Population Growth and Firm-Product Dynamics*

Michael Peters
Yale University and NBER

Conor Walsh
Columbia University

June 2022

Abstract

Population growth has declined markedly in almost all major economies. We argue this trend has important consequences for creative destruction, product concentration, and firm dynamics. We propose a rich model of growth with multi-product firms, and show that lower population growth reduces entry and creative destruction, increases product concentration, raises market power and firm size, and lowers aggregate growth. At the same time, lower population growth increases the mass of products available to consumers, making the short-run welfare impacts ambiguous. In an application to the US, the slowdown in population growth accounts for a substantial share of the fall in the entry and exit rates, and the increase in product concentration and firm size. By contrast, the impact on markups is modest. The effect on aggregate growth is initially positive, before turning negative thereafter.

Keywords: Growth, Firm Dynamics, Creative Destruction, Demographics, Dynamism, Markups

JEL Codes: O40, D22, J11, O47

*Peters: m.peters@yale.edu. Walsh: caw2226@columbia.edu. We are extremely grateful to Huiyu Li for her insightful discussion at the AEA meetings and Sara Moreira and David Berger for their help in constructing the RevLBD data. We thank Paco Buera, Francesco Caselli, Murat Celik, Chad Jones, Pete Klenow, Sam Kortum, Ilse Lindenlaub, Ben Pugsley, Maarten De Ridder, and Jesse Schreger for comments. We also thank seminar participants at the LSE, Stanford University, the University of Toronto, PSU, Yale University, CEMFI, Queen's University, the STLAR 2019 Conference, the SED 2019, VMAC, the "Taxation, Innovation and the Environment" conference at the College de France and the 2021 AEA meetings for helpful suggestions. Ferdinand Pieroth provided outstanding research assistance. Walsh gratefully acknowledges the support of the Washington Center for Equitable Growth. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the US Census Bureau. All results have been reviewed to ensure no confidential information is disclosed.

1 Introduction

Almost all major economies have experienced a substantial decline in population and labor force growth in recent decades.[1] The left panel of Figure 1 shows the pronounced downward trend in the US: since the late 1970s, labor force growth has fallen from 3% to 1%. Moreover, according to the BLS, this trend is projected to continue for the foreseeable future, driven largely by continuing declines in fertility. Global population growth has also been falling, driven mostly by drastic declines in Asia and Latin America. A world of low and falling population growth looks like it is here to stay.

In this paper we study the effect of falling population growth on firm dynamics, product concentration and aggregate economic performance. We do so in the context of a semi-endogenous growth model with multi-product firms. Our baseline model is an enhanced version of Klette and Kortum (2004), augmented by the possibility of population growth, new-variety creation, own-innovation, and a demand elasticity that exceeds unity. The model is rich enough to rationalize many first-order features of the firm-level data, yet has an analytic solution that allows us to express the process of firm dynamics and the aggregate growth rate directly as a function of population growth.

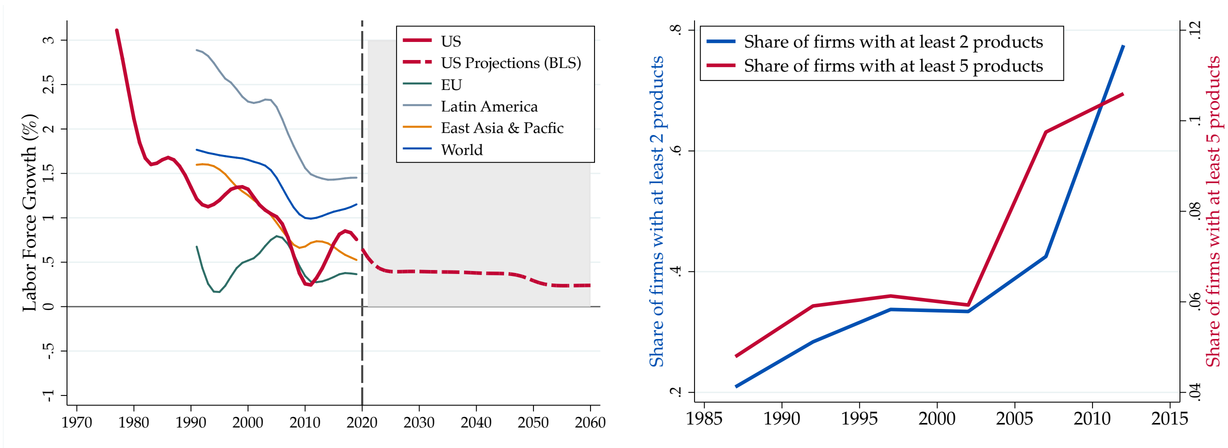
Our theory makes tight predictions for the effects of falling population growth. We first show that a slow-down in population growth reduces creative destruction and the creation of new products. The reason is the following. In the long-run, the number of products available to consumers has to grow at the same rate as the population. If that was not the case, firm profits would either grow without bound or converge to zero, both of which are inconsistent with free entry. Falling population growth thus goes hand in hand with a decline in product innovation, reducing both the creation of new varieties and the rate of creative destruction.

We then show this decline in churning leads to rising product concentration and larger firms. In particular, we show that, under standard assumptions, the decline in product creation and creative destruction is *only* accommodated through a decline in entry. By contrast, incumbent firms' innovation policies are, in general equilibrium, independent of the rate of population growth. This change in the composition of creative destruction implies that incumbent firms face less competition by entrants, which allows them to

[1] Throughout the paper, we often speak of population growth and labor-force growth interchangeably. We take these to be exogenous to market concentration and firm dynamics. Across the developed world, decreases in fertility in the 1960s and 1970s manifested in slower rates of growth in the labor force in the 1980s and 1990s - see De Silva and Tenreyro (2017).

Figure 1: Labor Force Growth and Product Concentration

(a) Falling Labor Force Growth (b) Rising Product Concentration



Notes: In the left panel we display historical and projected labor force growth for the US and for several regions across the globe. In the right panel we display the share of US manufacturing firms with at least two (shown blue, left axis) and at least five (shown in red, right axis) products, using data from the Census of Manufacturing. The number of product produced is self-reported by firms and assigned to 10-digit NAICS categories by the US Census. See Section 4.2 below for details.

accumulate more products over their life-cycle. As a consequence, concentration and firm size rise, entry and exit rates fall and firms get older.

In the right panel of Figure 1 we document this rise of product concentration for the US manufacturing sector. The share of manufacturing firms producing at least two products, shown in blue, rose from about 20% in the late 1980s to almost 80% in 2012. Similarly, the share of firm producing at least five products, shown in red, also more than doubled and reached more than 10% in 2012. Consistent with our theory, product markets in the US are increasingly dominated by large, multi-product firms.[2]

In addition to firm dynamics, our theory also makes clear predictions about the relationship between population growth and per-capita income, both in levels and in growth rates. As in many aggregate models of semi-endogenous growth, the long-run growth rate declines as population growth falls. However, we show that an important countervailing effect makes the relationship between population growth and welfare a priori ambiguous. By reducing creative destruction, falling population growth increases the value of incumbent firms, because future profits are discounted at a lower rate. Free entry therefore requires an increase in the economy-wide number of products to increase competition. Note that this rise in the number of *products* per worker coincides with a decline in the number of *firms* per worker, that is, an increase in average firm size, be-

[2]Below we report the full distribution. We also show these trends are not unique to the manufacturing sector, but are are present in other sectors.

cause the number of products per firm rises. Because additional varieties raise income per-capita, the welfare consequences of declining population growth hinge on the relative importance of these static variety gains relative to the dynamic losses from lower growth.

These results are robust to a variety of changes in the environment. Most importantly, we extend our model to a setting where firms compete a la Bertrand and market power is endogenous. While all our theoretical results above directly generalize to this setting, declining population growth interacts with firms' ability to charge markups in an interesting way. In our theory, more productive firms post higher markups, and productivity increases over the firms' life cycle. Because creative destruction reduces firms' chances of survival, it hinders incumbents from accumulating market power. Declining population growth, by lowering creative destruction, therefore reduces competition and increases markups. Hence, the trend of rising product concentration shown in Figure 1 goes hand in hand with large producers being able to sell their products at high markups.

To quantify the strength of this mechanism, we calibrate our model to data for the population of US firms. In addition to targeting standard moments such as the entry rate, average size, and life-cycle growth, we also link firm-level information on sales to the US Census. We can therefore explicitly target a measure of the life-cycle of firm-level markups for a majority of firms in the US. Exploiting information on the evolution of both markups and size at the firm-level allows us to separately identify own-innovation and variety creation at the firm-level.

With the calibrated model in hand, we ask a simple question: what are the implications of the decline in the rate of labor-force growth since 1980 shown in Figure 1? Our theory is tractable enough that we can solve for the transitional dynamics induced by this path, treating the projections of the BLS as the rational expectations of the agents in our theory. We find this decline has quantitatively large effects. Our model can explain a substantial share of the decline in the entry and exit rate, the increase in average firm size, and the degree of concentration. However, markups change little; our calibrated model implies markups increase by around 1%. The effect on income growth is more subtle. Whereas growth will inevitably decline in the long-run, the static effect of variety creation increases income growth for about one decade during the transition. However, the overall welfare consequences of falling population growth are negative.

Related Literature. We are not the first to connect the decline in population growth to changes in firm dynamics. Karahan et al. (2019) and Hathaway and Litan (2014) use geographic variation to provide direct empirical support that a lower rate of population growth reduces the start-up rate. Karahan et al. (2019) and Hopenhayn et al. (2018) study the relationship between population growth and firm dynamics in a neoclassical model, where firm productivity is exogenous and changes in demographics can only affect the incentives of entering firms. By contrast, our theory builds on models with endogenous firm dynamics, where population growth affects innovation incentives by both entrants and incumbents, and has novel implications for creative destruction, market power, and aggregate productivity growth.

Our theory builds on Schumpeterian models in the tradition of Aghion and Howitt (1992) and Klette and Kortum (2004). We augment these models by allowing for efficiency improvements of existing firms as in Atkeson and Burstein (2010), Luttmer (2007), Akcigit and Kerr (2018), or Cao et al. (2017), the creation of new varieties as in Young (1998), and endogenous markups as in Peters (2020) or Acemoglu and Akcigit (2012). Our model is thus akin to a version of Garcia-Macia et al. (2019) or Klenow and Li (2021), featuring both endogenous markups and endogenous innovation choices, and incorporating changes in the long-run growth in the labor force. To the best of our knowledge, our paper is the first that focuses squarely on how demographic changes affect creative destruction, aggregate growth, and the firm-size distribution in the context of firm-based models of growth.[3]

The relationship between economic growth and population growth has been been subject to an extensive literature. Many models of endogenous growth feature "strong scale effects" whereby economic growth depends on the population level. By contrast, models of semi-endogenous growth are characterized by "weak scale effects" and imply economic growth is determined by population growth.[4] In our model, growth is tied to the micro process of firm dynamics. This link puts tight restrictions on the relationship between economic growth and population growth. If growth depends on the level of the population, so does the firm size distribution. By contrast, if economic growth depends on population growth, the firm size distribution is also independent of the size of workforce and only a function of its growth rate. In order for the firm size distribution to be stationary in the presence of a growing population, growth thus (generically) needs to be semi-endogenous.[5] Moreover, the relationship between economic growth and population

[3] Engbom (2017) studies the implications of population aging in the context of a search model.

[4] Classic examples of the former are Aghion and Howitt (1992), Romer (1990), Klette and Kortum (2004) or Grossman and Helpman (1991). Examples of the latter are Jones (1995), Kortum (1997), Young (1998), or Peretto (1998). See Jones (2021) for a recent survey and Atkeson et al. (2018) for a discussion of the quantitative importance of this distinction.

[5] In our model, population growth is not the only determinant of the equilibrium growth rate, and our

growth is governed by parameters that we can discipline with firm-level data.

In our quantitative application, we focus on the case of the US. A growing literature highlights the decline of dynamism in the US. This literature, which is summarized in Akcigit and Ates (2019a), shows that the entry rate has fallen substantially (Alon et al., 2018; Decker et al., 2014; Karahan et al., 2019), that broad measures of reallocation have declined (Haltiwanger et al., 2015), that industries are becoming more concentrated (Kehrig and Vincent, 2017; Autor et al., 2020), and that markups and profits are rising (Edmond et al., 2018; De Loecker et al., 2020; Van Vlokhoven, 2021). In terms of explanations, the literature has proposed improvements in IT technology (Aghion et al. (2019); Lashkari et al. (2019)), a rise in the use of intangible capital (De Ridder (2019)), or changes in the process of knowledge diffusion (Akcigit and Ates (2019b), Olmstead-Rumsey (2020)). Our paper is complementary to these studies by highlighting that all these phenomena occurred within an environment of declining population growth and are key implications of the theory we propose. Falling population growth might therefore be an important secular determinant of firm-dynamics and aggregate growth in the decades to come.

2 The Baseline Model

Time is continuous and there is a mass L_t of identical individuals, each supplying one unit of labor inelastically. The rate of population growth $\dot{L}_t/L_t = \eta_t$, which we take as exogenous, is the crucial parameter of this paper. Households have preferences over a final consumption good c_t given by $U = \int_0^\infty e^{-(\rho-\eta_t)t} \ln(c_t)\, dt$.

The final consumption good is composed of a continuum of differentiated varieties, that (as in Klette and Kortum (2004)) may be produced by multiple firms:

$$Y_t = \left(\int_0^{N_t} \left(\sum_{f \in S_{it}} y_{fit} \right)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}} .$$

Here, N_t is the mass of available products indexed by i. This mass evolves endogenously through the creation of new and the destruction of old products. S_{it} denotes the set of firms with the knowledge to produce product i, which likewise evolves endogenously.

Firms can be active in multiple product markets. Each firm f is characterized by the set of products it produces, denoted by Θ_f, and the efficiency of producing these products,

economy features growth even in the presence of a stable population.

indexed by $\{q_{fi}\}_{i\in\Theta_f}$. We denote the number of products firm f produces by n_f. Production of each good uses only labor, and is given by $y_{fi} = q_{fi}l_{fi}$, where q_{fi} denotes the efficiency of firm f in producing product i.

Because the output of firms producing the same product is perfectly substitutable, each product is only produced by the most efficient firm. Suppose to begin with that the producing firm charges a constant markup over marginal cost $\mu = \frac{\sigma}{\sigma-1}$.[6] Aggregate output Y_t and equilibrium wages w_t are thus given by

$$Y_t = Q_t N_t^{\frac{1}{\sigma-1}} L_t^P \quad \text{and} \quad w_t = \mu^{-1} Y_t / L_t^P, \tag{1}$$

where $Q_t \equiv \left(\int q_i^{\sigma-1} dF_t\left(q\right)\right)^{1/(\sigma-1)}$ is a measure of average efficiency, F_t is the endogenous distribution of product efficiency, and L_t^P is the total amount of labor devoted to the production of goods. Equilibrium profits per product are given by

$$\pi_t\left(q\right) = \left(\mu-1\right)\left(\frac{q}{Q_t}\right)^{\sigma-1}\frac{L_t^P}{N_t}w_t. \tag{2}$$

Hence, profits are high if the product's efficiency q is large relative to average efficiency Q_t and if average employment per product, L_t^P/N_t, is large.

2.1 Product Innovation, Entry and Aggregate Growth

Suppose first that entry is the sole source of creative destruction and product creation. Potential entrants have access to a linear entry technology, where each worker generates a flow of φ_E new products. Conditional on successfully creating a new product, this product can either be a new variety, or it can improve upon an existing product from another firm. After entry, as in Atkeson and Burstein (2010) or Luttmer (2007), the firm's efficiency q grows at an exogenous own-innovation rate I, that is, $\dot{q}_{it} = Iq_{it}$.[7] We also assume that product lines die at an exogenous rate δ. Doing so helps ensure stationarity at low or negatives levels of population growth, but is otherwise inconsequential.

In the baseline we assume product creation is "undirected". With probability α, the new product represents a technological advance over a (randomly selected) existing product,

[6] This can be the case if firms have to pay an infinitesimal sunk cost before producing, as less productive firms will not enter (see Garcia-Macia et al. (2019)).

[7] For simplicity, we start by assuming I is exogenous and constant over time. In Section 2.6, we show how to endogenize this rate and how it depends on population growth.

increases the efficiency by a factor $\lambda > 1$, and forces the current producer to exit ("creative destruction"). With the complementary probability $1-\alpha$, the product will be new to society as a whole, and the mass of available products N_t grows ("variety creation"). The efficiency of new varieties is given by $q' = \omega Q_t$, where ω is drawn from a fixed distribution $\Gamma(\omega)$. Hence, as in Buera and Oberfield (2020), the efficiency of new varieties is determined both by the existing knowledge embedded in Q_t and by novel ideas. It is useful to define $\overline{\omega} \equiv \left(\int \omega^{\sigma-1} d\Gamma(\omega)\right)^{1/(\sigma-1)}$, which we also refer to as the mean efficiency of new products.

Let Z_t denote the aggregate flow of entry and $z_t = Z_t/N_t$ the entry intensity per product. The rates of variety creation ν_t and creative destruction τ_t are then given by

$$\nu_t = (1-\alpha) z_t \quad \text{and} \quad \tau_t = \alpha z_t. \tag{3}$$

Creative destruction τ and variety creation ν both depend on the entry intensity z and are thus closely linked. Our formulation of undirected innovation makes this link particularly stark. However, as we show in Section 2.6, the optimal level of creative destruction and variety creation positively co-move even in a more general setting where α is a choice variable.

Given ν_t and τ_t, we can compute the aggegate growth rate. The variety growth rate is simply the *net* rate of product creation

$$g_t^N \equiv \frac{\dot{N}_t}{N_t} = \nu_t - \delta = \frac{1}{1-\alpha} z_t - \delta. \tag{4}$$

Similarly, the efficiency growth rate is given by

$$g_t^Q \equiv \frac{\dot{Q}_t}{Q_t} = \frac{\overline{q}^{\sigma-1}-1}{\sigma-1} z_t + I, \quad \text{where} \quad \overline{q} = \left(\alpha\lambda^{\sigma-1} + (1-\alpha)\,\overline{\omega}^{\sigma-1}\right)^{\frac{1}{\sigma-1}}. \tag{5}$$

Average efficiency grows for two reasons. The first term is related to product creation and hence the entry rate z_t. The average efficiency gain of a new product, $\overline{q}$, is a CES-weighted average of the efficiency improvement of creative destruction λ and the relative efficiency of new varieties $\overline{\omega}$. Because $\lambda > 1$, creative destruction is always a source of efficiency growth. Whether the creation of new varieties raises or lowers efficiency growth depends on their initial efficiency $\overline{\omega}$. If new products are, on average, as productive as existing products, that is, $\overline{\omega} = 1$, the growth rate of average efficiency is independent of ν_t.[8] If, by contrast, new products are, on average, worse ($\overline{\omega} < 1$), faster product creation has a negative effect on efficiency growth. Average efficiency thus increases in z_t as long as α and $\overline{\omega}$ are sufficiently large. The second term in (4) captures the vertical component of life-cycle productivity growth I, which raises firm productivity and also translates into aggregate efficiency growth g^Q.

