Heterogeneous Markups, Growth and Endogenous Misallocation

Michael Peters*
Yale University

December 2016

Abstract

This paper proposes a parsimonious model of growth, where imperfect product markets allow firms to charge heterogeneous markups. Firms increase the productivity of their existing products to raise markups and grow in size by expanding into new markets. The interplay between these forces determines the stationary distribution of markups and firm size and the long-run rate of growth. Policies that raise firms' costs to enter new markets increase misallocation by lowering product market competition and reduce firm size by limiting the number of markets firms compete in. I apply the theory to panel data from manufacturing firms in Indonesia and show that seemingly large differences in the number of active firms and their average size are consistent with empirically plausible differences in firms' life cycle dynamics, entry rates and aggregate productivity growth between rich and poor countries.

*eMail: m.peters@yale.edu. I am especially grateful to Daron Acemoglu, Abhijit Banerjee and Rob Townsend for their invaluable guidance. I also thank Ufuk Akcigit, Susanto Basu, Penny Goldberg, Chang-Tai Hsieh, Sam Kortum, Giuseppe Moscarini, Ezra Oberfield and Aleh Tsyvinsk as well as numerous seminar participants, whose questions and suggestions benefited the paper substantially. I am also grateful for the financial assistance of the George and Obie Shultz fund to acquire the data necessary for this project.

1 Introduction

One of the major developments in the recent literature on growth and long-run economic development has been the focus on firm-level data. Some of the key stylized facts, which emerged from this literature, are (i) that firms in poor counties are small (but that there are lots of them), (ii) that the small size of such firms is less a reflection of their small scale when entering but rather due to slow growth dynamics over the life cycle and (iii) that poor countries suffer from persistent misallocation, whereby variation in firms' marginal products reduces aggregate productivity. At the same time, entry rates do not vary systematically across countries.

In this paper I show that these patterns are consistent with a parsimonious model of firm growth, where imperfect product markets allow firms to charge heterogeneous markups. I start from the premise that developing economies might suffer from what I call *market barriers,* i.e. frictions, which will hamper firms' ability to enter new markets. Such barriers can be related to policies like size-requirements or lengthy approval processes for production licenses. They could also be technological in nature, whereby the costs of breaking into markets firms previously did not cater to, are higher in developing countries. Such barriers reduce competition and increase markups and misallocation by rendering product markets captive to incumbent firms. At the same time, they also limit the extent of firm growth, keep firms small and hence increase the number of active firms. I find that this mechanism can quantitatively go a long way when calibrated to cross-country patterns on life cycle growth and entry rates. Importantly, I also show that these patterns are consistent with a stable world income distribution: despite very different processes of firm dynamics, the model is consistent with countries growing at essentially the same rate in long-run.

The theory has two key ingredients. I start from recent models of firm–based growth in the spirit of Klette and Kortum (2004). This is an attractive framework which provides a tractable way to connect firm-level growth and entry incentives to the resulting distribution of firm size and entry and exit patterns. I then augment this framework with a structure of imperfect product markets, where firms engage in non-competitive pricing and charge variable markups. In particular, firms can spend resources to increase the productivity in their existing markets to shield themselves from competition and sustain high markups in equilibrium. At the same time, firms face the constant threat of losing their markets to either new entrants or other firms, which try to expand their costumer base. It it this joint process of new firms entering the economy and existing firms expanding into new markets and engaging in markup increasing quality improvements, which will determine the long-run distribution of markups and firm size.

The model has an analytic solution and makes precise predictions how these equilibrium objects are related. In particular, the distribution of markups and firm size are fully characterized by two endogenous summary statistics. The key statistic determining the distribution of markups is the what I call *intensity of creative destruction,* which is simply the rate at which firms lose their exiting markets relative to the speed at which they increase their market power. In particular, I show that the stationary distribution of markups takes a pareto form, whose shape parameter is increasing in this endogenous statistic. If creative destruction is intense, the cross-sectional distribution of markups is compressed because entering and expanding firms introduce sufficient churning in the economy to keep monopoly power limited. If on the other hand creative destruction is limited, the distribution of markups has a fat tail because incumbent firms are unlikely to be replaced and have ample time to accumulate market power. A similar statistic exists for the properties of the firm size distribution, which is fully parametrized by the *intensity of market expansion,* which is simply the rate at which existing firms break into new markets relative to the equilibrium rate of entry. The higher this intensity, the larger will the average firm be and the less firms will be able to survive in equilibrium.

With these two summary statistics at hand, it is easy to analyze different policies. An equilibrium with both misallocation and many small firms requires a low intensity of creative destruction and a low intensity of market expansion. Market barriers have this exact implication as they reduce the extent to which existing firms expand into into novel product markets. This not only keeps markups high and dispersed, but it also reduces average firm size and the rate of employment growth over firms' life cycle. In equilibrium, this directly implies that the economy will be populated by many firms, most of which are small, and that resources will be misallocated across firms. In contrast, higher entry costs, i.e. frictions which raise the costs of *new* firms to enter the economy, have the opposite effect. While lower entry rates will also increase markups and hence cause misallocation, firms will grow faster as they face less competition from entrants. In equilibrium, entry costs will therefore increase average firm size and reduce the number of active producers.

I apply the theory to firm-level panel data for formal manufacturing firms in Indonesia. Two important empirical moments to discipline the equilibrium intensities of creative destruction and market expansion are the increase of markups and firm size over the firm's life cycle. I find that firms both increase their markups and their size as they age. To quantify the importance of market barriers and entry costs for the cross-country data, I calibrate the model and identify differences in market barriers and entry costs between the US and Indonesia from two cross-country moments: the equilibrium entry rate and the rate of life cycle growth. The calibration implies that both entry costs and market barriers are lower in the US. While I estimate entry to be about 15% less costly in the US, the difference in expansion barriers amounts to more than 30%. The quantitative effect of this variation is sizable. The model predicts that the aggregate importance of small producers declines by 75% and average firm size more than doubles. Moreover, fiercer competition reduces markups and misallocation. This in turn increases aggregate TFP by about 0.3% and reduces the "monopoly tax" on factor supplies (i.e. the labor wedge) by about 2%.

The implications for the equilibrium growth rate are subtle. Higher entry costs and expansion barriers in Indonesia reduce the extent of creative destruction. This reduces the equilibrium growth rate. At the same time, by increasing the survival probabilities for existing firms, such barriers increase the incentives for firms to increase productivity within their markets. This tends to increase the equilibrium growth rate, albeit at the cost of higher markups. In my calibration, these effects essentially cancel out, rendering the growth rate insensitive to entry costs and expansion barriers. Differences in firm size and misallocation are therefore consistent with a situation of balanced growth, where the distribution of income across countries does not diverge in the long-run. This is in stark contrast to a version of the model, which abstracts from the markup channel. If firms were not engaged in markup increasing innovation, the observed cross-country variation in entry rates and life cycle employment growth would imply large differences in aggregate productivity growth.

Finally I use the variation across different provinces in Indonesia to give some direct evidence for the importance of market barriers. I show that average firm size correlates positively with both the extent of life cycle employment growth and average markups and negatively with regional entry rates. This is consistent with a view where expansion barriers are an important source of variation across regions within Indonesia but inconsistent with varying costs of entry.

Related Literature This paper provides a parsimonious theory, where the extent of static misallocation and the long-run distribution of firm size are jointly determined. Empirically, these aspects of the firm-level data are negatively correlated. In particular, firms in poor countries are small but numerous (see e.g. Hsieh and Olken (2014), Poschke (2011) or Bento and Restuccia (forthcoming)), they experience slower growth as they age (Hsieh and Klenow (2014)) and misallocation of resources across such firms is more severe (see e.g. Hsieh and Klenow (2009), Restuccia and Rogerson (2008), Bartelsman et al. (2013) and the survey article by Hopenhayn (2012)).

A recent literature has connected these two observations by exploring how changes in exogenous distortions affect firms' innovation incentives and the patterns of firm entry. Hsieh and Klenow (2014), Buera and Jaef (2016) and Bento and Restuccia (forthcoming) for example show how "size-dependent policies", i.e. distortions, which are exogenously correlated with firm size, reduce firms' incentive to expand. Fattal Jaef (2011) and Yang (2012) study the aggregate effects of such distortions in environments with free entry. In this paper, I take a different approach. I construct a model, where misallocation is fully endogenous and hence emerges as an equilibrium outcome together with the firm size distribution and the rate of entry.[1]

My theory is an endogenous growth model in the Schumpeterian tradition of Aghion and Howitt (1992) and Grossman and Helpman (1991). In terms of modeling choices I build heavily on Klette and Kortum (2004). This framework is analytically attractive, can rationalize many features of the data and has been used to empirically study the extent of creative destruction in Denmark and the US (Lentz and Mortensen, 2008; Garcia-Macia et al., 2016), the welfare implications of industrial policies (Acemoglu et al., 2012), the importance of different margins of innovation behavior at the firm-level (Akcigit and Kerr, 2015) and differences in the process of firm dynamics across countries (Akcigit et al., 2015). In contrast to these papers, I show how to extend this framework in a tractable way to generate heterogeneous markups across producers. This extra margin does not only generate additional testable predictions but also has novel aggregate implications. In particular, my model falls outside of the class of models analyzed in Atkeson and Burstein (2015) as changes in innovation policy will not only affect aggregate productivity growth but also change the nature of competition on output markets.

While this is, to the best of my knowledge, the only paper that focuses on imperfect output markets in relation to the literature on growth and misallocation, there is a large and growing literature in the field of international trade stressing the importance of markups. On the theory side, Bernard et al. (2003), Melitz and Ottaviano (2008) or Atkeson and Burstein (2008) are examples of theories that generate heterogeneous markups. In contrast to the model of this paper, all these frameworks are static in the sense that firm efficiency is exogenous. That trade might have additional welfare gains by reducing markup heterogeneity and misallocation is explicitly stressed in Edmond et al. (2015), Epifani and Gancia (2011) and Holmes et al. (2013).[2]

There is also large empirical literature which shows that markups vary systematically in the cross-section of firms. In particular, markups are low for entering firms (Foster et al. (2008)), high for exporters (De Loecker and Warzynski, 2012) and increase in response to trade liberalizations, which reduce the price of imported inputs (De Loecker et al., 2016). Furthermore, firm-specific prices are argued to be an important source of variation in revenue-based productivity measures (De Loecker, 2011a). The dynamic model in this paper is consistent with these facts as it predicts that markups are increasing in age and size and that pass-through is imperfect.

In order to explain the negative correlation between firm size and misallocation at the country-level, my theory suggests an important role for frictions for existing firms to enter new product markets. This is consistent with a set of recent papers, which argue for the importance of costly consumer acquisition (e.g. Arkolakis (2008), Gourio and Rudanko (2014) and Perla (2016)). In a static model, Bento (2016) also shows that higher costs of entry for new firms will increase average firm size and stresses the importance of frictions for firms entering multiple markets. The finding that product market competition has ambiguous effects on the long-run growth rate is also present in Aghion et al. (2001) and Aghion et al. (2005).[3]

The rest of the paper proceeds as follows. In the next section I present the theory and show how misallocation

[1]The reduced form of my model is isomorphic to these papers and hence provides a micro-foundation for the equilibrium distribution of the modeling device of firm-specific wedges. There are, of course, other theories of misallocation. The vast majority of contributions focuses on frictions in firm' input choices, studying models with imperfect capital markets (Buera et al., 2011; Moll, 2014; Banerjee and Moll, 2010; Midrigan and Xu, 2010), contractual imperfections (Acemoglu et al., 2007) or capital adjustment costs Asker et al. (2014). This is in contrast to this paper, which argues that the equilibrium distribution of marginal products reflects monopolistic power and not binding input constraints. In the empirical analysis in Section 3 I will present evidence to distinguish some of these explanations.

[2]Arkolakis et al. (2012), however, find that heterogeneous markups might also reduce the gains from trade.

[3]The mechanism, however, is different. These papers stress the importance of "escape competition", whereby product market competition increases innovation incentives by firms. In my theory, the ambiguous effect of market barriers on the rate of growth is due to a within-firm composition effect, whereby firms shift their innovation activity from market expansion to productivity increases of their existing products.

and the firm size distribution are jointly determined in equilibrium. Section 3 contains the empirical analysis using Indonesian micro data and the calibration exercise to quantify the aggregate effects of market barriers and entry costs. Section 4 concludes.

2 Theory

2.1 The Environment

There is a measure one of infinitely lived households, supplying their unit time endowment inelastically. Individuals have preferences over the unique consumption good, which are given by

$$U = \int_{t=0}^{\infty} e^{-\rho t} ln\left(c_t\right) dt.$$

This final good, which I take to be the numeraire, is a Cobb-Douglas composite of a continuum of differentiated varieties

$$\mathrm{ln} Y_t = \int_0^1 \mathrm{ln}\left(\sum_{f \varepsilon S_{it}} y_{fit}\right) di, \tag{1}$$

where y_{fit} is the quantity of variety i bought from firm f and S_{it} denotes the number of firms active in the market for variety i at time t. Hence, different varieties i and i' are imperfect substitutes, whereas there is perfect substitutability between different brands within a variety. I will also refer to a variety as a market.

Firms can be active in multiple markets and the only source of heterogeneity across firms is their factor-neutral productivity to produce different varieties. In particular, a firm f producing variety i with current productivity q produces output according to

$$y_{fi} = q_{fi} l,$$

where l is amount of labor hired. The market for intermediate goods is monopolistically competitive, so that firms take aggregate prices as given. However, firms compete a la Bertrand with producers offering the same variety. This strategic interaction across producers will be the source of heterogeneous markups and aggregate misallocation.

Both the set of competing firms $\left[S_{it}\right]_i$ and firms' productivities $\left[q_{fit}\right]_{fi}$ evolve endogenously through (i) the entry of new producers into the economy, (ii) the expansion of existing firms into new markets, i.e. into varieties they did not produce before and (iii) productivity increases by current producers in markets they already serve. While the first two margins of growth are considered in Klette and Kortum (2004), the third aspect is novel. It is this intensive margin of innovation that allows firms to gain competitiveness relative to other firms in the market they serve and will give rise to heterogeneous markups across producers. At the aggregate level, this ingredient provides the link between growth, misallocation and the distribution of firm size.

2.2 Static Allocations, Markups and Misallocation

Consider first the static allocations for a given number of firms and distributions of productivity q. Given that production takes place with a constant returns to scale technology, firms compete in prices and different brands of variety i are perceived as perfect substitutes, in equilibrium only the most productive firm will be active. However, the presence of competing producers (even though they are less efficient) imposes a constraint on the leading firm's price setting. Because the demand function associated with (1) has a unitary price elasticity, the most efficient firm will resort to limit pricing. Letting q_{it} denote the productivity of the actual producer, i.e. the most efficient firm within the market, the equilibrium markup in market i is given by

$$\mu_{it} \equiv \frac{p_{it}}{w_t/q_{it}} = \frac{w_t/q_{it}^F}{w_t/q_{it}} = \frac{q_{it}}{q_{it}^F}, \tag{2}$$

where w_t denotes the equilibrium wage and w_t/q_{it}^F is the marginal cost of the second most productive firm, which I will refer to as the follower.[4] Intuitively, a bigger productivity advantage shields the current producer from competition and allows him to post a higher markup.

From (2) one can also derive the allocation of labor at the firm-level. Letting N_{ft} be the set of markets firm f is active in, total employment of firm f, l_{ft}, is given by

$$l_{ft} = \sum_{i\in N_{ft}} l_{fit} = \sum_{i\in N_{ft}} \frac{1}{q_{ift}} y_{ift} = \sum_{i\in N_{ft}} \frac{1}{q_{ift}} \frac{Y_t}{p_{it}} = \frac{Y_t n_{ft}}{w_t} \times \left(\frac{1}{n_{ft}} \sum_{i\in N_{ft}} \mu_{it}^{-1} \right) \equiv \frac{Y_t n_{ft}}{w_t} \times \mu_{ft}^{-1}, \tag{3}$$

where (1) implies that $p_{it} y_{ift} = Y_t$, $n_{ft} = |N_{ft}|$ is the number of markets the firm is active in and the last equality defines the average markup at the firm-level as $\mu_f^{-1} \equiv \frac{1}{n} \sum_{i\in N_{ft}} \mu_{it}^{-1}$. Hence, expanding into novel markets increases employment at the firm-level. Conversely, for a given number of markets n_{ft}, higher markups reduce firm-employment, so that variation in markups induce variation in employment holding the number of markets fixed.

It is this variation, which induces an inefficient allocation of resources and hence aggregate misallocation. To see this, note that firm f's revenue labor productivity is given by

$$MRPL_{ft} \equiv \frac{p_{ft} y_{ft}}{l_{ft}} = \frac{n_f Y_t}{l_{ft}} = w_t \times \mu_{ft}, \tag{4}$$

i.e. revenue productivity is not equalized but reflects the variation in equilibrium markups. In the framework of Hsieh and Klenow (2009) and Restuccia and Rogerson (2008), measured revenue productivity is proportional to $\frac{(1+\tau_{K,i})^\alpha}{1-\tau_{Y,i}}$, where $\tau_{K,i}$ and $\tau_{Y i}$ are exogenous firm-specific taxes on capital and output. Hence, firms charging a high markup have high productivity and would be identified as facing high distortionary taxes.

Given the above structure, the economy has a transparent aggregate representation. Letting $L_{P,t}$ denote the total mass of production workers, (3) implies that

$$L_{P,t} = \sum_f l_{ft} = \frac{Y_t}{w_t} \sum_f \sum_{i\in N_{ft}} \mu_{it}^{-1} = \frac{Y_t}{w_t} \times \left(\int_0^1 \mu_{it}^{-1} di \right). \tag{5}$$

Similarly, equilibrium wages are given by

$$w_t = exp\left(\int_0^1 ln\left(q_{it}^F\right) di \right) = exp\left(\int_0^1 ln\left(\frac{q_{it}}{\mu_{it}}\right) di \right) = Q_t \times exp\left(\int_0^1 ln\left(\mu_{it}^{-1}\right) di \right),$$

[4]It is at this point where the assumption of the aggregate production function being Cobb-Douglas simplifies the exposition. If the demand elasticity was to exceed unity, the firm might want to set the unconstrained monopoly price in case its productivity advantage over its closest competitor is big enough. In particular, if $\sigma > 1$ was the demand elasticity, the optimal price was $p = \frac{w}{q^F} min\left(\frac{\sigma}{\sigma-1}\frac{q^F}{q}, 1\right)$. In the limit where $\sigma \to 1$, we get $min\left(\frac{\sigma}{\sigma-1}\frac{q^F}{q}, 1\right) = 1$. This assumption that leading firms will always set the limit price will make the dynamic decision problem of firms very tractable.

where $\ln Q_t = \int_0^1 \ln q_{it} di$ is the usual CES efficiency index. Hence, aggregate output is given by

$$Y_t = Q_t \times \left(\frac{exp\left(\int_0^1 ln\left(\mu_{it}^{-1}\right) di\right)}{\int_0^1 \mu_{it}^{-1} di} \right) \times L_{P,t} \equiv TFP_t \times L_{P,t}, \tag{6}$$

so that aggregate TFP is the product of firms' physical productivity measure Q_t and a term reflecting firms' market power. In particular, (5) and (6) show that the aggregate implications of the underlying distribution of markups are summarized by the two sufficient statistics

$$M_t = \frac{exp\left(\int_0^1 ln\left(\mu_{it}^{-1}\right) di\right)}{\int_0^1 \mu_{it}^{-1} di} = \frac{exp\left(E\left[ln\left(\mu_{it}^{-1}\right)\right]\right)}{E\left[\mu_{it}^{-1}\right]} \tag{7}$$

$$\Lambda_t = \left(\int_0^1 \mu_{it}^{-1} di\right) = E\left[\mu_{it}^{-1}\right]. \tag{8}$$

While Λ_t measures the gap between the equilibrium wage and the social marginal products of labor (see (5)), M_t determines aggregate TFP (see (6)). Hence, Λ_t and M_t are akin to the *labor wedge* and the *efficiency wedge* in the terminology of Chari et al. (2007).

Equations (7) and (8) stress that the macroeconomic implications of market power are fully summarized by the marginal distribution of markups and that different moments of this distributions have different aggregate effects. In particular, (7) implies that $M_t \leq 1$ and that $M_t = 1$ if and only if markups are equalized across markets. Hence, aggregate TFP depends on the *dispersion* of markups in the sense that a common proportional increase in markups in every market, will leave the efficiency wedge unchanged. Conversely, the labor wedge will decline by the exact same amount and hence reflects the average level of market power (see (8)).[5] Note also that the canonical case of constant markups as generated by a CES demand system with differentiated products is a special case of this result: TFP will be identical to its competitive counterpart but monopolistic power reduces factor prices. As long as factors are in fixed supply, monopolistic power will not have any effects on efficiency.

2.3 Dynamics: Entry, Innovation and Market Expansion

Both the production possibility frontier (as summarized by Q_t) and the distribution of markups depend on the underlying distribution of productivity across firms. Following Aghion and Howitt (1992) and Grossman and Helpman (1991), I model firms' efficiencies as being ordered on a quality-ladder with proportional productivity improvements of size $\lambda > 1$.[6] In a given market i, productivity increases can stem from three distinct sources: (i) a new firm can enter market i with a new technology, (ii) an existing firm, who is not currently active in market i, can expand into this market and (iii) the current producer in market i can increase his productivity to gain additional monopoly power.

I assume that these different sources of growth are fully symmetric in that they improve upon the current frontier technology: if the current productivity in market i is given by q_{it}, the new productivity is given by λq_{it}. Treating productivity increases through current and new producers symmetrically is not only standard in most Schumpeterian models of growth, but is particularly appealing in the current context, in that it stresses the different allocational consequences of entry, market expansion and innovation. While new producers and incumbents increase the frontier technology by the same amount, the implications for equilibrium markups and allocational efficiency

[5] See also Epifani and Gancia (2011), who derive a similar result in an economy with international trade.

[6] Specifically, letting r denote the rung of the ladder, qualities are ordered according to $q_{r+1} = \lambda q_r$.

are very different. In particular, the expression for equilibrium markups in (2) implies that

$$\mu_{it} = \frac{q_{it}}{q_{it}^F} \equiv \lambda^{\Delta_{it}}, \tag{9}$$

where $\Delta_{it} \geq 1$ is the producer's productivity advantage over the competing firms in market i. Hence, in case the innovation stems from the current producer of variety i, the equilibrium markup in market i *increases* by a factor λ. In contrast, when productivity growth is induced by a novel producer (which can either be an entirely new firm or an existing firm, which did not produce in market i in the past), the equilibrium markup in market i *decreases* by a factor $\lambda^{\Delta(i,t)-1}$, as the new producer is only a single step ahead on the quality ladder.[7]

Given this structure, the state of the firm is a multi-dimensional object: the number of products n, the quality of each of these products $[q_j]_{j=1}^n$ and the quality-gaps in each product line $[\Delta_j]_{j=1}^n$. However, the Cobb-Douglas demand structure in (1) implies that equilibrium profits in market i are given by

$$\pi_{it} \quad = \quad \left(1-\mu_{it}^{-1}\right)Y_t = \left(1-\lambda^{-\Delta_{it}}\right)Y_t \equiv \pi_t\left(\Delta_i\right), \tag{10}$$

i.e. they only depend on the quality gap and not on the level of quality q_{fi}. Hence, I restrict attention to equilibria where firm behavior only depends on the payoff relevant state variables $\left(n, [\Delta_j]_{j=1}^n\right)$.

Given this state, the firm can spend resources to improve its productivity in the markets it is currently producing in and it can try to break into novel markets. I adopt the usual stochastic formulation, whereby the firm can chose the flow rate to increase the productivity of existing products, $[I_{it}]_{i=1}^n$, and to expand in a novel, randomly selected market, X_t, at a cost $\Gamma\left(X, [I_i]_{i=1}^n\right)$, which - consistent with Bollard et al. (2016) - I denote in units of labor. It is analytically convenient to formulate the problem in terms of the expansion intensity per currently served market, $x_t \equiv X_t/n$. Optimal behavior is then described by the value function $V_t\left(n, [\Delta_i]_{i=1}^n\right)$, which is given by

$$\begin{aligned} r_t V_t\left(n,[\Delta_i]_{i=1}^n\right) - \dot{V}_t\left(n,[\Delta_i]_{i=1}^n\right) \quad = \quad & \sum_{i=1}^n \pi_t\left(\Delta_i\right) - \sum_{i=1}^n \tau_t \left[V_t\left(n,[\Delta_i]_{i=1}^n\right) - V_t\left(n-1,[\Delta_i]_{j\neq i}\right)\right] + \\ & \max_{x,[I_i]_{i=1}^n} \left\{ \sum_{i=1}^n I_i \left[V_t\left(n,\left\{[\Delta_i]_{j\neq i}, \Delta_i+1\right\}\right) - V_t\left(n,[\Delta_i]_{i=1}^n\right)\right] + \right. \\ & \left. nx\left[V_t\left(\left\{n+1,[\Delta_i]_{i=1}^n, 1\right\}\right) - V_t\left(n,[\Delta_i]_{i=1}^n\right)\right] - \Gamma\left(nx,[I_i]_{i=1}^n\right)w_t \right\}. \end{aligned} \tag{11}$$

The value of the firm consists of three parts. First there is the total flow payoff, which is simply the sum of profits across all markets. Second, there is the possibility of losing any of the n existing markets to other firms. This happens at the endogenous rate τ, which will be determined in equilibrium and which I refer to as the rate of creative destruction. Finally, there is the option value of innovating. First of all, the firm has the option to increase the markup in each of its markets with flow rate I_i. Secondly, the firm can spend resources to break into a new market with flow rate nx. Note that the quality gap in novel markets is always equal to unity. Finally, the last term in (11) captures the cost of innovation and market expansion.

While the recursive formulation for V_t in (11) looks somewhat daunting, it turns out that (11) admits a simple closed form solution. I assume a particular functional form for the cost function $\Gamma(.)$, which is consistent with

[7] Note that the continuous time formulation of the model precludes the possibility that a variety experiences both entry and a productivity improvement by the current producer, which is of second order.

balanced growth and allows me to derive an analytic solution

$$\Gamma\left(nx,[I_i]_{i=1}^n,n\right) = \sum_{i=1}^n c^I\left(I_i;\Delta_i\right)+c^X\left(nx;n\right) \text{ where } c^I\left(I_i;\Delta_i\right)=\lambda^{-\Delta}\frac{I^\zeta}{\varphi_I} \text{ and } c^X\left(X;n\right)=\frac{n^{1-\zeta}X^\zeta}{\varphi_x}. \quad (12)$$

Here φ_I and φ_x parametrize the efficiency of the innovation and expansion technology and $\zeta > 1$ ensures that the cost function is convex so that there is a unique solution. In addition, both cost functions contain scaling variables, which make the model consistent with balanced growth (Sutton, 1997; Luttmer, 2010).[8]

As far as new entrants are concerned, I assume that potential entrants have access to a linear technology, whereby each unit of hired labor generates a flow of φ_z marketable ideas. As firms enter in a single market with a unitary quality gap, the equilibrium degree of entry z is described by the free entry condition

$$V_t\left(1,1\right)\leq\frac{1}{\varphi_z}w_t=0 \text{ with equality if } z>0. \quad (13)$$

For the remainder of the paper, I will focus on the case with positive entry, where the condition in (13) holds with equality.[9]

2.4 The Stationary Equilibrium

Given this set-up, I will now characterize the stationary (or balanced-growth-path) equilibrium of this economy, which is defined in the usual way.

Definition. A stationary (or balanced-growth-path) equilibrium is a set of allocations $[l_{it},I_{it},x_{it},z_t,y_{it},c_t]_{it}$ and prices $[w_t,r_t,p_{it}]_{it}$ such that (i) all aggregate variables grow at a constant rate, (ii) consumers chose $[y_{it},c_t]_{it}$ to maximize utility, (iii) firms chose $[I_{it},x_{it},p_{it}]$ optimally, (iv) the free entry condition is satisfied, (v) all markets clear and (vi) the cross-sectional distributions of markups and firm size are stationary.

