q_{t+1}(z,\bar{a}',n+1)}{q_t(z,\bar{a},n)}|\bar{a}'\geq \bar{a}^*(n+1,z),\bar{a},n\right)=$$

$$=\exp\left(\frac{\lambda^2\sigma_{\varepsilon}^2(\sigma-1)}{2\sigma(1+n\lambda)(1+(n+1)\lambda)}\right)\frac{\Phi\left(\frac{K(b^{\sigma},z|n)}{s_n}+s_n\right)}{\Phi\left(\frac{K(b^{\sigma},z|n)}{s_n}\right)}.$$

By the property of the standard normal cumulative distribution function, $\Phi\left(\frac{K(b^{\sigma},z|n)}{s_n}+s_n\right)/\Phi\left(\frac{K(b^{\sigma},z|n)}{s_n}\right)$ declines in K. Thus, conditional on age and survival, the expected growth rate declines in size.

C Solving the Stationary Equilibrium

The stationary equilibrium objects that need to be solved for are $e^{\underline{z}}, M, P, Y, w$.

Change of variables

$$
\begin{aligned}
u^{\sigma-1} &= \frac{(e^{z})^{\sigma-1}P^{\sigma-1}Y}{w^{\sigma-1}} \\
(u^{*})^{\sigma-1} &= \frac{(e^{\underline{z}})^{\sigma-1}P^{\sigma-1}Y}{w^{\sigma-1}}
\end{aligned}
$$

u^{*} is a solution to the firm's entry problem

$$V(u^{*}, b^{e}, 0) = 0$$

Next

$$M = \frac{L}{\tilde{r} - \tilde{\pi}},$$

where $\tilde{r}$ is the mean revenue of firms and $\tilde{\pi}$ is the mean profit level of firms.

$e^{\underline{z}}$ is a solution to

$$
\begin{aligned}
M &= J\left(\frac{e^{z_{min}}}{e^{\underline{z}}}\right)^{\xi} \times Mass \\
e^{\underline{z}} &= \left(\frac{J \times Mass}{M}\right)^{\frac{1}{\xi}} e^{z_{min}},
\end{aligned}
$$

where, $Mass$ is the mass of firms as determined by $m(z, b, n)$. Next

$$
\begin{aligned}
Y &= L + M\tilde{\pi}, \\
P &= \frac{u^{*}}{Y^{\frac{1}{\sigma-1}} e^{\underline{z}}}.
\end{aligned}
$$

D Data

The measure of a plant's sales is taken to be the real value of production (variable RP in the dataset). The real value of production is reported in thousands of pesos. Thus, all values are multiplied by 1000.

The first calibrating moment - the mean of the logarithm of sales - is constructed by taking a cross sectional mean of the logarithm of plants' sales for a given year. The reported value is the mean across annual observations between 1983 and 1991.

For the second calibrating moment - the share of sales from entrants - an entrant is defined as a plant which is observed selling a positive amount in a given period, and is not observed in the sample in the previous period. The reported value is the mean across annual observations between 1983 and 1991.

The third calibrating moment - the sales-weighted growth rate of entrants - is constructed by taking the sales-weighted mean of the cumulative growth rate of sales of surviving entrants. The reported value is the mean across annual observations between 1983 and 1991.

The fourth calibrating moment - the age coefficient - is taken to be the value of the coefficient β_2 in the regression below

$$\log\left(\frac{RP_{i,t+1}}{RP_{i,t}}\right) = \alpha + \beta_1 \log(RP_{i,t}) + \beta_2 \log(Age_{i,t}) + \epsilon_{i,t},$$

where $RP_{i,t}$ is the real value of production for plant i at time t. Variable *Age* measures the number of *consecutive* years a plant is observed in the sample. For example, for a plant that is observed in years 1984, 1986, and 1987, the plant's age in 1984 is 1, in 1986 is 1, in 1987 is 2. The age of a plant in 1983 (the start of our sample period) is determined by the difference between 1983 and the plants start year (variable X6 in the dataset). The regression is run on the subsample of plants with $Age_{i,t} \leq 20$. Results are reported in column (1) in Table 3 in the main text.