Using (4) and (5), the overall growth of labor productivity $Y_t/L_t^P = Q_t N_t^{\frac{1}{\sigma-1}}$, which we denote by g_t^y, depends on both efficiency growth and variety growth:

$$g_t^y = g_t^Q + \frac{1}{\sigma-1} g_t^N = \frac{\overline{q}^{\sigma-1} - \alpha}{\sigma-1} z_t + I - \frac{\delta}{\sigma-1}.$$

Note that aggregate productivity growth is always increasing in the rate of product creation z_t, because $\overline{q}^{\sigma-1} > \alpha$ even if even if $\overline{\omega} < 1$ (recall that $\lambda > 1$ and $\overline{\omega} > 0$). Intuitively, even if efficiency growth is decreasing in z_t, overall productivity growth increases once the variety gains are taken into account. Holding z_t constant, a higher rate of obsolescence δ reduces growth though the loss of varieties.

Optimal Entry and the Value Function. To solve for the equilibrium rate of entry, let $V_t(q)$ denote the value of producing a product with efficiency q at time t. This value function solves the HJB equation:

$$r_t V_t(q) - \dot{V}_t(q) = \pi_t(q) + \frac{\partial V_t(q)}{\partial q} I q - (\tau_t + \delta) V_t(q). \tag{6}$$

The value $V_t(q)$ is increasing in the current flow profits $\pi_t(q)$ and the rate of efficiency growth Iq, and decreasing in the risk of exit, which happens at the endogenous rate of creative destruction τ_t and the exogenous rate of obsolescence δ.

Given $V_t(q)$, we can compute the expected value of creating a new product. With probability α, the new entrant improves over a randomly selected product with efficiency q', and creatively destroys the current producer. This yields the value $V_t(\lambda q')$. With the complimentary probability, a new variety with quality ωQ_t and the associated value $V_t(\omega Q_t)$ is created. Integrating over the existing distribution $F_t(q')$ and the exogenous distribution of the efficiency of new varieties $\Gamma(\omega)$ yields the value of creating a new product

$$V_t^{PC} = \alpha \int V_t(\lambda q')\, dF_t(q') + (1-\alpha) \int V_t(\omega Q_t)\, d\Gamma(\omega). \tag{7}$$

[8] Note that $g^Q = \frac{1}{\sigma-1}\left[(\lambda^{\sigma-1} - 1)\tau_t + \left(\overline{\omega}^{\sigma-1} - 1\right)\nu_t\right]$. Substituting (3) yields (5).

To make progress, note that profits are homogeneous in $q^{\sigma-1}$ and so is the value function $V_t(q)$. This implies that the product creation value in (7) is given by $V_t^{PC} = V_t(\bar{q}Q_t)$, where $\bar{q}$ is given in (5). Hence, V_t^{PC} is simply the value of a product with "average" quality $\bar{q}Q_t$. The free entry condition thus reads

$$\frac{w_t}{\varphi_E} = V_t^{PC} = V_t(\bar{q}Q_t), \tag{8}$$

and hence requires $V_t(\bar{q}Q_t)$ to be tied to the equilibrium wage. Using (8), we can solve for $V_t(q)$ explicitly, both off and on the balanced growth path. As we show in Section A-1.1.3 in the Appendix, the solution to the differential equation in (6), $V_t(q)$, is given by

$$V_t(q) = \frac{\pi_t(q)}{r_t + \tau_t + \delta + (\sigma-1)\left(g_t^Q - I\right) - g_{wt}}. \tag{9}$$

The value of a firm is the present discounted value of profits, where the appropriate discount rate reflects four distinct considerations: the interest rate (r_t), the risk of firm death $(\tau_t + \delta)$, the fact that a higher growth rate of average efficiency reduces the firms' relative competitiveness $\left((\sigma-1)\left(g_t^Q - I\right)\right)$, and the rate of wage growth (g_w). Note that a higher growth of wages reduces the effective discount rate (and hence increases $V_t(q)$), because free entry ties the growth rate of wages to the growth of the value function. Faster wage growth is thus associated with capital gains.

2.2 Equilibrium

To characterize the equilibrium, define the two aggregate statistics $\mathcal{N}_t \equiv N_t/L_t$ and $\ell_t^P \equiv L_t^P/L_t$. We refer to the mass of products per capita $\mathcal{N}_t$ as the economy's *variety intensity* and to the share of working employed in production ℓ_t^P as the *production share*. These two aggregate statistics are sufficient to characterize the entire equilibrium path.

To determine this path, note first that labor market clearing requires that $L_t = L_t^P + N_t \frac{1}{\varphi_E} z_t$. Using that $z_t = \frac{1}{1-\alpha}\nu_t$, this can be written as

$$\frac{1-\ell_t^P}{\mathcal{N}_t} = \frac{1}{\varphi_E}\frac{\nu_t}{1-\alpha}. \tag{10}$$

Holding the variety intensity $\mathcal{N}_t$ constant, a higher production share ℓ_t^P reduces the creation of new varieties ν_t, as fewer resources are allocated toward research. Similarly, holding the allocation of workers constant, a higher variety intensity reduces the number of researchers per product and hence the rate of new variety creation. Equation (10) is the first key equation to characterize the equilibrium.

The second key equation is the free-entry condition. Substituting the solution for V_t in (9) into the free entry condition (8), and using the consumer Euler equation $g_y = r_t - \rho$, the expressions for τ_t and g^Q given in (3) and (5) and the fact that $g^w = g^y - g^{\ell^P}$ (see (1)), free entry requires that

$$\frac{1}{\varphi_E} = \frac{V_t\left(\bar{q}Q_t\right)}{w_t} = \frac{(\mu-1)\,\bar{q}^{\sigma-1}}{\rho+\delta+\left(\frac{\bar{q}^{\sigma-1}}{1-\alpha}-1\right)\nu_t+g_t^{\ell^P}}\frac{\ell_t^P}{\mathcal{N}_t}. \tag{11}$$

Now recall that $\dot{\mathcal{N}_t}/\mathcal{N}_t = \nu_t - \delta - \eta$. Substituting for ν_t shows that (10) and (11) are two differential equations in $\left\{\mathcal{N}_t, \ell_t^P\right\}_t$. Together with the initial condition $\mathcal{N}_0$ and the transversality condition, they fully determine the equilibrium path.[9]

The Balanced Growth Path. Consider first a BGP where the interest rate and the economy-wide growth rate are constant. This implies both variety creation ν and creative destruction τ are constant, and the population grows at a constant rate. Equations (10) and (11) then require that $\mathcal{N}_t$ and ℓ_t^P are constant. This has the important implication that the mass of varieties N_t has to grow at the rate of population growth:

$$\eta = g_N = \nu_t - \delta = (1-\alpha)\,z - \delta. \tag{12}$$

The aggregate quantity of product creation is thus directly tied to the growth rate of the labor force η. This link between population and product growth is a consequence of the free entry condition. If the number of products was growing faster than the population, profits per product would be declining. Eventually, entry would stop as the equilibrium wage would exceed the value of product creation. Conversely, if population growth was higher than the rate of new product creation, flow profits would perpetually rise. The free entry condition would then require a steady increase in the rate at which future profits are discounted. This, however, would eventually violate the economy's resource constraint.

With equation (12) at hand, we can analytically characterize the allocations along the BGP as a function of population growth:

Proposition 1. *On a BGP, the following holds:*

[9] See Section A-1.1.4 in the Appendix, where we derive this system of differential equations.

1. *The rates of variety creation ν, creative destruction τ, and entry z are given by*

$$\nu = \eta + \delta \qquad \tau = \frac{\alpha}{1-\alpha}\left(\eta+\delta\right) \qquad z = \frac{\eta+\delta}{1-\alpha}. \tag{13}$$

2. *Aggregate productivity growth g^y is given by*

$$g^y = \frac{\overline{q}^{\sigma-1}-1}{\sigma-1}\frac{1}{\alpha}\tau + I + \frac{1}{\sigma-1}\eta, \tag{14}$$

   where $\overline{q}^{\sigma-1} = \alpha\lambda^{\sigma-1} + (1-\alpha)\,\overline{\omega}^{\sigma-1}$ (see (5)).

3. *The production share ℓ^P and the variety intensity $\mathcal{N}$ are given by*

$$\mathcal{N} = \frac{\varphi_E\left(\mu-1\right)\overline{q}^{\sigma-1}}{\rho+\delta+\left(\mu\overline{q}^{\sigma-1}-(1-\alpha)\right)\frac{1}{\alpha}\tau} \quad \textit{and} \quad \ell^P = \frac{\rho+\delta+\left(\overline{q}^{\sigma-1}-(1-\alpha)\right)\frac{1}{\alpha}\tau}{\rho+\delta+\left(\mu\overline{q}^{\sigma-1}-(1-\alpha)\right)\frac{1}{\alpha}\tau}. \tag{15}$$

Proof. See Section A-1.1.2 in the Appendix. □

Proposition 1 contains three key theoretical results. First, a decline in population growth reduces variety creation, creative destruction and entry. In particular, because each firm only produces a single product, the entry rate $\mathcal{E}$ is trivially given by $\mathcal{E} = z$ and hence declines as population growth falls.

Second, aggregate growth g_y depends directly on the rate of population growth η. First, population growth determines variety creation ($\frac{1}{\sigma-1}\eta$). Second, population growth also affects creative destruction τ and hence the rate of efficiency growth g^Q. Although the effect of population growth on variety growth is always positive, its effect on efficiency growth depends on the average efficiency of newly created products $\overline{\omega}$ and the increment of creative destruction λ. The overall effect on income growth, however, is unambiguous. Upon substituting the expression for τ in (13), the change in income growth with respect to population growth is given by

$$\frac{dg^y}{d\eta} = \frac{\overline{q}^{\sigma-1}-\alpha}{\left(\sigma-1\right)\left(1-\alpha\right)} > 0.$$

Hence, as is typical in models of semi-endogenous growth, falling population growth reduces long-run income growth. Note that the relationship between population growth and income growth is determined by α, σ and $\overline{q}$, and hence governed by parameters that, as we show below, can be disciplined from firm-level data. For example, Jones (2021) or

Bloom et al. (2020) show that $g^y = \frac{1}{\beta}\eta$, where β parametrizes the extent to which ideas are getting harder to find. Hence, as far as the relationship between income and population growth is concerned, our model implies that $\frac{1}{\beta} = \frac{\overline{q}^{\sigma-1}-\alpha}{(\sigma-1)(1-\alpha)}$.

Equation (14) also highlights that growth is not bound to be zero if population growth is zero. First, the "vertical dimension" of life-cycle growth I is a source of aggregate efficiency growth. Second, even if $\eta = 0$, τ and ν are positive to replace products that become obsolete. If $\overline{q}^{\sigma-1} > 1$, such newly created products are (on average) better then exiting products. This form of selection is an additional force pushing for growth to be positive even if the population.is stable.

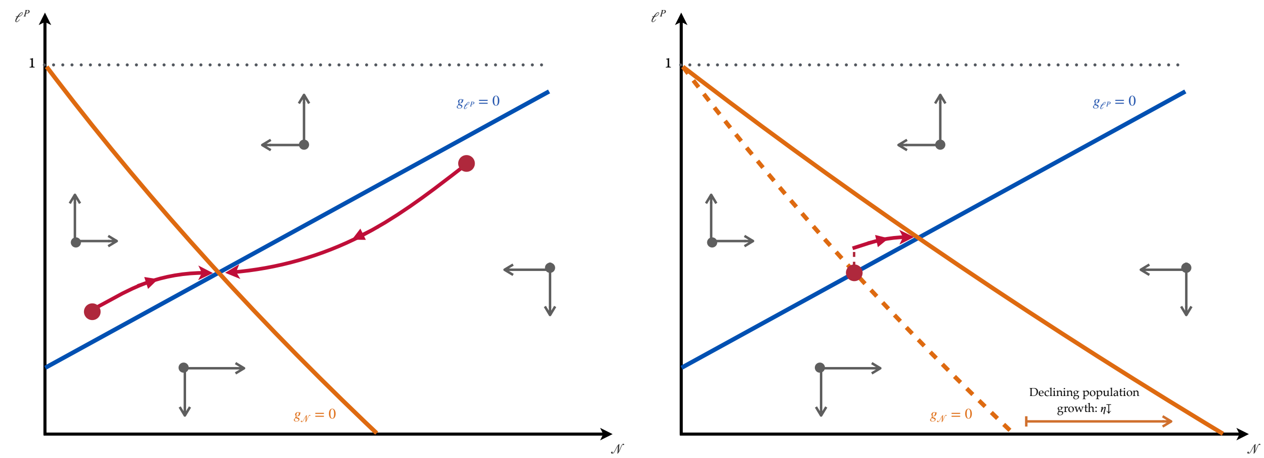
Third, the *level* of varieties relative to the population, $\mathcal{N}$, and the share of workers allocated to research, ℓ^P, are also functions of population growth η, which enters through the rate of creative destruction.. As seen in equation (15), a decline in population growth (and hence creative destruction) increases the variety intensity $\mathcal{N}$ if $\mu\overline{q}^{\sigma-1} > 1-\alpha$, that is if λ, α, μ and $\overline{\omega}$ are sufficiently large. To understand the role of this condition, note that $\tau + (\sigma-1)\left(g_Q - I\right)$ appears in the equilibrium discount rate of corporate profits (see (11)). On the one hand, lower population growth reduces creative destruction τ.This channel increases the value of entry because firms live longer and profits are discounted at a lower rate. On the other hand, lower population growth could increase average efficiency growth g_Q if the average efficiency of new products $\overline{q}^{\sigma-1}$ is sufficiently low. This channel would lower the value of entry because firms face more competition during their life-time. As long as $\mu\overline{q}^{\sigma-1} > 1-\alpha$ (which is the case for our estimated parameters), the creative-destruction effect dominates the efficiency-growth effect, and falling population growth increases the value of entry through a lower rate of discounting. Free entry therefore requires the *level* of flow profits to go down, which is achieved through an increase in the number of varieties per capita $\mathcal{N}$. This increase in the variety intensity is a static countervailing force to the negative growth implications of falling population growth.

Finally, (14) and (15) highlight that our model features weak scale effects: the population size L and the cost of entry φ_E do not affect the growth rate, but only the mass of varieties N_t. Hence, changes in the scale of the economy or the efficiency of product creation have level effect, not growth effects.

Transitional Dynamics. Equations (10) and (11) not only describe the BGP, but the entire equilibrium path. We can characterize this path with a phase diagram depicted in the left panel of Figure 2. The downward sloping schedule shown in orange depicts the locus of a stable variety intensity ($g_{\mathcal{N}} = 0$). This locus follows from the resource constraint: if ℓ_t^P is

Figure 2: Equilibrium path of $\{\ell_t^P, \mathcal{N}_t\}$

(a) Phase diagram (b) Dynamics after a fall in population growth



Note: The left panel shows the phase diagram for the equilibrium path of $(\ell^P, \mathcal{N})$. The right panel shows the response to a fall in population growth.

too high (low), there is too little (much) production creation and the variety intensity falls (rises). The upward sloping schedule shown in blue represents the locus of a constant production share ($g_{\ell^P} = 0$) and summarizes the free entry condition. If the variety intensity is too high (low), there is little creative destruction and the production share, and with it flow profits, has to fall (rise) to satisfy free entry. Hence, there is a unique stable arm (shown in red), that takes the economy to the BGP characterized in Proposition 1.

This phase diagram is not only useful to establish the stability and uniqueness of the equilibrium path, but also to analyze the impact of a fall in population growth. This experiment is shown in the right panel of Figure 2. A fall in population growth rotates the orange locus to the right: for a given variety intensity there are too many workers employed in the research sector, given that the entry rate has to fall eventually.[10] Hence, on impact, the production share ℓ_t^P jumps up. During the transition there is a continual rise in the variety intensity and a reallocation of workers out of the research sector. This initial rise in the production of goods and the increase in the number of products per capita constitute a source of welfare gains, especially in the short-run.

[10] In Section A-1.1.4 in the Appendix we show that we can express the free entry condition shown in blue without reference to η.

2.3 Product Creation by Incumbent Firms

So far, we assumed that entrants are the sole source of product creation. We now extend our model to also allow incumbents to engage in these activities. Doing so is crucial to understand the link between population growth and firm-dynamics. Note that in absence of incumbent innovation, firms only produce a single product and exit at rate $\tau + \delta$ irrespective of their size and age. Both of these predictions are, of course, empirically counterfactual. By allowing for product creation by incumbent firms, we capture the fact that firms, on average, grow as they age and that exit rates are declining in size and age.[11] Crucially, we show that falling population growth increases firms' life-cycle growth and raises product concentration (as documented in Figure 1). At the same time, we also show that the aggregate implications of population growth contained in Proposition 1 survive entirely unchanged.

Suppose that, in addition to the vertical dimension of efficiency growth I, firms can also grow horizontally by adding new products to their portfolio. Following Klette and Kortum (2004), we assume that firms choose the Poisson rate X with which they expand into new product lines. Such expansion activities are costly, and we denote these costs (in units of labor) as

$$c_t^X(X, n) = \frac{1}{\varphi_x} X^{\zeta} n^{1-\zeta} = \frac{1}{\varphi_x} x^{\zeta} n, \tag{16}$$

where $\zeta > 1$, n denotes the number of products the firm is currently producing and $x = X/n$ is the firms' innovation intensity. Conditional on successfully creating a product innovation, we treat incumbents entirely symmetrically to entrants, that is, a fraction α of new ideas improve upon an existing product and a fraction $1 - \alpha$ yields a new variety, whose average efficiency is $\overline{\omega} Q_t$. Letting $x_t = \frac{1}{N_t} \int x_{it} di$ denote the average expansion intensity by incumbent firms, the aggregate amounts of variety creation ν and creative destruction τ are given by

$$\nu_t = (1-\alpha)(z_t + x_t) \quad \text{and} \quad \tau_t = \alpha (z_t + x_t), \tag{17}$$

and thus reflect the activities of both entrants and incumbents.

Allowing for active innovative behavior by incumbent firms naturally changes the value

[11] If product creation is only due to entrants, firm growth tends to be negative. Note that employment is proportional to $(q/Q)^{\sigma-1}$ and thus grows at rate $(\sigma-1)(I - g_Q) = -\left(\overline{q}^{\sigma-1} - 1\right) z$. Hence, in the empirically relevant case, where product innovation contributes positively to average efficiency growth, that is $\overline{q}^{\sigma-1} > 1$, average employment declines conditional on survival.

function. Because firms can now produce multiple products, let $\{q_{fi}\}_{i\in\Theta_f}$ be the state variables at the firm-level and $V_t\left(\{q_{fi}\}_{i\in\Theta_f}\right)$ be the corresponding value function. As we show in Section A-1.1 in the Appendix, this value function is additively separable across products, i.e. $V_t\left(\{q_{fi}\}_{i\in\Theta_f}\right) = \sum_{i=1}^{n_f} V_t(q_i)$, and the product-level value function $V_t(q)$ solves the HJB equation

$$r_t V_t(q) - \dot{V}_t(q) = \pi_t(q) + \frac{\partial V_t(q)}{\partial q} Iq - (\tau_t + \delta) V_t(q) + \max_x \left\{ x V_t^{PC} - \frac{x^\zeta w_t}{\varphi_x} \right\}, \tag{18}$$

where the value of production creation V_t^{PC} is still given in (7).

Compared to the value function in the entry-only model (6), the possibility of expanding horizontally carries an option value, which is determined by the endogenous creation value V_t^{PC} and the expansion costs $\frac{1}{\varphi_x} x^\zeta w_t$. However, the solution to (18) is again very similar to the case without incumbent innovation. In particular, V_t is still homogenous in $q^{\sigma-1}$, so that $V_t^{PC} = V_t(\bar{q}Q_t)$.