The first useful result to characterize the equilibrium is that - along the BGP - the value function has a tractable closed form solution. In particular, I show in Section 6.1 in the Appendix that

$$V\left(n,[\Delta_i]_{i=1}^n\right)=\frac{\pi\left(1\right)+\left(\zeta-1\right)\frac{1}{\varphi_x}x^\zeta w_t}{\rho+\tau}\times n+\sum_{i=1}^n\frac{\pi\left(\Delta_i\right)-\pi\left(1\right)+\left(\zeta-1\right)\lambda^{-\Delta_i}\frac{I^\zeta}{\varphi_I}w_t}{\rho+\tau}, \quad (14)$$

where x and I are the optimal innovation and expansion rates and $\tau=z+x$ is the endogenous rate of creative destruction, i.e. the flow rate of which a firm in a given market is replaced by either a new entrant or an existing firm, who expands into the respective market. The value of the firm has an intuitive structure. In particular it consists of two additive parts. The first term is similar to the baseline model of Klette and Kortum (2004) and captures the value of serving a market with a quality gap of unity (and hence a markup of λ) plus the inframarginal rents of the concave expansion technology. This part of a firm's value scales linearly in the number of markets n. The second term captures the possibility of exploiting market-power. It consists of the flow value of being able to

[8] The term $\lambda^{-\Delta}$ in $c^I(.)$ implies that innovations are easier the bigger the within-market productivity advantage Δ and is similar in spirit to the assumption of knowledge capital made in Klette and Kortum (2004) or the setup in Atkeson and Burstein (2010). Intuitively: per-period profits are given by $\left(1-\lambda^{-\Delta}\right)Y$ (see (10)) and hence concave in Δ. For innovation incentives to be constant, the marginal costs of innovation have to be lower for more advanced firms. The leading term in (12) $\left(\lambda^{-\Delta}\right)$ is exactly the right normalization to balance those effects. Note that firms only generate a high productivity gap when they have multiple innovation *in a row*. Hence, (12) effectively posits that firms can build on their own innovations of the past. The term $n^{1-\zeta}$ in $c^X(.)$ serves a similar purpose and implies that the cost of expanding at rate x per market (i.e. $nx=X$) is linear in n.

[9] I show in Section 6.1 in the Appendix that a sufficient condition for the free entry condition to be satisfied along the balanced growth path is $\rho>\left(\left(\zeta-1\right)/\zeta\right)\left(\varphi_x/\left(\zeta\varphi_z\right)\right)^{\frac{1}{\zeta-1}}$.

sustain higher markups augmented by the possibility of increasing markups further, which again is captured by the rents from the innovation technology. Because, firms are long-lived, the value function is given by the net-present value of these flow payoffs, where the appropriate discount rate is not only the rate of time preference ρ, but it also contains the rate of creative destruction τ, to account for the risk of losing markets to other firms in the economy.

Proposition 1. *Consider the setup described above. There exists a unique stationary equilibrium, where the innovation and expansion rates I and x and the rate of entry z are constant. The rate of creative destruction, i.e. the rate at which the producer in a given market is replaced, is given by $\tau = z + x$. The economy-wide growth rate is given by $g = ln(\lambda) \times (I + \tau)$.*

Proof. See Section 6.1 in the Appendix. □

Proposition 1 establishes that this economy permits a unique stationary equilibrium. In that equilibrium there is a constant flow of entering firms z and existing producers innovate within their own markets and expand into new markets at constant rates. As in any Schumpeterian model, the economy is characterized by creative destruction, whereby new producers replace incumbent firms. In this economy, the flow rate of creative destruction is given by $\tau = z + x$.[10] The aggregate rate of growth is simply given by the rate of growth of technology Q_t (see (6)), because the efficiency wedge M_t is constant in a stationary equilibrium (see below). Because all three sources of innovation generate productivity improvements of the same size, g is proportional to the sum of creative destruction τ and firms' markup increasing innovation efforts I.

The cross-sectional distributions of markups and firm size Firms' equilibrium entry, expansion and innovation policies also determine the equilibrium distribution of markups and the distribution of firm size. In a stationary equilibrium, both these distributions are time-invariant. Consider first the distribution of markups. Conveniently, markups only depend on the distribution of quality gaps Δ across markets (see (9)). Hence, the cross-sectional distribution of markups is fully characterized by $\{\nu(\Delta, t)\}_{\Delta=1}^{\infty}$, where $\nu(\Delta, t)$ denotes the measure of markets with quality gap Δ at time t. These measures solve the set of differential equations

$$\dot{\nu}(\Delta, t) = \begin{cases} -(\tau + I)\,\nu(\Delta, t) + I\nu(\Delta - 1, t) & \text{if } \Delta \geq 2 \\ \tau(1 - \nu(1, t)) - I\nu(1, t) & \text{if } \Delta = 1 \end{cases}, \tag{15}$$

where $\dot{\nu}(\Delta, t)$ denotes the time derivative. Intuitively, there are two ways for a market i to leave state (Δ, t): the current producer could have an innovation (in which case the quality gap would increase from Δ to $\Delta + 1$) or a new producer could enter (in which case the quality gap would decrease to unity). The only way for a market to enter the state (Δ, t) is by being in state $\Delta - 1$ and then having the current producer experience an increase in productivity (which happens at rate I). The state $\Delta = 1$ is special, because all markets where the producing firm gets replaced enter this state. In addition, all market leave the state $(1, t)$ if the current producer increases his productivity.

Equation (15) is the key equation to characterize the equilibrium distribution of markups. Three properties are noteworthy. First of all, note that the distribution is fully determined from the two endogenous variables (I, τ) and is hence jointly determined with the economy-wide growth rate g. Secondly, the distribution of firm size is not required to solve for the distribution of markups across products. This is due to the fact that all firms innovate and expand at constant rates I and x per market. Finally, (15) highlights the pro-competitive effects of creative

[10] Recall that the product space has measure one. Because the aggregate rate of entry is given by z, this is also the rate at which each product is subject to entry by a new firm. Similarly, because each existing firm innovates at rate x *per product,* the aggregate rate at which existing firms expand into the new markets is also given by x.

destruction: while productivity increases by existing producers are markup increasing, creative destruction shocks shift the distribution of markups downwards.

In addition to this cross-sectional distribution of markups, the model also generates a distribution of firms over the number of markets they serve. As the stochastic process of firms losing markets and expanding into new markets is the same as in Klette and Kortum (2004), the firm size distribution takes exactly the same form.[11] More precisely, the mass of firms who are active in n markets at time t, $\omega(n,t)$, will be constant in the stationary equilibrium. Note that in contrast to Klette and Kortum (2004), total sales and employment are no longer proportional. While firm sales are proportional to the number of active markets n, firm employment is also affected by the firm's average markup.

Proposition 2. *Consider the economy above and let I, x and τ be the equilibrium rates of innovation, expansion and creative destruction in a stationary equilibrium. Let*

$$\vartheta_I \equiv \frac{\tau}{I} \quad \text{and } \vartheta_x \equiv \frac{x}{\tau}.$$

Then the following is true:

1. *Let $\theta \equiv \frac{ln(1+\vartheta_I)}{ln(\lambda)}$. The distribution of markups $\mu = \lambda^{\Delta}$ is given by*

$$G(\mu) = 1 - \mu^{-\theta}, \tag{16}$$

and the efficiency and labor wedge M and Λ (see (7) and (8)) are given by

$$\Lambda = \frac{\theta}{1+\theta} \quad \text{and} \quad M = e^{-1/\theta}\frac{1+\theta}{\theta}. \tag{17}$$

2. *The mass of firms serving n markets is given by*

$$\omega(n) = \frac{1}{n} \times \frac{1-\vartheta_x}{\vartheta_x} \times \vartheta_x^n.$$

The number of active firm F and the share of aggregate output accounted for by firms with at most n markets, S_n, is given by

$$F = \frac{1-\vartheta_x}{\vartheta_x} ln\left(\frac{1}{1-\vartheta_x}\right) \quad \text{and} \quad S_n = 1 - (\vartheta_x)^n.$$

Proof. See Section 6.2 in the Appendix. □

Proposition 2 contains the main theoretical result of this paper: the cross-sectional distribution of markups $G(\mu)$, the extent of misallocation M and Λ, the equilibrium mass of firms F and the shape of the firm size distribution S_k are jointly determined from firms' innovation and entry incentives and take a tractable form. In particular, two endogenous statistics are crucial. The first is the *intensity of creative destruction ϑ_I*, which measures the extent of firms being replaced by new producers relative to firms' innovative effort in their existing markets. The second is the *intensity of market expansion ϑ_x*, which measures the share of aggregate creative destruction accounted for by existing firms. Not only do these two endogenous statistics fully summarize the joint distribution of markups and firm size, but these two aspects neatly separate.

[11] See Section 6.2 in the Appendix for details.

The endogenous distribution of markups takes a pareto form, whose shape parameter θ is fully determined from the endogenous intensity of creative destruction ϑ_I. If creative destruction is intense, the shape parameter is large so that both markup heterogeneity and the average markup decline. If on the other hand creative destruction is of little importance, the resulting distribution of markups has a fat tail and both the average markup and their dispersion is large. Hence, (16) is very different from Bernard et al. (2003), who generate a Pareto distribution of markups from firms' *exogenous* productivity draws.[12] The macroeconomic consequences of this endogenous markup distribution are in turn fully summarized the two sufficient statistics M and Λ, which only depend on ϑ_I and have the closed form representation given in (17).[13] In particular, it is easy to verify that both M and Λ are increasing in ϑ_I. This captures the pro-competitive effect of creative destruction: by reducing equilibrium markups, creative destruction reduces misallocation and increases TFP and equilibrium factor prices.

The equilibrium firm size distribution in contrast does not depend on firms' intensive innovation I, but is entirely determined by the equilibrium intensity of market expansion ϑ_x. In particular, if existing firms' expansion activities are an important component of the process of creative destruction, the equilibrium distribution of firm size will be such that a large share of output is produced in large firms. This directly implies that the number of active firms will be small. Formally, both F and S_k are decreasing in the expansion intensity ϑ_x.[14]

Proposition 2 illustrates that there is no a priori reason for the degree of misallocation and the properties of the firm size distribution to be related. However, Proposition 2 makes precise predictions, which environments are characterized both by a multitude of small firms *and* a high degree of misallocation - a situation, which is consistent with the stylized facts on firm level outcomes in many developing countries. For this to be the case it has to be that both ϑ_I and ϑ_x are low in poor countries. In Section 2.5 below I will argue that this is the case if it is costly for existing firms to break into new markets, i.e. if φ_x is low. In contrast, frictions for new firms to enter the economy, i.e. barriers to entry, which lower φ_z, have qualitatively different effects.

The Dynamics of Markups and Firm Size Proposition 2 focuses on the cross-sectional implications of the theory. The model, however, also makes tight predictions for the resulting life cycle dynamics of markups and firm size and these moments will be informative to take the model to the data.. Consider first the evolution of markups. The underlying mechanism which generates the endogenous pareto tail in my model is akin to the city-size dynamics of Gabaix (1999), in that markups *within a market* have an intuitive life cycle interpretation: as long as other firms do not replace the current producer, markups stochastically increase. Once a new producer breaks into the respective market, markups are "reset" to λ and the process begins afresh. In fact, as shown in Section 6.4 the Appendix, the conditional distribution of quality gaps Δ as a function of the time a market is served by a particular producer, which I will refer to as "product age" a_P, is a Poisson distribution with parameter Ia_P, i.e. is given by

$$h_{\Delta+1}(a_P) = \frac{1}{\Delta!}(Ia_P)^{\Delta} e^{-Ia_P}. \tag{18}$$

Hence, conditional on not being replaced, the distribution of markups continuously shifts outwards as incumbent firms engage in productivity improvements to rack up their monopoly power. Equation (18) also implies that the average log markup in a market conditional on being served by the same firm for a_P years is given by

$$E[ln(\mu)|\text{product age}{=}a_P] = ln(\lambda)(1 + Ia_P), \tag{19}$$

[12] I also want to point out that (16) describes the distribution of markups across markets and not across firms. While the former is the welfare-relevant statistic, the latter is measured in firm level data. I will come back to this in the empirical analyses below.

[13] Note that Δ is not a continuous variable but only takes integer values. For simplicity I treat markups as continuous. See Section 6.2 in the Appendix for the closed form expressions for the discrete case.

[14] Recall that a stationary equilibrium requires that $x < \tau$ so that $\vartheta_x < 1$.

i.e. is increasing in age at a rate proportional to I.

Markets, however, are not served by the same firm for eternity. In particular, creative destruction ignited by other producers in the economy will limit how long existing firms can survive in a given market. Because producers in a given market are replaced at rate τ, the probability of serving a market for a_P years is given by $e^{-\tau\times a_P}$. Hence, the extent to which old, high-markup markets exist in the economy depends crucially on the degree of creative destruction τ. If τ is high, it is rare to see firms serving a particular product market for a long time. The long-run distribution of markups is shaped by the interplay of these two processes, which lead to a pareto distribution.[15]

To link these observations to the dynamics of markups at the firm-level it is important to recognize that firms are active in many markets. Hence, the evolution of firm-level markups is subtle. Consider a firms of age a_f. On the one hand, old firms tend to have high markups for the reason encapsulated in (19): old firms are the only firms with the potential of having had enough time to build up markups *within* a market over time. On the other hand, old firms *also* had ample time to expand into new markets. And as markups in new markets are lower than markups for the average variety the firm sells, expanding firms will tend to have low average markups. Hence, the model implies a "product life cycle" within the firm, where firms constantly accumulate market power in their existing products and add new, low markup products to their portfolio.

To see this more clearly, suppose that firms never horizontally expand (i.e. $x = 0$) and hence never serve more than a single market. In that case, the age of the firm a_f directly corresponds to the time a market has been served by a particular producer a_P. Hence, the average log markup by *firm* age is also given by (19). Allowing firms to expand horizontally into new markets breaks this tight link between markups and firm age. In particular, I show in Section 6.4 of the Appendix that the model implies that the average log markup as a function of firm age is given by

$$E\left[ln\left(\mu_f\right)|\text{firm age} = a_f\right] = ln\left(\lambda\right)\left(1 + I\times E\left[a_P|a_f\right]\right), \tag{20}$$

where

$$E\left[a_P|a_f\right] \quad = \quad \frac{\frac{1}{\tau}\left(1-e^{-\tau a_f}\right)-\frac{1}{x}e^{-\tau a_f}\left(1-e^{-xa_f}\right)}{1-e^{-(x+\tau)a_f}}\times\left(1-\frac{e^{-xa_f}}{\gamma\left(a_f\right)}ln\left(\frac{1}{1-\gamma\left(a_f\right)}\right)\right)+a_f\times\frac{e^{-xa_f}}{\gamma\left(a_f\right)}ln\left(\frac{1}{1-\gamma\left(a_f\right)}\right),$$

and $\gamma\left(a\right) = x\frac{1-e^{-(\tau-x)\times a}}{\tau-x\times e^{-(\tau-x)\times a}}$. The expression in (20) has the same structure as (19) - except that the mapping between firm age a_f and product age a_P is more complicated and depends on both the rate of market expansion x and the extent of creative destruction. In particular, the possibility of firms breaking into new markets implies that $E\left[a_P|a_f\right] \leq a_f$. Moreover, $lim_{x\to 0}E\left[a_P|a_f\right] = a_f$, so that (19) emerges as a special case.

The life cycle of employment also has a closed form expression. In particular, (3) implies that

$$E\left[ln\left(l_f\right)|a_f\right] = \ln\left(\frac{Y}{w}\right) + E\left[\ln\left(n\right)|a_f\right] - E\left[ln\left(\mu_f\right)|a_f\right], \tag{21}$$

where

$$E\left[\ln\left(n\right)|a_f\right] = \frac{1-\gamma\left(a_f\right)}{\gamma\left(a_f\right)}\sum_{i=1}^{\infty} ln\left(i\right)\times\gamma\left(a_f\right)^i.$$

Equation (21) again shows the two forces determining the life cycle of firm-employment. For a given number of markets n, older firm will be smaller as they have higher markups. At the same time however, older firms will cater to more markets as $E\left[\ln\left(n\right)|a_f\right]$ is increasing in a_f if $x > 0$.

The two relationships in (20) and (21) will be important to take the model to the micro data as they are

[15] In a recent paper, Jones and Kim (2016) exploit a similar structure to argue that creative destruction will limit income equality by reducing the time entrepreneurs have to accumulate firm-specific human capital.

informative about the different margins of firm growth. In particular, the relative speed at which older firms increase their markups and their size identifies the relative importance of firms expanding their scope of production *horizontally,* i.e. in novel product markets, or *vertically,* i.e. through productivity improvements in their existing markets. The more important firms' markup increasing innovations I relative to their expansion rate x, the steeper the age-profile of markups and the flatter the extent of life cycle employment growth.

To see this, note first that for a given age, the average markup is increasing in I and average employment is decreasing in I. Secondly, the effects of the rate of expansion x is more subtle. A decrease in the rate of market expansion x *raises* the extent of life cycle growth for markups but *lowers* the growth rate of employment. The effect on markups stems precisely from the composition effect mentioned above: if firms enter novel markets only very infrequently, a small fraction of their sales is accounted for by new and hence low-markup markets. As far as employment is concerned, two effects are at play: not only are firms only active in few markets, but in addition the average markups is also higher. Both of these effects reduce the extent to which firms increase their employment as they age. In the limit, as $x = 0$, firms only serve a single market irrespective of their age. Hence, $E\left[\ln(n)|a_f\right] = 0$ so that (21) implies that employment and markups are inversely proportional. As markups increase in age, firms would - conditional on survival - decrease in size while they *increase* their markups and profitability. Similarly, more creative destruction τ will lead to slower life-growth for both markups and firm size: if firms lose their existing markets at a fast rate, they stay small and will not have time to build-up market-specific monopoly power.[16]

2.5 The Effects of Entry Costs and Market Barriers

So far the theory provides a tractable dynamic model, where misallocation, growth, the firm size distribution and the life cycle dynamics of markups and firm size are jointly determined in equilibrium. In this paper I want to use this framework to study the consequences of two particular frictions, which are arguably important in developing economies. These are *market barriers,* i.e. frictions for existing firms to expand into new product markets, and *entry costs,* i.e. costs for new firms to enter the economy.[17]

For simplicity I model both of these frictions in the following reduced-form way. Consider first the case of market barriers and let $\phi_x \leq 1$ be the probability that an expanding firm replaces the existing producer conditional on having generated a superior technology. A low level of ϕ_x can for example reflect license requirements, bureaucratic red tape, which have to be overcome before a new firm can be active in a new market, or other policies aimed to shield existing producers from potential competitors. I adopt a similar strategy for the case of entry costs, i.e. new firms replace an existing firm with probability $\phi_z \leq 1$: the lower ϕ_z, the more costly it is for new firms to actually enter the economy.

This structure is convenient because it is nested in the theory laid out above. In particular, defining the realized expansion and entry flows rates above as $x = \phi_x \tilde{x}$ and $z = \phi_z \tilde{z}$, where $\tilde{x}$ and $\tilde{z}$ are the gross expansion and entry rates, the model above incorporates these frictions once we *define* the expansion and entry cost shifter φ_x and φ_z as

$$\varphi_x \equiv \tilde{\varphi}_x \times \phi_x^{\zeta} \text{ and } \varphi_z \equiv \tilde{\varphi}_z \times \phi_z, \tag{22}$$

[16] I was not able to show these comparative statics analytically. It is possible to analytically show that $E\left[\ln(n)|a_f\right]$ is increasing in x and decreasing in τ. For $E\left[ln(\mu_f)|a_f\right]$, I could not find any example where the average markup is increasing in τor x. This directly implies that $E\left[\ln(l)|a_f\right]$ is increasing in x. I also could not find any example where average employment was increasing in τ, i.e. where the effect of lower markups were to dominate the effect of losing markets at a faster rate.

[17] One widely used measure of entry costs is developed in Djankov et al. (2002). They measure the fees and time costs to legally operate a business for a variety of countries. Such variation in the regulation of entry has been linked to cross-country income differences in Barseghyan (2008), Barseghyan and DiCecio (2009) or Herrendorf and Teixeira (2011). There are also studies focusing on particular epsiodes of delicensing. The dismantling of India's Licence Raj, for example, has been studied in Aghion et al. (2008). Even though all these studies refer to "entry costs", the empirical variation is likely to capture both market and entry barriers in the sense of my theory.

where $\tilde{\varphi}_x$ and $\tilde{\varphi}_z$ denote the technological efficiency of the entry and expansion technology. Hence, expansion barriers and entry costs are akin to a reduction in expansion and entry efficiency. I can then use the results from above to derive the aggregate effects of changes in entry costs and market barriers.

Proposition 3. *Consider a stationary equilibrium in the economy above and suppose that $\zeta \geq \overline{\zeta} > 1$ and $\rho < \overline{\rho}$. Then,*

1. *Higher entry costs and market barriers reduce creative destruction, i.e. $\frac{\partial \tau}{\partial \phi_z} > 0$ and $\frac{\partial \tau}{\partial \phi_x} > 0$,*
2. *Higher entry costs and market barriers increase misallocation by reducing the intensity of creative destruction, i.e. $\frac{\partial \vartheta_I}{\partial \phi_z} > 0$ and $\frac{\partial \vartheta_I}{\partial \phi_z} > 0$,*
3. *Higher entry costs increase average firm size and reduce the number of firms by increasing the expansion intensity, i.e. $\frac{\partial \vartheta_x}{\partial \phi_z} < 0$,*
4. *Higher market barriers reduce average firm size and increase the number of firms by reducing the expansion intensity, i.e. $\frac{\partial \vartheta_x}{\partial \phi_x} > 0$,*
5. *The effect of entry costs and market barriers on the equilibrium growth rate is ambiguous.*

Proof. See Section 6.3 in the Appendix. The restrictions that $\zeta \geq \overline{\zeta}$ and $\rho < \overline{\rho}$ are a sufficient conditions. It can be shown that $\overline{\zeta} < 2$. ☐

Proposition 3 summarizes the consequences of market barriers and entry costs on the stationary equilibrium in this economy. To understand how firms' equilibrium incentives are affected by changes in such barriers, it is helpful to think about their dynamic optimality conditions. Consider first the optimal rate of market expansion x. Recall that the value function in (14) is additively separable across markets. Hence, the marginal value of expansion for an existing firm is exactly the same as the value of entry, which - given the linear entry technology - is equal to the entry costs. This directly implies that the equilibrium rate of expansion x equalizes the marginal cost of expansion and the marginal cost of entry and is therefore only a function of parameters

$$x = \left(\frac{\tilde{\varphi}_x \phi_x^{\zeta}}{\tilde{\varphi}_z \phi_z} \frac{1}{\zeta} \right)^{\frac{1}{\zeta - 1}}. \tag{23}$$

Naturally, x is decreasing in markets barriers and increasing in the costs of entry.

The incentives for firms to vertically innovate to increase their markups are more subtle. The value function implies that the marginal returns to increase markups in market i are given by

$$V_t\left(n, \left\{[\Delta_j]_{j \neq i}, \Delta_i + 1\right\}\right) - V_t\left(n, [\Delta_i]_{i=1}^n\right) \quad = \quad \frac{\pi_t(\Delta_i + 1) - \pi_t(\Delta_i)}{\rho + \tau} + \frac{(\zeta - 1)\left[c_I(I, \Delta_i + 1) - c_I(I, \Delta_i)\right] w_t}{\rho + \tau}.$$

The first term is the benefit of being able to post higher markups, which will result in higher profits. The second term reflects changes in the innovation technology: according to (12), increases in quality will increase the efficiency of future innovation, which represents a capital gain. Simplifying terms and setting this equal to the marginal cost of innovation implies that the optimality innovation rate I is determined from

$$\frac{\frac{\lambda - 1}{\lambda}}{\rho + \tau} \frac{Y_t}{w_t} = \frac{1}{\varphi_I} \left[\zeta I^{\zeta - 1} + \frac{\frac{\lambda - 1}{\lambda}(\zeta - 1)}{\rho + \tau} I^{\zeta} \right]. \tag{24}$$

This shows that firm's incentives to increase markups depend on two endogenous aggregate variables - the rate of creative destruction τ and size of the market $\frac{Y_t}{w_t}$ (relative to the cost of innovation). In particular, an increase in the rate of creative destruction τ will *reduce* the firm's incentives to increase its markup. The reason is that the expected time-horizon of monopolistic rents becomes shorter, i.e. the endogenous discount factor $\rho + \tau$ increases. Conversely, innovation incentives are high if aggregate demand Y_t is large relative to the cost of innovation w_t.

Equations (23) and (24) are useful to understand the economics behind Proposition 3. Consider first the case of entry costs. Higher entry costs reduce the mass of entering firms and the rate of equilibrium creative destruction τ, despite the increase in market expansion x. Holding I constant, this will increase firms' monopoly power and thereby the extent of misallocation by reducing the extent of churning in the market place. The corresponding change in I is ambiguous. On the hand, the decline in creative destruction raises firms' incentives to increase their markups. On the other hand, equilibrium requires that $\frac{Y}{w} = L_P \Lambda^{-1}$ (see (5)), where L_P is the amount of production workers. While higher entry costs will induce misallocation and hence lower Λ, they might also reduce the share of production workers and thereby aggregate demand. Regardless of the change in I, it is nevertheless the case that entry costs will unambiguously lower the intensity of creative destruction $\vartheta_I = \frac{\tau}{I}$ and hence increase misallocation. At the same time, higher entry costs will *reduce* the number of active firms in the economy by *increasing* their size. The lower rate of creative destruction will shift the firm size distribution to the right as existing firms have an easier time to accumulate markets. In addition, free entry requires that the value of entering a market has to increase, which increases firms' incentives to expand into new markets (see (23)). Hence, both the output share of small firms and the equilibrium mass of producers *decline* as entry costs increase.

In contrast, barriers to market expansion for existing firms have qualitatively different effects. Similar to entry costs, market barriers for existing firms also reduce both the level and the intensity of creative destruction τ and ϑ_I and thereby increase misallocation through high and dispersed markups. At the same time, frictions for existing firms to enter in new markets reduce the intensity of market expansion ϑ_x and thereby imply that average firm size decreases and the number of active firms increases.

Finally, Proposition 3 stresses that the relationship between entry costs and market barriers and the endogenous growth rate is subtle. The reason is precisely the possibility for firms to engage in markup increasing innovation. If creative destruction and vertical innovation are strong substitutes, it is possible that $I + \tau$, which - recall - determines the equilibrium growth rate, *increases* in response to higher entry of market barriers. In fact, this substitutability is not only a theoretical possibility but turns out to be quantitatively important in the calibrated economy.[18]

In summary, Proposition 3 suggests that in order to explain the joint behavior of misallocation, firm size and the number of firms across countries, variation in market barriers are theoretically attractive in that they simultaneously can account for all these features. In particular, they imply a negative correlation between misallocation and average firm size, making them a plausible candidate to explain the empirically observed differences in the market environment between rich and poor countries. Entry barriers in contrast face somewhat of an uphill battle, in that their first-order effect on important moments like average firm size or the number of producers is counterfactual: high costs of entry will increase misallocation but also make firms larger.

3 Empirical Analysis

I will now take this theory to plant-level data from the Indonesian manufacturing sector. The analysis has three parts. First, I provide direct evidence that the empirical patterns of firm-level markups are qualitatively consistent with theory. Secondly, I calibrate the model to moments of the Indonesian microdata and use information on entry

[18] While this result sounds similar to the finding in Aghion et al. (2001), who argue that product market competition increases growth through higher innovation incentives for incumbent firms (which they refer to as the effect of "escape competition"), the mechanism is different. In my model, the ambiguous effect on the aggregate growth is purely a composition effect, whereby an increase in market barriers (entry costs) reduces firm expansion x (entry z) but increases firms' incentive to raise markups I.

rates and life cycle employment growth for the US to measure differences in entry costs and expansion barriers and quantify the aggregate importance of such differences. Finally, I use regional variation across provinces in Indonesia to provide further evidence for the importance of market barriers.

3.1 The Data

The main data set for the empirical analysis is the Manufacturing Survey of Large and Medium-Sized Firms in Indonesia (Statistik Industri). This data has also been used in Amiti and Konings (2007), Blalock et al. (2008), Yang (2012) and Hsieh and Olken (2014). The Statistik Industri is an annual census of all formal manufacturing firms[19] in Indonesia and contains information on firms' revenue, employment, capital stock, intermediate inputs and other firm characteristics. I will focus on the time period between 1990 and 1998, i.e. the years prior to the Indonesian financial crisis. My final sample has about 180.000 observations.