Proposition 2. *The optimal rate of product innovation x is given by*

$$x = \left(\frac{\varphi_x}{\zeta} \frac{V_t(\bar{q}Q_t)}{w_t}\right)^{\frac{1}{\zeta-1}} = \left(\frac{1}{\zeta}\frac{\varphi_x}{\varphi_E}\right)^{\frac{1}{\zeta-1}}, \tag{19}$$

where $V_t(q_i)$ is defined in (18). On a BGP, $V_t(q)$ is given by

$$V_t(q) = \underbrace{\frac{\pi_t(q)}{\rho + \tau + \delta + (\sigma-1)(g^Q - I)}}_{\text{Production value}} + \underbrace{\frac{1}{\rho+\tau+\delta}\frac{\zeta-1}{\varphi_x} x^\zeta w_t}_{\text{Innovation value}}. \tag{20}$$

Proof. See Section A-1.1 in the Appendix. □

Proposition 2 contains two important results. First, (19) shows that the equilibrium rate of incumbent product innovation x is constant and a function of technological parameters only. It is independent of *any* general equilibrium variables and, in particular, does not depend on the rate of population growth η. The reason is that the free entry condition in (7) still applies, which ties the ex-ante value of product innovation $V_t(\bar{q}Q_t)$ to the entry costs $\frac{1}{\varphi_E} w_t$. Economically, it follows from the fact that incumbents' innovation technology has decreasing returns at the firm level, whereas entry, which operates at the aggregate level, has constant returns.[12] Hence, the free-entry condition pins down the value of

[12]Note that incumbent product creation also has constant returns in the aggregate: if the number of firms

product creation, and incumbent firms optimally choose the rate of product creation to equalize the marginal cost and the marginal benefits. This also implies that equation (19) holds both on and off the BGP and relies only on the free-entry condition to be binding.

Second, along a BGP, the value function $V_t(q)$ has an explicit solution. It is the sum of the net present value of flow profits (the "production value") and the option value of product innovation (the "innovation value"). The production value is exactly the same as in the entry-only model. The innovation value is, of course, absent in the entry-only model. If incumbent innovation gets prohibitively expensive, i.e. $\varphi_x \to 0$, the solution in (20) coincides with (9).

With the results of Proposition 2 in hand, it is also immediate that most results of Proposition 1 apply without any change in the presence of incumbent product creation. Importantly, however, the *composition* of creative destruction and variety creation between entrants and incumbents is now endogenous and depends on population growth.

Proposition 3. *Consider the model with product creation by incumbent firms. Then,*

1. *The expressions for creative destruction τ, variety creation ν and aggregate growth g^y are exactly the same as in the "entry-only" model characterized in Proposition 1;*
2. *The rates of entry z and incumbent product creation x are given by*

$$x = \left(\frac{1}{\zeta}\frac{\varphi_x}{\varphi_E}\right)^{\frac{1}{\zeta-1}} \quad \text{and} \quad z = \frac{\eta+\delta}{1-\alpha} - x. \tag{21}$$

Proof. See Section A-1.1.6 in the Appendix. There we also derive the equilibrium conditions for $\mathcal{N}$ and ℓ^P, which are very similar to (15). □

Proposition 3 highlights a key theoretical result of our analysis: allowing for product innovation by incumbents does not change *any* aggregate outcomes but only makes the *composition* of product creation and creative destruction endogenous. Crucially, because x is independent of population growth, the entirety of the decline in population growth is absorbed by the economy's extensive margin: Entrants do all the work and incumbents are insulated from demographics. This implies that the share of creative destruction and

were to double, the amount of product creation performed by incumbents would also double. In Section 2.6 below, we generalize our results to the case in which the entry process has decreasing returns in the aggregate. In that case, x is also affected by population growth.

variety creation due to incumbents, s_x, is given by (see (17) and (21))

$$s_x \equiv \frac{(1-\alpha)\,x}{\nu} = \frac{\alpha x}{\tau} = \frac{1-\alpha}{\eta+\delta}x = \frac{1-\alpha}{\eta+\delta}\left(\frac{1}{\zeta}\frac{\varphi_x}{\varphi_E}\right)^{\frac{1}{\zeta-1}}. \tag{22}$$

Naturally, incumbents' share of innovative activity is increasing in their relative efficiency φ_x/φ_E. More importantly, it is declining in population growth η. This compositional change, whereby falling population growth increases incumbent innovation *relative* to entrants is a key aspect of how population growth changes the process of firm dynamics and the firm-size distribution - see Section 2.4 below.

One key implication of our theory is that scale effects are absent for both the rate of growth and the equilibrium size distribution: along a BGP, the employment distribution is stationary and fully determined from the entry flow z, the rate of product innovation by incumbents x, and the rate of own-innovation I, all of which are independent of the level of the population L_t.[13] This symmetry for the role of scale effects is not a coincidence. The standard model of Klette and Kortum (2004) features strong scale effects and implies that the firm size distribution depends on the size the population.[14] By contrast, our model implies that differences in the level of the population affect the mass of varieties N_t, leaving the process of firm-dynamics and aggregate growth unchanged. In that sense, our model is akin to Young (1998), augmented with a full endogenous process of firm dynamics. Population growth is therefore consistent with a BGP and a stationary firm size distribution. Given that empirically the distribution of firm size is reasonably stable when compared to the large changes in the size of the population, firm-based models of growth point towards a world of semi-endogenous growth.

2.4 Population Growth and Firm Dynamics

To see how population growth affects the process of firm dynamics, recall that firms gain products at rate x and lose products at rate $\tau+\delta$. Then define the *net* rate of product

[13]This result does not hinge on taking I to be exogenous. In Section 2.6 we treat I as endogenous and show that it is independent of level of the population.

[14]The Klette and Kortum (2004) model is nested in our framework. It is a parametrization where the population is constant (i.e. $\eta = 0$), there is no own-innovation (i.e. $I = 0$) and the mass of varieties is exogenous (i.e. $\alpha = 1$ and $\delta = 0$). Incumbents' innovation efforts are still constant and given by (19). The rate of entry is given by $z = \tau - x$ and efficiency grows at rate $g^Q = \frac{\lambda^{\sigma-1}-1}{\sigma-1}\tau$. Creative destruction τ and the production share ℓ^P are then determined from the free entry condition and the resource constraint. It is easy to show that an increase in L increases z and hence τ. A larger population thus increases growth and creative destruction and reduces firm size.

accumulation $\psi \equiv x - (\tau + \delta)$. Using (13) to express τ in terms of the rate of population growth η yields $\psi = x - \frac{\alpha\eta+\delta}{1-\alpha}$: a decline in population growth increases the net rate of product accumulation, because firms face less of a threat of creative destruction.

Let $S(a)$ denote the share of firms that survive until age a, and $\overline{n}(a)$ the average number of products of a firm of age a, i.e. the firm's "product life cycle". As we show in Section A-1.1.8 in the Appendix, $S(a)$ and $\overline{n}(a)$ are given by

$$S(a) = \frac{\psi e^{\psi a}}{\psi - x(1 - e^{\psi a})} \quad \text{and} \quad \overline{n}(a) = 1 - \frac{x}{\psi}\left(1 - e^{\psi a}\right). \tag{23}$$

In Figure 3, we display $S(a)$ and $\overline{n}(a)$ graphically. Naturally, $S(a)$ is declining and satisfies $\lim_{a\to\infty} S(a) = 0$, because all firms exit eventually. Similarly, $\overline{n}(a)$ is increasing because surviving firms are selected on having had many successful product innovations and little creative destruction. More importantly, lower population growth *increases* firms' survival rates and *raises* the life-cycle profile of product growth. Hence, falling population growth increases the concentration of the product space in two ways. First, because lower population growth reduces creative destruction, it increases firms' chances of survival. As a consequence the age distribution shifts to the right. This increases concentration because older firms are larger. Second, as highlighted by the shift in the $\overline{n}(a)$ schedule, lower population growth also increases firm size *conditional* on age because, in line with the pattern shown in Figure 1, firms accumulate more products as they age. Both of these mechanisms push towards rising concentration and larger firms.

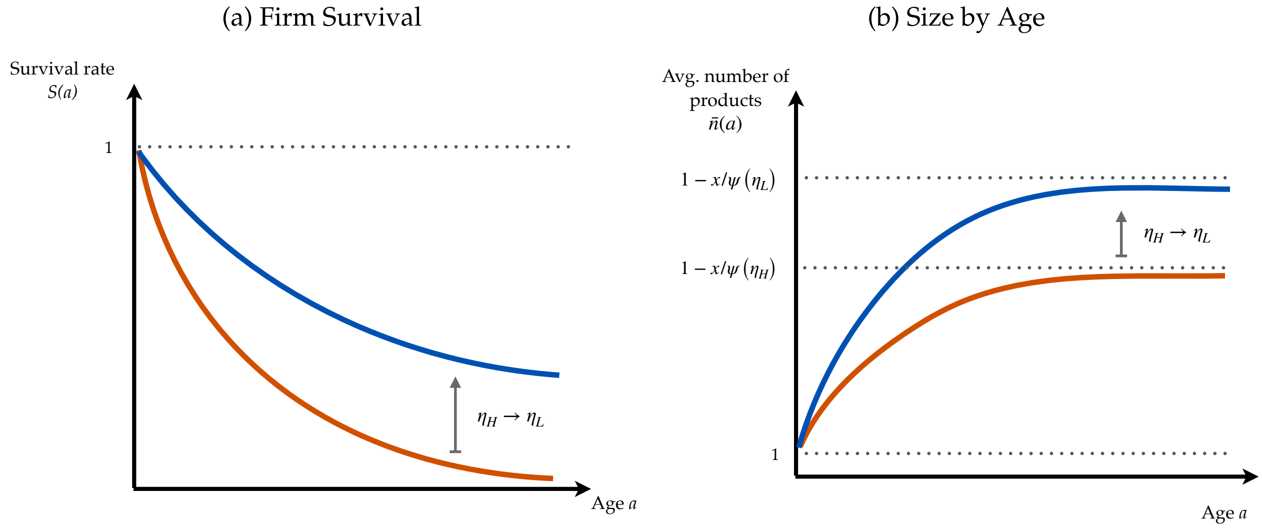
To see this link between population growth and product concentration formally, in Section A-1.1.8 in the Appendix we formally characterize the product distribution. In particular, using the results of Luttmer (2011) and Cao et al. (2017), we show that, as as long as $\eta > \psi > 0$, the distribution of the number of products n_f has a Pareto tail ϱ_n, which is given by

$$\varrho_n = \frac{\eta}{\psi} = \frac{(1-\alpha)\eta}{x(1-\alpha) - \delta - \alpha\eta}. \tag{24}$$

Hence, the tail of the product distribution is a closed-form expression of the rate of population growth η, and a decline in population growth reduces ϱ_n and increases concentration.

Equation (24) highlights that population growth affects the product distribution through two channels. Holding firms' net expansion rate ψ constant, lower population growth increases concentration because it reduces the rate at which new firms, which are, on average, small, enter. In addition, lower population growth endogenously increases the net

Figure 3: Falling Population Growth and Rising Concentration



Note: The figure shows the relationship between population growth η and firms' survival probabilities $S(a)$ in the left panel and the relationship between population growth η and the average number of products $\overline{n}(a)$ in the right panel.

accumulation rate ψ and concentration. Interestingly, this increase in product concentration goes hand in hand with an *increase* in the number of varieties per person $\mathcal{N}_t = N_t/L_t$: even though population growth reduces the number of firms per worker, it increases the number of products per worker because each existing firm offers a larger product portfolio. Hence, higher concentration at the firm-level can coexist with an expansion of product variety.

Even though population growth affects the tail of the product distribution, it might not affect the tail of the employment distribution. Recall that firm-level employment is given by

$$l_{ft} = \sum_{i=1}^{n_f} \left(\frac{q_i}{Q_t}\right)^{\sigma-1} \times \frac{L_t^P}{N_t} = \sum_{i=1}^{n_f} \left(\frac{q_i}{Q_t}\right)^{\sigma-1} \times \frac{\ell^P}{\mathcal{N}}, \tag{25}$$

that is, the firm-size distribution depends on the distribution of both the number of products n and of scaled efficiency q/Q. The tail of the employment distribution is thus given by $\varrho_l = \min\left\{\varrho_n, \frac{1}{\sigma-1}\varrho_q\right\}$, where ϱ_n is the tail of the product distribution given in (24) and ϱ_q is the tail of the scaled efficiency distribution. Intuitively, firms can be large in two ways: by having many products, or by having an extraordinarily good product.

As we show in Section A-1.1.8 in the Appendix, the distribution of relative efficiency also has a Pareto tail and the tail parameter ϱ_q depends on λ, α, σ, and $\overline{q}$ and is implicitly defined by:

$$\varrho_q \left(\frac{\overline{q}^{\sigma-1} - 1}{\sigma - 1} \right) = -1 + \alpha \lambda^{\varrho_q}. \tag{26}$$

Crucially, and in stark contrast to (24), the tail of the efficiency distribution ϱ_q is independent of population growth η. Even though declining population growth always increases concentration, whether this increase also shows up in the tail of the size distribution depends on the comparison of ϱ_n and ϱ_q. If $\varrho_n < \varrho_q$, the product distribution is the dominating force and lower population growth increases the thickness of the tail of the employment distribution. If $\varrho_n > \varrho_q$, the tail of the employment distribution is determined from the distribution of relative efficiency and independent of the rate of population growth. Which of these cases prevails is a quantitative question.

2.5 The Mechanism: Demand or Supply?

Falling population growth impacts the economy both through a decline in labor supply and via aggregate demand. It is thus natural to ask whether the resulting changes in firm dynamics and aggregate growth contained in Propositions 1 and 3 reflect supply, demand or both.

To distinguish the supply and demand channel, suppose there is a second sector that has access to a production technology $Y_t = \mathcal{A}_t H_t$, where H_t denotes the number of workers in sector 2 and $\mathcal{A}_t$ grows at an exogenous rate $g^{\mathcal{A}}$. The mass of sector-2-workers H_t grows at rate η^H and they can only provide labor to sector 2. All individuals have identical intra-temporal Cobb Douglas preferences $c_t = c_{1t}^{\vartheta} c_{2t}^{1-\vartheta}$, and spend a share ϑ on goods of sector 1. This setup, which can be also be thought of as an open economy extension where sector 2 is a foreign country, allows us to independently vary aggregate demand and labor supply. Changes in population growth of sector-1-workers, η, still have supply and demand effects. By contrast, population growth of sector-2-workers, η^H, has no effect on labor supply in sector 1.

In Section SM-2 in the Supplementary Material we characterize this model in detail and derive the analogue of Propositions 1 and 3 in this more general environment. First, we show that creative destruction τ, variety creationν, entry z, and incumbent innovation x are *exactly* the same as in our baseline economy. Hence, neither $g^{\mathcal{A}}$ nor η^H have any effect on these outcomes and, as a a consequence, on the resulting firm-size distribution.

By contrast, real consumption growth of workers in sector 1 is now

$$g^c = \vartheta \left(g^Q + \frac{1}{\sigma - 1} g^N \right) + (1 - \vartheta) \left(g^{\mathcal{A}} + \eta^H - \eta \right),$$

that is, an increase in $g^{\mathcal{A}}$ or η^H raises consumption growth through cheaper relative prices. These results highlight that the relationship between population growth, firm-dynamics and local productivity growth is a *supply side* phenomenon. However, domestic consumers can be somewhat shielded from falling local population growth by global productivity or population growth.

2.6 Discussion of Assumptions

Three assumptions made our theory particularly tractable. First, we assumed a linear entry technology. Second, product creation was undirected: a constant share α of product innovation results in creative destruction rather than new-variety creation. Third, we took the rate of own-innovation I to be exogenous. In this section, we show that our main results do not hinge on these modeling choices - see Section A-1.2 in the Appendix for details.

Decreasing Returns in the Entry Technology. Assume the productivity of entrant labor hired to produce new ideas is given by $\varphi_E(z_t) = \tilde{\varphi}_E z_t^{-\chi}$. Here, z_t is the aggregate entry rate that each entrant takes as given. For $\chi = 0$, this specification yields the constant-returns case analyzed above. For $\chi > 0$, the cost of entry rises with the aggregate entry rate. Free entry requires that $V_t(\bar{q}Q_t) = \frac{1}{\varphi_E(z_t)} w_t = \frac{1}{\tilde{\varphi}_E} z_t^{\chi} w_t$. Hence, the aggregate entry supply curve is increasing with an elasticity $1/\chi$.

Propositions 1 and 3 extend to this case in a straightforward case. Along a BGP, the rate of variety growth is still tied to the rate of population growth, i.e. $\nu = \eta + \delta$. This directly implies that the rate of creative destruction τ and the aggregate growth rate g^y are exactly the same as in Propositions 1 and 3. By contrast, the composition of product innovation depends on the strength of congestion χ. As we show in Section A-1.2 in the Appendix, z and x are then determined from

$$\frac{\eta + \delta}{1 - \alpha} = z + x \quad \text{and} \quad x = \left(\frac{\varphi_x}{\zeta \tilde{\varphi}_E} \right)^{\frac{1}{\zeta - 1}} z^{\frac{\chi}{\zeta - 1}}.$$

Comparing this expression with (21) in Proposition 3 highlights that the entry elasticity

χ determines how entry and incumbent innovation co-move. As long as $\zeta - 1 > \chi$, incumbents' innovation share $x/(z+x)$ is declining in η as in (22) and all our qualitative results apply. Below we show that the canonical free entry specification $\chi = 0$ provides a good quantitative fit to the data.

Endogenizing the Direction of Innovation α. Our second assumption concerns the direction of innovation α. In Section A-1.2 in the Appendix, we present a detailed analysis of an extension of our model, where entrants and incumbents can directly choose the flow rate at which they want to creatively destroy products (x_{CD} and z_{CD}) and at which they want to create new varieties (x_{NV} and z_{NV}).

This extension of our model is still very tractable. First, the value function takes a similar form to (20) in Proposition 2. Second, letting φ_N and φ_{CD} denote the relative costs of new variety creation and creative destruction, the optimal rates of incumbent innovation are given by

$$x_{NV} = \left(\frac{\varphi_N}{\zeta}\frac{V_t\left(\overline{\omega}Q_t\right)}{w_t}\right)^{\frac{1}{\zeta-1}} \text{ and } x_{CD} = \left(\frac{\varphi_{CD}}{\zeta}\frac{V_t\left(\lambda Q_t\right)}{w_t}\right)^{\frac{1}{\zeta-1}}. \tag{27}$$

Third, because entering firms have the same innovation technology as incumbents, $z_{NV} = x_{VN}z$ and $z_{CD} = x_{CD}z$, where z is the aggregate flow of entry.

Equation (27) highlights why variety creation and creative destruction are tightly linked: both depend on the same value function $V_t(q)/w_t$. In fact, in the special case where $\lambda = \overline{\omega}$, this model is exactly isomorphic to our baseline model because equation (27) implies $\frac{\alpha}{1-\alpha} = \frac{x_{CD}}{x_{NV}} = \left(\varphi_{CD}/\varphi_N\right)^{1/(\zeta-1)}$, and the expression for creative destruction τ is unchanged. In Section A-1.2 in the Appendix, we also analyze the general case of $\lambda \neq \overline{\omega}$, which implies α is no longer constant. However, quantitatively, falling population growth still reduces both creative destruction and the relative importance of entrants.

Endogenous Own-Innovation I**.** Our results also extend seamlessly to the case where I is endogenous. Because the proofs of Propositions 1 and 3 never used any properties of firms' vertical innovation choices I, the expressions for τ, z, x and g_y are exactly the same as in Propositions 1 and 3, except that I is no longer a parameter but a choice variable. In Section A-1.2 in the Appendix we solve for I explicitly, and show that I depends on the creative destruction rate τ and parameters other than population growth. Moreover, I is increasing in τ, that is, a decline in population growth *reduces* the rate of own-innovation. The endogenous response of incumbents' own-innovation efforts thus amplifies the negative growth consequences of falling population growth.[15]

[15]That I is increasing in τ seems surprising given that higher creative destruction reduces the expected

3 Endogenous Market Power

So far, we have assumed markups were constant and equal to the standard CES markup. We now generalize our model by assuming firms compete *a la* Bertrand within product lines. Doing so makes the distribution of markups endogenous, and allows us to study the effects of falling population growth on market power.

Given the CES structure of demand, each firm would like to charge a markup of $\frac{\sigma}{\sigma-1}$ over marginal cost. However, the presence of competing firms within their product line implies the most efficient producer might have to resort to limit pricing. The markup charged in product i, μ_i, is thus given by $\mu_i = \min\left\{\frac{\sigma}{\sigma-1}, \Delta_i\right\}$, where q_i^C is the efficiency of the next best competitor, and $\Delta_i \equiv q_i/q_i^C > 1$ is the firm's efficiency advantage relative to it competitors (we also refer to this as the "gap").

The static equilibrium allocations generalize in a straightforward way. Aggregate output and equilibrium wages are now given by $Y_t = \mathcal{M}_t Q_t N_t^{\frac{1}{\sigma-1}} L_t^P$ and $w_t = \Lambda_t Y_t / L_t^P$, where

$$\mathcal{M}_t = \frac{\left(\int \mu^{1-\sigma} \left(q/Q_t\right)^{\sigma-1} dF_t\left(q,\mu\right)\right)^{\frac{\sigma}{\sigma-1}}}{\int \mu^{-\sigma} \left(q/Q_t\right)^{\sigma-1} dF_t\left(q,\mu\right)} \quad \text{and} \quad \Lambda_t = \frac{\int \mu^{-\sigma} \left(q/Q_t\right)^{\sigma-1} dF_t\left(q,\mu\right)}{\int \mu^{1-\sigma} \left(q/Q_t\right)^{\sigma-1} dF_t\left(q,\mu\right)},$$

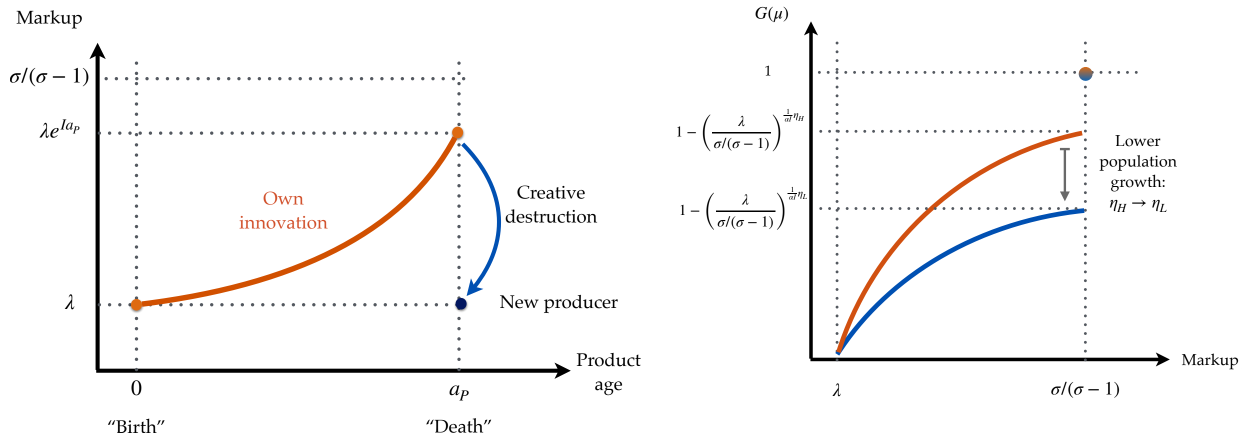
and $F_t(q,\mu)$ denotes the joint distribution of efficiency and markups. The two aggregate statistics $\mathcal{M}_t$ and Λ_t fully summarize the static macroeconomic consequences of monopoly power. Market power reduces both production efficiency (the misallocation term $\mathcal{M}_t$) and lowers factor prices relative to their social marginal product (the labor wedge Λ_t). Because the joint distribution of markups and efficiency, $F_t\left(q,\mu\right)$, is a function of the rate of population growth, declining population growth affects both $\mathcal{M}_t$ and Λ_t.

Most of our theoretical results directly carry over to this more general environment. Most importantly, the key results of Propositions 1 and 3 *exactly* hold in the model with Bertrand competition: the expressions for creative destruction τ and variety creation ν, the aggregate rate of growth g_y and entry and incumbent innovation z and x still hold in an environment with imperfect product markets and as a consequence the link between population growth and product concentration applies unchanged. In addition, although

life-span. Holding L_t^P/N_t fixed, this intuition is indeed correct. However, once the change in L_t^P/N_t is taken into account, the effect of creative destruction becomes positive. The reason is that free entry requires the average production value plus the innovation value to be equal to the entry costs. A lower rate of population growth increases the innovation value and reduces the production value. And as the returns to own-innovation only scale with the production value, they are lower in an environment with lower population growth.

Figure 4: Falling Population Growth and Rising Market Power

(a) The Life-Cycle of Product Markups (b) The Distribution of Markups



Notes: This left panel shows a stylized example of how markups evolve at the product level. When a firm takes over a product, markups increase through own-innovation. Once the product is lost to another firm, markups are reset to the baseline level of λ. The right panel shows what happens to the distribution of markups when population growth falls.

the value function is more involved, we show in Section A-1.3 in the Appendix that we can still derive an analytic expression that has a similar form to the one derived in the constant markup case.

Allowing for imperfect competition, however, yields additional insights, because our model features a crucial asymmetry between productivity growth due to creative destruction and own-innovation. Suppose the current producer of product i has an efficiency gap of Δ_i. If this firm is replaced by another producer, the efficiency gaps *reduces* to λ as the new firm's efficiency exceeds the one of the previous producer by the creative-destruction step size λ. By contrast, if the existing firm increases its efficiency through own-innovation, the markup increases at rate I (as long as $\Delta_i \leq \frac{\sigma}{\sigma-1}$). Therefore, own-innovation is akin to a positive drift for the evolution of markups, whereas creative destruction is similar to a "reset" shock, which keeps the accumulation of market power in check. This process is displayed in the left panel of Figure 4.

This stochastic process gives rise to a stationary distribution of markups. Newly created varieties do not face any competitor and charge a markup of $\frac{\sigma}{\sigma-1}$. Products that have been creatively destroyed at some point in the past are subject to Bertrand competition and the markup depends on Δ. Let N_t^{NC} denote the mass of products without any competitor, and let $N_t^C = N_t - N_t^{NC}$ be the mass of products that are subject to competition. In Section A-1.3.1 in the Appendix, we prove two results. First, we show that, along a BGP, the share of product without any competitor is given by $N_t^{NC}/N_t = 1-\alpha$; that is, it is simply given by the share of product creation that results in new varieties. Second, the marginal distribution of efficiency gaps among products with a competitor is a Pareto distribution with a tail parameter of $(\tau+\eta+\delta)/I$:

$$F^C(\Delta) = 1 - (\lambda/\Delta)^{\frac{\tau+\eta+\delta}{I}} = 1 - (\lambda/\Delta)^{\frac{1}{1-\alpha}\frac{\eta+\delta}{I}}, \tag{28}$$

where the second equality uses that $\tau = \frac{\alpha}{1-\alpha}(\eta+\delta)$. Equation (28) highlights that slower population growth increases the distribution of efficiency gaps in a first-order stochastic dominance sense. First, slower population growth shifts the product distribution toward old products, which, on average, have higher markups. In addition, because slower population growth also reduces creative destruction and product churning, this effect is amplified: the average efficiency gap is increasing even for a given cohort of firms.

To translate efficiency gaps into markups, recall that $\mu(\Delta) = \min\left\{\frac{\sigma}{\sigma-1}, \Delta\right\}$. A reduction in population growth therefore increases markups along the whole distribution and shifts more mass towards the maximum CES markup. Because higher markups reduce the labor share Λ and more dispersed markups reduce allocative efficiency $\mathcal{M}$, lower population growth tends to increase profits relative to factor payments and has adverse effects on static allocation efficiency. In the right panel of Figure 4, we depict how the distribution of markups changes in response to a decline in population growth from η_H to η_L.

The macroeconomic consequences of misallocation are summarized by $\mathcal{M}$ and Λ, which depend on the joint distribution between efficiency gaps Δ and efficiency q. To derive this distribution, define relative efficiency $\hat{q} = \ln(q/Q_t)^{\sigma-1}$ and let $\hat{\lambda} = \ln\lambda^{\sigma-1}$. Denote $F_t^C(\Delta,\hat{q})$ as the joint distribution of efficiency gaps and relative efficiency for products that have a next-best competitor. Similarly, denote $F_t^{NC}(\hat{q})$ as the distribution of relative efficiency for products that do not have a competitor. We show in Section A-1.3.1 in the Appendix that these objects satisfy the differential equations

$$\frac{\partial F_t^C(\Delta,\hat{q})}{\partial t} = \underbrace{-\frac{\partial F_t^C(\Delta,\hat{q})}{\partial\Delta} I\Delta - (\sigma-1)(I-g_t^Q)\frac{\partial F_t^C(\Delta,\hat{q})}{\partial\hat{q}}}_{\text{drift from own innovation}} \underbrace{-\left(\tau_t+\delta+\eta\right)F_t\left(\Delta,\hat{q}\right)}_{\text{product loss}}$$
$$+ \underbrace{\lim_{s\to\infty}\tau_t F_t^C\left(s,\hat{q}-\hat{\lambda}\right)}_{\text{creative destruction of } C \text{ products}} + \underbrace{\tau_t\frac{N_t^{NC}}{N_t^C}F_t^{NC}(\hat{q}-\hat{\lambda})}_{\text{creative destruction of } NC \text{ products}},$$

$$\frac{\partial F_t^{NC}(\hat{q})}{\partial t} = \underbrace{-\frac{\partial F_t^{NC}(\hat{q})}{\partial \hat{q}}(\sigma-1)\left(I-g_t^{Q}\right)}_{\text{drift from own innovation}} - \underbrace{(\tau_t+\delta+\eta)\,F_t^{NC}(\hat{q})}_{\text{product loss}} + \underbrace{\frac{(1-\alpha)}{\alpha}\tau_t\Gamma\left(\frac{\exp(\hat{q})}{\sigma-1}\right)}_{\text{new products}}.$$

These expressions highlight the separate roles of own-innovation and creative destruction in influencing the evolution of efficiency and markups. Own-innovation causes both production efficiency and the gap to drift upwards at the deterministic rate I, whereas creative destruction "resets" the mass in the distribution above Δ to have a gap of λ. Below, we quantify the impact of falling population growth on $F_t(\Delta, \hat{q})$ computationally both along the BGP and during the transition.

4 Quantitative Analysis: Calibration

To quantify the importance of declining population growth, we now calibrate our model to data from the US. We parametrize the model to a balanced growth path matching key moments of the data in 1980, when labor force growth was approximately 2%. We then study the aggregate impact of the historical and projected decline in population growth since 1980 by computing the dynamic response in our model.

4.1 Data

Our main dataset is the US Census Longitudinal Business Database (LBD). The LBD is an administrative dataset containing information on the universe of employer establishments since 1978. It contains information on the age, industry, employment, and payroll of each establishment, along with identifiers at the firm level that allow us to track the ownership of each establishment over time. We define the age of the firm in the LBD as the age of the oldest establishment that the firm owns. The birth of a new firm requires both a new firm ID in the Census and a new establishment record.

To measure firms' markups, we require information on sales. We augment the LBD data with information on firm revenue from administrative data contained in the Census Bureau's Business Register, following Moreira (2015) and Haltiwanger et al. (2016). The Business Register is the master list of establishments and firms for the US Census, and we are able to match approximately 70% of the records to the LBD.

In Table A-1, we provide some basic summary statistics on the firms in our dataset. In total our data comprises about 3.61 million firms in 1980 and 4.95 million firms in 2010.

During that time period, concentration rose substantially. Average firm employment increased by around 10%, from 20 to 22 employees, the aggregate employment share of firms with less than 20 employees declined from 21.5% to 18.8%, and that of very large firms (with more than 10,000 employees) increased from 25.7% to 27%. Furthermore, firms became substantially older: the employment share of firms less than five years old declined from 38% to 30%. Qualitatively, all these facts are implications of our theory. Below we show the observed decline in population growth goes a long way toward replicating these patterns quantitatively.

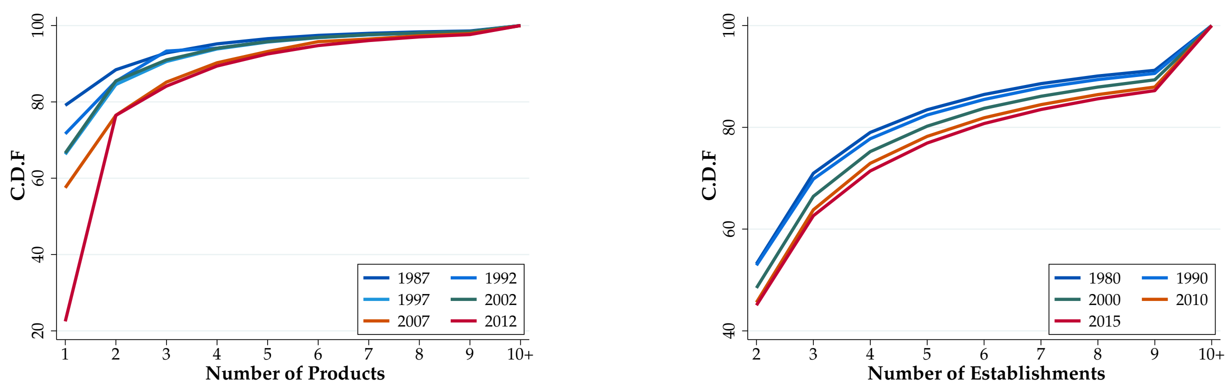
In addition, we use data on the number of products per firm from the Census of Manufacturers between 1987 and 2012. As part of the survey, firms self-report the type of products they sell and the Census Bureau assigns these products to 10-digit NAICS categories. We only use the simple count of products per firm.

4.2 Product Concentration: Direct Evidence

A key implication of our theory is that falling population growth reduces creative destruction and increases product concentration. In Figure 1 we already showed that the importance of multi-product firms in the US manufacturing sector rose sharply since the 1980s. In this section we show that this trend is apparent for the whole product distribution and that it is not unique to the manufacturing sector.

In the left panel of Figure 5 we depict the distribution of the number of products per manufacturing in different years, using the Census of Manufacturing. As our theory predicts, the distributions of products per firm shifted steadily outwards in a first-order stochastic dominance sense. We also extend this analysis to sectors outside of manufacturing, using data from the LBD. To be able to observe a 'product', we focus on firms that provide non-tradable services, such as restaurants, retail, construction, or accommodations. From the point of view of such firms, each distinct establishment in space serving a distinct market (Starbucks on 116th street vs. Starbucks in Madison Square) can be thought of as a distinct product in our theory. In the right panel, we plot the distribution of the number of establishments per firm, focusing on firms with at least two establishments. As for products in manufacturing, the plant distribution in the non-tradable service sector shifted steadily to the right. While in 1980, more than 50% of multi-establishment non-tradable firms had only two establishments, this share declined by 10 percentage points by 2015. By contrast, the share of firms with more than 10 establishments increased by roughly

Figure 5: Rising Product Concentration



Note: Panel (a) shows the number of products per firm in the Census of Manufacturing. Panel (b) displays the number of establishments per firm in non-tradable industries (Restaurants, Retail, Construction, Accommodation) from the LBD.

5%. Through the lens of our theory, these patterns are a natural consequence of declining creative destruction by entrants relative to incumbents and a key implication of declining population growth.

4.3 Calibration

Our model is parsimoniously parametrized and rests on 11 parameters:

$$\Psi = \left\{ \underbrace{\alpha, \zeta, \varphi_E, \varphi_x, I, \bar{\omega}, \lambda}_{\text{Innovation \& Entry technology}}, \underbrace{\delta}_{\text{Exog. exit}}, \underbrace{\eta}_{\text{Pop. growth}}, \underbrace{\rho, \sigma}_{\text{Preferences}} \right\}.$$

We set three of them exogenously. We fix the elasticity of substitution between products σ at 4, following Garcia-Macia et al. (2019), set the discount rate ρ to 0.95, and assume a quadratic innovation cost function (i.e. $\zeta = 2$) as in Acemoglu et al. (2018).

The rate of labor force growth η is directly observed in the data and is our key parameter for the comparative statics. The remaining seven parameters are calibrated internally. First, we target three moments from the cross-sectional size distribution in 1980: the entry rate, average firm size and the Pareto tail of the employment distribution. Second, we utilize two moments of firm-growth, namely the dynamics of sales and markups over firms' life-cycle. Finally, we rely on two aggregate moments: the aggregate growth rate and the average markup. In Table 1 we report the parameters and the main moments

Table 1: Model Parameters

	Structural Parameters Description	Value	Moments	Data	Model
η	Labor force growth in 1980	0.02	Data from BLS	2%	2%
λ	Step size on quality ladder	1.11	Aggregate poductivity growth	2%	2%
I	Rate of own innovation	0.023	Markup growth by age 10 (RevLBD)	10.2%	10.2%
φ_X	Cost of inc. product creation	0.04	Sales growth by age 10 (RevLBD)	58%	58%
φ_E	Cost of entry	0.12	Avg. firm size (BDS)	20.7	20.7
δ	Exogenous rate product death	0.06	Entry rate in 1980 (BDS)	11.6 %	11.6 %
α	Share of creative destruction	0.59	Average profit share	25%	25%
$\bar{\omega}$	Relative efficiency of new products	0.45	Pareto tail of LBD employment distribution in 1980	1.1	1.1
ζ	Curvature of innovation cost	2	Set exogenously		
σ	Demand elasticity	4	Set exogenously		
ρ	Discount rate	0.05	Set exogenously		

Note: This table reports the calibrated parameters for the full model. Data for the firm lifecycle comes from the US Census Longitudinal Database, augmented with revenues from tax-information using the Census Bureau's Business Register. Data for average firm size and the firm entry rate in 1980 are taken from the public-use Business Dynamics Statistics.

we target. While all moments are targeted simultaneously, there is nevertheless a tight mapping between particular moments and parameters which highlights how the different parameters are identified.

Innovation efficiency of incumbent firms: I and φ_x. We identify the relative efficiency of vertical own-innovation and horizontal expansion from the life-cycle profiles of sales and markups. Because markup growth is driven by incumbents' own-innovation activities (see Figure 4), this moment is informative about the rate of efficiency improvement I. Sales growth is additionally affected by the rate of incumbent product creation, which depends directly on the cost of product expansion φ_x.