The Statistik Industri data aims to focus on large, formal producers and therefore has a size threshold of 20 employees. In the context of a developing economy like Indonesia, this is a heavily selected sample of firms. Hsieh and Olken (2014) for example analyze data from the Indonesian economic census, which covers all producers, and find that the share of firms with less than 10 workers is essentially indistinguishable from 100 percent. At the same time, the (few) firms in the Statistik Industri data are sufficiently large to account for roughly 40% of total employment. Table 1 contains some descriptive statistics and shows that the average plant has about 140 employees. It is also the case that the firm size distribution is skewed - while the median plant has only 45 employees, the 90% quantile of the distribution is 350. Compared to the US manufacturing sector, plants in Indonesia are of course still small. In the US, one third of all establishments have more than 20 employees and such plants account for more than 90% of total employment. Moreover, the top 3.5% of plants have more than 250 employees and account for almost half of manufacturing employment.[20]

For the purposes of this paper, this focus on large producers has advantages and disadvantages. On the positive side, this data covers firms for which considerations of productivity improvements, strategic pricing and market expansion are probably more relevant. The majority of micro-firms in Indonesia are arguably subsistence entrepreneurs, which are unlikely to engage in such activities and which do not compete in the same product markets as large, formal employers.[21] Additionally, the data has a panel dimension, which allows me to use the information contained in the dynamics of markups and firm size over the life cycle to calibrate the structural parameters. Hence, I will not have to rely on the cross-sectional age-size or age-markup relationship to test the implications of the theory. To the best of my knowledge, there is no dataset covering the universe of Indonesian firms, which has a panel dimension.

The main drawback of this selection criterion is that it complicates the measurement of entry and exit as I only see firms appearing and disappearing from the data. Table 1 shows that there are on average 10.4% firms entering and 8.2% of firms exiting the data. Naturally, these firms are much smaller than the average firm so that the population of entrants (exiting firms) accounts for 5% (4.5%) of aggregate employment in the data. Interestingly, they account for an even smaller fraction of sales in the economy, reflecting the fact that they have smaller markups (as predicted by theory).

To map these moments to the theory, note that the relevant notion of firm age is the time a particular firm has been active in (potentially many) markets $i \in [0, 1]$. So if we think of the relevant set of product markets as

[19] To be absolutely precise, the data is collected at the plant level. As the majority of plants reports to be single branch entities, I will for the following refer to each plant as a firm.

[20] See Table OA-2 in Section OA-2.1 in the Online Appendix, for all the details.

[21] There is mounting evidence for the importance of "stagnant" entrepreneurs in developing economics - see e.g. Schoar (2010), Hurst and Pugsley (2012), Akcigit et al. (2015) or Hsieh and Olken (2014). That these firms do not compete in the same product markets as their larger, formal counterparts is e.g. argued in La Porta and Shleifer (2009) or La Porta and Shleifer (2014).

the markets formal firms compete in, a new firm in the Census is indeed an entrant in the sense of the theory. I therefore consider two strategies to calibrate the model. For my benchmark calibration I treat new plants in the Census as entrants. This allows me to measure the entry-rate directly from the data. In an alternative strategy, I treat the measure of entrants as unobserved and model the empirical selection criterion by size (i.e. the size cutoff) explicitly. Specifically, I explicitly calibrate the model to match the flow of firms passing the size threshold and therefore appearing in the data.

3.2 The Markups of Indonesian Manufacturing Firms

One theoretical advance of this paper is to construct a model of heterogeneous markups. Hence, in this section I will provide direct evidence that important qualitative properties of the markups of Indonesian manufacturing firms are consistent with the theory.

3.2.1 Measuring Markups

To measure markups, I follow the approach pioneered by Jan De Loecker in various contributions (De Loecker et al., 2016; De Loecker and Warzynski, 2012; De Loecker, 2011b) and hence relegate most of the details to the Appendix. The main benefit of this approach is that it allows one to measure firm-specific markups without having to take a stand on many aspects of the theory. In particular, one does not need assume a particular structure of output markets.

Both to test the theory and to calibrate the model, I only require information on the variation of markups - either in the cross-section of firms or along a firms' life cycle. This implies that I can employ a measure of markups, which does not require an estimate of firms' production functions (or more precisely the output elasticities). To see why, recall that the model implies that firm $f's$ markup μ_f is simply inversely proportional to the labor share (see (3)). That one can infer markups from factor shares does not require the stringent assumptions on functional forms or the structure of output markets of my theory. More specifically, consider a firm f, which is a price-taker in input markets. The optimality conditions from the firms' cost-minimization problem then imply that the markup μ_f satisfies the equation

$$\mu_f = \theta_{l,f} \times s_{l,f}^{-1}, \tag{25}$$

where $\theta_{l,f} = \frac{\partial ln(y_f)}{\partial ln(l)}$ is the output elasticity of labor and $s_{l,f} = \frac{wl}{py}$ is the firm's labor share in value added (or more generally, any expenditure share of a flexible input). In my model, the output elasticity of labor is unity, so that (25) implies (3). Note that the derivation of (25) did not use any information on the structure of demand, on how firms compete or how innovation and/or entry occurs.

If θ_l was known, one could directly infer firms' markups from their observed labor shares. If θ_l is not know, but assumed to be constant across firms (i.e. stemming from a Cobb-Douglas production function), (25) still identifies firms' markups up to a constant of proportionality. This is sufficient to study both the time-series and cross-sectional properties of markups (see also De Loecker and Warzynski (2012) for a similar approach). In this spirit, my baseline measure of firms' markups μ_f is the residual from the regression

$$ln\left(s_{l,ft}^{-1}\right) = \delta_s + \delta_t + u_{ft}, \tag{26}$$

i.e. $\hat{\mu}_{ft} = \hat{u}_{ft}$. Here, δ_s is a set of 5-digit fixed effects and δ_t is a set of year fixed effects. Under the assumption that $\theta_{l,f}$ does not vary within industries (and years), the variation of $\hat{\mu}_{ft}$ is exactly the same as if I had estimated

θ_l at the 5-digit level in a first stage and then calculated μ_f according to (25) using the estimate $\hat{\theta}_l$.[22] However, in the empirical analysis below, I also do consider richer specifications, where I explicitly control for firms' input choices like the capital- or material intensity, to allow for variation in output elasticities across firms within 5-digit industries. Furthermore, I also report results where I measure markups from firms' material shares instead of labor shares.

3.2.2 The cross-section and the life cycle of markups

Given the measure of markups contained in (26) I can now confront the empirical patterns of markups with the theory. The main predictions of the theory are that markups should differ across producers and that these should be correlated with both firm age and size. In Figure 1, I show to what extent this is the case in the Indonesian data. In the left panel I depict the evolution of log markups over firms' life cycle. I want stress that this pattern is estimated from the time-series variation and *not* from the cross-sectional age-size relationship. Hence, as in the theory there is a attrition as firms exit the market and the size of the dots reflects the size of the surviving firms in the cohort. More specifically, I focus on all firms that entered the data after 1990 (which allows me to measure plant-age in a consistent way) and then calculate the average markup by age relative to entering firms. As predicted by theory, markups increase in age. In particular, they show a somewhat concave profile and seem to level off at the end (even though markups for old firms are not very precisely estimated). Quantitatively, markups of 7 year old firms are on average 8% larger compared to recent entrants. In terms of the theory, this schedule refers exactly to the expression characterized in (20) and I will use it as an explicit moment to calibrate the model.

In the right panel of Figure 1, I display the correlation of markups and firm size (as measured by employment). This figure uses the entire cross-sectional data, i.e. does not condition on plants having entered the data after 1990. There is a sizable positive correlation, implying that larger firms post higher markups. Quantitatively, plants in the ninth decile of the employment distribution have about 25% higher markups as firms in the second decile. In terms of the theory, this positive correlation is driven by the underlying relationship between size and age: larger firms are older and hence had more time to increase markups within their markets *and* to increase in size by expanding into new markets. Through the lens of the theory, this positive correlation between firm size and market power implies that firms grow both horizontally and vertically. Without the possibility of expansion, firms had only a single product and the employment-age relationship would be *negative* - see (21). Similarly, if firms were to only grow horizontally by adding new markets, markups and age should not be systematically related. Hence, the joint behavior of markups and firm size by age is informative about firms' "mix" of horizontal and vertical growth and will therefore be an important calibration target for the model.

Given the importance of these two moments, let me provide some robustness of the patterns in Figure 1. Consider first the life cycle growth of markups. I will focus on regressions of the form

$$ln\left(\mu_{ft}\right) = \delta_t + \delta_s + \beta \times age_{ft} + \alpha \times ln\left(k_{ft}/l_{ft}\right) + h'_{ft}\gamma + u_{ft}, \tag{27}$$

where k/l denotes the firms' capital-labor ratio, h contains additional firm-characteristics and δ_t and δ_s denote year and 5-digit industry fixed effects. As in Figure 1 I focus on the unbalanced panel of firms entering the economy after 1990. Hence, the coefficient on age is identified from the rate of life cycle growth and not from the cross-sectional relationship between markups and age. The results are contained in Table 2.

Columns 1 and 2 contain the specifications with and without five digit sector-fixed effects - this makes little difference for the estimated age-profile. In column 3, I include firms' capital-labor ratio to control for a correlation

[22] Furthermore, my data is standard in the sense that it does not contain information on firm-specific prices. Hence, to estimate the output elasticity θ (which corresponds to *physical* output) one needs to impose additional structure.

between capital-intensity and firm size (and hence age). Doing so reduces the estimated age-coefficient slightly. In column 4, I show that both entrants and exiting firms have lower markups. This is consistent with the model: both entering and exiting firms are small and therefore - on average - young and hence they should have lower markups on average.[23] In column 5, I therefore include firm age directly, which should account for the low markup of entrants and it does. The results for exiting firms, however, are not consistent with theory. In the model, exit is *only* a function of the number of markets n and not the age of the firm. However, the model implies that markups are decreasing in size holding age fixed. Intuitively: by having been unsuccessful in expanding into new markets, small old firms did not add low markup market into their portfolio. Hence, conditional on age, exiting firms should actually have higher markups. While the negative coefficient does increase once age is controlled for, exiting firms still have lower markups given their age. This is consistent with for example Foster et al. (2001, p. 309), who find that "low-productivity plants are more likely to exit even after controlling for other factors such as establishment size and age." To the extent that markups are increasing in physical productivity, this finding is consistent with the result in column 5.

In column 6 I explicitly control for age and firm revenue (as measured by value added), which in the theory is proportional to the number of markets n. Empirically, both older and larger firms have higher markups, even though the theory implies that markups should be declining in n conditional on age. In column 6 I control directly for selection by conditioning on survival. In the theory, there is no selection in that the distribution of markups conditional on age is the same for all firms. In the data, the growth of markups shown in Figure 1 could stem from a higher exit hazard of firms, that systematically have low markups. Column 6 shows that markups are increasing over time even for those firms that do survive until the end of the sample. Finally, column 7 contains a specification with firm fixed effects, which again yields a quantitatively similar estimate for the life cycle growth of markups.[24]

The theory abstracts from any distortions to firms' input choices. This allows me to the use data on firms' factor shares to measure markups. As stressed by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), this interpretation might be misleading if firms are subject to frictions, which distort their input choices - see equation (4). That revenue productivity is increasing in firm size forms the basis of a large literature on size-dependent policies (see e.g. Guner et al. (2008), Fattal Jaef (2011) or Bento and Restuccia (forthcoming)). While firm-specific market power is a natural mechanism to generate this pattern, there could be other possibilities for why large firms appear to be constrained - see e.g. the discussion in Hsieh and Olken (2014). Note however, that benchmark dynamic models of financial constraints imply exactly the opposite pattern as reported above: borrowing constraints tend to bind in the early stages of the life cycle and get relaxed as the firms ages (see e.g. Clementi and Hopenhayn (2006)). In Section OA-2.3 in the Appendix, I present additional evidence on the comparison between factor shares reflecting markups or distorted marginal products. In particular, I show that exporters or firms relying on FDI have in fact higher revenue productivity and that firms, who report to be capital-constrained, have lower revenue productivity. These pattern are consistent with factor-shares reflecting market power, but harder to rationalize with standard models of credit-constraints.

[23]See also Foster et al. (2008), who - using US data - report qualitatively similar findings for particular industries selling homogenous products.

[24]In Section OA-2.3 in the Appendix I present additional robustness checks for these results. In particular, I show that the results do not substantially depend on wether or not I correct the measure of markups for measurement error as suggested in De Loecker and Warzynski (2012). I also consider the case of material shares in total sales. Markups should distort all factors within the firm equally, so that (25) should also hold for $s_{M,ft} = \frac{m_f}{py_f}$, i.e. the share of sales going to materials. The results are very similar to the ones reported in Table 2 when firm's material-labor ratio $ln\left(\frac{m}{l}\right)$ is controlled for. This is important, as there is a strong positive correlation between $ln\left(\frac{m}{l}\right)$ and age, i.e. older firms seem to rely relatively more on materials. If this change in production factors is not controlled for, the correlation between material shares and age is negative. I also redo the analysis shown in Table 2 when I control for $ln\left(\frac{m}{l}\right)$. This does not change the results substantially.

3.3 The Aggregate Effects of Market Barriers and Entry Costs

I will now use a calibrated version of the model to quantify the aggregate effects of market barriers ϕ_x and entry costs ϕ_z on (i) the extent of misallocation, (ii) the distribution of firm size and (iii) the long-run growth rate. I adopt the following strategy. First I calibrate the structural parameters to moments of the Indonesian micro-data and use the calibrated economy to revisit the results of Proposition 3, in particular the effect of entry costs and market barriers on the economy-wide growth rate, which was ambiguous in the theory. Then, I will use information on entry rates and life cycle employment growth from the US manufacturing sector to identify the underlying differences in market barriers and entry costs across countries and quantify the implied changes in the stationary firm size distribution, the extent of misallocation and the long-run growth rate.

3.3.1 Calibration

The model is very parsimonious. Given a rate of time preference ρ, which I will set exogenously, the theory is fully parametrized by five parameters: the innovation step-size λ, the cost shifters for innovation, market-expansion and entry φ_I, φ_z and φ_x and the curvature of the innovation and expansion technology ζ. Recall that φ_x and φ_z contain the institutional aspects of entry costs ϕ_z and market barriers ϕ_x (see (22)). My calibration strategy is as follows. As shown in Proposition 2 and in the expression for the life cycle of firm size and markups, all micro-moments related to the firm size distribution or the process of firm-dynamics *only* depend on the three endogenous outcomes (I, x, τ) and one exogenous parameter λ. Hence, the model provides a direct mapping from the data to (I, x, τ) and λ. I then use the equilibrium conditions to find the required structural parameters to yield (I, x, τ) as equilibrium outcomes consistent with optimal behavior and market clearing. In particular, for a given cost elasticity ζ, the uniqueness of the equilibrium implies that there is a unique mapping from the calibrated allocations (I, x, τ) to $(\varphi_I, \varphi_z, \varphi_x)$. Credibly identifying the curvature parameter ζ is difficult without exogenous variation in innovation costs. Hence, I will take this parameter from the literature and provide robustness.

I will use the following four moments to identify the structural parameters. First of all I will use the data contained in Figure 1, i.e. the life cycle of markups. As seen from the closed-form expression in (20), this moment will identify I for a given expansion rate x, rate of creative destruction τ and step-size λ. Given τ and I, I will discipline λ to match a given aggregate growth rate of productivity g (see Proposition (2)). To calibrate x, I force the model to be consistent with the observed employment life cycle growth in the data, which in the model is given by expression (21). This leaves one moment for the equilibrium flow rate of entry z (or alternatively the rate of creative destruction $\tau = x + z$). The most natural moment is of course the equilibrium rate of entry, which is given by (see Proposition 2)

$$Entry\ Rate = \frac{z}{F} = \frac{z}{\frac{1-\vartheta_x}{\vartheta_x} \times ln\left(\frac{1}{1-\vartheta_x}\right)} = \frac{\tau - z}{ln\left(\frac{\tau}{z}\right)}. \tag{28}$$

For a given rate of creative destruction, the entry rate is increasing in z. Given the minimum size requirement in the data, this strategy is consistent with the model if one assumes that the relevant aggregate product market (i.e. the continuum of markets $i \in [0,1]$) consists of markets catered to by the population of manufacturing firms with at least 20 employees.

To see how important that restriction is, I will also consider an alternative calibration strategy, where I explicitly take the truncation of the data into account. In particular, suppose that firms are only sampled in the census data once they cater to at least n_0 markets. The mass of firms crossing the threshold at each point in time is therefore given by $E^C = x \times (n_0 - 1) \times \omega_{n_0 - 1}$, where ω_{n_0} is the firm size distribution characterized in Proposition 2 and $x \times (n - 1_0)$ is the expansion rate of firm with $n_0 - 1$ products. Moreover, the model also implies a closed form expression for the share of sales these firms account for, i.e. for the moment reported in Table 1. This moment can be used to identify the flow rate of entry z, because the share of sales of incoming firms is increasing in z holding the rate of creative destruction constant. Intuitively, the higher the share of creative destruction accounted for by entering firms $\frac{z}{\tau}$, the thinner the tail of the firm size distribution. Hence, the share of sales accounted for by firms of size n_0 relative to larger firms will increase. While I focus on a calibration strategy based on the entry rate in (28) in the main body of the paper, this alternative strategy yields very similar results and I report the results in Section 6.6 of the Appendix.

Table 3 summarizes the calibration results. In terms of data moments, I first require the model to match the fact that markups increase by 0.08 log points over a 7 year horizon (displayed in Figure (1)). I also calibrate the model to the observed life cycle profile of employment growth. Empirically, the relationship between age and log employment is almost linear, reflecting an almost constant growth rate of employment within a given cohort conditional on survival. Quantitatively, firms in Indonesia increase their employment by roughly 0.5 log points (i.e. a factor of 1.6) in the first 7 years of their life (see Figure 2 below, where I depict the employment life cycle from both the data and the model and Section 6.6 in the Appendix). Hence, in contrast to firms in India, which - according to Hsieh and Klenow (2014) - experience essentially no growth as they age, Indonesian firms do grow over time conditional on survival. However, it is important to note that both the sample of firms and the methodology underlying my calibration differs from the exercise in Hsieh and Klenow (2014). First of all, the Indonesian data is biased towards bigger, formal firms. Secondly, my moments are estimated from panel data and not inferred from the cross-sectional age-size relationship. This turns out to be important as the cross-sectional age-size relationship in Indonesia is also relatively flat, which could be due to measurement error in firm age.[25] To identify the flow rate of entry z, I require the model to be consistent with the observed entry rate (i.e. the rate of new firms appearing in the Census) of 10.4%. Finally, I discipline λ to match a given aggregate rate of productivity growth of 3%.[26]

The results of this exercise are contained in Table 3. I report the data and model moments, the structural parameters and the implied endogenous innovation outcomes. The model can be calibrated to the calibration targets exactly. To get a better sense of the link between the model and data, consider Figure 2, where I depict the calibrated and observed life cycle patterns for markups (left panel) and employment (right panel). While the model is only calibrated to match the data for 7 year old firms, it captures the general age pattern for both markups and employment reasonably well. As shown in the theory, the model implies that markups are monotone in age, conditional on survival. Hence, the model is unable to generate the decrease in markups for old firms. Even in the data, however, this decrease is imprecisely estimated. In terms of the employment life cycle, the model replicates the data well.

The model also has implications for various non-targeted moments. Given the calibrated endogenous outcomes τ and x, the model implies that entrants account for $\frac{z}{\tau} \approx 13.5\%$ of the aggregate creative destruction. This is on the low end of other findings in the literature. Foster et al. (2001, p. 309) for example find that about "15% of job creation is accounted for by entry and exit". The model also implies that the share of growth accounted for by firms increasing productivity in their existing market relative to breaking into new market, $\frac{I}{I+x}$, is given by about 73%. This is qualitatively consistent with Garcia-Macia et al. (2016), who estimate that about 75%-80% of incumbent growth is due to own-quality improvements.[27] Note that these two facts imply that the share of aggregate growth

[25]See Section OA-2.3 in the Appendix, where I replicate the cross-sectional age-size relationship using the methodology by Hsieh and Klenow (2014). This difference between the "true" life cycle and the cross-sectional patters do not seem to be unique to the Indonesian context, but are also present in for example Chile (see Buera and Jaef (2016))

[26]A detailed description of the construction of all these data moment is contained in Section 6.6 of the Appendix. There I also present additional regression evidence for the life cycle of employment.

[27]Their inference however, does not stem from the age-profile of markups but rather from the distribution of job creation and destruction and the exit-size relationship of incumbent firms.

accounted for by entering firms, $\frac{z}{I+\tau}$, is very low - it is only 4%. There are two main reasons why this is the case. First of all, I abstracted from the entry of new varieties. If new varieties are more likely to be produced by entering firms, the entry share in aggregate growth will increase. Secondly, I restrict the step size of innovation, λ, to be the same across all sources. If entrants were to enter with technologies, which represented a drastic innovation, a given entry rate would be consistent with a larger share of growth. While such margins could be added to the theory, they would complicate the characterization of the stationary distribution of markups.

The model also implies that - given the pareto distribution of markups characterized in Proposition 2 - average markups across products are given by $\frac{\theta}{\theta-1}$, where θ is the endogenous pareto-tail characterized in Proposition 2. In the calibrated model, I find that average markups are 12%. These numbers are broadly consistent with Akcigit and Kerr (2015), who report a profitability of 10.9%. They are also in the range of the estimates of De Loecker and Warzynski (2012), albeit at the lower end. Depending on the specification used, they estimate markups between 10% and 28%. In my model, average-markups are directly implied from the process of firm-dynamics and a given aggregate growth rate. Intuitively: for average markups to be higher (given the estimated life cycle slope) the quality ladder step size λ would need to be higher. This however, would also imply a higher aggreate growth rate.

As a final check of the theory, we can also look at the extent of exit. The theory implies that the share of firms surviving until age a_f is given by[28]

$$S\left(a_f\right) = \frac{(\tau - x) \times e^{-(\tau-x)\times a_f}}{\tau - x \times e^{-(\tau-x)\times a_f}}. \tag{29}$$

Through the lens of the model, which is stationary, we can measure $S\left(a_f\right)$ in two ways: either we can look at a cross-section at time t and "count" the firms of different ages. Or we can follow a cohort through its life cycle and keep track of the surviving firms. In Figure 3, I depict the model's implication for $S\left(a_f\right)$ and the data - both measured from the cross-section in 2000 and from time-series of the cohort, which entered in 1991. Two properties stand out. First of all, the survival probabilities estimated from the cross-section and from the panel turn out to be extremely similar. Secondly, while the model is qualitatively consistent with the data, it slightly over-predicts the extent of "shake-out", i.e. firms exit at too fast a rate.[29]

3.3.2 Quantifying the Effects of Market Barriers and Entry Costs

Proposition 3 characterized the qualitative effects of market barriers and entry costs. In this section I use the calibrated model and data from the US to gauge wether these effects are quantitatively important. To build intuition for the quantitative analysis, consider Figure 4, where I depict the relationship between different equilibrium outcomes and market barriers (left panel) and entry costs (right panel). In particular I consider two calibration moments (the entry rate and the extent of life cycle employment growth), the two summary statistics for the endogenous distribution of markups and firm size derived in Proposition 2 (the number of active firms F, which only depends on ϑ_x, and the endogenous tail parameter of the markup distribution, which only depends on ϑ_I) and the aggregate growth rate. I normalize all variables to unity in the initial equilibrium.

Consider first the left panel, i.e. the effects of increasing market barriers ϕ_x. As shown in Proposition 3, if firms are prohibited from growing, the firm size distribution shifts to the left and the number of firms increases. At the

[28]Letting $p_n(a)$ be the probability that firm of age a is active in n markets, we have $S(a) = 1 - p_0(a)$. It can be shown that $p_0(a) = \frac{\tau}{x}\gamma_a$, where γ_a is given defined in (20). See Section OA-1.2.3 in the Online Appendix or the Appendix of Klette and Kortum (2004).

[29]The reason is the following: in the model, firms *only* margin of employment growth is to enter in new markets and to replace other producers - productivity growth in existing markets actually reduce employment through increasing markups. This sharp distinction is conceptually and analytically useful. It is, however, restrictive. For example, if the elasticity of demand exceeded the Cobb-Douglas case of unity, increases in quality could also led to increases in employment as firms would pass-through some of their cost-reduction into prices consumers face. In that case, the model could rationalize a given slope of the age-employment schedule with margins other than creative destruction. See also Garcia-Macia et al. (2016) and Luttmer (2010).

same time, misallocation gets severe in that the tail parameter of the stationary markup distribution θ declines. The economy-wide growth rate in contrast is hardly affected, because the decline in creative destruction τ is essentially compensated by an increase in I. In terms of the empirical moments used to calibrate the model, an increase in market barriers will reduce the extent of life cycle growth by reducing firms' expansion rate x (despite the fact that aggregate creative destruction also declines). Also note that the entry rate increases. Recall that in a stationary equilibrium, the mass of entrants has to be equal to the mass of exiting firms, i.e. the mass of single-product firms experiencing a destruction shock. Even though creative destruction in the aggregate declines, there is still more exit, simply because in an economy populated by small firms it is more likely that a producer exiting from a particular market coincides with an exiting firm.

The right panel of the figure contains the same statistic for an increase in entry costs. Except for the case of the markup distribution, the effects of entry costs are the polar opposite of the case of market barriers: higher entry costs increase the rate of life cycle growth, they reduce the number of firms by making firms bigger and they (naturally) lower the entry rate. As for the case of expansion barriers, the effect on the aggregate growth rate is very small.

These patterns are useful to relate differences in entry costs and expansion barriers to differences in firm level data. Consider for example a developed economy like the US. While firms in the US grow at a faster rate, the equilibrium entry rate is not much higher - if anything it is probably lower. Together with Figure 4, these two facts directly imply that the US economy has to be characterized by lower market barriers.

To quantify the aggregate importance of such barriers, I will therefore recalibrate the parameters governing market barriers (ϕ_x) and entry costs (ϕ_z) to match the entry rate and the life cycle growth rate of employment in the US. The remaining parameters are left unchanged. As a benchmark for the rate of life cycle growth in the US, I consider the results of Hsieh and Klenow (2014), who report that US firms grow by a factor 2 at the 10 year horizon.[30] This corresponds to an annual rate of employment growth of about 7% conditional on survival.[31] As for the entry rate, I target a value of 8% for the baseline results, which is consistent with Karahan et al. (2015), who report a start-up rate of between 8% and 11% for the whole economy and Akcigit et al. (2015), who calculate an entry rate of 7.5% in the US manufacturing sector. However, I will show explicitly how sensitive the results are with respect to these choices. For ease of comparison, I used the calibration to the Indonesia economy reported in Table 3 to also calculate the implied employment growth for firms in Indonesia at the 10 year horizon, which turns out to be 1.7. Hence, the US entry rate is lower and the rate of life cycle growth higher.

The results of this exercise are contained in Table 4. The left panel contains the calibration results, i.e. the respective moments and the resulting structural parameters. The first result, which emerges is that while both entry costs and market barriers are lower in the US, the market barrier margin is particularly relevant. While the entry technology in the US is about 15% more productive, the costs for existing firms to break into new markets are about a third lower. The right panel contains the resulting equilibrium implications. I report various objects, which concern the (i) distribution of firm size, (ii) the distribution of markups and (iii) the dynamic aspects of the economy.

As far as the firm size distribution is concerned, differences in entry costs and market barriers are quantitatively important. As firms have more opportunities to expand their scale of production, the firm size distribution shifts to the right and the economy sustains less firms in equilibrium. The equilibrium number of firms declines by 60%, i.e. average firm size more than doubles. This reallocation comes especially at the expense of small firms so that

[30] More specifically, they show that 10-14 year old plants are twice as large as plants less than 5 years old.

[31] This is somewhat lower than the findings of Akcigit and Kerr (2015), who estimate a rate of unconditional employment growth of 7.5% for the universe of patenting US firms, a sample of firms, which is more selected than the population of manufacturing firms studied in Hsieh and Klenow (2014).

the share of small firms, which I take here to be firms who serve only a single market, declines by more than 70%. Hence, seemingly small differences in the rate of entry and life cycle growth can have very large effects on the cross-sectional distribution of firm size.

This shift towards large firms is accompanied with pro-competitive effects. Even though markups are strongly increasing in size in the cross-section, lower entry costs and market barriers increase average firm size and simultaneously reduce markups as firms' increase their markups at a lower rate. In particular, while seven year old firms in Indonesia have about 8% higher markups than current entrants (see Figure 1), this difference declines by two percentage points in the US.[32] This reduction in markup growth reduces both average markups and the dispersion of markups. As pointed out above: the welfare relevant distribution of markups across markets does not coincide with the distribution of measured markups, as firms are active in multiple product markets. In particular, an increase in firm size almost mechanically implies a reduction in measured revenue productivity dispersion as the variation *within* firms increases. This divergence is quantitatively important as the decline in the standard deviation of log markups across firms exceeds the decline in "true" markup dispersion by 50%.[33] The last two columns show that the decline in misallocation increases TFP modestly by 0.2% and that the reductions in monopoly power are akin to a 1.8% decline in taxes on static factors.[34]

Finally, consider the growth implications. The most striking result is that the economy-wide growth rate g hardly changes - if anything it slightly *declines* once entry barriers and expansion barriers are dismantled. The reason is the equilibrium effect on firms' markup-increasing productivity investments. While creative destruction increases by 15%, firms' incentives to increase productivity within their existing markets decline by 11%. That these different margins of growth are negatively related is not surprising - recall the optimality condition in (24), which showed that the marginal value of higher markups is discounted at rate $\rho + \tau$. This competition effect, which is present in most models of Schumpeterian growth, is - in this calibration - sufficiently strong that $I + \tau$ declines, even though τ increases.

The results in Table 4 have three important implications. First of all, seemingly *large* changes in the stationary firm size distribution and the number of active firms are fully consistent with empirically plausible *small* differences in observable entry rates, employment life cycle growth and the increase in markups by age. Secondly, such large differences do *not* imply that countries are predicted to grow at vastly different rates. A growth differences of the one reported in Table 4 only accumulates to a productivity level difference of 2% after 20 years. Hence, the model exactly predicts that firm size distributions across countries could be vastly different while the distribution of income across countries might - for all practical purposes - be relatively stable.[35] Finally, the results suggest that frictions

[32]This is consistent with the results reported in Hsieh and Klenow (2014), who argue that the increase in revenue-productivity by age, which in my model is proportional to markups, is steeper in India relative to the US. See also Bento and Restuccia (forthcoming) or Fattal Jaef (2011), who assume a particular elasticity of distortions to productivity and then quantify the effects of this elasticity on aggregate outcomes.

[33]I do not have an analytic formula for the stationary distribution of firm-level markups. I therefore calculate the dispersion in measured markups numerically. To account for sampling variation, I draw 300 distributions of firm-level markups from the stationary distribution and report the median of these numbers. In Section OA-2.7 in the Online Appendix I report the entire distribution for both the standard deviation and the interquartile range of log markups for both countries.

[34]The implied TFP losses seem small, especially compared to the much bigger numbers reported in Hsieh and Klenow (2009). There are two reasons. First of all, by reducing misallocation in India to US standards, they consider a much bigger liberalization experiment as the empirical productivity dispersion in India is 50% higher. This in itself will cause the implied efficiency losses to be bigger. As noted by Jones (2013): "Small departures from the optimal allocation of labor have tiny effects on TFP (an application of the envelope theorem), but significant misallocation can have very large effects." Secondly, they consider an elasticity of substitution across products of three, whereas I impose a unitary demand elasticity. To see that Table 4 is then directly comparable to their results, recall that in their set-up, the change in aggregate TFP is approximately given by $dln\left(TFP\right) = -\frac{\sigma}{2} dvar\left(ln\left(ARPL\right)\right)$, where σ is the elasticity of substitution across varieties. For $\sigma = 1$ and $dvar\left(ln\left(ARPL\right)\right) = 0.105^2 - 0.0847^2$, one exactly recovers a TFP loss of 0.2%.

[35]There are, of course, differences in the rate of growth across countries. These are, however, at least partially driven by transitional dynamics. Also: the results in Table 4 do not imply that the US economy is predicted to grow at a rate of 3%. What Table 4 shows is that differences in firm dynamics can lead to very different distribution of firm size without large differences in the rate of growth.

for existing firms to expand into new markets are more important than differences in entry costs to understand the empirical patterns of manufacturing firms in developing countries. Not only do they readily imply that many firms are small and experience little growth as they age, but they simultaneously predict that misallocation is more severe.

The Importance of Heterogeneous Markups Allowing for the channel of heterogeneous markups is crucial for these results. To see this, consider a version of this model, where firms cannot increase the productivity in their existing markets, i.e. $\varphi_I = 0$ so that $I = 0$. In that special case, the model reduces to the baseline model of Klette and Kortum (2004), where markups are constant and equal to λ. In the last column of Table 4, I report the results from a comparable calibration, which abstract from heterogeneous markups. These models again match the same entry rates, the same extent of life cycle growth and - for the case of Indonesia - the same aggregate growth rate. They of course do not match the life cycle growth of markups displayed in Figure 1 and do not feature any endogenous misallocation of resources.

The last column in Table 4 shows that this model would predict changes in the number of firms and the share of small firms of a similar magnitude.[36] However, the growth implications are very different as this model predicts that growth would increase by 0.6%. In the baseline model, the endogenous response of incumbent firms ensures that large changes in the firm size distribution do not lead to large changes in long-run growth. In the economy without markups, this countervailing effect is missing and the theory would imply that the slightly faster life cycle growth of firms in the US would come hand-in-hand with a much higher growth rate.

Robustness In Section OA-2.3 in the Appendix I conduct various robustness checks of these results. First of all, I show that neither the implied changes in the firm size distribution, nor the implications for the aggregate growth rate depend substantially on the choice of the curvature parameter ζ. The resulting changes in the number of firms or the importance of small firms are all very similar to the ones reported in Table 4 and the quantitative effect of the changes in entry and expansion barriers on the equilibrium growth rate is very small (even though for large values of ζ, the predicted growth rate in the US is actually slightly higher). I also study the sensitivity with respect to the underlying moments, i.e. the extent of life cycle growth and the extent of entry. The baseline calibration for the US assumed an entry rate of 8% and that employment grows by a factor of two during the first 10 years of a firm's life cycle. The elasticity of average firm size and and the share of small firms with respect to these two moments is quite sizable. If one were to assume that the extent of life cycle growth in the US was 2.5 instead of 2, the number of firms and the share of small firms would fall by 90%. In contrast, the effect on the growth rate is still almost indistinguishable from zero. Again this is very different for the economy, which abstracts from the markup channel. As in the last column in Table 4, changes in the firm size distribution and changes in the aggregate growth are very tightly linked: if life cycle growth in the US was 2.5 instead of 2, the *differences* in aggregate productivity growth would be 3%.

3.4 Market barriers and entry costs in Indonesia

The analysis above suggests that differences in market barriers could be an important determinant of firm size, firm growth and misallocation. In this last section I will provide some direct corroborating evidence for such frictions. To do so I exploit regional variation across markets in Indonesia. As I do not have direct information on the type

[36] It it also the case that the change in entry and expansion barriers is qualitatively similar. Entry barriers decline by 5% and market barriers decline by 32%.

of barriers different firms might face, I use the theory to suggest an empirical strategy based on the *joint* patterns of various firm-level outcomes.

The basic intuition is simple and follows from the model's comparative static result shown in Figure 4. If different regions in Indonesia differed only in their market barriers ϕ_x, locations with low market barriers should see fewer and bigger firms, lower entry rates and a steeper schedule of life cycle employment growth. Additionally, product markets in such regions should also be characterized by lower markups. If in contrast entry costs ϕ_z were the dominant source of variation across regions, it would *also* be the case that firm size should be negatively correlated with regional entry rates and positively correlated with the slope of life cycle growth - however, the underlying source of cost variation would be exactly reversed. Now large firms should reside in regions with high entry costs and one would expect a positive correlation between firm size and the prevailing markups.

I implement this strategy in the following way. The Indonesian micro-data allows me to link individual firms to their geographic location. I define a geographical region as a province, of which there are 27 in the data. Because I do not have information on where firms sell their products, I need to assume that firms are predominantly active in their own province. Provinces obviously differ in their industrial composition. As industries differ in their average size, I conduct the entire analysis at the region-industry level and control for the common industry component using fixed effects. Hence, the variation of interest is geographical in nature. More specifically, I calculate my outcomes of interest, i.e. average firm size, entry and exit rates, average markups and the employment life cycle growth rates for each province-industry-year cell and then consider regressions of the form

$$y_{rst} = \delta_s + \delta_r + \beta \times AvgSize_{rst} + \gamma \times ln\left(pop_r\right) + \alpha \times Ag_r + u_{rst}, \tag{30}$$

where δ_s and δ_t are sector and time fixed effects, $AvgSize_{rst}$ is the average size of producers active in region r, in sector s in time t and y_{rst} are the different outcome variables mentioned above. Moreover, I also control for the size of the population in region r and the regional agricultural share to account for the effects of market size.[37] Given the focus on the regional variation, I cluster all standard errors at the province level, to allow for correlation in the error term across industries within a province.

My preferred measure of size is firms sales (rather then employment), as the theory predicts that average sales only depend on the expansion intensity $\vartheta_x = \frac{x}{\tau}$ (see Proposition 2). To calculate the employment life cycle, I again rely on the panel dimension and adopt the same methodology as for Figure 2.[38] To not identify the parameters from sparsely populated region-industry-year cells, I estimate (30) only for the set of observations, which contain at least 50 observations and I weigh the regression by the number of observations in each bin.

Table 5 contains the results. Consider first the first three columns, which show that average firm size in a region is negatively correlated with entry and exit rates and positively correlated with the extent of life cycle growth. These correlation are consistent with the model and hold regardless of wether the source of variation across regions stems from entry costs or market barriers. In the remaining columns I turn to the analysis of markups. Columns four and five show that average size is negatively correlated with markups as measured by either the average markup of the 90%-quantile of the within region-industry. This is consistent with the model if regional firm size is driven by differences in market access but not consistent with an explanation based on entry costs. In columns six and seven I focus directly on the effects of entry and show that regional entry rates are also negatively correlated with

[37]To measure the geographical characteristics, I exploit information from the Village Potential Statistics (PODES) dataset in 1996. The PODES dataset contains detailed information on all of Indonesia's 65.000 villages. Using the village level data, I then aggregate this information to the province level and match these to the firm-level data. In particular I measure the size of the population and the share of the population living in villages within a region, which are predominantly agricultural.

[38]In particular, for each cohort I calculate employment growth at the three year horizon after controlling for a full set of 5-digit product and year fixed effects and than average these growth rates at the industry-province level. I have to opt for a somewhat short horizon of three years to not lose too many observations given that I only have data for the years 1991 to 1998.

the level of markups albeit not positively so. An explanation based on market barriers (entry costs) would require this coefficient to be positive (negative).

Finally, the last two columns allow for both effects to be at play. Because entry *rates* and average size both depend on the same endogenous outcome $\frac{x}{\tau}$, I find it useful to measure the flow rate of entry z as the entry rate relative to the average size of firms (i.e. $z = \frac{EntryRate}{AvgSize}$ as implied by (28)). Hence, the last column essentially contains a specification where markups are regressed on $\frac{x}{\tau}$ (as measured by the average size) and z. The theory predicts both coefficients to be negative. Holding z constant, an increase in $\frac{x}{\tau}$ will increase creative destruction $\tau = z + x$ and hence increase the intensity of creative destruction $\vartheta_I = \frac{\tau}{I}$. The same is true for the effect of z holding $\frac{x}{\tau}$ constant. Empirically, both coefficients are negative, but the majority of variation is driven by regional differences in firm size.[39] While these patterns are suggestive of market barriers potentially playing an important role in the determination of misallocation, firm dynamics and growth, they highlight the need to directly measure why the costs of entering products markets within countries might be systematically related to the level of development.

4 Conclusion

There is ample evidence that firms in developing countries are small (but numerous), that these firms grow slowly as they age and that misallocation across firms is persistent. In this paper I offer an explanation for these patterns. I construct a tractable growth model where firms charge heterogeneous markups and show that these empirical regularities naturally follow if firms in poor countries face high costs to expand into new product markets. Such market barriers not only keep firms small, but they also increase markups and misallocation by rendering product markets uncompetitive. This is in stark contrast to entry costs, i.e. costs for new firms to enter the economy, which also increase misallocation but which at the same time raise average firm size.

In order to gauge the quantitative importance of this mechanism, I calibrate the model to firm-level panel data from Indonesia. First I provide direct evidence that both employment and markups are systematically increasing as firms age. Secondly, I use the regional variation across Indonesian provinces and show that the observed patterns of markups, firm size and entry and exit rates are consistent with variation in market barriers, but hard to reconcile with differences in entry costs. Finally, I show that the quantitative importance of this mechanism is sizable. The observed rates of entry and life cycle growth in the US imply that firms in the US are about 30% more efficient to break into new markets than firms in Indonesia. In contrast, entry costs are only 15% lower. These differences reduce average firm size by a factor of more than two, increase the importance of small firms by a factor of four and modestly increase allocational efficiency through lower markups. Importantly, these cost differences do not yield counterfactual growth implications as the equilibrium growth rate hardly changes. Large differences in firm size and misallocation are therefore consistent with a stable distribution of income across countries.

References

Acemoglu, Daron, Pol Antras, and Elhanan Helpman, "Contracts and Technology Adoption," *American Economic Review*, 2007, *97* (3), 916–943.

_ **, Ufuk Akcigit, Nicholas Bloom, and William Kerr**, "Innovation, Reallocation and Growth," 2012. Working Paper.

[39] In section OA-2.3 in the Appendix I show that the results are robust to (a) measuring firms size by employment, (b) not weighting the regressions, (c) different choices to the minimum number of firms for a given cell to appear in the final sample, (d) not clustering the standard errors and (e) different measures of entry.

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

_ **, Christopher Harris, Peter Howitt, and John Vickers**, "Competiton, Imitation and Growth with Step-By-Step Innovation," *Review of Economic Studies*, 2001, *68*, 467–492.

_ **, Nick Bloom, Richard Blundell, Rache Griffith, and Peter Howitt**, "Competition and Innovation: An Inverted U-Relationship," *Quarterly Journal of Economics*, 2005, pp. 701–26.

_ **, Robin Burgess, Steve Redding, and Fabrizio Zilibotti**, "The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India," *The American Economic Review*, 2008, *94* (4), 1397–1412.

Akcigit, Ufuk and William Kerr, "Growth through Heterogeneous Innovations," 2015. Working Paper.

_ **, Harun Alp, and Michael Peters**, "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries," 2015. Working Paper.

Amiti, Mary and Jozef Konings, "Trade Liberalization, Intermediate Inputs and Productivity," *American Economic Review*, 2007, *97* (5), 1611–1638.

Arkolakis, Costas, "Market Penetration Costs and the New Consumer Margin in International Trade," *Journal of Political Economy*, 2008, *118* (6), 1151–1199.

_ **, Arnaud Costinot, Dave Donaldson, and Andres Rodriguez-Clare**, "The Elusive Pro-Competitive Effects of Trade," 2012. Working Paper.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 2014, *122* (5), 1013–1063.

Atkeson, Andrew and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 2008, *98* (5), 1998–2031.

_ **and** _ , "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 2010, *118* (3), 433–484.

_ **and** _ , "Aggregate Implications of Innovation Policy," 2015. Working Paper.

Banerjee, Abhijit and Benjamin Moll, "Why Does Misallocation Persist?," *American Economic Journal: Macroeconomics*, 2010, *2* (1), p. 189–206.

Barseghyan, Levon, "Entry Costs and Cross-Country Differences in Productivity and Output," *Journal of Economic Growth*, 2008, *13* (2), 145–167.

_ **and Riccardo DiCecio**, "Entry Costs, Misallocation and Cross-Country Income and TFP Differences," Working Paper, Federal Reserve Bank of St. Louis 2009.

Bartelsman, Eric J., John C. Haltiwanger, and Stefano Scarpetta, "Cross-Country Differences in Productivity: The Role of Allocative Efficiency," *American Economic Review*, 2013, *103* (1), 305–334.

Bento, Pedro, "Competition, Innovation, and the Number of Firms," 2016. Working Paper.

_ **and Diego Restuccia**, "Misallocation, Establishment Size, and Productivity," *American Economic Journal: Macroeconomics*, forthcoming.

Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, "Plants and Productivity in International Trade," *American Economic Review*, 2003, *93* (4), 1268–1290.

Blalock, Garrick, Paul J. Gertler, and David I. Levine, "Financial Constraints on Investment in an Emerging Market Crisis," *Journal of Monetary Economics*, 2008, *55* (3), 568–591.

Bollard, Albert, Peter J. Klenow, and Huiyu Li, "Entry Costs Rise with Development," 2016. Working Paper.

Buera, Francisco and Roberto N. Fattal Jaef, "The Dynamics of Development: Innovation, and Reallocation," 2016. Working Paper.

_ , Joe Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 2011, *101* (5), 1964–2002.

Chari, V.V., Patrick J. Kehoe, and Ellen R. McGrattan, "Business Cycle Accounting," *Econometrica*, 2007, *75* (3), 781–836.

Clementi, Gian Luca and Hugo Hopenhayn, "A Theory of Financing Constraints and Firm Dynamics," *The Quarterly Journal of Economics*, 2006, *121* (1), 229–265.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-De-Silvaes, and Andrei Shleifer, "The Regulation of Entry," *Quarterly Journal of Economics*, 2002, *117* (1), 1–37.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "Competition, Markups and the Gains from International Trade," *American Economic Review*, 2015, *105* (10), 3183–3221.

Epifani, Paolo and Gino Gancia, "Trade, Markup Heterogeneity and Misallocations," *Journal of International Economics*, 2011, *83* (1), 1–13.

Foster, Julia, John C. Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

Foster, Lucia, John Haltiwanger, and C. J. Krizan, "Aggregate Productivity Growth. Lessons from Microeconomic Evidence," in Edwin R. Dean Charles R. Hulten and Michael J. Harper, eds., *New Developments in Productivity Analysis*, University of Chicago Press, 2001, chapter 8, pp. 303–372.

Gabaix, Xavier, "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 1999, *114* (3), 739–767.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Pete Klenow, "How Destructive is Innovation?," 2016. Working Paper.

Gourio, Francois and Leena Rudanko, "Customer Capital," *Review of Economic Studies*, 2014, *81* (3), 1102–1136.

Grossman, Gene and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, 1991, *58* (1), 43–61.

Guner, N., G. Ventura, and Y. Xu, "Macroeconomic implications of size-dependent policies," *Review of Economic Dynamics*, 2008, *11* (4), 721–744.

Herrendorf, Berthold and Arilton Teixeira, "Barriers to Entry and Development," *International Economic Review*, 2011, *52* (2), 573–602.

Holmes, Thomas J., Wen-Tai Hsu, and Sanghoon Lee, "Allocative Efficiency, Mark-ups, and the Welfare Gains from Trade," *Working Paper*, 2013.

Hopenhayn, Hugo, "On the Measure of Distortions," *Working Paper*, 2012.

Hsieh, Chang-Tai and Benjamin A. Olken, "The Missing Missing Middle," *Journal of Economic Perspectives*, 2014.

_ and Peter Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (2), 771–807.

_ and _ , "The Life Cycle of Plants in India and Mexico," *Quarterly Journal of Economics*, 2014, *129* (3), 1035–1084.

Hurst, Erik and Ben Pugsley, "What Do Small Businesses Do," *Brookings Papers on Economic Activity*, 2012.

Jaef, Roberto N. Fattal, "Entry, Exit and Misallocation Frictions," 2011. Working Paper.

Jones, Charles, "Misallocation, Economic Growth and Input-Output Economics," in Daron Acemoglu, M. Arrelano, and E. Dekel, eds., *Advances in Econometrics: 10th World Congress*, Vol. 2, Cambridge University Press, 2013.

_ and Jihee Kim, "A Schumpeterian Model of Top Income Inequality," 2016. Working Paper.

Karahan, Fatih, Benjamin Pugsley, and Aysegül Sahin, "Understanding the 30-year Decline in the Start-up Rate: A General Equilibrium Approach," 2015. Working Paper.

Klette, Tor Jakob and Samuel Kortum, "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

Lentz, Rasmus and Dale T. Mortensen, "An Empirical Model of Growth through Product Innovation," *Econometrica*, 2008, *76* (6), 1317–1373.

Loecker, Jan De, "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 2011, *79* (5), 1407–1451.

_ , "Recovering Markups from Production Data," *International Journal of Industrial Organization*, 2011, *29*, 350–355.

_ and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 2012, *102* (6), 2437–2471.

_ , Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik, "Prices, Markups and Trade Reform," *Econometrica*, 2016, *84* (2), 445–510.

Luttmer, Erzo G. J., "Models of Growth and Firm Heterogeneity," *Annual Review of Economics*, 2010, *2*, 547–576.

Melitz, Marc and Giancarlo Ottaviano, "Market Size, Trade and Productivity," *Review of Economic Studies*, 2008, *75* (1), 295–316.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and Misallocation: Evidence From Plant-Level Data," 2010. NBER Working Paper 15647.

Moll, Benjamin, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, 2014, *104* (10), 3186–3221.

Perla, Jesse, "Product Awareness, Industry Life-Cycles, and Aggregate Profits," 2016. Working Paper.

Porta, Rafael La and Andrei Shleifer, "The Unofficial Economy and Economic Development," in "Brookings Papers on Economic Activity" 2009, pp. 275–352.

_ **and** _ , "Informality and Development," *Journal of Economic Perspectives*, 2014, pp. 109–126.

Poschke, Markus, "The firm size distribution across countries and skill-biased change in entrepreneurial technology," 2011. Working Paper.

Restuccia, Diego and Richard Rogerson, "Policy Distortions and Aggregate Productivity with Heterogenous Establishments," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Schoar, Antoinette, "The Divide between Subsistence and Transformational Entrepreneurship," in "Innovation Policy and the Economy, Volume 10," University of Chicago Press, 2010, pp. 57–81.

Sutton, John, "Gibrat's Legacy," *Journal of Economic Literature*, 1997, *35*, 40–59.

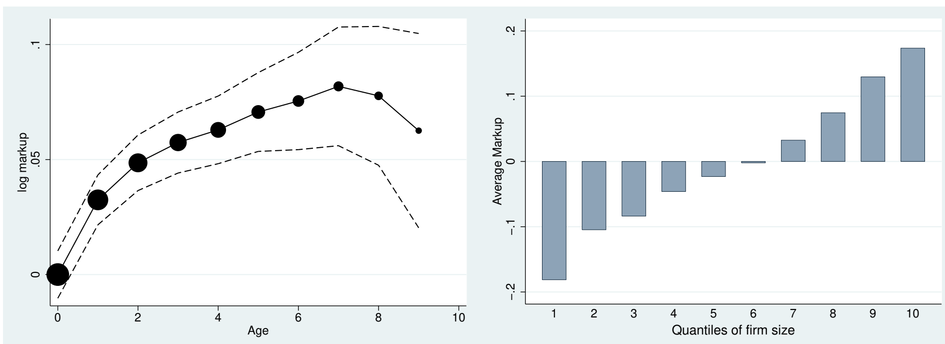
Yang, Mu-Jeung, "Micro-level Misallocation and Selection: Estimation and Aggregate Implications," *Working Paper*, 2012.

5 Tables and Figures

Firm size distribution				Entrants			Exiting firms		
Mean	Quantiles			Entry rate	Share of		Exit rate	Share of	
	25%	50%	90%		employment	sales		employment	sales
143	27	45	351	10.4%	5.0%	3.9%	8.2%	4.4%	3.3%

Notes: The table contains descriptive statistics on the sample of manufacturing plants in Indonesia. Columns 1 - 4 contain selected statistics about the distribution of employment. Columns 5 and 7 contain the entry and exit rate. Columns 6 and 8 report the employment share of entering and exiting firms. All results are simple averages over the time of the sample, i..e 1991 to 1997. Table OA-1 in the Online Appendix contains the annual results.

Table 1: The manufacturing sector in Indonesia



Notes: The figure shows the life cycle of markups (left panel) and the cross-sectional relationship between markups and firm size. To calculate the markup life cycle, I focus on the unbalanced panel of firms entering the economy after 1991. I calculate log markups within 5-digit-industry-year cells, then calculate the average by the age of the cohort and normalize log markups of entering cohorts to zero. Because of attrition, the size of the cohort is declining in age. The dots reflect the size of the cohort. I also depict the 10% confidence intervals around the estimated average profile. To calculate log markups by size, I simply average log markups within the respective quantiles of the cross-sectional employment distribution.

Figure 1: Markups along the life cycle and by firm size

Targets	Data	Model		Calibrated parameters		Endogenous outcomes	
Life-Cycle of Markup	0.082	0.082	φ_I	Productivity of innovation	8.896	I	0.561
Life-Cycle of Employment	0.494	0.494	φ_x	Productivity of expansion	1.077	x	0.207
Entry rate	10.4%	10.4%	φ_z	Productivity of entry	2.590	z	0.032
Aggregate productivity growth	3%	3%	λ	Innovation step size	1.038	τ	0.240
-	-	-	ζ	Curvature of cost function	2		
-	-	-	ρ	Rate of time preference	0.05		

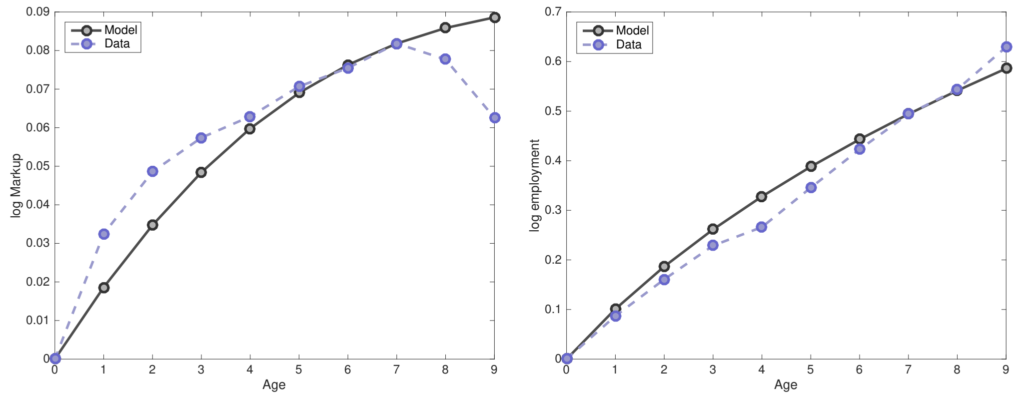
Notes: This table contains the calibrated parameters, the targeted data moments, the resulting moments in the model and the equilibrium outcomes for I, x and z. Given (I, x, z) all other equilibrium allocations can be calculated. I measure the life cycle of markups and employment as the log difference between entrants and seven year old firms - see the left panel of Figure 1 for the case of markups and Section OA-2.3 in the Appendix for the case of employment. The entry rate stems from Table 1. The two parameters ζ and ρ are set exogenously. As explained in the text, the mapping from the data moments to (I, x, z, λ) does not depend on ζ and ρ. See Section 6.6 in the Appendix for details.

Table 3: Calibration

	log markups							
Age	0.0184***	0.0137***	0.0103***		0.0134***	0.00825***	0.0108***	0.0169***
	(0.00132)	(0.00122)	(0.00117)		(0.00158)	(0.00238)	(0.00127)	(0.00184)
$\ln \frac{k}{l}$			0.121***	0.124***	0.123***	0.112***	0.127***	0.131***
			(0.00207)	(0.00222)	(0.00222)	(0.00496)	(0.00237)	(0.00305)
Entry				-0.0250***	0.00777			
				(0.00638)	(0.00744)			
Exit				-0.0176**	-0.0144*			
				(0.00835)	(0.00836)			
log value added						0.0232***		
						(0.00769)		
Industry FE	N	Y	Y	Y	Y	Y	Y	-
Firm FE	N	N	N	N	N	N	N	Y
N	55212	55212	55212	48556	48556	55212	42434	55212
R^2	0.012	0.219	0.287	0.293	0.294	0.106	0.298	0.757

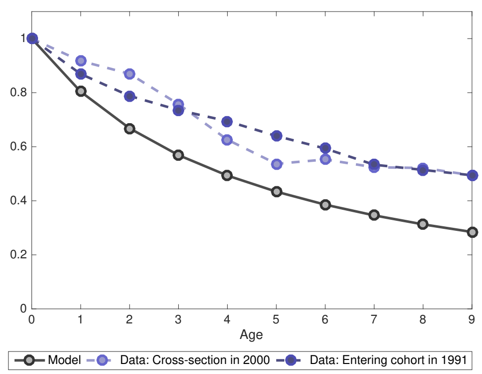
Notes: Robust standard errors in parentheses with ***, **, and * respectively denoting significance at the 1%, 5% and 10% levels. I focus on the unbalanced panel of firms, who enter the market after 1990. I use the data from 1991 to 2000. All specifications include year fixed effects. $ln\,(k/l)$ denotes the (log) capital-labor ratio at the firm level. 'Entry" and "Exit" are indicator variables for wether the firm enters (exit) the market in a given year. In column 7 I focus on the balanced panel, i.e. only consider firms that survive to the end of my sample period. The specifications with industry fixed effects control for industry affiliation at the 5 digit level. Column 8 contains firm fixed effects.