As we show in detail in Section A-2.2 in the Appendix, we can derive the life-cycle profiles of sales and markups (essentially) explicitly. This is not only convenient from a quantitative standpoint but also clarifies our identification strategy. The main insight in deriving these moments is to first express markups and sales of a given product as a function of the "product age" a_P, that is the amount of time a given product has been produced by the same firm. Average sales as a function of product age a_P are given by

$$s_P(a_P) \equiv E\left[\left.\frac{p_i y_i}{Y}\right| a_p\right] = E\left[\mu_i^{1-\sigma}\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\middle| a_p\right] = \mu(a_p)^{1-\sigma} e^{(\sigma-1)\left(I-g^Q\right)a_p}\overline{q}^{\sigma-1},$$

where $\mu(a_P) = \min\left\{\frac{\sigma}{\sigma-1}, \Delta(a_P)\right\} = \min\left\{\frac{\sigma}{\sigma-1}, \lambda e^{Ia_P}\right\}$, and the remaining terms are average relative quality.[16]

With this expression in hand, we can calculate the life-cycle of sales and markups at the *firm*-level. Average sales and markups as a function of firm age a_f are given by

$$s_f\left(a_f\right) = E\left[\sum_{n=1}^{N_f} s_P(a_P)\middle| a_f\right] \quad \text{and} \quad \mu_f\left(a_f\right) = E\left[\left(\sum_{i=1}^{N_f}\mu(a_p)^{-1}\frac{s_P(a_P)}{\sum_{i=1}^{N_f} s_P(a_P)}\right)^{-1}\middle| a_f\right],$$

where the expectations are taken with respect to the conditional distribution of N_f and a_P, conditional on a_f. As we show in Section A-2.2 in the Appendix, we can calculate these conditional distributions essentially explicitly. This allows us to compute the life-cycle profiles of sales and markups without having to simulate the model.

Empirically, we measure markups at the firm level by the inverse labor share $\mu_f = \frac{py_f}{wl_f}$, where py_f is the total revenue of the firm, and wl_f is the total wage bill. While this approach allows us, in principle, to measure markups for the population of US firms, we only use firms' markup *growth* to calibrate our model. More specifically, letting $\mu_{f,t}$ be the markup of firm f at time t, we run a regression of the form

$$\ln \mu_{f,t} = \sum_{a=0}^{20} \gamma_a^{\mu} \mathbb{I}_{Age_{ft}=a} + \theta_f + \theta_t + \epsilon_{f,t}, \tag{29}$$

where $\mathbb{I}_{Age_{ft}=a}$ is an indicator for whether the firm is of age a and θ_f and θ_t are firm and time fixed effects respectively. Hence, γ_a^{μ} provides a non-parametric estimate of the rate of markup growth. We calibrate our model to the growth rate at the 10-year horizon, γ_{10}^{μ}. Because we explicitly control for a firm fixed effect when estimating (29), we do not have to take a stand on firms' output elasticities as long as they are constant with age.

We follow the same approach when we estimate the life-cycle of sales; that is, we also estimate (29) using log sales as the dependent variable and target γ_{10}^{py} in our quantitative

[16]Because own-quality q increases at rate I while average quality Q increases at rate g^Q, $e^{(\sigma-1)\left(I-g^Q\right)a_p}$ is the relative drift of these random variables. The last term reflects the initial average quality when the firm adds the product to its portfolio.

model. In the LBD, firms increase their average markup by roughly 10 percentage points and grow in size by about 80% by age 10.

Entry costs and product obsolescence: φ_E and δ. We choose φ_E and δ to jointly match the entry rate and average firm size. The free condition determines market size L_t^P/N_t as a function of entry efficiency φ_E. This in turn is a key component of average firm employment. We thus choose φ_E to match an average firm employment of 20.04 in 1980 from the BDS. The exogenous rate of obsolescence δ directly influences the exit and hence - in a BGP - the entry rate of firms. We target the entry rate in 1980 of 11.6%.

Productivity growth through innovation: λ and $\overline{\omega}$. The parameters λ and $\overline{\omega}$ determine the relative quality of creatively destroyed products and newly generated varieties. We infer these parameters from the aggregate growth rate and the tail of the firm size distribution. That λ and $\overline{\omega}$ directly affect the growth rate is apparent from Proposition 1. For the tail of the firm size distribution, we find in our calibration that $\varsigma_n > \frac{1}{\sigma-1}\varsigma_q$, i.e. the tail of the employment distribution is given by $\frac{1}{\sigma-1}\varsigma_q$, where ς_q is given in (26). Given α and σ, this tail only depends on λ and $\overline{\omega}$. For our calibration, we chose λ and $\bar{\omega}$ to target a rate of productivity growth of 2% and a tail parameter of the firm size distribution of 1.1 (Luttmer, 2007).[17]

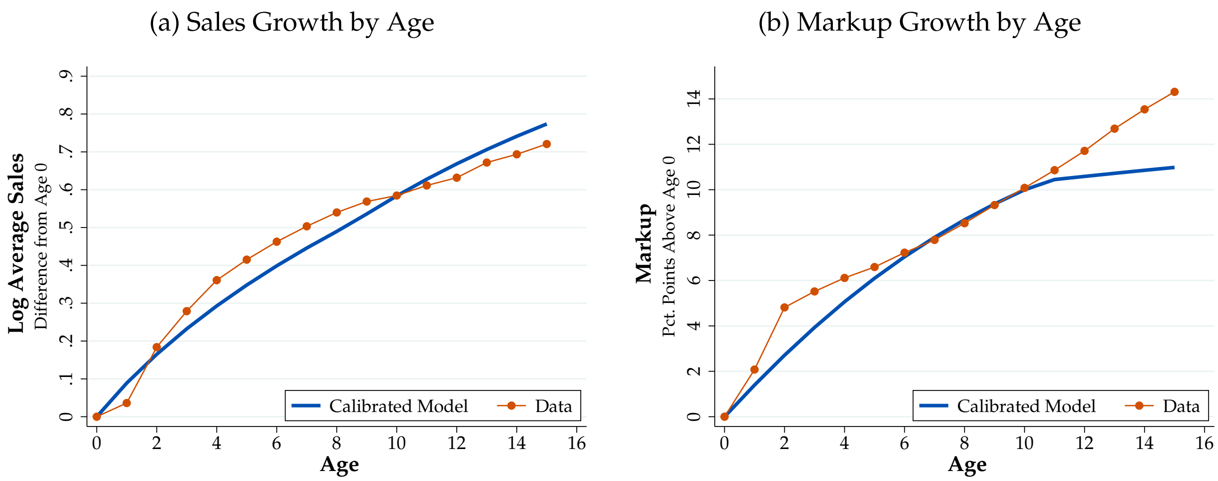
New varieties vs. creative destruction: α. The share of new products in innovation, $1-\alpha$, plays an important role for the level of markups in the economy. The higher α, the lower the economy-wide markup, because the higher the share of products that are subject to Bertrand competition. We target an economy-wide profit share of 25%.

4.4 Estimates and Model Fit

As seen in Table 1, our model is able to match the targeted moments perfectly. To rationalize the fact that firm-level markups grow by around 10 percentage points at the 10 year horizon, our model implies a rate of own-innovation of around 2.3%. For a creative destruction event, we estimate a productivity increase of 11%. This is required to match an annual aggregate growth rate of 2%. On average, about 60% of product innovations result in creative destruction, and 40% generate a new variety. The initial ion efficiency of these new products, $\overline{\omega}$, is estimated to be low, about 50% of the average product in the economy. This relatively low value is required to match the thickness of the tail of

[17] While none of the moments we target depend on higher moments of the initial efficiency distribution $\Gamma(\omega)$, the shape of $F(\hat{q}, \Delta)$ is affected by this choice. We choose Γ to be a Pareto with mean $\overline{\omega}$ and a tail index of 4, which rationalizes the relatively low dispersion of entrant size in the LBD.

Figure 6: Lifecycle Growth in Firm Sales and Markups



Note: Panel (a) in this figure compares the lifecycle of firm sales in the model to the estimated lifecycle in the data. The data lifecycle plots the age coefficients from estimating equation (29) in the LBD. The sample size is 35, 300, 000, where this number has been rounded to accord with Census Bureau disclosure rules. Panel (b) does the same for relative markups.

the employment distribution. These estimates imply that the average efficiency of new product, $\overline{q}$, is given by 0.95.

These parameters also determine the long-run growth impact of falling population growth:

$$dg_y/d\eta = \frac{\overline{q}^{\sigma-1} - \alpha}{(\sigma-1)(1-\alpha)} \approx 0.2. \tag{30}$$

Hence, a one percentage point decline in the rate of population growth reduces economic growth by 0.2 percentage point. This result is consistent with the findings of Bloom et al. (2020). In their model, $dg_y/d\eta = 1/\beta$ where β is the "degree of diminishing returns" (i.e. $\dot{A}_t/A_t = A_t^{-\beta} L_t^R$). They estimate $1/\beta = 0.33$ for the aggregate economy.

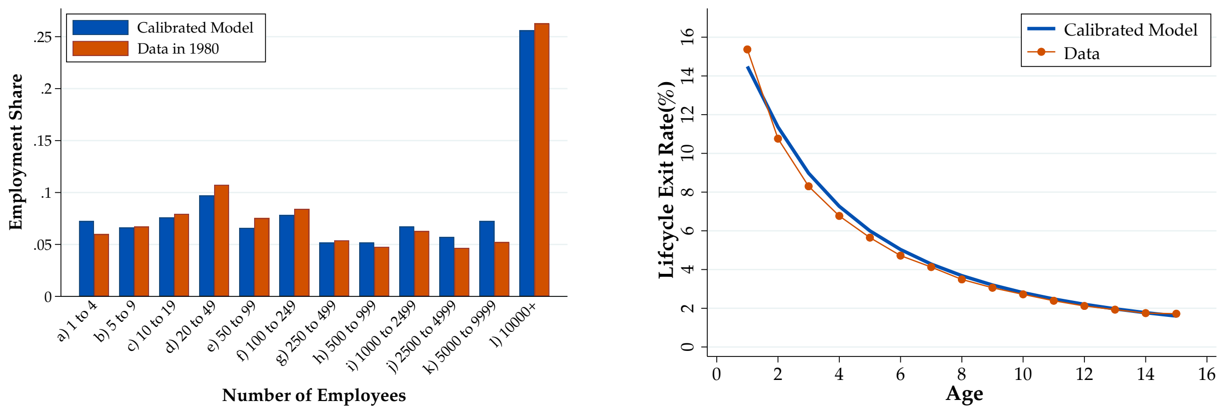
Our model also matches a variety of additional non-targeted moments despite its parsimonious parametrization. Consider first the sales and markup life cycle. In Figure 6 we show the model's performance by plotting the estimated coefficients γ_a^{μ} and γ_a^{py} from specification (29) estimated in the model and in the data. As highlighted in Table 1, we calibrate our model to match γ_{10}^{py} and γ_{10}^{μ}. Figure 6 shows that the model's implication for the whole age profile of sales (in the left panel) and markups (in the right panel) is quite close to what is observed in the data.[18]

In Figure 7, we confront our model's predictions for the size distribution and firms' exit

[18]In Figure A-2 in the Appendix we show the joint density of markups and efficiency, illustrating the positive correlation between markups and efficiency induced by survival and own-innovation.

Figure 7: Size Distribution and Exit Hazards: Model vs Data

(a) Size Distribution (b) Exit Rates By Age



Notes: This figure plots the employment shares by firm size (left panel) and the lifecycle exit rates (right panel) in the calibrated model (blue) and the data in 1980 (orange). The data in the left panel is from the BDS. The exit rates by age in the data are taken from the increments in a Kaplan-Meier survival function estimated on all firms in the LBD born between 1980 and 1990.

hazard with the data. Whereas we have explicitly targeted average size and the Pareto tail, the left panel shows that our model matches the full non-parametric firm size distribution very well.[19] Note, in particular, that it replicates the aggregate importance of very large firms with more than 1000 employees, which account for 25% of aggregate employment. A central reason our model successfully replicates the firm-size distribution is that it provides a good fit for the empirically observed exit hazards, despite the fact that we do not target them in the estimation - see right panel of Figure 7. In our theory, exit rates are declining in age because older firms have more product lines, and owning more products progressively lowers the likelihood that they will all be destroyed within a particular year.[20]

[19]For replicability we chose size bins that are also available in the publicly available data from the BDS.

[20]In Section A-2.3 in the Appendix we also analyze our model's predictions for firms' exit rates by size. Our model overestimates the extent to which large firms exit. The reason is that (in our calibration) the thick tail of the employment distribution is driven by the distribution of product quality. Hence, large firms are firms with a few superstar products, not those with many products. And because creative destruction is independent of product quality, such firms are as likely to exit as other firms. In Section SM-3 in the Supplementary Material we also show how we can extend our model to allow for type heterogeneity along the lines of Sterk et al. (2021) to address this counterfactual prediction.

5 The Aggregate Impact of Falling Population Growth

We now use our model to quantify the consequences of falling population growth shown on Figure 1, which, for convenience, we replicate in panel (a) of Figure 8. While the US labor force grew by 2% in 1980, this rate declined to about 1% two decades later. The official BLS projections paint an even direr picture for the years ahead: by 2050 it projects the US labor force to grow at around 0.25%. To study the implications of this trend, we start with the calibrated BGP in 1980, and then feed this path of population growth into the model, assuming that all agents have rational expectations about this path. All other parameters are held constant.

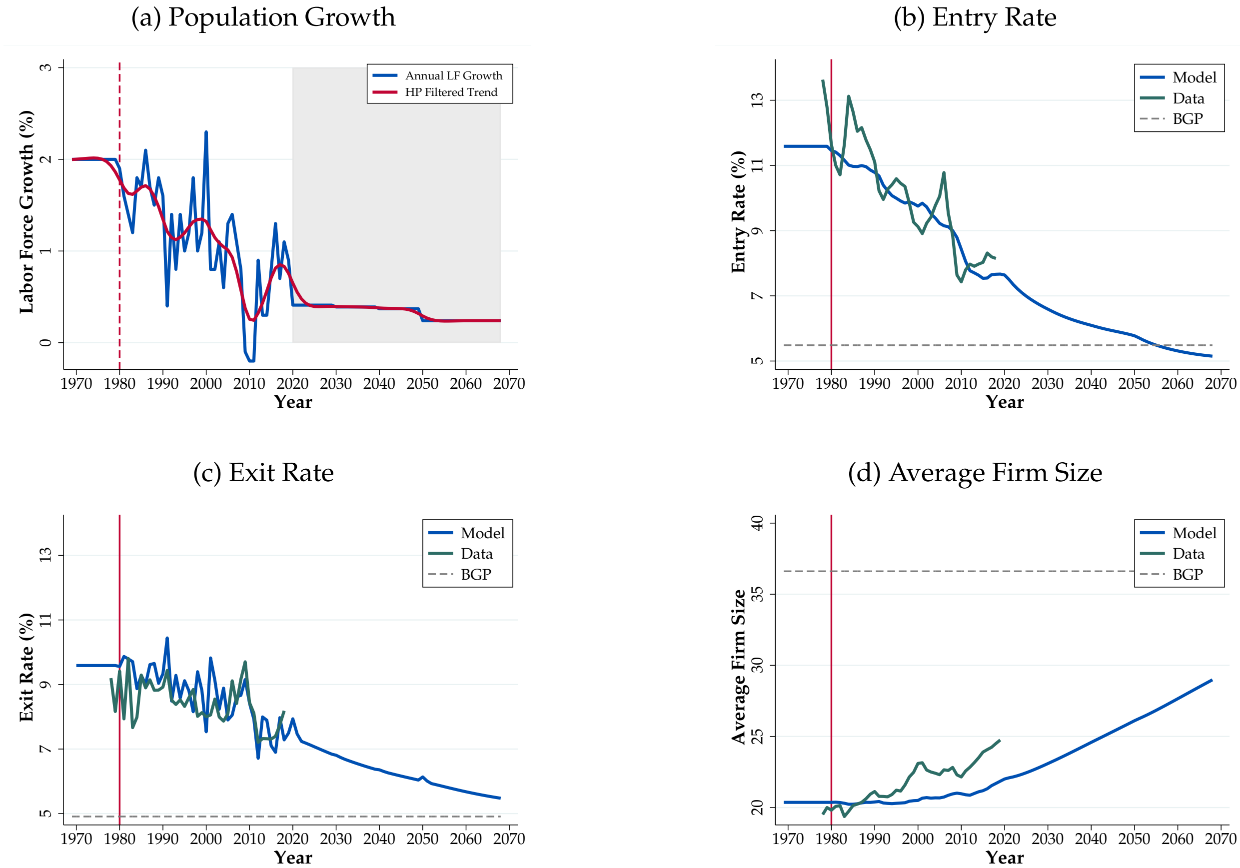
5.1 Declining Population Growth and Changing Firm Dynamics

We start by considering the impact on firm dynamics. We focus first on the entry rate, the exit rate and average firm size. In Figure 8, we plot both the data and the implications of our theory. Consider first the data, shown in green. The entry rate, shown in panel (b), declined markedly in the last 30 years from around 12% in the 1980s to around 8% in the mid 2000s. Note this series of the entry rate tracks the evolution of population growth closely, and the contemporaneous correlation is 0.74. Similarly, the exit rate, shown in panel (c), declined from 9% in 1980 to almost 7% in 2015. Average firm size, shown in panel (d), rose from 20 to 23 employees, increasing by around 15%.

In blue, we superimpose the predictions of our theory. Recall that we used both the entry rate and average size in 1980 as a calibration target, and hence match these numbers by construction. The exit rate, by contrast, is not targeted. The subsequent fall in entry and exit and the rise in average size are the sole consequence of the observed and projected decline in population growth.

Figure 8 shows that the decline in population growth goes a long way toward explaining the observed trends. For the entry and exit rates, our model matches the US experience almost perfectly. For average size, our model predicts a somewhat slower increase than what is observed in the data, and highlights that the long-run increase in firm size will take many decades to settle at a higher value once labor force growth stabilizes. The increase in concentration is also similar to what is observed in the data, with the employment share of large firms (defined by the BDS to be 10,000 employees or more) increasing by 1% by 2015, roughly in line with the data (see Table A-1).

Figure 8: Declining Population Growth and Changing Firm Dynamics



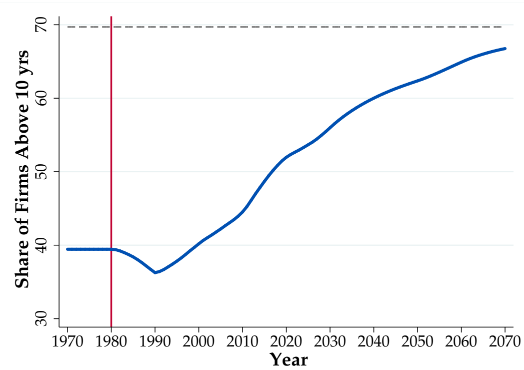
Note: Panel (a) displays the growth rate of the labor force in the US, with the raw series in blue and a HP-filtered trend component in red. The data is sourced from the BLS, accessed through FRED. Grey shading indicates projections. The remaining panels show the the dynamic response of the entry rate (panel (a)), the exit rate (panel (b)), and average firm size (panel (c)) to the path of population growth shown in panel (a) . The same objects from the BDS data are shown in green.