Table 2: The life cycle of markups



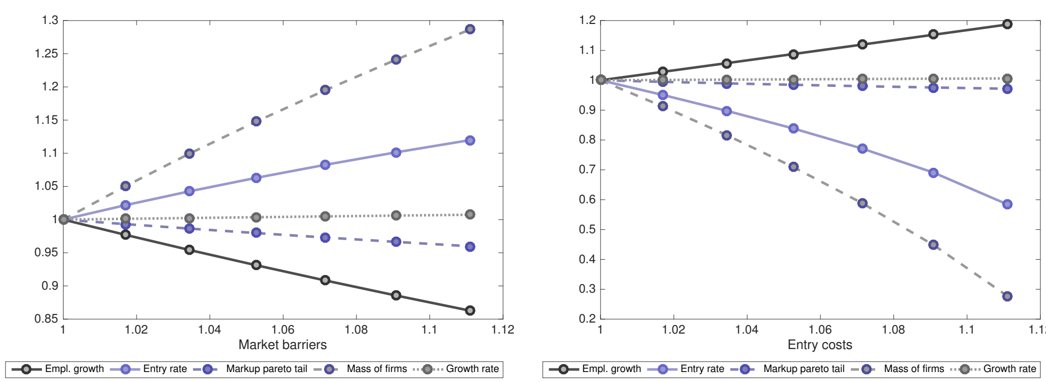
Notes: The figure displays the model's prediction and the data for the life cycle of markups (left panel) and employment growth (right panel). The model is calibrated according to the parameters in Table 3. The data is taken from Figure 1 for the case of markups and from Section 6.6 in the Appendix for the case of employment.

Figure 2: The life cycle of markups and firm size (Model vs. Data)



Notes: The figure displays the model's prediction and the data for the share of surviving firms by age. The model is calibrated according to the parameters in Table 3 and the survival function is given by (29). For the data, the figure contains both the share of firms by age in 2000 cross-section (relative to the number of firms at the time of entry) and the share of firms of the entering cohort in 1991 by age.

Figure 3: Firm survival as a function of age (Model vs. Data)



Notes: The figure shows the effect of market barriers, ϕ_x, (left panel) and entry costs, ϕ_z, (right panel) on (i) the life cycle rate of employment growth ("Empl. growth"), (ii) the entry rate, (iii) the tail parameter of the endogenous distribution of markups $\theta = ln\left(1+\vartheta_I\right)/ln\left(\lambda\right)$ (see Proposition 2), (iv) the equilibrium number of firms $F = \frac{1-\vartheta_x}{\vartheta_x} ln\left(\frac{1}{1-\vartheta_x}\right)$ (see Proposition 2) and (v) the growth rate $g = ln\left(\lambda\right)\left(I+\tau\right)$ (see Proposition 2). I normalize the initial equilibrium the model is calibrated to to unity and then increase market barriers and entry costs.

Figure 4: The effects of market barriers and entry costs

Calibration	Indonesia	US	Equilibrium implications	Indonesia	US	Change	Change without markups
Entry Rate	10.4%	8%	*The Distribution of Firm Size*				
Empl. Life-Cycle	1.7	2	Number of firms	0.313	0.122	-60.1%	-61.1%
Expansion Barrier	1	0.67	Output share of small firms	0.136	0.035	-73.9%	-75.6%
Entry Barriers	1	0.86	*Markups and Misallocation*				
			Life cycle of markups	8.2%	6.2%	0.02	-
			Average markup	11.73%	9.25%	-0.025	-
			Dispersion in log markups	10.5%	8.47%	-19.3%	-
			Measured markup dispersion	8.1%	5.86%	-27.5%	-
			Efficiency wedge M	0.995	0.997	0.2%	-
			Labor wedge Λ	0.887	0.903	1.8%	-
			Innovation, Expansion and Entry and Growth				
			Rate of growth g	3%	2.91%	-0.0009	0.0063
			Rate of innovation I	0.561	0.498	-11.1%	-
			Rate of market expansion x	0.207	0.267	28.6%	42.3%
			Rate of creative destruction τ	0.240	0.277	15.3%	20.8%

Notes: The first panel contains the calibration moments. The entry rate is simply the share of firms, which are entrants. For the US, the employment life cycle is calculated as average employment of firms between 10 and 14 years old relative to firms with age less than 5 and stems from Hsieh and Klenow (2014). The parameters for the Indonesian economy are contained in 3. For the US economy, I recalibrate the relative efficiency of expansion and entry, i.e. $\frac{\phi_z^{US}\varphi_z^{US}}{\phi_z^{IND}\varphi_z^{IND}} = \frac{1}{0.86}$ and $\frac{\left(\phi_x^{\zeta}\right)^{US}\varphi_z^{US}}{\left(\phi_x^{\zeta}\right)^{IND}\varphi_z^{IND}} = \frac{1}{0.67}$. The remaining parameters are the same as in Table 3. The employment life cycle for Indonesia at the 10 year horizon reported in the first panel is calculated from the calibrated model to ease comparison. The closed-form expression for the endogenous outcomes in the second panel are contained in Propositions 2, (20) and (21). The column "Change" reports the change in the respective equilibrium outcomes. The last panel ("Change without markups") reports the implied changes in the equilibrium outcomes for the economy without endogenous markups.

Table 4: Changes in market barriers and entry costs: Indonesia and the US

	Entry rate	Exit rate	LC empl growth	Markups Avg.	q^{90}	Avg.	q^{90}	Avg.	q^{90}
Avg va	-0.019***	-0.011***	0.132***	-0.054***	-0.049***			-0.057***	-0.053***
	(0.004)	(0.003)	(0.034)	(0.014)	(0.010)			(0.015)	(0.010)
Entry rate						-0.181	-0.169		
						(0.113)	(0.165)		
$\ln z$								-0.017	-0.016
								(0.011)	(0.014)
Controls	Local population; Local agricultural share; Industry FE; Year FE								
N	455	463	463	462	462	454	454	454	454
R^2	0.369	0.320	0.640	0.378	0.369	0.348	0.353	0.388	0.374

Notes: Standard errors are clustered at the level of a province and contained in parentheses with ***, **, and * respectively denoting significance at the 1%, 5% and 10% levels. Regression are run at the province-industry level, where industries are measured at the 3-digit level. The variables are all measured within these province-industry cells. The entry and exit rates are measured as the share of entering and exiting firms. The employment life cycle is measured as the growth of cohort employment over the 3 year horizon (see also Figure 2). Log markups are measured as the residual from a regression of log inverse labor shares on a set of year and 5-digit industry fixed effects and $ln\left(k/l\right)$ (see (26)) and "Avg" is the mean log markup and q^{90} is the 90% quantile. "Avg va" is the average log value added within a industry-region-year cell. The flow rate of entry z is measured as in the theory, i.e. from $\frac{\text{Entry rate}}{\text{Avg va}}$ (see (28)). All regressions contain a full set of industry and year fixed effects and control for the log of the province population and the share of villages within the province, which are agricultural. I only consider province-industry cells with at least 50 observations and all regressions are weighted using the number of observations within each cell as a weights.

Table 5: Firm size, entry and markups across regions in Indonesia

6 Appendix

6.1 Proof of Proposition 1

I prove Proposition 1 in two steps. I first derive the value function (14). Then I show that the conditions in Proposition 1 uniquely define the optimal choices for (I, x, z).

6.1.1 Solving for the value function $V(.)$ in (14)

To derive (14), conjecture first that the value function takes an additive form[40]

$$V_t\left(n, [\Delta_i]_{i=1}^n\right) = V_t^P(n) + \sum_{i=1}^n V_t^M(\Delta_i). \tag{31}$$

Equation (11) then implies that V_t^P and V_t^M are defined by the differential equations

$$rV_t^M(\Delta) - \dot{V}_t^M(\Delta) = \pi_t(\Delta) - \pi_t(1) - \tau V_t^M(\Delta) + \max_I \left\{ I\left[V_t^M(\Delta+1) - V_t^M(\Delta)\right] - c^I(I,\Delta)\, w_t \right\}, \tag{32}$$

and

$$\begin{aligned} rV_t^P(n) - \dot{V}_t^P(n) \quad &= \quad n \times \pi_t(1) + \sum_{i=1}^n \tau\left[V_t^P(n-1) - V_t^P(n)\right] \\ &\quad + \max_X \left\{ X\left[V_t^P(n+1) + V_t^M(1) - V_t^P(n)\right] - c^X\left(\frac{X}{\phi_e}, n\right) w_t \right\}. \end{aligned} \tag{33}$$

Now consider a steady state where both value functions grow at rate g and assume the cost functions in (12). Then we can write (32) as

$$(r+\tau-g)\, V_t^M(\Delta) = \pi_t(\Delta) - \pi_t(1) + \max_I \left\{ I\left[V_t^M(\Delta+1) - V_t^M(\Delta)\right] - \frac{1}{\varphi_I}\lambda^{-\Delta} I^{\zeta} w_t \right\}. \tag{34}$$

Conjecture that

$$V_t^M(\Delta) = \kappa_t - \alpha_t \lambda^{-\Delta}. \tag{35}$$

Then $V_t^M(\Delta+1) - V_t^M(\Delta) = \frac{\lambda-1}{\lambda}\alpha_t\lambda^{-\Delta}$. This implies that optimal innovation rate I solves

$$I_t{}^{\zeta-1} = \left(\frac{\lambda-1}{\lambda}\frac{\varphi_I}{\zeta}\frac{\alpha_t}{w_t}\right)^{\frac{1}{\zeta-1}}. \tag{36}$$

Suppose that $\frac{\alpha_t}{w_t} = \text{const}$ (which I will verify below). (36) then implies that $I_t(\Delta) = I$. Using (36), (35) and the Euler equation $\rho = r - g$ yields

$$(\rho+\tau)\left[\kappa_t - \alpha_t\lambda^{-\Delta}\right] = \left(\frac{1}{\lambda} - \lambda^{-\Delta}\right) Y_t + \frac{\zeta-1}{\varphi_I}\lambda^{-\Delta} I^{\zeta} w_t,$$

[40] The analysis in this section only contains the most important steps. A detailed derivation is contained in Section OA-1.1 of the Online Appendix.

so that $\kappa_t = \frac{\frac{1}{\lambda}Y_t}{\rho+\tau}$ and $\alpha_t = \frac{Y_t - \frac{\zeta-1}{\varphi_I}I^\zeta w_t}{\rho+\tau}$. Hence, (35) yields

$$V_t^M(\Delta) = \frac{\left(\frac{1}{\lambda} - \lambda^{-\Delta}\right)Y_t + \lambda^{-\Delta}\frac{\zeta-1}{\varphi_I}I^\zeta w_t}{\rho+\tau} = \frac{\pi_t(\Delta) - \pi_t(1) + (\zeta-1)c_I(I,\Delta)w_t}{\rho+\tau}.$$

Note also that this implies that

$$V_t^M(1) = \frac{(\zeta-1)w_t\frac{1}{\lambda}\frac{1}{\varphi_I}I^\zeta}{\rho+\tau}. \tag{37}$$

Now turn to $V_t^P(n)$. Define $X = xn$ and conjecture that $V_t^P(n) = n \times v_t$, where v_t grows at rate $g = r - \rho$. Hence,

$$(\rho+\tau)v_t = \pi_t(1) + \max_x\left\{x\left[v_t + V_t^M(1)\right] - \frac{1}{\varphi_x}\left(\frac{x}{\phi_x}\right)^\zeta w_t\right\}$$

The optimality condition for x reads

$$v_t + V_t^M(1) = \frac{\zeta}{\varphi_x\phi_e^\zeta}x^{\zeta-1}w_t. \tag{38}$$

As v_t and $V_t^M(\Delta)$ both grow at rate g, this implies that x is indeed constant. In particular, given x, v_t is given by

$$(\rho+\tau)v_t = \pi_t(1) + (\zeta-1)\frac{1}{\varphi_x\phi_e^\zeta}x^\zeta w_t. \tag{39}$$

To solve for v_t, let $v_t = \overline{v} \times w_t$. The unknowns $(x, \overline{v})$ are then determined from (38) ,(39) and (37) as

$$\begin{aligned}\frac{\zeta}{\varphi_x\phi_e^\zeta}x^{\zeta-1} &= \overline{v} + \frac{V_t^M(1)}{w_t} = \overline{v} + \frac{(\zeta-1)\frac{1}{\lambda}\frac{1}{\varphi_I}I^\zeta}{\rho+\tau} \\ \overline{v} &= \frac{\frac{\lambda-1}{\lambda}\frac{Y_t}{w_t} + \frac{\zeta-1}{\varphi_x\phi_e^\zeta}x^\zeta}{\rho+\tau}.\end{aligned}$$

The final value function $V_t^P(n)$ is given by

$$V_t^P(n) = n \times \frac{\pi_t(1) + (\zeta-1)c^X(1,x)w_t}{\rho+\tau}.$$

Substituting into (31) yields (14).

6.1.2 Existence and Uniqueness

We now prove existence and uniqueness of the equilibrium. We need to solve for the tupel (I, x, z). Alternatively, we can solve for (I, x, τ) and then solve for $z = \tau - x$. From the static allocations we know that $Y_t\Lambda_t = w_t L_t^P$. Note that Λ_t is a known function of τ/I (see Proposition 2) and hence I write it as $\Lambda\left(\frac{\tau}{I}\right)$. To solve for L_t^P we need the labor market clearing condition, which is given by $1 = L_t^P + L_t^I + L_t^x + L_t^z$,where L_t^j denotes the total amount of labor used for the respective sources of innovation, expansion and entry. Note that $L_t^z = \frac{1}{\varphi_I\phi_z}z$, $L_t^x = \frac{1}{\varphi_x\phi_x^\zeta}x^\zeta$ and

$$L_t^I = \int_{j=0}^1 c^I(I,\Delta_j)\,dj = \int_{j=0}^1 \frac{1}{\varphi_I} \times I^\zeta\lambda^{-\Delta_j}\,dj = \frac{1}{\varphi_I} \times I^\zeta \times \Lambda\left(\frac{\tau}{I}\right).$$

Hence, the equilibrium is defined by the four equations

$$
\begin{aligned}
1 &= \Lambda\left(\frac{\tau}{I}\right) \times \frac{Y_t}{w_t} + \frac{1}{\varphi_I} \times I^{\zeta} \times \Lambda\left(\frac{\tau}{I}\right) + \frac{1}{\varphi_I \phi_z} \times (\tau - x) + \frac{1}{\varphi_x \phi_x^{\zeta}} \times x^{\zeta} && (40)\\
\frac{1}{\phi_z \varphi_z} &= \frac{\frac{\lambda-1}{\lambda}\frac{Y_t}{w_t} + \frac{\zeta-1}{\varphi_x \phi_x^{\zeta}} x^{\zeta} + \frac{\zeta-1}{\varphi_I}\frac{1}{\lambda} I^{\zeta}}{\rho + \tau} && (41)\\
\frac{Y_t}{w_t} &= \frac{\lambda}{\lambda-1}\frac{\zeta}{\varphi_I} I^{\zeta-1}(\rho+\tau) + (\zeta-1)\frac{1}{\varphi_I} I^{\zeta} && (42)\\
\frac{Y_t}{w_t} &= \frac{\zeta}{\varphi_x \phi_x^{\zeta}} x^{\zeta-1}(\rho+\tau)\frac{\lambda}{\lambda-1} - \frac{\lambda}{\lambda-1}\frac{\zeta-1}{\varphi_x \phi_x^{\zeta}} x^{\zeta} - \frac{1}{\lambda-1}\frac{\zeta-1}{\varphi_I} I^{\zeta}. && (43)
\end{aligned}
$$

To solve for the unknowns $\left(\frac{Y}{w}, I, \tau, x\right)$, note first that (43) and (41) imply that

$$
x^{\zeta-1} = \frac{\varphi_x \phi_x^{\zeta}}{\varphi_z \phi_z}\frac{1}{\zeta}. \tag{44}
$$

This determines x in terms of parameters. We can then use (41), (42) and (40) to arrive at two equations in the two unknowns (τ, I)

$$
\begin{aligned}
1 &= \Lambda\left(\frac{\tau}{I}\right) \times \left(\frac{\lambda}{\lambda-1}\frac{\zeta}{\varphi_I} I^{\zeta-1}(\rho+\tau) + \zeta\frac{1}{\varphi_I} I^{\zeta}\right) + \frac{1}{\varphi_z \phi_z} \times \tau - h(\phi). && (45)\\
\frac{1}{\phi_z \varphi_z} &= \frac{\zeta}{\varphi_I} I^{\zeta-1} + \frac{\frac{\zeta-1}{\varphi_I} I^{\zeta}}{\rho+\tau} + \frac{h(\phi)}{\rho+\tau}, && (46)
\end{aligned}
$$

where

$$
h(\phi) = \left(\frac{\zeta-1}{\zeta}\right)\left(\frac{\phi_x}{\phi_z}\right)^{\frac{\zeta}{\zeta-1}}\left(\frac{\varphi_x}{\zeta \varphi_z^{\zeta}}\right)^{\frac{1‘}{\zeta-1}} \geq 0. \tag{47}
$$

Given a solution (I, τ) and x from (44), we can calculate $\frac{Y}{w}$ from (42) and $z = \tau - x$. Hence, we only have to show that (45) and (46) have a unique solution. Rewriting (45) and (46) in terms of $\vartheta_I = s$ yields

$$
\begin{aligned}
1 &= \Lambda(s) \times \left(\frac{\lambda}{\lambda-1}\frac{\zeta}{\varphi_I} I^{\zeta-1}(\rho+sI) + \zeta\frac{1}{\varphi_I} I^{\zeta}\right) + \frac{1}{\varphi_z \phi_z} \times sI - h(\phi) && (48)\\
\frac{1}{\phi_z \varphi_z} &= \frac{\zeta}{\varphi_I} I^{\zeta-1} + \frac{\frac{\zeta-1}{\varphi_I} I^{\zeta}}{\rho+sI} + \frac{h(\phi)}{\rho+sI}, && (49)
\end{aligned}
$$

where $\Lambda'(s) > 0$. Let us write the first equation as

$$
1 = H(I,s) \equiv \Lambda(s) \times \left(\frac{\lambda}{\lambda-1}\frac{\zeta}{\varphi_I} I^{\zeta-1}(\rho+sI) + (\zeta-1)\frac{1}{\varphi_I} I^{\zeta}\right) + \frac{1}{\varphi_I} \times I^{\zeta} \times \Lambda(s) + \frac{1}{\varphi_z \phi_z} \times sI - h(\phi).
$$

Then $\frac{\partial H(I,s)}{\partial s} > 0$ and $\frac{\partial H(I,s)}{\partial I} > 0$. Hence, (48) defines a downward sloping continuous schedule in the (s, I) space, which we call $I^{LM}(s)$. Moreover, $\lim_{s\to\infty} I^{LM}(s) \to 0$ and $\lim_{s\to 0} I^{LM}(s) \to \infty$. Now write the second equation, which stems from the free entry condition as

$$
\frac{1}{\phi_z \varphi_z} = \frac{\zeta}{\varphi_I} I^{\zeta-1} + \frac{1}{\rho+sI}\left(\frac{\zeta-1}{\varphi_I} I^{\zeta} + h(\phi)\right) = G\left(I^{FE}(s), s\right). \tag{50}
$$

Clearly, $\frac{\partial G(I,s)}{\partial s} < 0$. Also

$$\begin{aligned}\frac{\partial G}{\partial I} &= \frac{(\zeta-1)\zeta}{\varphi_I} I^{\zeta-1}\frac{1}{I} - (\rho+sI)^{-2} s\left[(\zeta-1)\frac{1}{\varphi_I}I^{\zeta} + h(\phi)\right] + (\rho+sI)^{-1}(\zeta-1)\zeta\frac{1}{\varphi_I}I^{\zeta-1} \\ &= (\rho+sI)^{-2}\left[\frac{(\zeta-1)\zeta}{\varphi_I} I^{\zeta-1}\frac{(\rho+sI)^2}{I} - s\left[(\zeta-1)\frac{1}{\varphi_I}I^{\zeta} + h(\phi)\right] + (\rho+sI)(\zeta-1)\zeta\frac{1}{\varphi_I}I^{\zeta-1}\right].\end{aligned}$$

Given the definition of h in (47), it can be shown that $\frac{\partial G(I,s)}{\partial I} > 0$. Hence, $\frac{\partial I^{FE}}{\partial s} > 0$. Note that $I^{FE}(s)$ has to be bounded for (50) to be satisfied. Hence, $0 \leq I^{FE}(s) \leq I^{max}$. This also implies that $\lim_{s\to\infty} sI^{FE}(s) = \infty$, so that $\lim_{s\to\infty} I^{FE}(s) \to \left(\frac{\varphi_I}{\phi_z\varphi_z}\frac{1}{\zeta}\right)^{\frac{1}{\zeta-1}} = I^{max}$. Now consider the case of $s \to 0$. As $I^{FE}(s)$ is declining in s it has to be that $sI^{FE}(s) \to 0$. Let $I^{FE}(0)$ be that limit. (50) then implies that $I^{FE}(0)$ is implicitly defined by $\frac{1}{\phi_z\varphi_z} - \frac{h(\phi)}{\rho} = \frac{\zeta}{\varphi_I}I^{FE}(0)^{\zeta-1} + \frac{\frac{\zeta-1}{\varphi_I}I^{FE}(0)^{\zeta}}{\rho}$. As $I^{FE}(0) > 0$, this requires that $\frac{1}{\phi_z\varphi_z} > \frac{h(\phi)}{\rho}$. From (47) we can write this condition as

$$\rho > \phi_z\varphi_z\left(\frac{\zeta-1}{\zeta}\right)\left(\frac{\phi_x}{\phi_z}\right)^{\frac{\zeta}{\zeta-1}}\left(\frac{\varphi_x}{\zeta\varphi_z^{\zeta}}\right)^{\frac{1`}{\zeta-1}} = \left(\frac{\zeta-1}{\zeta}\right)\left(\frac{\phi_x^{\zeta}}{\phi_z}\times\frac{\varphi_x}{\zeta\varphi_z}\right)^{\frac{1}{\zeta-1}}. \tag{51}$$

As long as (51) is satisfied, there is a unique solution (I, s) for the system of equations (48) and (49). Hence, there is a unique $\tau = s \times I$. The optimal expansion rate x is given by (44).

6.2 Proof of Proposition 2

Consider first the distribution of quality gaps $\nu(\Delta, t)$. In a stationary equilibrium we have $\dot{\nu}(\Delta, t) = 0$. (15) then implies that

$$\nu(\Delta) = \left(\frac{I}{\tau+I}\right)^{\Delta}\frac{\tau}{I} = \left(\frac{1}{1+\frac{\tau}{I}}\right)^{\Delta}\frac{\tau}{I} = \left(\frac{1}{1+\vartheta_I}\right)^{\Delta}\vartheta_I$$

Hence, $P(\Delta \leq d) = 1 - \left(\frac{1}{1+\vartheta_I}\right)^{d} = 1 - e^{-ln(1+\vartheta_I)\times d}$. This implies that log markups $ln(\mu) = \Delta ln(\lambda)$ are exponentially distributed with parameter θ.Similarly, $F(\mu; x) = P(\lambda^{\Delta} \leq \mu) = 1 - \mu^{-\theta}$ as required in (16). To derive (17), note that[41]

$$\begin{aligned}\Lambda &= \int \mu^{-1}\theta\mu^{-(\theta+1)}d\mu = \frac{\theta}{\theta+1} \\ M &= exp(-Eln(\mu))\Lambda^{-1} = exp(-\theta^{-1})\frac{\theta+1}{\theta}.\end{aligned}$$

Now consider the distribution of firms across the number of products they produce. Let $\omega(n) \equiv F \times \tilde{\omega}(n)$, where $\tilde{\omega}(n)$ denotes the measure of firms producing $n$products, i.e. $\sum_{n=1}^{\infty}\tilde{\omega}(n) = 1$. We know from Klette and Kortum (2004) that

$$\tilde{\omega}(n) = \frac{\frac{1}{n}\left(\frac{x}{\tau}\right)^{n-1}}{\sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{x}{\tau}\right)^{j-1}}. \tag{53}$$

In a stationary equilibrium, the mass of entering and exiting firms has to be equal so that

[41]Note that these expression are derived taking the distribution of markups as continuous, even though the model implies that they are discrete. Taking this discreteness explicitly into account yields

$$\Lambda^{Dis} = \sum_{i=1}^{\infty}\lambda^{-i}\mu(i) = \frac{\frac{\tau}{I}}{\lambda\left(\frac{\tau}{I}+1\right)}\sum_{i=0}^{\infty}\left(\frac{1}{\lambda\left(\frac{\tau}{I}+1\right)}\right)^{i} = \frac{\frac{\tau}{I}}{\lambda\frac{\tau}{I}+\lambda-1}. \tag{52}$$

Similarly, $\int_0^1 \Delta(\nu)d\nu = \sum_{i=1}^{\infty} i\mu(i) = \frac{\tau}{I}\sum_{i=1}^{\infty} i\left(\frac{1}{1+\frac{\tau}{I}}\right)^{i} = \frac{1+\frac{\tau}{I}}{\frac{\tau}{I}}$, so that $M^{Dis} = \frac{1}{\Lambda\left(\frac{\tau}{I}\right)}\lambda^{-\int_0^1\Delta(\nu)d\nu} = \lambda^{-\frac{1+\frac{\tau}{I}}{\frac{\tau}{I}}}\frac{\lambda-1+\lambda\frac{\tau}{I}}{\frac{\tau}{I}}$.

$$F=\frac{z}{\tau}\times\sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{x}{\tau}\right)^{j-1}=\frac{z}{x}\times\sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{x}{\tau}\right)^{j}=\frac{z}{x}\times ln\left(\frac{z+x}{z}\right)=\frac{1-\vartheta_x}{\vartheta_x}ln\left(\frac{1}{1-\vartheta_x}\right). \tag{54}$$

The share of products produced by firms with at most k products is given by

$$S_k = \sum_{n=1}^{k}F\tilde{\omega}(n)\,n=\left(\frac{z}{\tau}\times\sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{x}{\tau}\right)^{j-1}\right)\times\frac{\sum_{n=1}^{k}\frac{1}{n}\left(\frac{x}{\tau}\right)^{n-1}n}{\sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{x}{\tau}\right)^{n-1}}=\frac{z}{\tau}\times\frac{\tau}{x}\sum_{n=1}^{k}\left(\frac{x}{\tau}\right)^{n}=1-\vartheta_x^k.$$

6.3 Proof of Proposition 3

I prove the different parts in turn.

1. Write the equilibrium conditions in (45) and (46) as

$$1 = \frac{\tau}{\lambda\tau+I(\lambda-1)}\times\left(\frac{\lambda}{\lambda-1}\frac{\zeta}{\varphi_I}I^{\zeta-1}(\rho+\tau)+\zeta\frac{1}{\varphi_I}I^{\zeta}\right)+\frac{1}{\varphi_z\phi_z}\times\tau-h(\phi)=H^{\tau}(I,\tau)-h(\phi) \tag{55}$$

$$\frac{1}{\phi_z\varphi_z} = \frac{\zeta}{\varphi_I}I^{\zeta-1}+\frac{\frac{\zeta-1}{\varphi_I}I^{\zeta}}{\rho+\tau}+\frac{h(\phi)}{\rho+\tau}=G^{\tau}(I,\tau). \tag{56}$$

It is easy to see that $\frac{\partial H^{\tau}}{\partial\tau}>0$, $\frac{\partial G^{\tau}}{\partial\tau}<0$ and $\frac{\partial G^{\tau}}{\partial I}>0$. Also note that

$$\frac{\partial H^{\tau}}{\partial I} = \frac{\tau(\lambda-1)\frac{\zeta}{\varphi_I}I^{\zeta}\left[(\rho+\tau)\frac{\lambda}{(\lambda-1)I}\left(\frac{(\zeta-1)\lambda\tau}{(\lambda-1)I}+\zeta-2\right)+\frac{\zeta\lambda\tau}{(\lambda-1)I}+\zeta-1\right]}{\left[\lambda\tau+I(\lambda-1)\right]^2}.$$

The denominator is positive for $\zeta\geq 2$ and negative for $\zeta=1$. Hence, that is some $\zeta\geq\bar{\zeta}$ such that $\frac{\partial H^{\tau}}{\partial I}>0$. A sufficient condition is $\zeta\geq 2$. Hence, (55) implies a schedule $I^{H}(\tau)$, which decreasing and (56) implies a schedule $I^{G}(\tau)$, which is increasing. Furthermore, note that (47) implies that

$$\frac{\partial h(\phi)}{\partial\phi_x} = \frac{\zeta}{\zeta-1}\frac{h(\phi)}{\phi_x}>0 \tag{57}$$
$$\frac{\partial h(\phi)}{\partial\phi_z} = -\frac{\zeta}{\zeta-1}\frac{h(\phi)}{\phi_z}<0.$$

Hence, an increase in ϕ_x or ϕ_z shifts the $I^{H}(\tau)$ curve up and the $I^{G}(\tau)$ curve down. This shows that higher entry costs and higher market barriers reduce τ.

2. Now consider the effect of ϕ_z and ϕ_x on ϑ_I. Consider again two equations (48) and (49). There I showed that these equations define the increasing schedule I^{FE} (from (49)) and the decreasing schedule I^{LM} (from (48)). The same argument then shows that entry costs and market barriers reduce ϑ_I.

3. From (23) it is immediate that higher entry costs increase x. Hence, higher entry costs increase ϑ_x.

4. Use (47) to write $\tau=\frac{\tau}{x}\times x=\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)$. Then we can write (55) and (56) as

$$
\begin{aligned}
1 &= \frac{\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)\times\left(\frac{\lambda}{\lambda-1}\frac{\zeta}{\varphi_I}I^{\zeta-1}\left(\rho+\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)\right)+\zeta\frac{1}{\varphi_I}I^{\zeta}\right)}{\lambda\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)+I(\lambda-1)}+\left(\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}-1\right)h(\phi)\\
\frac{1}{\phi_z\varphi_z} &= \frac{\zeta}{\varphi_I}I^{\zeta-1}+\frac{\frac{\zeta-1}{\varphi_I}I^{\zeta}}{\rho+\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)}+\frac{h(\phi)}{\rho+\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)}.
\end{aligned}
$$

Note that market barriers ϕ_x only enter through the h function. Again let the first schedule be $H\left(I,\vartheta_x^{-1}\right)$ and the second schedule be $G\left(I,\vartheta_x^{-1}\right)$. It is easy to see that $\frac{\partial G}{\partial I}>0$ and $\frac{\partial G}{\partial\vartheta_x^{-1}}<0$. Hence, $G\left(I,\vartheta_x^{-1}\right)$ implies un upward-sloping schedule $I^G\left(\vartheta_x^{-1}\right)$. It is also clear that $\frac{\partial H}{\partial\vartheta_x^{-1}}>0$ as ϑ^{-1} is proportional to τ. Furthermore, under the same conditions as above, i.e. $\zeta\geq\overline{\zeta}$, we have $\frac{\partial H}{\partial I}>0$. Hence, the H schedule defines a downward sloping locus $I^H\left(\vartheta_x^{-1}\right)$ in the $\left(I,\vartheta_x^{-1}\right)$ space. Now note that $\frac{\partial H}{\partial h}>0$. As $\frac{\partial h}{\partial\phi_x}$, an increase in ϕ_x will shift the $I^H\left(\vartheta_x^{-1}\right)$ schedule down. Similarly,

$$
\frac{\partial G}{\partial h}=\frac{\rho+\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)-\left(\frac{\zeta-1}{\varphi_I}I^{\zeta}+h(\phi)\right)\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z}{\left(\rho+\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)\right)^2}=\frac{\rho-\frac{\zeta-1}{\varphi_I}I^{\zeta}\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z}{\left(\rho+\frac{1}{\vartheta_x}\frac{\zeta}{\zeta-1}\varphi_z\phi_z h(\phi)\right)^2},
$$

For ρ small, this implies that $\frac{\partial G}{\partial h}<0$. Hence, an increase in ϕ_x will shift the $I^G\left(\vartheta_x^{-1}\right)$ up. This implies that an increase in ϕ_x will increase ϑ_x. Conversely, market barriers will reduce ϑ_x

5. The calibrated model in Section 3.3 is an example where the growth consequences are ambiguous.

6.4 The life cycle properties of markups

In this section I derive the life cycle properties of markups.

The distribution of markups as a function of product age: equation (18) I first show that the distribution of quality gaps Δ as a function of age conditional on survival, $\zeta_\Delta(a)$, is given by $\zeta_{\Delta+1}(a)=\frac{1}{\Delta!}(Ia)^{\Delta}e^{-Ia}$. Let $p_\Delta(a)$ denote the probability of the product having a quality gap Δ at age a when it was introduced at time 0. The corresponding flow equation are

$$
\dot{p}_\Delta(a)=\begin{cases}(1-p_0(a))\tau & \text{for }\Delta=0\\ -p_1(a)(I+\tau) & \text{for }\Delta=1\\ p_{\Delta-1}(a)I-p_\Delta(a)(I+\tau) & \text{for }\Delta\geq 2\end{cases}.
$$

The solution to this set of differential equations is given by

$$
\begin{aligned}
p_0(a) &= 1-e^{-\tau\times a}\\
p_{i+1}(a) &= \left(\frac{1}{i!}\right)I^i a^i\left(e^{-(I+\tau)a}\right)\text{ for }i\geq 0.
\end{aligned}
$$

The distribution of markups conditional on survival is then

$$
p_{i+1}^S(a)\equiv\frac{p_{i+1}(a)}{1-p_0(a)}=\left(\frac{1}{i!}\right)I^i a^i\left(e^{-Ia}\right).
$$

This is a Poisson distribution with parameter Ia, so that $E[\Delta|a]=Ia$. (19) then follows because $ln(\mu)=ln(\lambda)\Delta$.

The expected log markup by age: equation (20) From (3) we know that firm-level markups are given by $\mu_f = \frac{py_f}{wl_f} = \frac{1}{\frac{1}{n}\sum_{j=1}^{n_i}\lambda^{-\Delta_j}}$. Hence,

$$ln\left(\mu_f\right) = -ln\left(\frac{1}{n}\sum_{j=1}^{n_i}\lambda^{-\Delta_j}\right) \approx ln\left(\lambda\right) \times \left[\frac{1}{N_f} \times \sum_{j=1}^{N_f}\Delta_j\right]. \tag{58}$$

so that $E\left[ln\left(\mu_f\right)|\text{Age} = \text{a}\right] = log\left(\lambda\right) \times E_n\left[E\left[\frac{1}{n}\sum_{j=1}^{n}\Delta_j|\text{Age} = a, N = n\right]|\text{Age} = a\right]$. Define the random variable $B = \{0, 1, 2, ..., n\}$ by

$$B = \begin{cases} 0 & \text{if none of the } n \text{ products was the initial product of the firm} \\ k & \text{if product } k \text{ was the initial product of the firm} \end{cases}.$$

Then

$$E\left[\frac{1}{n}\sum_{j=1}^{n}\Delta_j|\text{Age} = a, N = n\right] = \sum_{k=0}^{n} E\left[\frac{1}{n}\sum_{j=1}^{n}\Delta_j|\text{Age} = a, N = n, B = k\right] \times P\left(B = k|\text{Age} = a, N = n\right).$$

To simplify notation I will simply denote the conditioning as a, n and k respectively. Then

$$\sum_{k=0}^{n} E\left[\frac{1}{n}\sum_{j=1}^{n}\Delta_j|a, n, k\right] P\left(B = k|a, n\right) = E\left[\Delta_j|a, n, \text{ni}\right] + \frac{1 - P\left(B = 0|a, n\right)}{n}\left(E\left[\Delta_j|a, n, \text{i}\right] - E\left[\Delta_j|a, n, \text{ni}\right]\right) \tag{59}$$

where $E\left[\Delta|a, n, \text{ni}\right]$ denotes the conditional expected value of Δ conditional on the fact that the product is *not* an initial product and $\sum_{k=1}^{n} P\left(B = k|a, n\right) = 1 - P\left(B = 0|a, n\right)$. Now let us solve for $E\left[\Delta|a, n, \text{ni}\right]$, $E\left[\Delta|a, n, \text{i}\right]$ and $P\left(B = 0|a, n\right)$ in turn.

Recovering $E\left[\Delta|a, n, \mathbf{ni}\right]$ and $E\left[\Delta|a, n, \mathbf{i}\right]$ As before let U denote the age of the product so that

$$E\left[\Delta|a, n, \text{ni}\right] = E_u\left\{E\left[\Delta|U = u\right]|a, n, \text{ni}\right\} = E_u\left\{\sum_{i=1}^{\infty} i \times p\left(i, u\right)|a, n, \text{ni}\right\}, \tag{60}$$

where the second line uses the fact that conditional on *product* age, no other characteristic matters and $p\left(i, u\right)$ is the probability of having a quality gap i conditional on the product having an age of u. As shown above this distribution follows a Poisson distribution of the form

$$p\left(i, u\right) = \left(\frac{1}{\left(i - 1\right)!}\right)\left(Iu\right)^{i-1} \times \exp\left(-Iu\right). \tag{61}$$

Hence, $\sum_{i=1}^{\infty} i \times p\left(i, u\right) = Iu + 1$. (60) therefore implies that

$$E\left[\Delta|a, n, \text{ni}\right] = 1 + I \times \int_{u=0}^{a} u \times f_{U|A,\text{ni}}\left(u|a, \text{ni}\right) du, \tag{62}$$

where $f_{U|A,\mathrm{ni}}$ is the density of the conditional age distribution of a product. In Section OA-1.2.1 in the Online Appendix I show that this density is given by

$$f_{U|A,\mathrm{ni}}\left(u|a,\mathrm{ni}\right)=\frac{\tau e^{-\tau\times u}+xe^{-(x+\tau)a}e^{xu}}{1-e^{-(x+\tau)a}}. \tag{63}$$

From (62) and (63) one can show that

$$E\left[\Delta|a,n,\mathrm{ni}\right]=1+I\times\frac{\frac{1}{\tau}\left(1-e^{-\tau a}\right)-\frac{1}{x}e^{-\tau a}\left(1-\exp\left(-xa\right)\right)}{1-\exp\left(-\left(x+\tau\right)a\right)}. \tag{64}$$

Turning to $E\left[\Delta|a,n,\mathrm{i}\right]$, it is clear that the *initial* product of a firm of age a is simply a years old. Hence,

$$E\left[\Delta|a,n,\mathrm{i}\right] \quad = \quad 1+Ia. \tag{65}$$

Solving for $P\left(B=0|a,n\right)$ Note first that $P\left(B=0|a,n\right)=\frac{P(B=0,a,n)}{P(a,n)}$. We are going to construct $P\left(B=0,a,n\right)$. Let us denote this probability by $Q\left(n,t\right)$ where t is the age of the firm. This probability evolves according to the differential equation

$$\dot{Q}\left(n,t\right) \quad = \quad x\left(n-1\right)Q\left(n-1,t\right)+\tau\left(n+1\right)Q\left(n+1,t\right)-n\left(x+\tau\right)Q\left(n,t\right)+\tau\left(p\left(n+1,t\right)-Q\left(n+1,t\right)\right) \tag{66}$$

where $p\left(n,t\right)$ denotes the probability of having n products at time t. Note also that $\dot{Q}\left(0,t\right)=\tau p\left(1,t\right)$. Define the function

$$H_Q\left(z,t\right)\equiv\sum_{n=0}^{\infty}Q\left(n,t\right)z^n. \tag{67}$$

Then $\frac{\partial H_Q(z,t)}{\partial z}=\sum_{n=1}^{\infty}nQ\left(n,t\right)z^{n-1}$ and $\frac{\partial H_Q(z,t)}{\partial t}=\dot{Q}\left(0,t\right)+\sum_{n=1}^{\infty}\dot{Q}\left(n,t\right)z^n$. Using (66) it follows that,

$$\begin{aligned}\frac{\partial H_Q\left(z,t\right)}{\partial t} &= \tau p\left(1,t\right)+\sum_{n=1}^{\infty}\left[x\left(n-1\right)\times Q\left(n-1,t\right)+\tau\left(n+1\right)\times Q\left(n+1,t\right)-n\left(x+\tau\right)Q\left(n,t\right)\right]z^n \\ &\quad +\tau\sum_{n=1}^{\infty}p\left(n+1,t\right)z^n-\tau\sum_{n=1}^{\infty}Q\left(n+1,t\right)z^n \\ &= \frac{\tau}{z}\left(H_P\left(z,t\right)-H_Q\left(z,t\right)\right)+\left(xz^2-\left(x+\tau\right)z+\tau\right)\frac{\partial H_Q\left(z,t\right)}{\partial z},\end{aligned}$$

where, as in (67), we defined

$$H_P\left(z,t\right)=\sum_{n=0}^{\infty}p\left(n,t\right)z^n, \tag{68}$$

Now define

$$\Psi\left(z,t\right)\equiv H_P\left(z,t\right)-H_Q\left(z,t\right). \tag{69}$$

Then

$$\dot{\Psi}\left(z,t\right) \quad = \quad \dot{H}_P\left(z,t\right)-\dot{H}_Q\left(z,t\right)=\left(xz^2-\left(x+\tau\right)z+\tau\right)\frac{\partial\Psi\left(z,t\right)}{\partial z}-\frac{\tau}{z}\Psi\left(z,t\right), \tag{70}$$

where $\dot{H}_P\left(z,t\right)=\left(xz^2-\left(x+\tau\right)z+\tau\right)\frac{\partial H_P(z,t)}{\partial z}$ follows the derivations in Klette and Kortum (2004). To solve for $\Psi\left(z,t\right)$ we need an initial condition. As every firm enters with a single product, we know that $p\left(1,t\right)=1$ and

$p(n,t) = 0$ for $n \neq 1$. Similarly, $Q(n,0) = 0$ for all n. Hence, (67), (68) and (69) imply that

$$\Psi(z,0) = \sum_{n=0}^{\infty} p(n,0) z^n - \sum_{n=0}^{\infty} Q(n,0) z^n = z, \tag{71}$$

which is the required initial condition. The solution to (70) with the initial condition in (71) is given by (see Section OA-1.2.2 in the Online Appendix for the proof)

$$\Psi(z,t) = \frac{(\tau - x) z \times e^{-\tau t}}{x(z-1) \times e^{-(\tau - x)t} - (xz - \tau)}. \tag{72}$$

From Klette and Kortum (2004, p. 1014) we know that $H_P(z,t)$ takes a similar form

$$H_P(z,t) = \frac{\tau(z-1) \times e^{-(\tau-x)t} - (xz - \tau)}{x(z-1) \times e^{-(\tau-x)t} - (xz-\tau)}.$$

(69) and (67) therefore imply that

$$H_Q(z,t) = \Psi(z,t) - H_P(z,t) = \frac{\tau(z-1) \times e^{-(\tau-x)t} - (xz-\tau) - (\tau - x) z \times e^{-\tau t}}{x(z-1) \times e^{-(\tau-x)t} - (xz-\tau)}. \tag{73}$$

From the definition of H_Q in (67) we can recover $Q(n,t)$ as the coefficients of the Taylor approximation around $z = 0$. In Section OA-1.2.3 in the Online-Appendix, I show that

$$Q(n,t) = \left(1 - \frac{\tau e^{-xt} - x e^{-\tau t}}{\tau - x}\right) \times p(n,t), \tag{74}$$

where $p(n,t)$ is described by $p(0,t) = \frac{\tau}{x}\gamma(t)$, $p(1,t) = (1-\gamma(t))(1-p(0,t))$ and $p(n,t) = \gamma(t)^{n-1} p(1,t)$ and the function $\gamma(t)$ is given by

$$\gamma(t) = \frac{x\left(1 - e^{-(\tau-x)t}\right)}{\tau - x \times e^{-(\tau-x)t}}. \tag{75}$$

Equation (74) has the important implication that the *conditional* probability of not having an initial product at time t is independent of n, i.e.

$$P(\text{not initial}|t,n) = \frac{Q(n,t)}{p(n,t)} = 1 - \frac{\tau e^{-xt} - x e^{-\tau t}}{\tau - x}.$$

Hence,

$$1 - P(B = 0|a,n) = \frac{\tau e^{-xa} - x e^{-\tau a}}{\tau - x}. \tag{76}$$

Note that $P(\text{not initial}|0,n) = 0$ and $\lim_{t\to\infty} P(\text{not initial}|t,n) = 1$ as required.

Substituting (64), (65) and (76) into (59) yields

$$E[a_P|a_f] \equiv E_n\left[E\left[\frac{1}{n}\sum_{j=1}^{n}\Delta_j|a,n\right]|a\right] = E[\Delta|a,\text{ni}] + (1 - P(B=0|a)) \times (E[\Delta|a,\text{i}] - E[\Delta|a,\text{ni}]) \times \sum_{n=1}^{\infty}\frac{1}{n} f_{N|A}(n|a),$$

where $f_{N|A}(n|a)$ is the conditional distribution of n conditional on a. This object is given by $f_{N|A}(n|a) = \frac{p(n,a)}{1-p(0,a)} = \gamma(a)^{n-1} \times (1 - \gamma(a))$. Hence,

$$E[ln(\mu)|a] = ln(\lambda)(1 + I \times E[a_P|a_f]),$$

where

$$E\left[a_P|a_f\right] = \frac{\frac{1}{\tau}\left(1-e^{-\tau a}\right)-\frac{1}{x}e^{-\tau a}\left(1-e^{-xa}\right)}{1-e^{-(x+\tau)a}} + \left(a - \frac{\frac{1}{\tau}\left(1-e^{-\tau a}\right)-\frac{1}{x}e^{-\tau a}\left(1-e^{-xa}\right)}{1-e^{-(x+\tau)a}}\right) \tag{77}$$

$$\times\left(\frac{\tau e^{-xa}-xe^{-\tau a}}{x\left(1-e^{-(\tau-x)a}\right)}\right)\times ln\left(\frac{\tau - x\times e^{-(\tau-x)a}}{\tau-x}\right). \tag{78}$$

This is the required expression in (20).

The distribution of size conditional on age: equation (21) Note that $E\left[ln\left(l\right)|a\right] = E\left[ln\left(\frac{nY}{w\mu_f}\right)|a\right] = ln\left(\frac{Y}{w}\right)+E\left[ln\left(n\right)|a\right]-E\left[ln\left(\mu_f\right)|a\right]$. To calculate $E\left[ln\left(n\right)|a\right]$ note that the distribution of n conditional on age is given by $f_{N|A}\left(n|a\right)=\gamma\left(a\right)^{n-1}\times\left(1-\gamma\left(a\right)\right)$. Hence,

$$E\left[ln\left(n\right)|a\right]=\left(\frac{1-\gamma\left(a\right)}{\gamma\left(a\right)}\right)\sum_{n=1}^{\infty}ln\left(n\right)\times\gamma\left(a\right)^{n}, \tag{79}$$

where $\gamma\left(t\right)=\frac{x\left(1-e^{-(\tau-x)t}\right)}{\tau-x\times e^{-(\tau-x)t}}$. It can also be shown that $\frac{\partial E[ln(n)|a]}{\partial\gamma}>0$, that $\frac{\partial\gamma(a)}{\partial\tau}<0$ and that $\frac{\partial\gamma(a)}{\partial x}>0$. Hence, $\frac{\partial E[ln(n)|a]}{\partial x}>0$ and $\frac{\partial E[ln(n)|a]}{\partial\tau}<0$.

6.5 Measuring markups

To measure markups I closely follow the approach of De Loecker and Warzynski (2012). The crucial empirical object to implement (26) is the firms' labor share $s_{l,ft}=\frac{w_t l_{ft}}{p_f y_{ft}}$. As pointed out by De Loecker and Warzynski (2012), the level of production y_{ft} might contain both unanticipated shocks to and measurement error. Hence, they propose to consider a regression of

$$\ln y_{ft}=\phi\left(l_{ft},k_{ft},m_{ft},z_{ft}\right)+\varepsilon_{ft}, \tag{80}$$

where $\phi\left(.\right)$ is estimated flexibly. Given the estimate $\hat{\phi}\left(.\right)$ one can recover an estimate of the measurement error $\hat{\varepsilon}_{ft}$ and form $s_{l,ft}=\frac{w_t l_{ft}}{p_f\frac{y_{ft}}{exp\left(\hat{\varepsilon}_{ft}\right)}}$ (see De Loecker and Warzynski (2012, Equation 16)). Note that this correction is in terms of physical output. As in their application, I only have access to revenue and not physical output and hence I treat deflated sales as a measure of physical quantity. I therefore measure the cost share $s_{l,ft}$ as

$$s_{l,ft}=\frac{wl_{ft}}{va_{ft}/exp\left(\hat{\varepsilon}_{ft}\right)},$$

where va_{ft} is observed value added, $\hat{\varepsilon}_{ft}$ is the residual from (80), where I take va_{ft} is the dependent variable and take $\phi\left(.\right)$ a second order polynomial in all (log) inputs and their interaction terms. For the specification with intermediate inputs instead of labor, the procedure is analogous.

6.6 The employment life cycle in Indonesia

The calibration uses the employment life cycle in Indonesia as an explicit calibration target. Focusing on the unbalanced panel of firms entering the economy after 1991, firms grow by about 8% a year. For the calibration I therefore use the estimated profile depicted in Figure 2 and target the log difference in employment for 7 year old firms. In Table 6 I report additional regression results of predicting firm employment from firm age. The specification is exactly the same as (27), except that I do not control for firms' capital intensity. Column 3 shows

	log employment					
Age	0.0930***	0.0814***		0.0923***	0.0811***	0.0317***
	(0.00193)	(0.00178)		(0.00248)	(0.00192)	(0.00210)
Entry			-0.187***	0.0466***		
			(0.00821)	(0.0101)		
Exit			-0.359***	-0.337***		
			(0.0106)	(0.0105)		
Industry FE	N	Y	Y	Y	Y	N
Firm FE	N	N	N	N	N	Y
N	76106	76106	64958	64958	59602	76106
R^2	0.035	0.218	0.212	0.231	0.239	0.927

Notes: Robust standard errors in parentheses with ***, **, and * respectively denoting significance at the 1%, 5% and 10% levels. I focus on the unbalanced panel of firms, who enter the market after 1990. I use the data from 1991 to 2000. All specifications include year fixed effects. 'Entry" and "Exit" are indicator variables for wether the firm enters (exit) the market in a given year. In column 5 I focus on the balanced panel, i.e. only consider firms that survive to the end of my sample period. The specifications with industry fixed effects control for industry affiliation at the 5 digit level. Column 5 contains firm fixed effects.

Table 6: The employment life cycle in Indonesia

that entrants and exiting firms are substantially smaller than the average firm. Column 4 shows that entrants are not too small given their age (in fact, they are slightly bigger) but that exiting firms are much smaller. This is exactly what the model predicts, because exiting firms are selected on n conditional on age, while entrants are not. Column 5 shows that the estimated age profile is quite similar once I condition on survival until the end of the sample. Finally, the last column controls for firm fixed effects. This lowers the age coefficient substantially.

Online Appendix for "Heterogeneous Markups, Growth and Misallocation"

[NOT FOR PUBLICATION]

OA-1 Theoretical results

OA-1.1 Solving for the value function $V(.)$ defined in (11)

Using (11) and (12) we can write the value function V as

$$\begin{aligned} rV_t(n,[\Delta_i]_{i=1}^n) - \dot{V}_t(n,[\Delta i]_{i=1}^n) &= \sum_{i=1}^n \pi_t(\Delta_i) + \sum_{i=1}^n \tau\left[V_t\left(n-1,[\Delta_i]_{j\neq i}\right) - V(n,[\Delta_i]_{i=1}^n)\right] \\ &+ \max_{[I_i]_{i=1}^n}\left\{\sum_{i=1}^n I_i\left[V_t\left(n,\left\{[\Delta_i]_{j\neq i},\Delta_i+1\right\}\right) - V_t(n,[\Delta_i]_{i=1}^n)\right] - \sum_{i=1}^n c^I(I_i,\Delta_i)w_t\right\} \\ &+ \max_X\left\{X\left[V_t\left(\{n+1,[\Delta_i]_{i=1}^n,1\}\right) - V_t(n,[\Delta_i]_{i=1}^n)\right] - c^X\left(\frac{X}{\phi_e},n\right)w_t\right\}. \end{aligned}$$

Note that (10) implies that $\pi_t(\Delta) = \left(1-\lambda^{-\Delta}\right)Y_t$. Conjecture that the value function take an additive form

$$V_t(n,[\Delta_i]_{i=1}^n) = V_t^P(n) + \sum_{i=1}^n V_t^M(\Delta_i). \tag{OA-1}$$

Then we get

$$\begin{aligned} \dot{V}(n,[\Delta_i]_{i=1}^n) &= \dot{V}_t^P(n) + \sum_{i=1}^n \dot{V}_t^M(\Delta_i) \\ V_t\left(n-1,[\Delta_i]_{j\neq i}\right) - V_t(n,[\Delta_i]_{i=1}^n) &= V_t^P(n-1) - V_t^P(n) - V_t^M(\Delta_i) \\ V_t\left(n,\left\{[\Delta_i]_{j\neq i},\Delta_i+1\right\}\right) - V_t(n,[\Delta_i]_{i=1}^n) &= V_t^M(\Delta_i+1) - V_t^M(\Delta_i) \\ V_t\left(\{n+1,[\Delta_i]_{i=1}^n,1\}\right) - V_t(n,[\Delta_i]_{i=1}^n) &= V_t^P(n+1) + V_t^M(1) - V_t^P(n). \end{aligned}$$

Substituting this into (OA-1) yields

$$\begin{aligned} r\left[V_t^P(n) + \sum_{i=1}^n V_t^M(\Delta_i)\right] - \dot{V}_t^P(n) - \sum_{i=1}^n \dot{V}_t^M(\Delta_i) &= \sum_{i=1}^n \pi_t(\Delta_i) + \sum_{i=1}^n \tau\left[V_t^P(n-1) - V_t^P(n) - V_t^M(\Delta_i)\right] \\ &+ \max_{[I_i]_{i=1}^n}\left\{\sum_{i=1}^n I_i\left[V_t^M(\Delta_i+1) - V_t^M(\Delta_i)\right] - \sum_{i=1}^n c^I(I_i,\Delta_i)w_t\right\} \\ &+ \max_X\left\{X\left[V_t^P(n+1) + V_t^M(1) - V_t^P(n)\right] - c^X\left(\frac{X}{\phi_e},n\right)w_t\right\}. \end{aligned}$$

Writing$\pi_t(\Delta) = \pi_t(\Delta) - \pi_t(1) + \pi_t(1)$ and rearranging terms we get

$$\begin{aligned}
& rV_t^P(n) - \dot{V}_t^P(n) + \sum_{i=1}^{n}\left[rV_t^M(\Delta_i) - \dot{V}_t^M(\Delta_i)\right] \\
= \; & \sum_{i=1}^{n}\left\{(\pi_t(\Delta_i) - \pi_t(1)) - \tau V_t^M(\Delta_i) + \max_{I_i}\left\{I_i\left[V_t^M(\Delta_i+1) - V_t^M(\Delta_i)\right] - c^I(I_i,\Delta_i)\,w_t\right\}\right\} \\
& n \times \pi_t(1) + \sum_{i=1}^{n}\tau\left[V_t^P(n-1) - V_t^P(n) - V_t^M(\Delta_i)\right] \\
& + \max_X\left\{X\left[V_t^P(n+1) + V_t^M(1) - V_t^P(n)\right] - c^X\left(\frac{X}{\phi_e}, n\right)w_t\right\}.
\end{aligned}$$

Hence, we find the value function in (OA-1) separately. First we can solve the differential equation

$$rV_t^M(\Delta) - \dot{V}_t^M(\Delta) = \pi_t(\Delta) - \pi_t(1) - \tau V_t^M(\Delta) + \max_I\left\{I\left[V_t^M(\Delta+1) - V_t^M(\Delta)\right] - c^I(I,\Delta)\,w_t\right\}. \tag{OA-2}$$

From here, $V_t^M(\Delta)$ is known. In particular $V_t^M(1)$ is known. Then we can solve

$$\begin{aligned}
rV_t^P(n) - \dot{V}_t^P(n) \;=\; & n \times \pi_t(1) + \sum_{i=1}^{n}\tau\left[V_t^P(n-1) - V_t^P(n)\right] \\
& + \max_X\left\{X\left[V_t^P(n+1) + V_t^M(1) - V_t^P(n)\right] - c^X\left(\frac{X}{\phi_e}, n\right)w_t\right\}.
\end{aligned} \tag{OA-3}$$

If we have a solution to (OA-2) and (OA-3) we can construct $V_t\left(n, [\Delta_i]_{i=1}^n\right)$ from (OA-1). Now consider a steady state, assume the cost functions in (12) and suppose that

$$\begin{aligned}
\dot{V}_t^M(\Delta) &= gV_t^M(\Delta) \\
\dot{V}_t^P(\Delta) &= gV_t^P(\Delta).
\end{aligned}$$

Solving (OA-2) Under these restrictions we can write (OA-2) as

$$(r+\tau-g)\,V_t^M(\Delta) = \pi_t(\Delta) - \pi_t(1) + \max_I\left\{I\left[V_t^M(\Delta+1) - V_t^M(\Delta)\right] - \frac{1}{\varphi_I}\lambda^{-\Delta}I^{\zeta}w_t\right\}. \tag{OA-4}$$