In Figure 8, we only display the implications of our theory until 2070. Given the population growth path shown in panel (a), our model has not reached a new BGP at this point. Hence, we also plot the long-run implications as dashed lines. The entry and exit rates adjust relatively quickly and are already quite close to their long-run BGP values by 2070. By contrast, our model predicts that average size has some way left to run due to the slow-moving firm size distribution and will increase substantially in the long-run.

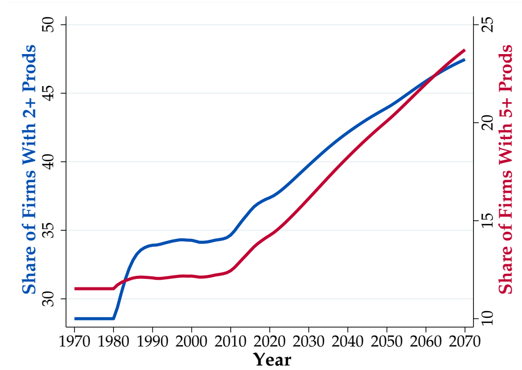
As also highlighted in Hopenhayn et al. (2018), the four series shown in Figure 8 are linked via a fundamental accounting equation: population growth is the sum of the entry rate $\mathcal{E}_t$ and the change in average firm size $\mathcal{S}_t$ minus the exit rate $\mathcal{X}_t$, $\eta_t = \dot{\mathcal{S}}_t/\mathcal{S}_t + \mathcal{E}_t - \mathcal{X}_t$. Because our theory correctly predicts an increase in average size and a fall in the exit rate, the decline in the entry rate is substantially *larger* than the decline in population growth.

Figure 9: Declining Population Growth, Firm Aging and Product Concentration

(a) The shifting age distribution



(b) Rising Product Concentration



Note: The left panel shows the share of firms older than 10 years old. The right panel shows the share of firms with at least two and at least five products.

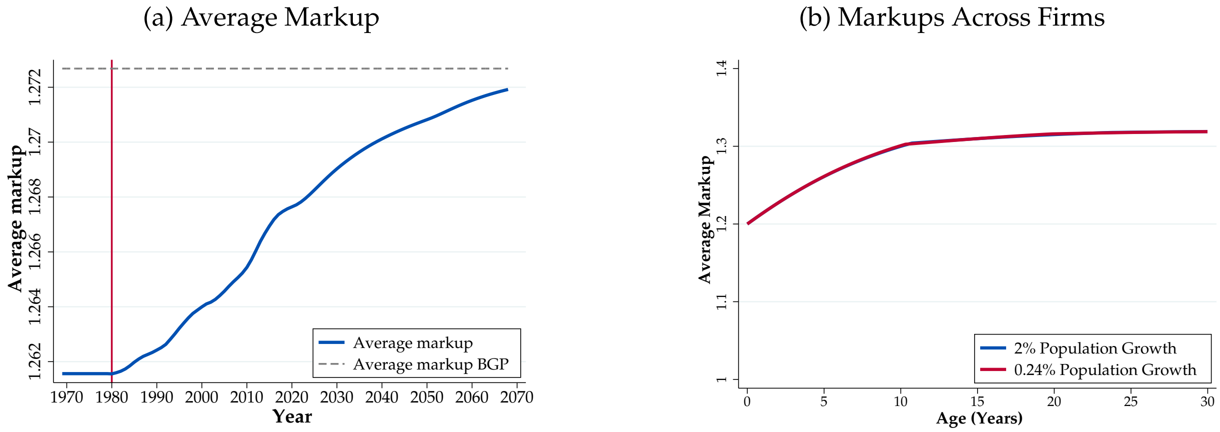
Along a BGP, where average size is constant, $\eta = \mathcal{E} - \mathcal{X}$. The fact that, empirically, the gap between the entry rate and the exit declined markedly since 1980, is a further indication that falling population growth is a key determinant of the observed changes in the US firm size distribution.

In Figure 9 we trace out the evolution of the age distribution and the rise in product concentration. In the left panel we depict the share of firms older than ten years. Falling population growth leads to firm aging as the fall in entry reduces the inflow of young firms and lower creative destruction increases firms' chance of survival. Quantitatively, our model suggests that the share of old firms increases from 40% to around 70% in the new BGP. The US economy has experienced a fair amount of firm aging, and the quantitative magnitude of the change appears similar.[21] In the right panel, we depicts the rise in product concentration. We focus on the two statistics shown in Figure 1: the share of firms with at least two and five products. Even though the units are not directly comparable, like in the data, both increase steadily in response to falling population growth. This rise in product concentration is an immediate reflection of the decline in creative destruction, which allows incumbent firms to accumulate products at a faster rate as they age.

Our model predicts that average size is increasing both because of a shift in the age distribution towards older firms and because lower population growth increases firm size conditional on age. Quantitatively, however, much of the increase depicted in Figure 8 is

[21]Because the age information in the LBD is censored in 1977, we cannot directly compute the share of firms older than 10 years old in 1980. Between 1990 and 2010, the share of firms older than ten years increased by 22 percentage points, from 29% to 51%.

Figure 10: Declining Population Growth and Rising Market Power



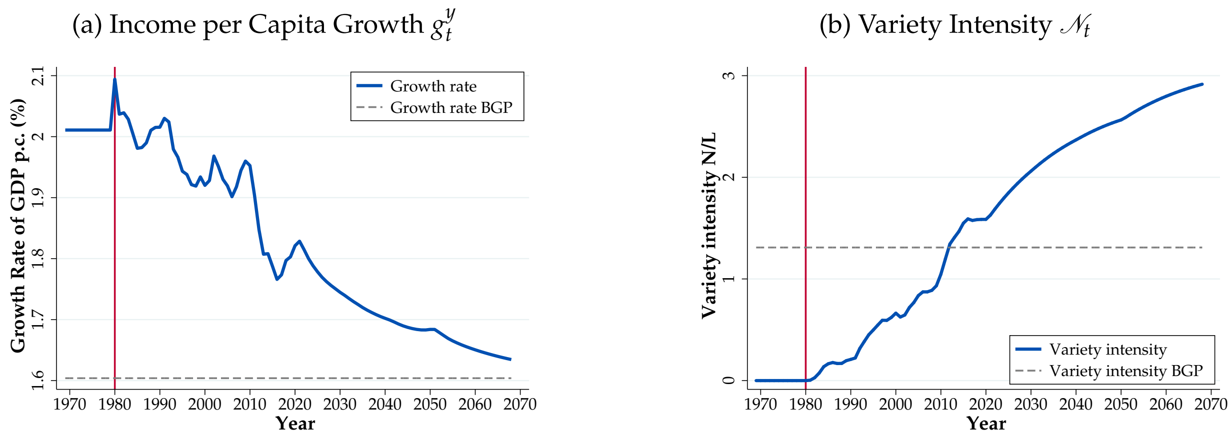
Notes: The left panel shows the transition path of the average product markup as labor force growth changes according to the path in Figure 8. The right panel shows average markups by age before and after the transition.

due to shifts in the age distribution even though firms' innovation spending is endogenous and responds to changes in population growth. In Figure A-4 in the Appendix, we show the change in sales growth and exit by age. These objects do change as population growth declines, but only modestly. This dominant role of the age distribution is consistent with the data, where size or exit rates by age also changed little (see Karahan et al. (2019) and Hopenhayn et al. (2018)).

Finally, in Figure 10, we report the implied change in the cost-weighted average markup. As implied by our theoretical results, the decline in population growth increases markups. Quantitatively, the increase in market power is modest: the average product markup increases by about 1%.

As shown in the right panel, our model implies this increase in markups occurs mostly across firms and is a reflection of the fact that firms become older. Within firms, products tend to become older because products are destroyed less frequently. On its own, this fact raises average markups. However, firms also accumulate more products, which are, on average, younger, and hence have lower markups. Quantitatively, these two forces almost exactly offset one another, leaving the life-cycle of markups essentially unchanged. Hence, as for the effects on firm size and exit, the rise in markups reflects compositional changes whereby large and old firms with high markups increase their market share. This pattern is qualitatively consistent with the findings reported in Kehrig and Vincent (2017) and Autor et al. (2020).

Figure 11: Declining Population Growth and Income per Capita



Note: The figure displays the dynamic response of the aggregate growth rate (left panel) and the variety intensity $\mathcal{N}_t = N_t/L_t$ to the path of population growth shown in Figure 8.

5.2 Declining Population Growth, Income per Capita and Welfare

We now turn to the normative implications. In Figure 11 we depict the growth rate of income per capita (left panel) and the change in the variety intensity $\mathcal{N}_t$ (right panel). Interestingly, the effect of population growth on output growth is not monotone. On impact, a population growth decline increases output growth for about one decade. This is due to an increase in the variety intensity $\mathcal{N}_t$: even though the number of firms per worker declines, the mass of products available to consumers increases by about 30% in the long-run. This amounts to a static productivity increase of about $\frac{1}{\sigma-1}\ln(1.3) \approx$ 9%. But because this increase in variety is transitory, output growth eventually declines and stabilizes at a lower level as in most models of semi-endogenous growth. In our calibration, the long-run growth rate declines from around 2% to 1.6%, consistent with (30).[22] This decline in growth stems to a large extent from falling variety growth. In fact, efficiency growth g^Q rises slightly in response to the decline in population growth, because we estimate the efficiency of new varieties $\overline{\omega}$ to be relatively low. The creative destruction hazard τ declines from about 20% to 17%.

The competing forces of rising variety and declining growth make the welfare consequences of falling population growth in principle ambiguous. We measure welfare in consumption equivalent terms, asking by how much would we need to change the level

[22]We also conducted our analysis for the case of endogenous own-innovation I. As shown in Section 2.6, this amplifies the effect of population growth on productivity growth. Quantitatively, we find that long-run growth declines to 1.2% instead of 1.6%.

of consumption per capita in each period in the old BGP to achieve the same level of welfare as the transition for a member of the representative household who discounts the future at rate ρ. We find that the long-run decline in economic growth dominates: falling population growth amounts to a 1.9% reduction in the level of consumption each period.

6 Conclusion

Most countries have experienced declining rates of fertility and a slowdown in population growth in recent decades. There is little reason to think this trend is going to reverse any time soon.

In this paper we have shown that this trend has important implications for the process of firm dynamics and aggregate productivity. We analyzed a model of semi-endogenous growth with multi-product firms that is rich enough to rationalize many first-order features of the micro-data, but nevertheless lends itself to an analytical characterization. Declining population growth reduces creative destruction and product creation, causes a decline in entry, and increases average firm size, markups and market concentration. Furthermore, lower population growth reduces economic growth in the long-run, but has positive effects on productivity in the short-run though a surge in variety creation.

In our application to the US, we draw three main quantitative conclusions. First, the population growth channel can account for a large share of the change in entry, exit rates and product concentration since the 1980s, and is thus likely to be an important contributor for the decline in dynamism in the US and the rest of the developed world. Second, even though the decline in population growth is predicted to lower economic growth in the long-run, economic growth remains higher for about one decade. Third, even though lower population growth increases market power and markups, we estimate this effect to be quantitatively small. Hence, the rise in markups and the fall in the labor share are unlikely to be driven by falling fertility.

References

Acemoglu, D. and U. Akcigit (2012): "Intellectual Property Rights Policy, Competition and Innovation," *Journal of the European Economic Association*, 10, 1–42. 4

ACEMOGLU, D., U. AKCIGIT, H. ALP, N. BLOOM, AND W. KERR (2018): "Innovation, reallocation, and growth," *American Economic Review*, 108, 3450–91. 28, SM-6

AGHION, P., A. BERGEAUD, T. BOPPART, P. J. KLENOW, AND H. LI (2019): "A Theory of Falling Growth and Rising Rents," Federal Reserve Bank of San Francisco. 5

AGHION, P. AND P. HOWITT (1992): "A Model of Growth through Creative Destruction," *Econometrica*, 60, 323–351. 4

AKCIGIT, U. AND S. T. ATES (2019a): "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," Tech. rep., National Bureau of Economic Research. 5

——— (2019b): "What Happened to US Business Dynamism?" Tech. rep., National Bureau of Economic Research. 5

AKCIGIT, U. AND W. R. KERR (2018): "Growth through heterogeneous innovations," *Journal of Political Economy*, 126, 1374–1443. 4

ALON, T., D. BERGER, R. DENT, AND B. PUGSLEY (2018): "Older and Slower: The Startup Deficit's Lasting Effects on Aggregate Productivity Growth," *Journal of Monetary Economics*, 93, 68–85. 5

ATKESON, A. AND A. BURSTEIN (2010): "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 118, 433–484. 4, 6

ATKESON, A., A. BURSTEIN, AND M. CHATZIKONSTANTINOU (2018): "Transitional Dynamics in Aggregate Models of Innovative Investment," Tech. rep., National Bureau of Economic Research. 4

AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. VAN REENEN (2020): "The fall of the labor share and the rise of superstar firms," *The Quarterly Journal of Economics*, 135, 645–709. 5, 37

BLOOM, N., C. I. JONES, J. VAN REENEN, AND M. WEBB (2020): "Are Ideas Getting Harder to Find?" *American Economic Review*, 110, 1104–44. 12, 32

BUERA, F. J. AND E. OBERFIELD (2020): "The global diffusion of ideas," *Econometrica*, 88, 83–114. 7

Cao, D., H. R. Hyatt, T. Mukoyama, and E. Sager (2017): "Firm Growth Through New Establishments," *Available at SSRN 3361451*. 4, 18

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*, 135, 561–644. 5

De Ridder, M. (2019): "Market Power and Innovation in the Intangible Economy," *Working Paper*. 5

De Silva, T. and S. Tenreyro (2017): "Population control policies and fertility convergence," *Journal of Economic Perspectives*, 31, 205–28. 1

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014): "The Role of Entrepreneurship in US Job Creation and Economic Dynamism," *Journal of Economic Perspectives*, 28, 3–24. 5

Edmond, C., V. Midrigan, and D. Y. Xu (2018): "How Costly Are Markups?" Tech. rep., National Bureau of Economic Research. 5

Engbom, N. (2017): "Firm and worker dynamics in an aging labor market," *Working Paper*. 4

Garcia-Macia, D., C.-T. Hsieh, and P. J. Klenow (2019): "How Destructive is Innovation?" *Econometrica*, 87, 1507–1541. 4, 6, 28

Grossman, G. and E. Helpman (1991): "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, 58, 43–61. 4

Haltiwanger, J., R. Decker, and R. Jarmin (2015): "Top Ten Signs of Declining Business Dynamism and Entrepreneurship in the US," in *Kauffman Foundation New Entrepreneurial Growth Conference*, Citeseer. 5

Haltiwanger, J., R. Jarmin, R. Kulick, and J. Miranda (2016): "High Growth Young Firms: Contribution to Job, Output and Productivity Growth," *CES 16-49*. 26

Hathaway, I. and R. E. Litan (2014): "What's Driving the Decline in the Firm Formation Rate? A Partial Explanation," *The Brookings Institution*. 3

Hopenhayn, H., J. Neira, and R. Singhania (2018): "From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share," Tech. rep., National Bureau of Economic Research. 4, 35, 37, A-23

JONES, C. (1995): "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 103, 759–784. 4

JONES, C. I. (2021): "The Past and Future of Economic Growth: A Semi-Endogenous Perspective," Tech. rep., National Bureau of Economic Research. 4, 11

KARAHAN, F., B. PUGSLEY, AND A. ŞAHIN (2019): "Demographic origins of the startup deficit," Tech. rep., National Bureau of Economic Research. 3, 4, 5, 37, A-23

KEHRIG, M. AND N. VINCENT (2017): "Growing Productivity Without Growing Wages: The Micro-level Anatomy of the Aggregate Labor Share Decline," *Economic Research Initiatives at Duke (ERID) Working Paper*. 5, 37

KLENOW, P. J. AND H. LI (2021): "Innovative growth accounting," *NBER Macroeconomics Annual*, 35, 245–295. 4

KLETTE, T. J. AND S. KORTUM (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112, 986–1018. 1, 4, 5, 14, 17, A-6

KORTUM, S. S. (1997): "Research, patenting, and technological change," *Econometrica*, 1389–1419. 4

LASHKARI, D., A. BAUER, AND J. BOUSSARD (2019): "Information Technology and Returns to Scale," *Working Paper*. 5

LUTTMER, E. G. (2007): "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 122, 1103–1144. 4, 6, 31

——— (2011): "On the mechanics of firm growth," *The Review of Economic Studies*, 78, 1042–1068. 18, A-8, SM-6

MOREIRA, S. (2015): "Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles," *Working Paper*. 26

OLMSTEAD-RUMSEY, J. (2020): "Market Concentration and the Productivity Slowdown," *Working Paper*. 5

PERETTO, P. F. (1998): "Technological change and population growth," *Journal of Economic Growth*, 3, 283–311. 4

PETERS, M. (2020): "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 88, 2037–2073. 4

Romer, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, 71–102. 4

Sterk, V., P. Sedláček, and B. Pugsley (2021): "The nature of firm growth," *American Economic Review*, 111, 547–79. 33, A-22, SM-6

van Vlokhoven, H. (2021): "The Rise in Profits and Fall in Firm Entry: A Tale of the Life Cycle of Profits," *Available at SSRN 3866852*. 5

Young, A. (1998): "Growth Without Scale Effects," *Journal of Political Economy*, 106, 41–63. 4, 17

Online Appendix for *"Population Growth and Firm Dynamics"*

[FOR ONLINE PUBLICATION]

A-1 Theory

A-1.1 Characterization of the Baseline Model

This section contains the derivation of all results for the baseline model characterized in Section 2. The household side is characterized by usual Euler equation $\frac{\dot{c}_t}{c_t} = r_t - \rho$ and the transversality condition $\lim_{t\to\infty}\left[e^{-\int_0^t (r_s-\eta)ds} a_t\right] = 0$, where a_t denotes per-capita assets of the representative household. Our assumption $\rho > \eta$ implies that the transversality condition is satisfied along a BGP.