Conjecture that

$$V_t^M(\Delta) = \kappa_t - \alpha_t\lambda^{-\Delta}. \tag{OA-5}$$

Then

$$V_t^M(\Delta+1) - V_t^M(\Delta) = -\alpha_t\lambda^{-\Delta}\frac{1}{\lambda} + \alpha_t\lambda^{-\Delta} = \frac{\lambda-1}{\lambda}\alpha_t\lambda^{-\Delta}.$$

This implies that optimal innovation Isolves

$$\frac{\lambda-1}{\lambda}\alpha_t\lambda^{-\Delta} = \frac{\zeta}{\varphi_I}\lambda^{-\Delta}I_t(\Delta)^{\zeta-1}w_t. \tag{OA-6}$$

Hence,

$$I_t{}^{\zeta-1} = \left(\frac{\lambda-1}{\lambda}\frac{\varphi_I}{\zeta}\frac{\alpha_t}{w_t}\right)^{\frac{1}{\zeta-1}}. \tag{OA-7}$$

Suppose that $\frac{\alpha_t}{w_t} = \text{const}$ (which I will verify below). (OA-7) then implies that $I_t(\Delta) = I$. Using (OA-7), (OA-5), the Euler equation $\rho = r - g$ and

$$\pi_t(\Delta) - \pi_t(1) = \left(\frac{1}{\lambda} - \lambda^{-\Delta}\right) Y_t$$

in (OA-4) yields

$$(\rho + \tau)\left[\kappa_t - \alpha_t \lambda^{-\Delta}\right] = \left(\frac{1}{\lambda} - \lambda^{-\Delta}\right) Y_t + \frac{\zeta - 1}{\varphi_I} \lambda^{-\Delta} I^{\zeta} w_t.$$

Hence, we need that κ_t and α_t satisfy

$$\begin{aligned} \kappa_t &= \frac{\frac{1}{\lambda} Y_t}{\rho + \tau} \\ \alpha_t &= \frac{Y_t - \frac{\zeta - 1}{\varphi_I} I^{\zeta} w_t}{\rho + \tau}. \end{aligned} \tag{OA-8}$$

Note that w_t and Y_t grow at rate g so that $\frac{\alpha_t}{w_t}$ is constant as required. Substituting in (OA-5) yields

$$V_t^M(\Delta) = \frac{\left(\frac{1}{\lambda} - \lambda^{-\Delta}\right) Y_t + \lambda^{-\Delta} \frac{\zeta - 1}{\varphi_I} I^{\zeta} w_t}{\rho + \tau} = \frac{\pi_t(\Delta) - \pi_t(1) + (\zeta - 1) c_I(I, \Delta) w_t}{\rho + \tau}.$$

From (OA-6) and (OA-8) we also know that the optimal innovation rate I solves the equation

$$\frac{\zeta}{\varphi_I} I^{\zeta - 1} (\rho + \tau) \frac{\lambda}{\lambda - 1} + (\zeta - 1) \frac{1}{\varphi_I} I^{\gamma} = \frac{Y_t}{w_t}. \tag{OA-9}$$

Note also that this implies that

$$V_t^M(1) = \frac{(\zeta - 1) w_t \frac{1}{\lambda} \frac{1}{\varphi_I} I^{\zeta}}{\rho + \tau}. \tag{OA-10}$$

Now turn to $V_t^P(n)$. Using (12), defining $X = xn$ and conjecturing

$$V_t^P(n) = n \times v_t, \tag{OA-11}$$

where v_t grow at rate g, (OA-3) reduces to

$$(\rho + \tau) V_t^P(n) \quad = \quad n \times \pi_t(1) + n \times \max_x \left\{ x\left[v_t + V_t^M(1)\right] - \frac{1}{\varphi_x}\left(\frac{x}{\phi_x}\right)^{\zeta} w_t \right\},$$

where we again used that $r - g = \rho$. Hence,

$$(\rho + \tau) v_t = \pi_t(1) + \max_x \left\{ x\left[v_t + V_t^M(1)\right] - \frac{1}{\varphi_x}\left(\frac{x}{\phi_x}\right)^{\zeta} w_t \right\}$$

The optimality condition for x reads

$$v_t + V_t^M(1) = \frac{\zeta}{\varphi_x \phi_e^{\zeta}} x^{\zeta - 1} w_t. \tag{OA-12}$$

As v_t and $V_t^M(\Delta)$ both grow at rate g, this implies that x is indeed constant. In particular, given x, v_t is given by

$$(\rho + \tau) v_t = \pi_t(1) + (\zeta - 1) \frac{1}{\varphi_x \phi_e^{\zeta}} x^{\zeta} w_t. \tag{OA-13}$$

To solve for v_t, let $v_t = \overline{v} \times w_t$. The unknowns $(x, \overline{v})$ are then determined from (OA-12) ,(OA-13) and (OA-10) as

$$\frac{\zeta}{\varphi_x \phi_e^{\zeta}} x^{\zeta-1} = \overline{v} + \frac{V_t^M(1)}{w_t} = \overline{v} + \frac{(\zeta-1)\frac{1}{\lambda}\frac{1}{\varphi_I} I^{\zeta}}{\rho+\tau} \tag{OA-14}$$

$$\begin{aligned}\overline{v} &= \frac{1}{\rho+\tau}\left(\frac{\pi_t(1)}{w_t} + \frac{\zeta-1}{\varphi_x \phi_e^{\zeta}} x^{\zeta}\right) \\ &= \frac{\frac{\lambda-1}{\lambda}\frac{Y_t}{w_t} + \frac{\zeta-1}{\varphi_x \phi_e^{\zeta}} x^{\zeta}}{\rho+\tau}.\end{aligned} \tag{OA-15}$$

Note that this implies that $\overline{v} > 0$. Note also that (OA-12) implies that x is increasing in $V_t^M(1)/w_t$ holding $\overline{v}$ fixed - this reflects the option value of entering in a new product and then building up markups. The final value function $V_t^P(n)$ is given by

$$\begin{aligned}V_t^P(n) &= n \times \frac{\frac{\lambda-1}{\lambda} Y_t + \frac{\zeta-1}{\varphi_x \phi_e^{\zeta}} x^{\zeta} w_t}{\rho+\tau} \\ &= n \times \frac{\pi_t(1) + (\zeta-1) c^X(1,x) w_t}{\rho+\tau}.\end{aligned} \tag{OA-16}$$

Th overall value of the firm is from (OA-1) given by

$$\begin{aligned}V_t\left(n, [\Delta_i]_{i=1}^n\right) &= n \times \frac{\pi_t(1) + (\zeta-1) c^X(1,x) w_t}{\rho+\tau} + \sum_{i=1}^{n}\left(\frac{\pi_t(\Delta) - \pi_t(1) + (\zeta-1) c_I(I,\Delta_i) w_t}{\rho+\tau}\right) \\ &= \frac{\sum_{i=1}^n \pi_t(\Delta) + n(\zeta-1) c^X(1,x) w_t + \sum_{i=1}^n (\zeta-1) c_I(I,\Delta_i) w_t}{\rho+\tau},\end{aligned} \tag{OA-17}$$

where x and I are the optimal innovation and expansion rates.

OA-1.2 Details for the derivation of $E[ln(\mu)|\textbf{Age}]$

OA-1.2.1 The distribution of product age conditional on firm age

Consider a firm of age a. Let $Q(t)$ be the probability that the firm owns the product at time t if it was born at time 0 and the production was not the initial product. Then first note that

$$P[U \leq u | \text{Age} = a] = \frac{\int_0^u \exp(-\tau j)\, x\,(1 - Q(a-j))\, dj}{Q(a)} \text{ for } u < a, \tag{OA-18}$$

because the probability of the product being j years old is simply the probability that the product is "born" at time $a-j$, which happens with probability $x(1-Q(a-j))$, times the probability that the product survives for j years. The division by $Q(a)$ is necessary, to arrive at the conditional probability.

To evaluate (OA-18), we need the function $Q(t)$. Now, $Q(t)$ solves the differential equation

$$\dot{Q}(t) = (1 - Q(t))\, x - Q(t)\, \tau,$$

as τ is the flow rate of losing a product and x is the flow rate of "winning" a product. Given an initial condition

$Q(0)$, this differential equation has the solution

$$Q(t) = \frac{x}{x+\tau} + \exp(-(x+\tau)t)\left(Q(0) - \frac{x}{x+\tau}\right).$$

Because we are interested in the case, where the firm did not own the product initially, we have $Q(0) = 0$, so that

$$Q(t) = \frac{x}{x+\tau}(1 - \exp(-(x+\tau)t)).$$

Using (OA-18), we get

$$F_{U|A}(u;a) = \frac{1 - \exp(-\tau u) - \exp(-(\tau+x)a)[1 - \exp(xu)]}{1 - \exp(-(x+\tau)a)}.$$

Hence, the conditional density is

$$f_{U|A}(u;a) \quad = \quad \frac{\partial F_{U|A}(u;a)}{\partial u} = \frac{\tau e^{-\tau \times u} + x e^{-(x+\tau)a} e^{xu}}{1 - e^{-(x+\tau)a}}.$$

This is the expression in (63).

OA-1.2.2 Solving for $\Psi(z,t)$ in (70)

We have to solve for $\Psi(z,t)$ described by

$$\dot{\Psi}(z,t) \quad = \quad \left(xz^2 - (x+\tau)z + \tau\right)\frac{\partial \Psi(z,t)}{\partial z} - \frac{\tau}{z}\Psi(z,t), \tag{OA-19}$$

with the initial condition $\Psi(z,0) = z$. The general solution to (OA-19) is given by

$$\Psi(z,t) = z \times \left(\frac{(\tau - xz)^x}{(1-z)^\tau}\right)^{\frac{1}{\tau - x}} \times c\left(\frac{z-1}{xz-\tau} \times e^{-(\tau-x)t}\right), \tag{OA-20}$$

where c is a differentiable function. Letting

$$y(z,t) = \frac{z-1}{xz-\tau} \times e^{-(\tau-x)t}, \tag{OA-21}$$

the initial condition implies that

$$\Psi(z,0) = z \times \left(\frac{(\tau - xz)^x}{(1-z)^\tau}\right)^{\frac{1}{\tau-x}} \times c(y(z,0)) = z.$$

From (OA-21), this implies that $c(.)$ has to solve

$$\left(\frac{(\tau - xz(y(z,0)))^x}{(1 - z(y(z,0)))^\tau}\right)^{\frac{1}{\tau-x}} \times c(y(z,0)) = 1, \tag{OA-22}$$

where $z(y(z,0))$ is defined by

$$y(z,0) = \frac{z(y(z,0)) - 1}{xz(y(z,0)) - \tau}.$$

Hence,

$$z\left(y\left(z,0\right)\right)=\frac{\tau y\left(z,0\right)-1}{xy\left(z,0\right)-1}.$$

This implies that

$$\begin{aligned}\tau-xz\left(y\left(x,0\right)\right) &= \tau-x\frac{\tau y\left(x,0\right)-1}{xy\left(x,0\right)-1}=\frac{x-\tau}{xy\left(x,0\right)-1}\\ 1-z\left(y\left(x,0\right)\right) &= 1-\frac{\tau y\left(x,0\right)-1}{xy\left(x,0\right)-1}=\frac{x-\tau}{xy\left(x,0\right)-1}\times y\left(x,0\right).\end{aligned}$$

Hence, (OA-22) reads

$$\left(\frac{x-\tau}{xy\left(x,0\right)-1}\right)^{-1}y\left(x,0\right)^{-\frac{\tau}{\tau-x}}\times c\left(y\left(x,0\right)\right)=1,$$

which implies that

$$c\left(y\left(x,0\right)\right)=\frac{x-\tau}{xy\left(x,0\right)-1}y\left(x,0\right)^{\frac{\tau}{\tau-x}}.$$

Substituting back into (OA-20) yields

$$\begin{aligned}\Psi\left(z,t\right) &= z\times\left(\frac{\left(\tau-xz\right)^{x}}{\left(1-z\right)^{\tau}}\right)^{\frac{1}{\tau-x}}\times\frac{x-\tau}{xy\left(x,t\right)-1}\times y\left(x,t\right)^{\frac{\tau}{\tau-x}}\\ &= z\times\left(\frac{\left(\tau-xz\right)^{x}}{\left(1-z\right)^{\tau}}\right)^{\frac{1}{\tau-x}}\times\frac{x-\tau}{x\frac{z-1}{xz-\tau}\times e^{-\left(\tau-x\right)t}-1}\times\left(\frac{z-1}{xz-\tau}\right)^{\frac{\tau}{\tau-x}}\times e^{-\tau t}\\ &= z\times\left(\tau-xz\right)^{-1}\times\frac{x-\tau}{x\frac{z-1}{xz-\tau}\times e^{-\left(\tau-x\right)t}-1}\times e^{-\tau t}\\ &= \frac{z\left(\tau-x\right)}{x\left(z-1\right)\times e^{-\left(\tau-x\right)t}-\left(xz-\tau\right)}\times e^{-\tau t}.\end{aligned}$$

This is the required solution in (72).

OA-1.2.3 Recovering $Q\left(n,t\right)$ from (73)

We recover $Q\left(n,t\right)$ from

$$H_{Q}\left(z,t\right)=\sum_{n=0}^{\infty}Q\left(n,t\right)z^{n}=\frac{\tau\left(z-1\right)\times e^{-\left(\tau-x\right)t}-\left(xz-\tau\right)-\left(\tau-x\right)z\times e^{-\tau t}}{x\left(z-1\right)\times e^{-\left(\tau-x\right)t}-\left(xz-\tau\right)}.$$

In particular, for $n>0$ we get

$$Q\left(n,t\right)=\frac{1}{n!}\frac{\partial^{n}}{\partial z^{n}}\left.H_{Q}\left(z,t\right)\right|_{z=0}. \tag{OA-23}$$

For the state $n=0$ we of course have

$$Q\left(0,t\right)=p\left(0,t\right).$$

Define the two function

$$\begin{aligned}g\left(z\right) &= \tau\left(z-1\right)\times e^{-\left(\tau-x\right)t}-\left(xz-\tau\right)-\left(\tau-x\right)z\times e^{-\tau t}\\ h\left(z\right) &= x\left(z-1\right)\times e^{-\left(\tau-x\right)t}-\left(xz-\tau\right).\end{aligned}$$

Note that

$$\begin{aligned} g'(z) &= \tau \times e^{-(\tau-x)t} - x - (\tau - x) e^{-\tau t} \\ h'(z) &= x \times e^{-(\tau-x)t} - x, \end{aligned} \tag{OA-24}$$

so that

$$\begin{aligned} g'(z) - h'(z) &= (\tau - x) \times e^{-(\tau-x)t} - (\tau - x) e^{-\tau t} \\ &= (\tau - x)\left(e^{-(\tau-x)t} - e^{-\tau t}\right) \\ &= (\tau - x) e^{-(\tau-x)t}\left(1 - e^{-xt}\right). \end{aligned}$$

Also

$$\begin{aligned} h(z) - g(z) &= (x - \tau)(z - 1) \times e^{-(\tau-x)t} + (\tau - x) z \times e^{-\tau t} \\ &= (\tau - x) e^{-(\tau-x)t}\left[-(z-1) + z \times e^{-xt}\right]. \end{aligned}$$

Then

$$\begin{aligned} \frac{\partial}{\partial z} H_Q(z,t) &= \frac{g'(z) h(z) - g(z) h'(z)}{h(z)^2} \\ &= \frac{(\tau - x) e^{-(\tau-x)t}\left[1 - e^{-xt}\right] h(z) + h'(z)(h(z) - g(z))}{h(z)^2} \\ &= \frac{(\tau - x) e^{-(\tau-x)t}\left[1 - e^{-xt}\right] h(z) + h'(z)(\tau - x) e^{-(\tau-x)t}\left[-(z-1) + z \times e^{-xt}\right]}{h(z)^2} \\ &= (\tau - x) e^{-(\tau-x)t}\frac{\left[1 - e^{-xt}\right] h(z) - h'(z)\left[(z-1) - z \times e^{-xt}\right]}{h(z)^2}. \end{aligned}$$

Now note that

$$\left[1 - e^{-xt}\right] h(z) - h'(z)\left[(z-1) - z \times e^{-xt}\right] = \tau\left(1 - e^{-xt}\right) - x\left(1 - e^{-\tau t}\right).$$

Hence,

$$\frac{\partial}{\partial z} H_Q(z,t) = (\tau - x) e^{-(\tau-x)t} \times \left[\tau\left(1 - e^{-xt}\right) - x\left(1 - e^{-\tau t}\right)\right] \times h(z)^{-2}.$$

Importantly, the first term does not depend on z, so we can write

$$\frac{\partial}{\partial z} H_Q(z,t) = m(t) \times h(z)^{-2}.$$

The higher derivatives are given by

$$\frac{\partial^2}{\partial z^2} H_Q(z,t) = m(t) \times \left[-2h(z)^{-3} h'(z)\right] = 2m(t) h_z(t) h(z)^{-3},$$

where

$$h_z(t) \equiv -h'(z) = x - x \times e^{-(\tau-x)t} = x\left(1 - e^{-(\tau-x)t}\right) > 0,$$

which does not depends on z (see (OA-24)). Similarly,

$$\frac{\partial^3 H_Q(z,t)}{\partial z^3} = 6m(t)[h_z(t)]^2 h(z)^{-4}.$$

Hence, the general solution is

$$\frac{\partial^n H_Q(z,t)}{\partial z^n} = (n!)\, m(t)\, h_z(t)^{n-1} h(z)^{-(n+1)}.$$

Evaluation this expression around $z=0$ yields

$$\begin{aligned}
\left.\frac{\partial^n H_Q(z,t)}{\partial z^n}\right|_{z=0} &= (n!)\, m(t)\, h_z(t)^{n-1} \frac{1}{h(0)^{(n+1)}} \\
&= (n!)\, m(t) \frac{x^{n-1}\left(1-e^{-(\tau-x)t}\right)^{n-1}}{\left(\tau - x \times e^{-(\tau-x)t}\right)^{n+1}} \\
&= (n!)\, m(t) \frac{1}{x\left(1-e^{-(\tau-x)t}\right)} \frac{1}{\left(\tau - x\times e^{-(\tau-x)t}\right)} \times \left[\frac{x\left(1-e^{-(\tau-x)t}\right)}{\left(\tau - x\times e^{-(\tau-x)t}\right)}\right]^n \\
&= (n!) \frac{(\tau-x)e^{-(\tau-x)t} \times \left[\tau\left(1-e^{-xt}\right) - x\left(1-e^{-\tau t}\right)\right]}{x\left(1-e^{-(\tau-x)t}\right)\left(\tau - x\times e^{-(\tau-x)t}\right)} \times \left[\frac{x\left(1-e^{-(\tau-x)t}\right)}{\left(\tau - x\times e^{-(\tau-x)t}\right)}\right]^n.
\end{aligned}$$

Using (OA-23). we get that

$$Q(n,t) = \frac{(\tau-x)e^{-(\tau-x)t} \times \left[\tau\left(1-e^{-xt}\right) - x\left(1-e^{-\tau t}\right)\right]}{x\left(1-e^{-(\tau-x)t}\right)\left(\tau - x\times e^{-(\tau-x)t}\right)} \times \left[\frac{x\left(1-e^{-(\tau-x)t}\right)}{\left(\tau - x\times e^{-(\tau-x)t}\right)}\right]^n.$$

This implies that for $n>1$ we can write $Q(n,t)$ recursively as

$$\begin{aligned}
Q(n,t) &= \frac{(\tau-x)e^{-(\tau-x)t} \times \left[\tau\left(1-e^{-xt}\right) - x\left(1-e^{-\tau t}\right)\right]}{x\left(1-e^{-(\tau-x)t}\right)\left(\tau - x\times e^{-(\tau-x)t}\right)} \times \left[\frac{x\left(1-e^{-(\tau-x)t}\right)}{\left(\tau - x\times e^{-(\tau-x)t}\right)}\right]^{n-1} \frac{x\left(1-e^{-(\tau-x)t}\right)}{\left(\tau - x\times e^{-(\tau-x)t}\right)}. \\
&= \gamma(t) \times Q(n-1,t),
\end{aligned}$$

where

$$\gamma(t) = \frac{x\left(1-e^{-(\tau-x)t}\right)}{\tau - x\times e^{-(\tau-x)t}}.$$

Note also that[42]

$$Q(0,t) = p(0,t) = \frac{\tau\left(1-e^{-(\tau-x)t}\right)}{\tau - x\times e^{-(\tau-x)t}} = \frac{\tau}{x}\gamma(t).$$

Finally,

$$Q(1,t) = \frac{(\tau-x)e^{-(\tau-x)t} \times \left[\tau\left(1-e^{-xt}\right) - x\left(1-e^{-\tau t}\right)\right]}{\left(\tau - x\times e^{-(\tau-x)t}\right)^2}.$$

Note that

$$1-\gamma(t) = 1 - \frac{x\left(1-e^{-(\tau-x)t}\right)}{\tau - x\times e^{-(\tau-x)t}} = \frac{\tau - x}{\tau - x\times e^{-(\tau-x)t}}$$

[42] Another way to see that is

$$Q(0,t) = H_Q(0,t) = \frac{\tau - \tau\times e^{-(\tau-x)t}}{\tau - x\times e^{-(\tau-x)t}} = \frac{\tau}{x}\gamma(t).$$

Note also that

$$1 - p(0,t) = 1 - \frac{\tau\left(1 - e^{-(\tau - x)t}\right)}{\tau - x \times e^{-(\tau - x)t}} = \frac{(\tau - x)\, e^{-(\tau - x)t}}{\tau - x \times e^{-(\tau - x)t}}.$$

Hence,

$$(1 - \gamma(t))(1 - p(0,t)) = e^{-(\tau - x)t} \times \frac{(\tau - x)^2}{\left(\tau - x \times e^{-(\tau - x)t}\right)^2}$$

so that

$$\begin{aligned}
\frac{Q(1,t)}{(1 - \gamma(t))(1 - p(0,t))} &= \frac{\frac{(\tau - x)e^{-(\tau - x)t} \times \left[\tau\left(1 - e^{-xt}\right) - x\left(1 - e^{-\tau t}\right)\right]}{\left(\tau - x \times e^{-(\tau - x)t}\right)^2}}{e^{-(\tau - x)t} \times \frac{(\tau - x)^2}{\left(\tau - x \times e^{-(\tau - x)t}\right)^2}} \\
&= \frac{\tau\left(1 - e^{-xt}\right) - x\left(1 - e^{-\tau t}\right)}{\tau - x} \\
&= 1 - \frac{\tau e^{-xt} - x e^{-\tau t}}{\tau - x}.
\end{aligned}$$

Hence, we get that

$$\begin{aligned}
Q(0,t) &= p(0,t) \\
Q(1,t) &= \left(1 - \frac{\tau e^{-xt} - x e^{-\tau t}}{\tau - x}\right)\left((1 - \gamma(t))(1 - p(0,t))\right) \\
&= \left(1 - \frac{\tau e^{-xt} - x e^{-\tau t}}{\tau - x}\right) p(1,t) \\
Q(n,t) &= \gamma(t)\, Q(n-1,t).
\end{aligned}$$

But we also have that

$$p(n,t) = \gamma(t)\, p(n-1,t).$$

Hence, we can write this (for $n > 1$)

$$Q(n,t) = \left(1 - \frac{\tau e^{-xt} - x e^{-\tau t}}{\tau - x}\right) \times p(n,t).$$

This has the important implication that conditional probability of not having an initial product is independent of n, because

$$P(\text{not initial}|t,n) = \frac{Q(n,t)}{p(n,t)} = 1 - \frac{\tau e^{-xt} - x e^{-\tau t}}{\tau - x}.$$

OA-2 Empirical results

OA-2.1 Descriptive Statistics

Tables OA-1 and OA-2 contain additional descriptive statistics for the cross-sectional patters of plant employment for Indonesia and the US. Table OA-1 contains moments of the firm size distribution and data for the importance of entering and exiting plants from the Indonesian data. It is interesting to see that entry rates drop substantially during the time of the financial crisis after 1997. Exit rates are somewhat higher initially in 1997 but are if anything lower in 1998 and 1999. In Table OA-2 I report data for the US manufacturing sector in 1995, i.e. for the same time period as the Indonesian data. The table shows the well-known skew for the US plant size distribution, whereby

	1991	1992	1993	1994	1995	1996	1997	1998	1999
					Firm size distribution				
Mean	133	143	152	156	143	138	135	140	143
25%	27	27	28	28	27	26	26	26	26
50%	45	45	48	50	45	43	43	43	44
90%	323	349	382	387	355	340	336.5	352	365
					Entering firms				
Entry rate	14.2%	11.6%	8.6%	9.1%	13.5%	10.9%	5.9%	5.6%	4.7%
Empl. share	8.6%	6.2%	4.3%	4.0%	5.2%	4.6%	3.2%	2.8%	2.0%
					Exiting firms				
Exit rate	8.5%	7.5%	6.6%	5.6%	8.2%	9.8%	10.5%	4.5%	5.1%
Empl. share	4.7%	5.0%	3.6%	3.3%	4.4%	5.3%	4.6%	1.5%	2.7%

Source: Statistik Industri data.

Table OA-1: The manufacturing sector in Indonesia

	Plant size								
	1-4	5-9	10-19	20-49	50-99	100-249	250-499	500-999	1000+
Total employment	243,189	416,887	756,241	1,688,650	1,852,659	3,170,629	2,366,066	1,862,374	2,748,872
Share of employment	1.6%	2.8%	5.0%	11.2%	12.3%	21.0%	15.7%	12.3%	18.2%
Number of establishments	102,150	61,420	54,036	52,817	25,748	20,215	6,849	2,754	1,335
Share of establishments	31.2%	18.8%	16.5%	16.1%	7.9%	6.2%	2.1%	0.8%	0.4%

Source: Business Dynamics Statistics.

Table OA-2: The US manufacturing sector in 1995

3.5% of establishments with more than 250 employees account for almost half of manufacturing employment.

OA-2.2 The cross-sectional age-size relationship

In the empirical analysis of this paper I exclusively used the panel dimension of the data to estimate the life cycle patterns for markups and employment. In the absence of panel data, other papers have focused on the cross-sectional age-size relationship - see for example Hsieh and Klenow (2014) or Akcigit et al. (2015). For comparison, I also include the cross-sectional age-size profile for various years of the Indonesian data in Figure 5 below. The relationship is much flatter and falls in between the cases of India and Mexico analyzed in Hsieh and Klenow (2014). This suggests that cohort effects upon entry are quantitatively important. Note that Hsieh and Klenow (2014) also find that the life cycle imputed by synthetic cohort is somewhat larger in both the US and India (though not for the case of Mexico).

OA-2.3 Robustness for Table 2 and additional results

In Table OA-3 I report the robustness of the estimated life cycle in Table 2 in various specifications. In particular I consider (i) the share of labor in sales and (ii) the share of intermediate inputs in sales as alternative measures. I also report the results without the above correction for measurement error. Finally, in the specification for firms' material share I consider the specification

$$ln\left(\mu_{ft}\right) = \delta_t + \delta_s + \beta \times age_{ft} + \alpha \times ln\left(k_{ft}/l_{ft}\right) + \psi \times ln\left(m_{ft}/l_{ft}\right) + x'_{ft}\gamma + u_{ft}, \qquad \text{(OA-25)}$$

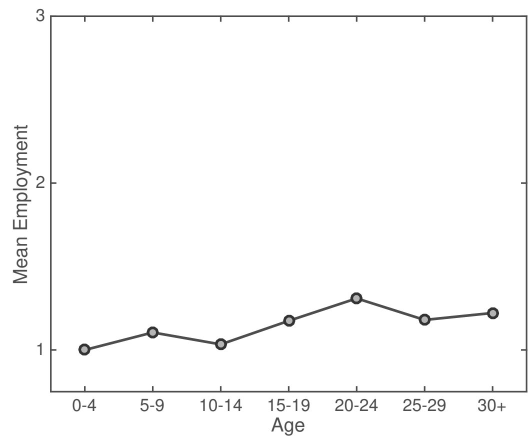
where m/l is the observed material-labor ratio. As seen in Table OA-3, this is important as age is strongly correlated with m/l.