A-1.1.1 Static Equilibrium

Consider first the static equilibrium allocations, in particular (1). Letting μ_i denote the markup in product i, the equilibrium wage is given by

$$w_t = \left(\int_0^{N_t} \mu_i^{1-\sigma} q_i^{\sigma-1} di\right)^{\frac{1}{\sigma-1}} = N_t^{\frac{1}{\sigma-1}} \left(\int \mu^{1-\sigma} q^{\sigma-1} dF_t(q,\mu)\right)^{\frac{1}{\sigma-1}}.$$

Similarly, aggregate output Y_t is given by

$$Y_t = N_t^{\frac{1}{\sigma-1}} \frac{\left(\int \mu^{1-\sigma} q^{\sigma-1} dF_t(q,\mu)\right)^{\frac{\sigma}{\sigma-1}}}{\int \mu^{-\sigma} q^{\sigma-1} dF_t(q,\mu)} L_t^P. \tag{A-1}$$

Defining $Q_t = \left(\int q^{\sigma-1} dF_t(q)\right)^{\frac{1}{\sigma-1}} = \left(E\left[q^{\sigma-1}\right]\right)^{\frac{1}{\sigma-1}}$ we can write (A-1) as $Y_t = N_t^{\frac{1}{\sigma-1}} Q_t \mathcal{M}_t L_t^P$ where

$$\Lambda_t = \frac{\int \mu^{-\sigma} (q/Q_t)^{\sigma-1} dF_t(q,\mu)}{\int \mu^{1-\sigma} (q/Q_t)^{\sigma-1} dF_t(q,\mu)} \text{ and } \mathcal{M}_t = \frac{\left(\int \mu^{1-\sigma} (q/Q_t)^{\sigma-1} dF_t(q,\mu)\right)^{\frac{\sigma}{\sigma-1}}}{\int \mu^{-\sigma} (q/Q_t)^{\sigma-1} dF_t(q,\mu)}. \tag{A-2}$$

For the case of $\mu_i = \mu$, $\mathcal{M}_t$ and Λ_t reduce to $\mathcal{M}_t = 1$ and $\Lambda_t = 1/\mu$ as required in (1). Product-level sales and profits are given by

$$py_i = \mu_i^{1-\sigma}\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\left(\frac{1}{\mathcal{M}_t\Lambda_t}\right)^{\sigma-1}\frac{Y_t}{N_t} \quad \text{and} \quad \pi_i = \left(1-\frac{1}{\mu_i}\right)py_i. \tag{A-3}$$

If markups are constant, (A-3) reduces to $py_i = \left(\frac{q_i}{Q_t}\right)^{\sigma-1}\frac{Y_t}{N_t}$ and $\pi_i = \left(\frac{\mu-1}{\mu}\right)\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\frac{Y_t}{N_t}$.

A-1.1.2 Aggregate Growth Rate

Given τ_t and $\nu_t = g_{Nt} + \delta$, the rate of quality growth is given by

$$g_Q = \frac{\dot{Q}_t}{Q_t} = \left(\frac{\lambda^{\sigma-1}-1}{\sigma-1}\right)\tau_t + \frac{\left(\overline{\omega}^{\sigma-1}-1\right)}{\sigma-1}\nu_t + I. \tag{A-4}$$

The growth rate of labor productivity is given by

$$g_t^{LP} = \frac{d}{dt}\ln\left(Q_t N_t^{\frac{1}{\sigma-1}}\right) = g_t^Q + \frac{1}{\sigma-1}g_t^N = I + \left(\frac{\lambda^{\sigma-1}-1}{\sigma-1}\right)\tau_t + \frac{\overline{\omega}^{\sigma-1}}{\sigma-1}\nu_t - \frac{\delta}{\sigma-1}. \tag{A-5}$$

A-1.1.3 Derivation of value function $V_t(q)$ in the model with entry (Equation (9))

Conjecture that the value function takes the form $V_t(q) = q^{\sigma-1}U_t$, so that $\dot{V}_t(q) = g_U V(t)$. The HJB equation in (6) then implies that

$$U_t = \frac{(\mu-1)\left(\frac{1}{Q_t}\right)^{\sigma-1}\frac{L_t^P}{N_t}w_t}{r_t+\tau_t+\delta-g_U}. \tag{A-6}$$

The free entry condition in (7) is thus given by

$$\frac{1}{\varphi_E}w_t = V^{Entry} = \frac{\overline{q}^{\sigma-1}(\mu-1)\frac{L_t^P}{N_t}w_t}{r_t+\tau_t+\delta-g_U} = \overline{q}^{\sigma-1}Q_t^{\sigma-1}U_t.$$

This implies that U_t grows at rate $g_U = g_w - (\sigma-1)g_Q$. Substituting this in (A-6) yields (9).

A-1.1.4 Proof of Proposition 1

Equations (13) and (14) follow directly from $\eta = (1-\alpha)z - \delta$ in (12). To derive (15), consider the free entry condition in (A-7). The Euler equation implies $g_{Y/L} = r_t - \rho$.

Also, $g_w = g_{Y/L} - g_{\ell^P}$. Finally, (A-4) implies that

$$\rho + \tau_t + \delta + (\sigma - 1)\, g_Q + g_{\ell^P} = \rho + \frac{\bar{q}^{\sigma-1}}{1-\alpha}\delta + \left(\frac{\bar{q}^{\sigma-1}}{1-\alpha} - 1\right)(g_{\mathcal{N}} + \eta) + g_{\ell^P}.$$

The free entry condition in (A-7) thus implies that

$$\frac{1}{\varphi_E} = \frac{(\mu - 1)\,\bar{q}^{\sigma-1}}{\rho + \frac{\bar{q}^{\sigma-1}}{1-\alpha}\delta + \left(\frac{\bar{q}^{\sigma-1}}{1-\alpha} - 1\right)(g_{\mathcal{N}} + \eta) + g_{\ell^P}} \frac{\ell_t^P}{\mathcal{N}_t}. \tag{A-7}$$

Similarly, the resource constraint in (10) can be written as

$$\left(\frac{1-\ell_t^P}{\mathcal{N}_t}\right) = \frac{1}{\varphi_E}\frac{\nu_t}{1-\alpha} = \frac{1}{\varphi_E}\frac{g_{\mathcal{N}} + \delta}{1-\alpha} = \frac{1}{\varphi_E}\frac{g_{\mathcal{N}} + \eta + \delta}{1-\alpha}. \tag{A-8}$$

These are two differential equations in $\mathcal{N}_t$ and ℓ_t^P. Together with the initial condition $\mathcal{N}_0$ and the consumers' transversality condition as a terminal condition, they determine the path $\{\mathcal{N}_t, \ell_t^P\}_t$.

Along the BGP, $g_{\mathcal{N}} = g_{\ell^P} = 0$. Equations (A-7) and (A-8) can then be solved for $\mathcal{N}$ and ℓ^P given in (15). In addition, (A-7) and (A-8) also characterize the transitional dynamics depicted in Figure 2. Rearranging terms in (A-8) and substituting for $g_{\mathcal{N}}$ in (A-7) yields

$$\begin{aligned} g_{\ell^P} &= \varphi_E (1-\alpha) \frac{\left(\frac{\mu\bar{q}^{\sigma-1}}{1-\alpha} - 1\right)\ell_t^P - \left(\frac{\bar{q}^{\sigma-1}}{1-\alpha} - 1\right)}{\mathcal{N}_t} - \rho - \delta \\ g_{\mathcal{N}} &= \varphi_E (1-\alpha)\left(\frac{1-\ell_t^P}{\mathcal{N}_t}\right) - \eta - \delta. \end{aligned}$$

This dynamic system is depicted in the phase diagram in Figure 2.

A-1.1.5 Proof of Proposition 2

We first derive the value function stated in Proposition 2. Upon rewriting the innovation value Ξ_t as

$$\Xi_t = n \times \max_x \left\{ x\left(\alpha \int V_t([q_i], \lambda q)\, dF_t(q) + (1-\alpha)\int V_t([q_i], \omega Q_t)\, d\Gamma(\omega) - V_t([q_i])\right) - \frac{x^\zeta}{\varphi_x} w_t \right\},$$

it is immediate that the value function is additive, i.e. $V_t([q_i]) = \sum_{i=1}^{n} V_t(q_i)$. The HJB equation associated with $V_t(q_i)$ is given by

$$r_t V_t(q) - \dot{V}_t(q) = \pi_t(q) + I\frac{\partial V_t(q)}{\partial q} q - (\tau + \delta) V_t(q) + \Xi_t, \tag{A-9}$$

where $\Xi_t = \max_x \left\{ x\left(\alpha V_t^{CD} + (1-\alpha) V_t^{NV}\right) - \frac{1}{\varphi_x} x^{\zeta} w_t \right\}$ with $V_t^{CD} = \int V_t(\lambda q) dF_t(q)$ and $V_t^{NV} = \int V_t(\omega Q_t) d\Gamma(\omega)$.

Suppose the value function takes the following form $V_t(q) = q^{\sigma-1} U_t + M_t$, where M_t and U_t grow at some rate g_M and g_U respectively. Then $I\frac{\partial V_t(q)}{\partial q} q = I(\sigma-1) q^{\sigma-1} U_t$. (A-9) can then be written as

$$(r_t + \tau + \delta - g_U) q^{\sigma-1} U_t + (r + \tau + \delta - g_M) M_t = \left((\mu-1)\left(\frac{1}{Q_t}\right)^{\sigma-1} \frac{L_t^P}{N_t} w_t + I(\sigma-1) U_t\right) q^{\sigma-1} + \Xi_t. \tag{A-10}$$

It is easy to show that along a BGP this implies that

$$U_t = \frac{(\mu-1)\left(\frac{1}{Q_t}\right)^{\sigma-1} \frac{L_t^P}{N_t} w_t}{\rho + \tau + \delta + (\sigma-1)(g_Q - I)} \quad \text{and} \quad M_t = \frac{\Xi_t}{\rho + \tau + \delta},$$

as $\Xi_t \propto w_t$. To see this note that $\Xi_t = \frac{\zeta-1}{\varphi_x} x^{\zeta} w_t$,where

$$x = \left(\frac{\varphi_x}{\zeta}\right)^{\frac{1}{\zeta-1}} \left(\alpha \frac{V_t^{CD}}{w_t} + (1-\alpha)\frac{V_t^{NV}}{w_t}\right)^{\frac{1}{\zeta-1}}. \tag{A-11}$$

The value function is therefore given by

$$V_t(q) = \frac{\pi_t(q)}{\rho + \tau + \delta + (\sigma-1)(g_Q - I)} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta} w_t}{\rho + \tau + \delta}.$$

Note also that $V_t^{CD} = \int V_t(\lambda q) dF_t(q) = V_t(\lambda Q_t)$ and $V_t^{NV} = V_t(\overline{\omega} Q_t)$.

A-1.1.6 Characterization of Equilibrium

The equilibrium is characterized by the following conditions. First, the evolution of aggregate productivity is given by in (A-4). Second, the rate of creative destruction is given by $\tau = \frac{\alpha}{1-\alpha} \nu_t$, where $\nu_t = (1-\alpha)(z_t + x)$. Third, labor market clearing requires $L_t = L_{Pt} + L_{Rt}$, where $L_{Rt} = N_t \frac{1}{\varphi_E}\left(z_t + \frac{1}{\zeta} x\right)$. Fourth, the Euler equation is given by

$r = \rho + g_c$, where g_c is the growth rate of per capita consumption. Wages and output are given by $Y_t = N_t^{\frac{1}{\sigma-1}} Q_t L_t^P$ and $w_t = \frac{1}{\mu} Y_t / L_t^P$. Market clearing requires $C_t = Y_t$. Hence, the growth rate of per capita consumption is given by

$$g_c = g_Y - \eta = g_w + g_{L^P} - \eta, \tag{A-12}$$

where $g_w = \frac{1}{\sigma-1} g_N + g_Q$ (see (A-5)). Thus, $r = \rho + g_w + g_{L^P} - \eta$. Finally, (A-10) and free entry implies that $\frac{1}{\varphi_E} = \bar{q}^{\sigma-1} u_t + m_t$, where

$$m_t = \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta}}{\rho + g_{L^P} - \eta + \tau + \delta - g_m} \text{ and } u_t = \frac{(\mu-1)\, \ell_t^P / \mathcal{N}_t}{\rho + g_{L^P} - \eta + \tau + \delta - g_u + (\sigma-1)(g_Q - I)}. \tag{A-13}$$

Then we can write the free entry condition as

$$\frac{1}{\varphi_E} = \frac{\bar{q}^{\sigma-1}(\mu-1)}{\rho + g_\ell + \delta - g_u + \left(\frac{\bar{q}^{\sigma-1}}{1-\alpha} - 1\right)\nu_t} \frac{\ell_t^P}{\mathcal{N}_t} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta}}{\rho + g_\ell + \frac{\alpha}{1-\alpha}\nu_t + \delta - g_m}.$$

Hence, the equilibrium is characterized by a path $\{\ell_t^P, \mathcal{N}_t\}_t$ that satisfies the free entry condition and labor market clearing

$$\frac{1}{\varphi_E} = \frac{\bar{q}^{\sigma-1}(\mu-1)}{\rho + g_\ell + \delta - g_u + \left(\frac{\bar{q}^{\sigma-1}}{1-\alpha} - 1\right)\nu_t} \frac{\ell_t^P}{\mathcal{N}_t} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta}}{\rho + g_\ell + \frac{\alpha}{1-\alpha}\nu_t + \delta - g_m} \tag{A-14}$$

$$\frac{1-\ell_t^P}{\mathcal{N}_t} = \frac{1}{\varphi_E}\left(\frac{\nu_t}{1-\alpha} - \frac{\zeta-1}{\zeta} x\right), \tag{A-15}$$

where g_u and g_m are the growth rates of u_t and m_t given in (A-13). For a given initial condition $\mathcal{N}_0$ and the terminal condition that $\ell_t^P \to \bar{\ell}^P$ and $m_t \to m$ and $u_t \to u$ one can solve for the dynamic path $\{\ell_t^P, \mathcal{N}_t\}_t$.

A-1.1.7 Balanced Growth Path

Along a BGP, income per capita grows at a constant rate. (A-12) implies that

$$g_c = g_w + g_{L^P} - \eta = \left(\frac{\bar{q}^{\sigma-1} - \alpha}{\sigma - 1}\right) \frac{\nu_t}{1-\alpha} - \frac{1}{\sigma-1}\delta + I + g_{\ell^P}.$$

Along the BGP it also has to be the case that $\ell^P = L_t^P / L_t$. Hence, g^N is constant along a BGP, i.e. $g_N = \nu - \delta = \eta$. With ℓ^P and $\mathcal{N}$ constant, $g_u = g_m = 0$ along the BGP. Hence,

$\left(\mathcal{N},\ell^{P}\right)$ are given by

$$\frac{1}{\varphi_{E}} = \frac{\overline{q}^{\sigma-1}\left(\mu-1\right)}{\rho+\left(\frac{\overline{q}^{\sigma-1}}{1-\alpha}-1\right)\eta+\frac{\overline{q}^{\sigma-1}}{1-\alpha}\delta}\frac{\ell_{t}^{P}}{\mathcal{N}_{t}}+\frac{\frac{\zeta-1}{\varphi_{x}}x^{\zeta}}{\rho+\frac{\alpha\eta+\delta}{1-\alpha}} \tag{A-16}$$

$$\frac{1-\ell^{P}}{\mathcal{N}} = \frac{1}{\varphi_{E}}\left(\frac{\eta+\delta}{1-\alpha}-\frac{\zeta-1}{\zeta}x\right). \tag{A-17}$$

These equations have a unique solution for $\mathcal{N}>0$ and $\ell^{P}\in(0,1)$.

A-1.1.8 Population Growth and Firm Dynamics (Section 2.4)

In this section we derive the relationship between population growth η and the process of firm dynamics. In particular, we derive (i) the survival function $S\left(a\right)$ in (23), (ii) the average number of products by age $\overline{n}\left(a\right)$ in (23), and (iii) the Pareto tail of the product distribution ζ_{n} in (24).

Firm survival $S\left(a\right)$ and the average number of products $\overline{n}\left(a\right)$ Let $p_{n}\left(a\right)$ be the probability that a firm has n products at age a. This evolves according to

$$\dot{p}_{n}\left(a\right)=\left(n-1\right)xp_{n-1}\left(a\right)+\left(n+1\right)\left(\tau+\delta\right)p_{n+1}\left(a\right)-n\left(x+\tau+\delta\right)p_{n}\left(a\right).$$

Because exit is an absorbing state, $\dot{p}_{0}\left(a\right)=\left(\tau+\delta\right)p_{1}\left(a\right)$. The solution to this set of differential equations is (see Klette and Kortum (2004))

$$p_{0}\left(a\right) = \frac{\tau+\delta}{x}\gamma\left(a\right) \tag{A-18}$$

$$p_{1}\left(a\right) = \left(1-p_{0}\left(a\right)\right)\left(1-\gamma\left(a\right)\right)$$

$$p_{n}\left(a\right) = p_{n-1}\left(a\right)\gamma\left(a\right) \tag{A-19}$$

where

$$\gamma\left(a\right)=\frac{x\left(1-e^{-(\tau+\delta-x)a}\right)}{\tau+\delta-x\times e^{-(\tau+\delta-x)a}}. \tag{A-20}$$

Given that $\frac{1-\alpha}{\alpha}\tau=\delta+\eta$, the net rate of accumulation ψ is given by

$$\psi \equiv x-\tau-\delta=x-\frac{\alpha}{1-\alpha}\left(\eta+\delta\right)-\delta=x-\frac{\alpha\eta+\delta}{1-\alpha}. \tag{A-21}$$

Hence, ψ is decreasing in η. From this solution for $p_n(a)$ we can calculate both the survival rate and the cross-sectional age distribution.

Let $S(a)$ denote share of firms that survive until age a. Then

$$S(a) = 1 - p_0(a) = \frac{\psi e^{\psi a}}{\psi - x(1 - e^{\psi a})}, \tag{A-22}$$

which is equation (23) in the main text.

To derive $\overline{n}(a)$, let $\overline{p}_n(a)$ denote the share of firms of age a with n production conditional on survival. Then, $\overline{p}_n(a) = \frac{p_n(a)}{1-p_0(a)}$ for $n \geq 1$. Using (A-18)-(A-19), $\overline{p}_n(a) = \gamma(a)^{n-1}(1-\gamma(a))$. Then,

$$\overline{n}(a) = E[N| A_f = a] = \sum_{n=1}^{\infty} n\overline{p}_n(a) = (1-\gamma(a))\sum_{n=1}^{\infty} n\gamma(a)^{n-1} = \frac{1}{1-\gamma(a)}.$$

Using (A-20), this implies $\overline{n}(a) = 1 - \frac{x}{\psi}\left(1 - e^{\psi a}\right)$, which is the expression in (23).

The Pareto tail of the product distribution ϱ_n. To derive the tail of the product distribution, let $\omega_t(n)$ be the mass of firms with n products at time t. For $n \geq 2$,

$$\dot{\omega}_t(n) = \omega_t(n-1)(n-1)x + \omega_t(n+1)(n+1)(\tau+\delta) - \omega_t(n)n(\tau+x+\delta).$$

For $n = 1$ we have $\dot{\omega}_t(1) = Z_t + \omega_t(2)2(\tau+\delta) - \omega_t(1)(\tau+x+\delta)$. Along the BGP the mass of firms grows at rate η, that is $\dot{\omega}_t(n) = \eta\omega_t(n)$. Denoting $\chi(n) = \frac{\omega_t(n)}{N_t}$,

$$\chi(2) = \frac{\chi(1)(\tau+x+\delta+\eta) - z}{2(\tau+\delta)} \tag{A-23}$$

and

$$\chi(n+1) = \frac{\chi(n)n(\tau+x+\delta) + \chi(n)\eta - \chi(n-1)(n-1)x}{(n+1)(\tau+\delta)} \quad \text{for } n \geq 2. \tag{A-24}$$

Given $\chi(1)$, these equations fully determine $[\chi(n)]_{n\geq 2}$ as a function of (x, z, τ). We can then pin down $\chi(1)$ from the consistency condition that

$$\sum_{n=1}^{\infty} \chi(n)n = \sum_{n=1}^{\infty} \frac{\omega_t(n)}{N_t}n = \frac{\sum_{n=1}^{\infty}\omega_t(n)n}{N_t} = 1. \tag{A-25}$$

Hence, equations (A-23), (A-24) and (A-25) fully determine the firm-size distribution $[\chi(n)]_{n\geq 1}$.