In Table OA-4 I report additional cross-sectional correlates of markups. I consider regressions akin to (27) (for

Measure for μ: Labor share in value added, uncorrected					
Age	0.00901***	0.00454***	0.00432**	0.00451***	0.0168***
	(0.00130)	(0.00144)	(0.00198)	(0.00157)	(0.00230)
Entry			-0.0180**		
			(0.00872)		
Exit			-0.0541***		
			(0.00937)		
N	76076	57281	50275	44024	57281
Measure for μ: Material share in sales, corrected					
Age	0.00715***	0.00567***	0.00729***	0.00531***	0.0101***
	(0.000411)	(0.000313)	(0.000430)	(0.000335)	(0.000457)
Entry			0.00727***		
			(0.00192)		
Exit			-0.0283***		
			(0.00221)		
N	55230	55230	48574	42442	55230
Measure for μ: Material share in sales, uncorrected					
Age	0.00196***	0.00120	0.00166	0.000781	0.00937***
	(0.000718)	(0.000787)	(0.00107)	(0.000850)	(0.00119)
Entry			-0.00235		
			(0.00495)		
Exit			-0.0481***		
			(0.00554)		
N	72227	55230	48574	42442	55230
Measure for μ: Labor share in value added, corrected, controlling for $ln(m/l)$					
Age	0.00690***	0.00596***	0.00829***	0.00614***	0.00868***
	(0.00109)	(0.00108)	(0.00145)	(0.00116)	(0.00170)
Entry			0.00775		
			(0.00694)		
Exit			0.0220***		
			(0.00812)		
N	55212	55212	48556	42434	55212
Measure for μ: Material share in sales, corrected, not controlling for $ln(m/l)$					
Age	-0.00120	-0.00146*	-0.000884	-0.00171**	-0.00411***
	(0.000779)	(0.000779)	(0.00106)	(0.000833)	(0.00119)
Entry			0.00739		
			(0.00503)		
Exit			0.0307***		
			(0.00583)		
N	55230	55230	48574	42442	55230
Measure for μ: Material share in sales, uncorrected, not controlling for $ln(m/l)$					
Age	-0.00569***	-0.00568***	-0.00624***	-0.00611***	-0.00464***
	(0.000897)	(0.00104)	(0.00142)	(0.00112)	(0.00157)
Entry			-0.00223		
			(0.00670)		
Exit			0.00901		
			(0.00737)		
N	72227	55230	48574	42442	55230
Industry FE	Y	Y	Y	Y	-
Firm FE	N	N	N	N	Y
$ln(k/l)$	N	Y	Y	Y	Y

Notes: Robust standard errors in parentheses with ***, **, and * respectively denoting significance at the 1%, 5% and 10% levels. I focus on the unbalanced sample of firms, who enter the market after 1990. I use the data from 1991 to 2000. All specifications include year fixed effects. $ln\,(k/l)$ denotes the (log) capital-labor ratio at the firm level. 'Entry" and "Exit" are indicator variables for wether the firm enters (exit) the market in a given year. In column 6 of the top panel and column 5 of the bottom panel I focus on the balanced panel, i.e. only consider firms that survive to the end of my sample period. The last column of the bottom panel considers log sales as a measure of size. The specifications with industry fixed effects control for industry affiliation at the 5 digit level.

Table OA-3: The Life-cycle of markups (Robustness)



Notes: The figure displays the cross-sectional age-size relationship for different years in the sample. Specifically, the figure displays average employment in each age-bin relative to the youngest age bin, i.e. firms of age less than 5 years old. I use the entire cross-sectional data and simply average employment by age for each year of the available data.

Figure 5: The cross-sectional age-size relationship

the case of measuring markups by the inverse labor share) and (OA-25) (for the case of using the share of material spending) and use a variety of characteristics other than age. To save space, Table OA-4 reports (for all the different specifications of interest) the coefficients on the respective characteristic. Each columns corresponds to a separate regression. The first four columns correspond to cross-sectional estimates for the entire sample from 1991 to 1997 (i.e. before the crisis). In particular I confirm the result of De Loecker and Warzynski (2012) that exporters have significantly higher markups.[43] I also explore the information in firms' financial balance sheet and show that firms who receive some part of their capital from FDI (column 2), the capital market (column 3) and from foreign loans (column 4) have higher markups. These patterns is consistent with these firms being technologically more advanced and hence having more market power. In contrast, it is harder to reconcile with markups reflecting binding financial constraints as these firms are arguable less constrained and hence should have low marginal products. In the last three columns I focus on the existence of constraints more directly. In 1996 the census contained direct questions on the type of constraints different firms are facing. Columns 5-7 therefore estimate the relative markup of firms who report that they face a constraint they could not overcome (column 5), that this constraint is particularly related to capital (column 6) and that they did not plan to expand because of a scarcity of capital (column 7). While some coefficients are not statistically significant depending on the specification, Table OA-4 shows that these firms are estimated to have *lower* marginal products. This is consistent with these firms being low productivity producers who post low markups but is harder to reconcile with a model of financial constraints.

OA-2.4 Robustness for Table 5

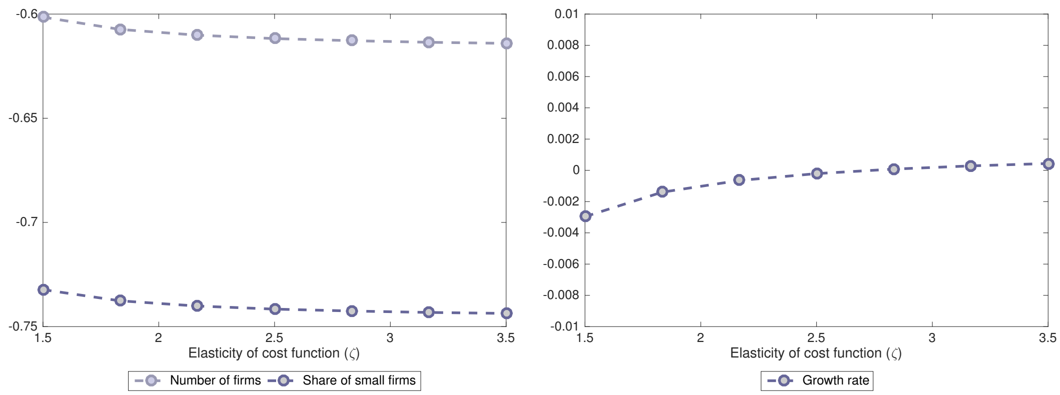
Table OA-5 contains numerous robustness checks for the empirical results reported in Table 5. In the baseline specification I only included industry-region-year cells with at least 50 firms, used value added as a measure of firm size and weighted all regressions by the number of firms in each cell. Table OA-5 shows that the results are qualitatively and quantitatively robust to these choices. Panel A and B are very similar to the baseline specification. Panel C shows that the effects become statistically insignificant once one raises the cutoff to 100 firms in a given year-industry-province cell. The reason is that Indonesia is quite concentrated and so that the number of provinces

[43] The results reported in Table OA-4 are also quantitatively consistent with De Loecker and Warzynski (2012). When they measure markups from the labor elasticity and correct the output data for measurement error, they estimate an exporter premium of 0.078.

Cross-sectional markup heterogeneity: $\ln \mu_{ft} = \delta_t + \delta_s + \beta \times FirmCharacteristic_{ft} + \alpha \times X_{ft} + u_{ft}$							
	Exporter	FDI	Capital market	Foreign loan	Contrained?	Capital constrained?	Financial growth barrier?
	Measure for μ: Labor share in value added, corrected						
β	0.088***	0.012	0.050*	0.058***	-0.016**	-0.011	-0.036**
	(0.004)	(0.012)	(0.027)	(0.011)	(0.008)	(0.010)	(0.017)
N	134,126	134,126	134,126	134,126	16,375	16,375	16,375
R^2	0.33	0.33	0.33	0.33	0.34	0.34	0.34
	Measure for μ: Labor share in value added, uncorrected						
β	0.167***	0.174***	0.065**	0.175***	-0.055***	-0.024**	-0.020
	(0.005)	(0.017)	(0.033)	(0.014)	(0.009)	(0.012)	(0.023)
N	138,953	138,953	138,953	138,953	16,375	16,375	16,375
R^2	0.18	0.17	0.17	0.17	0.20	0.19	0.19
	Measure for μ: Material share in sales, uncorrected						
β	0.084***	0.101***	0.008	0.091***	-0.020***	-0.015***	-0.027***
	(0.001)	(0.003)	(0.008)	(0.003)	(0.002)	(0.003)	(0.005)
N	134,235	134,235	134,235	134,235	16,375	16,375	16,375
R^2	0.89	0.88	0.88	0.88	0.89	0.89	0.89
	Measure for μ: Material share in sales, corrected						
β	0.108***	0.198***	0.027	0.150***	-0.037***	-0.013*	-0.008
	(0.003)	(0.009)	(0.020)	(0.007)	(0.005)	(0.007)	(0.012)
N	134,235	134,235	134,235	134,235	16,375	16,375	16,375
R^2	0.61	0.60	0.60	0.60	0.62	0.62	0.62

Notes: Robust standard errors are shown in parentheses. ***, ** and * denotes significance at the 1%, 5% and 10% level respectively. The table reports the results of regressing log markups on various firm characteristics. Each column represents a separate regression and I report the coefficient on the respective firm characteristic. All regressions include a full set of 5-digit product fixed effects, a set of year fixed effects and $ln\left(\frac{k}{l}\right)$, i.e. the firm's (log) capital-labor ratio. Capital is measured as the total value of assets reported in the industrial census. The first two panels measure markups from the inverse labor share in value added. The last two panels measure markups from the inverse material share in sales. In the last two panels I also control for the material labor ration $ln\,(m/l)$ as in (OA-25). "Exporter", "FDI", "Capital market" and "Foreign loans" are dummy variable indicating whether the firm exports and finances its investment through FDI, foreign loans or funds from the Indonesian capital market. The last three columns use the special census supplement of the census in 1996. The survey asked whether firms were facing any constraints, whether this constraint was related to missing capital and whether missing capital was the main obstacle in firm expansion.

Table OA-4: Determinants of labor productivity: Imperfect input markets and borrowing constraints



Notes: The figure displays the change in the number of firms and the share of small firms (left panel) and the change in the growth rate (right panel) for different values of ζ. Specifically, for each ζ, I recalibrate the model to the US and Indonesian moments contained in Table 4 and calculate the change in the equilibrium outcomes. The baseline specification is $\zeta = 2$.

Figure 6: Sensitivity with respect to ζ

drops substantially between Panel B and Panel C. Panel D shows that the effects of firm size on markups is weaker once firm size is measured by employment. Panel E shows that the effects are almost the same when the regressions are not weighted.

OA-2.5 Robustness of the calibration results

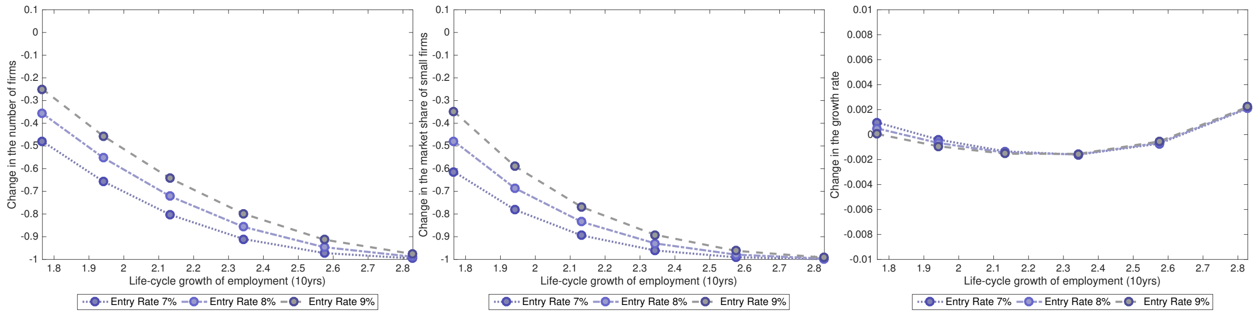
Sensitivity with respect to ζ In Figure 6 below I study the sensitivity of the results with respect to the elasticity ζ. More specifically, I consider a range of ζ between 1.5 and 3 and redo the analysis around Table 4. I report three moments: the change in the number of firms, the change in the share of small firms and the change in the growth rate. For the baseline, I assumed a value of $\zeta = 2$. In the left panel of Figure 6 I report the change in the number of firms and the share of small firms. It is seen that this moment is not substantially affected by the particular value of ζ: as for the baseline case, the number of firms declines by 74%, the share of small firms by 85%. In the right panel, I depict the change in the growth rate. Two results stand out: First of all, depending on the value of ζ the growth rate can either increase of decrease in response to a reduction of entry and expansion barriers. Secondly, the implied change in the growth rate are relatively small. Hence, the result that large changes in the firm size distribution are consistent with an essentially stable distribution of income is robust.

Sensitivity with respect to the calibrated moments In Figure 7 I study the sensitivity of these implications with respect to the underlying moments, i.e. the extent of life cycle growth and the extent of entry. More specifically, Figure 7 contains the change in the equilibrium number of firms, the change in the aggregate sales share of single-market producers and the change in the rate of growth as a function of employment growth and for different values of the entry rate. The baseline calibration for the US assumed an entry rate of 8% and that employment grows by a factor of two during the first 10 years of a firm's life cycle. The left and middle panel show the elasticity of average firm size and and the share of small firms with respect to these two moments is quite sizable. If one were to assume that the extent of life cycle growth in the US was 2.5 instead of 2, the model would predict a five to eight-fold difference in these moments. The right panel show the implications for the aggregate growth rate of manufacturing productivity. In contrast to the firm-level outcomes the model does not predict sizable growth implications. Not only is the effect on the growth rate ambiguous in that for example a rate of life cycle growth of 2.5 instead of 2 would increase US growth relative to Indonesia, but even for extremely large differences in firm-dynamics, the

	Entry rate	Exit rate	LC empl growth	Markups					
				Avg.	q^{90}	Avg.	q^{90}	Avg.	q^{90}
Panel A: Excluding cells with less than 60 firms									
Avg va	-0.024***	-0.011**	0.145***	-0.057***	-0.052***			-0.060***	-0.055***
	(0.006)	(0.004)	(0.040)	(0.018)	(0.014)			(0.019)	(0.015)
Entry rate						-0.193	-0.160		
						(0.125)	(0.176)		
$\ln z$								-0.017	-0.014
								(0.011)	(0.014)
N	381	385	385	384	384	380	380	380	380
R^2	0.387	0.350	0.673	0.432	0.412	0.395	0.393	0.434	0.415
Panel B: Excluding cells with less than 75 firms									
Avg va	-0.020*	-0.017***	0.188***	-0.046*	-0.038*			-0.047*	-0.038*
	(0.008)	(0.004)	(0.044)	(0.021)	(0.016)			(0.022)	(0.016)
Entry rate						-0.250*	-0.199		
						(0.123)	(0.172)		
$\ln z$								-0.021	-0.016
								(0.012)	(0.017)
N	306	308	308	307	307	305	305	305	305
R^2	0.421	0.368	0.757	0.472	0.445	0.460	0.443	0.478	0.450
Panel C: Excluding cells with less than 100 firms									
Avg va	-0.020**	-0.017**	0.223***	-0.035	-0.025			-0.035	-0.024
	(0.006)	(0.006)	(0.052)	(0.022)	(0.016)			(0.023)	(0.015)
Entry rate						-0.180**	-0.094		
						(0.069)	(0.130)		
$\ln z$								-0.020	-0.014
								(0.010)	(0.017)
N	228	230	230	229	229	227	227	227	227
R^2	0.469	0.387	0.810	0.439	0.492	0.433	0.493	0.447	0.498
Panel D: Measuring firm size by employment									
Avg empl	-0.042***	-0.025***	0.272***	-0.003	-0.024			-0.008	-0.031*
	(0.005)	(0.006)	(0.067)	(0.020)	(0.016)			(0.020)	(0.015)
Entry rate						-0.181	-0.169		
						(0.113)	(0.165)		
$\ln z$								-0.010	-0.011
								(0.011)	(0.014)
N	455	463	463	462	462	454	454	454	454
R^2	0.377	0.327	0.662	0.335	0.349	0.348	0.353	0.343	0.353
Panel E: No weights									
Avg va	-0.017***	-0.010**	0.090**	-0.060***	-0.053***			-0.063***	-0.058***
	(0.004)	(0.004)	(0.031)	(0.016)	(0.010)			(0.016)	(0.010)
Entry rate						-0.159	-0.234		
						(0.128)	(0.180)		
$\ln z$								-0.014	-0.019
								(0.012)	(0.016)
N	455	463	463	462	462	454	454	454	454
R^2	0.304	0.285	0.532	0.350	0.290	0.323	0.278	0.366	0.298

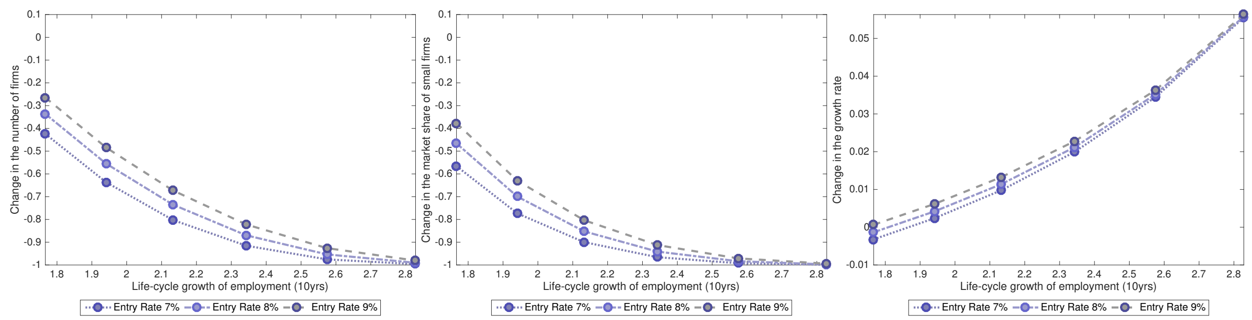
Notes: Standard errors are clustered at the level of a province and contained in parentheses with ***, **, and * respectively denoting significance at the 1%, 5% and 10% levels. Regression are run at the province-industry level, where industries are measured at the 3-digit level. The variables are all measured within these province-industry cells. The entry and exit rates are measured as the share of entering and exiting firms. The employment life cycle is measured as the growth of cohort employment over the 3 year horizon (see also Figure 2). Log markups are measured as the residual from a regression of log inverse labor shares on a set of year and 5-digit industry fixed effects and $ln(k/l)$ (see (26)) and "Avg" is the mean log markup and q^{90} is the 90% quantile. "Avg va" is the average log value added within a industry-region-year cell. The flow rate of entry z is measured as in the theory, i.e. from $\frac{\text{Entry rate}}{\text{Avg va}}$ (see (28)). All regressions contain a full set of industry and year fixed effects and control for the log of the province population and the share of villages within the province, which are agricultural. I only consider province-industry cells with at least 50 observations and all regressions are weighted using the number of observations within each cell as a weights.

Table OA-5: Firm-Size, Entry and Markups across Regions in Indonesia (Robustness)



Notes: This figure displays the model's predictions for the change in the number of firms (left panel), the share of sales of small, i.e. one-product firms (middle panel) and the equilibrium growth rate (right panel) as a function of the life cycle employment growth over 10 years and the equilibrium entry rate. The baseline results in Table 4 refer to the case of a life cycle growth of 2 and an entry rate of 8%.

Figure 7: Sensitivity of results to underlying moments



Notes: This figure displays the model's predictions for the change in the number of firms (left panel), the share of sales of small, i.e. one-product firms (medium panel) and the equilibrium growth rate (right panel) as a function of the life cycle employment growth over 10 years and the equilibrium entry rate for an economy without heterogeneous markups, i.e. where firms' own-quality improvement are restricted to $I = 0$. The baseline results in Table 4 refer to the case of a life cycle growth of 2 and an entry rate of 8%.

Figure 8: Sensitivity of results to underlying moments in the economy without heterogeneous markups

differences in aggregate productivity growth do not exceed 0.5%.

The Importance of Markups Figure 8 below replicates the exercise in the economy without heterogeneous markups reported in the last column of Table 4 for various combinations of the two calibration targets, i.e. the equilibrium entry rate and the extent of life cycle growth. The structure of Figure 8 parallels the one of Figure 7, i.e. I report the change in the number of firms (left panel), the change in the aggregate output share of small, one-product firms (medium panel) and the change in the aggregate growth rate (right panel). Comparing Figures 8 and 7 it is seen that the implications for the number of firms and the importance of small firms are roughly similar, i.e. depending on the precise calibration target, both moments decline by a factor of 2 to 5. However, the models vastly differ in the implications for the economy-wide growth rate. As in Table 4, elasticity of aggregate growth with respect to the employment life cycle is very large once markups are constrained to be constant.

OA-2.6 Identifying the model based on the sales share of entering firms

Here I conduct an alternative calibration, where instead of taking the rate of new firms in the Census as the entry rate, I explicitly treat the data as truncated and match the share of sales of firms crossing the size-threshold and hence entering in the dataset. Suppose that firms enter the dataset once they cater to n_0 markets.[44] The mass of firms entering the dataset is therefore given by $E^C = x \times (n_0 - 1) \times \omega_{n_0-1}$, where ω_{n_0} is the firm size distribution characterized in Proposition 2 and $x \times (n - 1_0)$ is the expansion rate of firm with $n_0 - 1$ products. The share of sales of these firms relative to all firms *in the data*, σ^E, is given by $\sigma^E = \frac{Y_t n_0 E^C}{\sum_{n \geq n_0} Y_t n \omega_n} = \frac{n_0 E^C}{1 - S_{n_0-1}}$, where S_k is the share of sales accounted for by firms with at most k products. In Proposition 2 I showed that $S_n = 1 - \vartheta_x^n$. Hence, we can write σ^E as

$$\sigma^E = n_0 x \frac{1 - \vartheta_x}{\vartheta_x} = x \frac{1 - \vartheta_x}{\vartheta_x} \times \frac{ln\left(1 - S_{n_0}\right)}{ln\left(\vartheta_x\right)} = z \times \frac{ln\left(1 - S_{n_0}\right)}{ln\left(\frac{\tau - z}{\tau}\right)}. \tag{OA-26}$$

Because σ^E is directly observed in the data (see Table 1), (OA-26) can be used to identify the flow rate of entry z for a given choice of S_{n_0}. Note that S_{n_0} is the share of sales of firms that are sampled in the Census. As for the value of S_{n_0}, note that Hsieh and Olken (2014) report that firm with more than 10 employees account for 45%. If informal firms are labor-intensive, the share of sales such firms account for will be larger. To infer S_{n_0} one has to take a stand to what extent informal and formal manufacturing firms compete for the same customers. While firms with 10 employees are likely to be active in the same product market as firms with 20 workers, this is probably not the case for many self-employed small-scale entrepreneurs without any personnel. Hence, I will assume that $S_{n_0} = 0.85$, i.e. that 85% of the relevant product market are accounted for by the firms I observe in the data. As seen from (OA-26), the decisive data moment is $\frac{\sigma^E}{ln\left(1 - S_{n_0}\right)}$, i.e. the actual share S_{n_0} was say higher, it is as if the observed sales share σ^E was too low.

In Table OA-6 I report the main counterfactual results based on this new calibration. In particular, I first calibrate the model to the Indonesian data by matching $\sigma^E = 0.033$ (see Table 1). Then I recalibrate the entry and expansion barriers to match the same US moments as in the baseline calibration. Hence, if the economy calibrated to match σ^E had implied an entry rate of 10.4%, the results in Table OA-6 would be numerically the same as for my baseline calibration. While this is not exactly the case, the results are quite similar. In particular, it is still the case that the effect on the firm size distribution is large, while the effect on the aggregate growth rate is very small. Note also that this calibration implies that the US economy benefits particularly from lower market barriers. The implied changes in market barriers and entry costs are larger than in the baseline calibration.

OA-2.7 Dispersion in measured productivity

The model predicts that the distribution of markups across markets is pareto with parameter θ. The distribution of log markups is therefore given by

$$P\left[ln\left(\mu\right) \leq \kappa\right] = P\left[\mu \leq e^{\kappa}\right] = 1 - e^{-\theta\kappa},$$

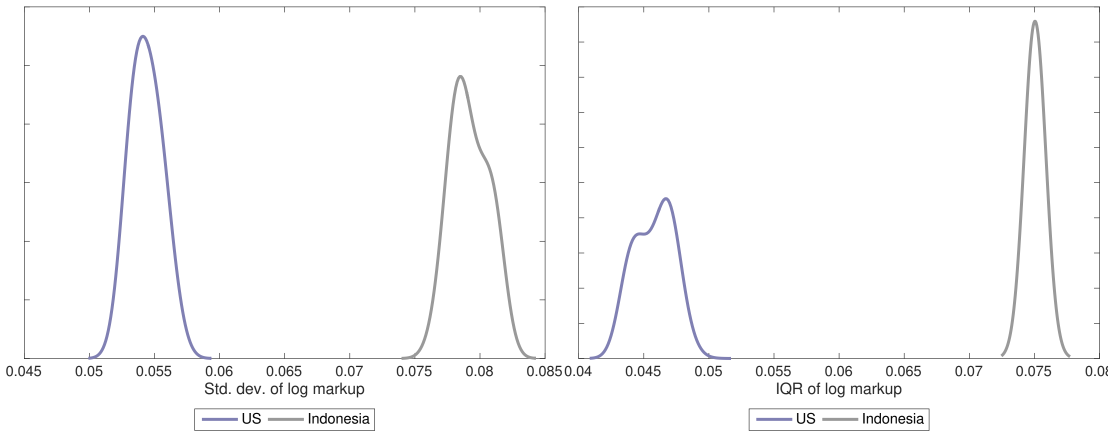
i.e. log markups are exponentially distributed with parameter θ. The standard deviation of log markups across markets is therefore simply θ^{-1}. Because firms are active in multiple markets, the standard deviation of markups across firms is not given by θ^{-1}. I was not able to derive a closed form expression for the dispersion of log markups across firms. However, given the calibrated values for I, x and τ, it is straightforward to simulate this distribution. In Figure 9 I depict the distribution of the estimated standard deviation (left panel) and interquartile range (right panel) for log markups at the firm level for the US and Indonesia. I depict the distribution to account for sampling variation in the stationary distribution. The distribution is estimated by simulating the stationary distribution 300 times and the calculating the respective dispersion moment. For the results in Table 4 I report the median from

[44] Note that, strictly speaking, in the data the sample criterion is based on employment and not sales (which are proportional to n). Specifying the truncation in terms of sales is much more tractable because one does not need to know the distribution of markups for firms crossing the threshold.

Calibration	Indonesia	US	Equilibrium implications	Indonesia	US	Change
Entry Rate	-	8%	*The Distribution of Firm Size*			
Entry sales share	3.3%	-	Number of firms	0.222	0.1259	-43.3%
Empl. Life-Cycle	1.7	2	Output share of small firms	0.0812	0.0367	-54.8%
Expansion Barrier	1	0.72	*Markups and Misallocation*			
Entry Barriers	1	0.54	Life cycle of markups	0.0818	0.0539	-0.028
			Average markup	13.08%	8.67%	-33.7%
			Dispersion in log markups	11.57%	7.98%	-31%
			Measured markup dispersion	8.48%	5.19%	-38.7%
			Efficiency wedge M	0.9939	0.9971	0.32%
			Labor wedge Λ	0.8734	0.9018	3.3%
			Innovation, Expansion and Entry and Growth			
			Rate of growth g	3	2.92	-0.0008
			Rate of innovation I	0.4113	0.3317	-19.3%
			Rate of market expansion x	0.1949	0.2645	35.7%
			Rate of creative destruction τ	0.2121	0.2745	29.4%

Notes: The first panel contains the calibration moments. The entry rate is simply the share of firms, which are entrants. The entry sales share is the share of sales accounted for by firms entering the Census. For the US, the employment life cycle is calculated as average employment of firms between 10 and 14 years old relative to firms with age less than 5 and stems from Hsieh and Klenow (2014). The parameters for the Indonesian economy are contained in 3. For the US economy, I recalibrate the relative efficiency of expansion and entry, i.e. $\frac{\phi_z^{US}\varphi_z^{US}}{\phi_z^{IND}\varphi_z^{IND}} = \frac{1}{0.86}$ and $\frac{\left(\phi_x^{\zeta}\right)^{US}\varphi_z^{US}}{\left(\phi_x^{\zeta}\right)^{IND}\varphi_z^{IND}} = \frac{1}{0.67}$. The remaining parameters are the same as in Table 3. The employment life cycle for Indonesia at the 10 year horizon reported in the first panel is calculated from the calibrated model to ease comparison. The closed-form expression for the endogenous outcomes in the second panel are contained in Propositions 2, (20) and (21). The column "Change" reports the change in the respective equilibrium outcomes.

Table OA-6: Comparing Indonesia and the US



Notes: The figure shows the distribution of the standard deviation of log markups (left panel) and the interquartile range of log markup (right panel) at the firm level. To construct these distributions, I simulate the stationary distribution of the economy 300 times and then calculate the respective dispersion moment. The underlying parameters are the ones reported in Table 4.

Figure 9: Measured productivity dispersion

this distribution.