The distribution described by (A-23), (A-24) and (A-25) has a Pareto tail as long as $\eta > x-\tau-\delta > 0$. Applying Proposition 3 in Luttmer (2011), the tail index of the product distribution is given by $\varrho_n = \frac{\eta}{x-\tau-\delta}$. Using that $\tau = \frac{\alpha}{1-\alpha}(\eta+\delta)$ we get that $\varrho_n = \frac{(1-\alpha)\eta}{x(1-\alpha)-\delta-\alpha\eta} = \frac{\eta}{\eta-z}$, where the second equality uses that $z = \frac{\eta+\delta}{1-\alpha} - x$. Also $\frac{\partial \varrho_n}{\partial \eta} > 0$, that is a decline in population growth reduces the Pareto tail towards unity and increases concentration.[23]

The Pareto tail of the efficiency distribution ϱ_q. In this section we derive the marginal distribution of efficiency q. In particular we derive (26), which we use to calibrate $\overline{\omega}$. Define $\hat{q}_t$ as the relative productivity of a product $\hat{q}_t \equiv \ln(q_t/Q_t)^{\sigma-1}$. The drift of $\hat{q}_t$ (conditional on survival) is given by

$$\frac{\partial \hat{q}_t}{\partial t} = (\sigma-1)\,I - (\sigma-1)\,d\ln Q_t = -\left(\frac{\alpha\left(\lambda^{\sigma-1}-1\right)}{1-\alpha} + \overline{\omega}^{\sigma-1} - 1\right)(\eta+\delta)\,, \tag{A-26}$$

where the second equality uses (14).

Let $F_t(\hat{q})$ denote the share of products at time t with $\hat{q}_i \leq \hat{q}$. This cdf evolves according to the differential equation

$$\frac{\partial F_t(\hat{q})}{\partial t} = -\underbrace{\frac{\partial F_t(\hat{q})}{\partial \hat{q}}\frac{\partial \hat{q}_t}{\partial t}}_{\text{Drift of } \hat{q}} + \underbrace{\tau\left(F_t(\hat{q}-\hat{\lambda}) - F_t(\hat{q})\right)}_{\text{Creative destruction}} - \underbrace{(\delta+\eta)\left(F_t(\hat{q}) - \Gamma\left(\exp\left(\frac{\hat{q}}{\sigma-1}\right)\right)\right)}_{\text{Product loss vs new product creation}},$$

where $\hat{\lambda} = \ln \lambda^{\sigma-1}$. In the steady state, $\frac{\partial F_t(\hat{q})}{\partial t} = 0$ so that

$$\frac{dF(\hat{q})}{dq}\frac{\partial \hat{q}_t}{\partial t} = \tau\left(F_t(\hat{q}-\hat{\lambda}) - F_t(\hat{q})\right) - (\delta+\eta)\left(F_t(\hat{q}) - \Gamma\left(\exp\left(\frac{\hat{q}}{\sigma-1}\right)\right)\right). \tag{A-27}$$

Guess that F is exponential in the tail with index ϱ_q, that is $\lim_{\hat{q}\to\infty} e^{\varrho_q \hat{q}}(1-F(\hat{q})) = a$ for

[23] Note that $\frac{\partial \varrho_n}{\partial \eta} = (1-\alpha)\frac{x(1-\alpha)-\delta}{(x(1-\alpha)-\delta-\alpha\eta)^2} > 0$. Using that $\tau = \frac{\alpha}{1-\alpha}(\eta+\delta)$, it follows that $x-\tau-\delta = \frac{1}{1-\alpha}\left(x(1-\alpha)-\alpha\eta-\delta\right)$. Hence, $x-\tau-\delta > 0$ implies that $x(1-\alpha)-\delta > \alpha\eta > 0$.

some a and ϱ_q. If we assume that Γ has a thin tail[24] then as $\hat{q} \to \infty$, (A-27) implies that

$$\lim_{\hat{q}\to\infty}\left(ae^{-\varrho_q\hat{q}}\varrho_q\frac{\partial\hat{q}_t}{\partial t}\right)=\lim_{\hat{q}\to\infty}\left[(\delta+\eta+\tau)-\tau e^{\varrho_q\hat{\lambda}}\right]ae^{-\varrho_q\hat{q}}-(\delta+\eta).$$

Hence, the tail coefficient ϱ_q solves the equation $-\varrho_q\frac{\partial\hat{q}_t}{\partial t}=-(\delta+\eta+\tau)+\tau e^{\varrho_q\hat{\lambda}}$. Substituting (A-26) and noting that $\tau=\frac{\alpha}{1-\alpha}(\eta+\delta)$ yields

$$\varrho_q\left(\alpha\lambda^{\sigma-1}+(1-\alpha)\overline{\omega}^{\sigma-1}-1\right)=-1+\alpha\lambda^{\varrho_q(\sigma-1)}.$$

This is equation (26) in the main text.

A-1.2 Model Extensions (Section 2.6)

A-1.2.1 Endogenizing the Direction of Innovation α.

Suppose that the firm can independently chose the flow of new varieties x_N and creative destruction x_{CD}. The value function is then given by

$$r_tV_t(q)-\dot{V}_t(q)=\pi_t(q)+I\frac{\partial V_t(q)}{\partial q}q-(\tau_t+\delta)V_t(q)+\Xi_t$$

where

$$\Xi_t\equiv\max_{x_N}\left\{x_NV_t^N-\frac{1}{\varphi_N}x_N^{\zeta}w_t\right\}+\max_{x_{CD}}\left\{x_{CD}V_t^{CD}-\frac{1}{\varphi_{CD}}x_{CD}^{\zeta}w_t\right\}. \tag{A-28}$$

φ_{CD} and φ_N parametrize the efficiency of creative destruction and new variety creation and V_t^N and V_t^N denote the value of creative destruction and new variety creation respectively. Along the BGP, the solution of $V_t(q)$ is given by

$$V_t(q)=\frac{(\mu-1)}{\rho+(g_N-\eta)+(g_Q-I)(\sigma-1)+\tau+\delta}\left(\frac{q}{Q_t}\right)^{\sigma-1}\frac{L_t^P}{N_t}w_t+\frac{\Xi_t}{r+\tau+\delta-g_{\Xi_t}}.$$

[24] Formally, assume that for any κ, we have $\lim_{\hat{q}\to\infty}e^{\kappa\hat{q}}(1-\Gamma\left(\exp\left(\frac{\hat{q}}{\sigma-1}\right)\right))=0.$

The optimal innovation rates associated with (A-28) are given by

$$x_{NV} = \left(\frac{\varphi_N}{\zeta} \frac{V_t^{NV}}{w_t} \right)^{\frac{1}{\zeta-1}} \text{ and } x_{CD} = \left(\frac{\varphi_{CD}}{\zeta} \frac{V_t^{CD}}{w_t} \right)^{\frac{1}{\zeta-1}}. \tag{A-29}$$

Note that this implies that the endogenous share of product creation directed to creative destruction is given by $\frac{\tilde{\alpha}}{1-\tilde{\alpha}} = \left(\frac{\varphi_{CD} V_t^{CD}}{\varphi_N V_t^N} \right)^{\frac{1}{\zeta-1}}$, i.e. the relative "bias" of innovation depends on the relative valuations. This also implies that

$$\Xi_t = \left(\frac{\zeta - 1}{\varphi_{NV}} x_{NV}^{\zeta} + \frac{\zeta - 1}{\varphi_{CD}} x_{CD}^{\zeta} \right) w_t, \tag{A-30}$$

where x_{NV} and x_{CD} are constant (see below). Hence, the value of product creation grows at rate w_t, i.e. $g_{\Xi_t} = g_w = r - \rho$. Similarly, along the BGP we have $g_N = \eta$, so that

$$V_t(q) \quad = \quad \frac{(\mu - 1)}{\rho + (g_Q - I)(\sigma - 1) + \tau + \delta} \left(\frac{q}{Q_t} \right)^{\sigma - 1} \frac{L_t^P}{N_t} w_t + \frac{\Xi_t}{\rho + \tau + \delta}.$$

As before, V_t^N and V_t^{CD} are given by $V_t(\lambda Q_t)$ and $V(\omega Q_t)$.

We assume the following process of entry. As in the baseline model, the economy has access to a linear entry technology whereby each worker generates a flow of φ_E new firms. These firms then have access to the same innovation technology as incumbents to eventually start producing either a creatively destroyed product or a new variety. In the event that no product is discovered, the potential firm exits.

Because new firms have - after paying the entry costs $\frac{1}{\varphi_E} w_t$ - the same opportunity as incumbents, their direction of innovation (i.e. new varieties versus creative destruction) is exactly the same as the one of incumbent firms. Hence, if z new firms are created (per product N_t), the total amount of creative destruction and variety creation by entrants is given by $z x_{CD}$ and $z x_{NV}$ respectively. It also implies that the free entry condition is given by $\frac{1}{\varphi_E} w_t = \Xi_t$.

BGP equilibrium The BGP equilibrium in this economy is fully characterized by innovation choices x_{NV} and x_{CD}, the entry flow z, value functions V^{NV}/w_t and V^{CD}/w_t, the rate of creative destruction τ and the mass of production labor per product L_t^P/N_t. As before, $g_N = \eta$ so that $\nu = x_{NV}(1+z) = \eta + \delta$ and $\tau = x_{CD}(1+z)$. The first order condition for x_{NV} and x_{CD} are given in (A-29). Equation (A-30) implies that the free entry condition is given by $\frac{1}{\varphi_E} = \frac{\Xi_t}{w_t} = \frac{\zeta-1}{\varphi_{NV}} x_{NV}^{\zeta} + \frac{\zeta-1}{\varphi_{CD}} x_{CD}^{\zeta}$. The value function is given by

$$\frac{V_t(q)}{w_t} = \frac{(\mu-1)\left(\frac{q}{Q_t}\right)^{\sigma-1}}{\lambda^{\sigma-1}\tau + (\omega^{\sigma-1}-1)\eta + \omega^{\sigma-1}\delta} \frac{L_t^P}{N_t} + \frac{1}{\rho+\tau+\delta}\frac{1}{\varphi_E}. \tag{A-31}$$

We can simplify this system further and express the BGP equilibrium in terms of x_{NV} and τ. Note first that $\tau = \frac{x_{CD}}{x_{NV}}(\eta+\delta)$. From (A-29) we get that $x_N^{\zeta}\frac{\zeta-1}{\varphi_N} = \frac{\zeta-1}{\zeta}\frac{V_t^N}{w_t}x_N$ and $x_{CD}^{\zeta}\frac{\zeta-1}{\varphi_N} = \frac{\zeta-1}{\zeta}\frac{V_t^{CD}}{w_t}x_{CD}$. Free entry therefore requires that

$$\frac{1}{\varphi_E} = \frac{\zeta-1}{\zeta}\left(\frac{V_t^{CD}}{w_t}x_N + \frac{V_t^{NV}}{w_t}x_{CD}\right).$$

Using the expressions for $\frac{V_t^{CD}}{w_t}$ and $\frac{V_t^{NV}}{w_t}$, to solve for $\frac{L_t^P}{N_t}$, substituting in (A-29), and using $\tau = \frac{x_{CD}}{x_{NV}}(\eta+\delta)$, we can express the entire equilibrium in terms of x_{NV} and τ. In particular, one can show that

$$x_N = \left(\frac{\varphi_N}{(\zeta-1)\varphi_E}\right)^{1/\zeta}\left(1+\left(\frac{\tau}{\eta+\delta}\right)^{\zeta}\frac{\varphi_N}{\varphi_{CD}}\right)^{-1/\zeta}.$$

Finally, τ is determined from

$$\left(\left(\frac{\lambda}{\overline{\omega}}\right)^{\sigma-1}-1\right)\frac{1}{\eta+\delta}\frac{\tau}{\rho+\delta+\tau} = \frac{\zeta}{(\zeta-1)^{\frac{\zeta-1}{\zeta}}}\left(\frac{\varphi_E}{\varphi_N}\right)^{\frac{1}{\zeta}}\frac{\left(\frac{\lambda}{\overline{\omega}}\right)^{\sigma-1}\frac{\tau}{\eta+\delta} - \left(\frac{\tau}{\eta+\delta}\right)^{\zeta}\frac{\varphi_N}{\varphi_{CD}}}{\left(1+\left(\frac{\tau}{\eta+\delta}\right)^{\zeta}\frac{\varphi_N}{\varphi_{CD}}\right)^{\frac{\zeta-1}{\zeta}}}, \tag{A-32}$$

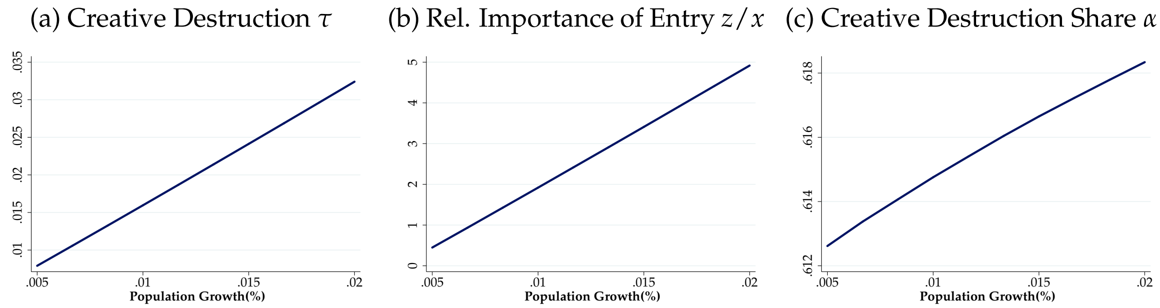
and hence depends directly on η.

Assume first that creative destruction and new variety creation leads to the same quality improvement, i.e. $\lambda = \overline{\omega}$. This implies that

$$\alpha = \frac{x_{CD}}{x_{CD}+x_{NV}} = \frac{(\varphi_{CD})^{\frac{1}{\zeta-1}}}{(\varphi_{CD})^{\frac{1}{\zeta-1}} + (\varphi_{NV})^{\frac{1}{\zeta-1}}},$$

that is α is indeed constant. Along the BGP, we still have $x_{NV}(1+z) = (1-\alpha)x(1+z) = \eta+\delta$. Similarly, creative destruction is given by $\tau = \alpha x(1+z)$. Hence, as in the baseline model, $\tau = \frac{\alpha}{1-\alpha}(\eta+\delta)$, i.e. lower population growth reduces creative destruction. Using the free entry condition yields $x^{\zeta} = \frac{1}{\zeta-1}\frac{1}{\varphi_E}\left(\frac{(1-\alpha)^{\zeta}}{\varphi_{NV}} + \frac{\alpha^{\zeta}}{\varphi_{CD}}\right)^{-1}$. As in the baseline model,

Figure A-1: The Effects of Population Growth (Endogenous α)



Note: This figure plots model outcomes in a calibrated version of the extended model with endogenous innovation direction as a function of the rate of population growth.

x is constant and fully determined from parameters governing the relative innovation technologies. And with x constant, we have $zx = \frac{\eta+\delta}{1-\alpha} - x$, i.e. the total entry flow per product, zx, is a decreasing function of population growth: in equilibrium entrants bear the brunt of declining population growth.

If $\lambda \neq \omega$, τ is determined from (A-32). One can show that there is at least one solution τ and that a fall in η lowers τ if there is a unique solution. In Figure A-1 we show creative destruction τ, the relative importance of entry $z = \frac{z_{NV}}{x_{NV}} = \frac{z_{CD}}{x_{CD}}$ and the share of creative destruction α as a function of population growth. The comparative static results shown in Figure A-1 accord well with our baseline model. Creative destruction is an increasing function of population growth, falling population reduces z, the flow of entrant product innovation relative to incumbents and the creative destruction share α is increasing in η. The result that α is increasing in η is driven by our estimates that $\overline{\omega} < \lambda$. This implies that V_t^{CD} puts a higher weight on the production as opposed to the innovation value. And because a fall in population growth reduces the innovation value, it raises the value of new variety creation relative to the value of creative destruction, leading to a decline in the creative destruction share α.

A-1.2.2 Endogenous own-innovation I

Suppose now that firms can chose the rate of own-innovation I. Assume that the cost function (in terms of labor) of achieving a drift I is given by $c\,(I; q/Q) = \left(\frac{q}{Q_t}\right)^{\sigma-1} \frac{1}{\varphi_I} I^{\zeta}$. Hence, the cost of innovation are convex in I and depend on firms' relative efficiency q/Q_t to make the model consistent with balanced growth and Gibrat's law for large firms.

Most results of the baseline model generalize in a straightforward way. In particular, Proposition 1 is exactly the same in this more general framework, except I in the expression for the growth rate is no longer a parameter but a choice variable. The value function is still additive across products and the value of a given product with efficiency q is given by

$$V_t(q) = \frac{\pi_t(q)}{\rho+\delta+\tau+\left(g^Q - \frac{\zeta-1}{\zeta}I\right)(\sigma-1)} + \frac{\frac{\zeta-1}{\varphi_x}x^\zeta w_t}{\rho+\tau+\delta}. \tag{A-33}$$

Hence, the only difference to the baseline model is the term $\frac{\zeta-1}{\zeta}$ in front of I in the discount rate. Using (A-33), the optimal innovation rate is given by

$$I = \left(\frac{\varphi_I}{\zeta}\left(\frac{1}{\varphi_E} - \frac{\frac{\zeta-1}{\varphi_x}x^\zeta}{\rho+\tau+\delta}\right)\frac{\sigma-1}{\overline{q}^{\sigma-1}}\right)^{\frac{1}{\zeta-1}}. \tag{A-34}$$

Because x is still given by the expression in (19), (A-34) determines I as a function of τ and exogenous parameters. Moreover, I is increasing in τ and hence declining in η.

A-1.3 Characterization of the Model with Bertrand Competition

In this section we derive the results for the model with Bertrand competition described in Section 3. The only difference relative to the baseline case is that the static profit function is given by (A-3), i.e.

$$\pi(q_i,\Delta_i) = \left(1-\frac{1}{\mu(\Delta_i)}\right)\mu(\Delta_i)^{1-\sigma}\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\frac{1}{(\mathcal{M}_t\Lambda_t)^{\sigma-1}}\frac{Y_t}{N_t}. \tag{A-35}$$

The value function is still additive across products, i.e. $V_t([q_i,\Delta_i]) = \sum_{i=1}^n V_t(q_i,\Delta_i)$. The HJB equation for $V_t(q_i,\Delta_i)$ is given by

$$(r_t+\tau_t+\delta)V_t(q,\Delta) - \dot{V}_t(q,\Delta) = \pi_t(q,\Delta) + I\left\{\frac{\partial V_t(q,\Delta)}{\partial q}q + \frac{\partial V_t(q,\Delta)}{\partial \Delta}\Delta\right\} + \Xi_t, \tag{A-36}$$

where $\Xi_t = \max_x\left\{x\left(\alpha V_t^{CD} + (1-\alpha)V_t^{NV}\right) - \frac{1}{\varphi_x}x^\zeta w_t\right\}$ with $V_t^{CD} = \int V_t(\lambda q,\lambda)\,dF_t(q)$ and $V_t^{NV} = \int V_t(\omega Q_t, \frac{\sigma}{\sigma-1})\,d\Gamma(\omega)$. Note that for notational simplicity we denote the quality gap for the creation of a new variety by $\frac{\sigma}{\sigma-1}$ to indicate that new varieties are able to charge the monopolistic markup.

Note first that free entry and the definition of Ξ_t still implies that $x = \left(\frac{\varphi_x}{\varphi_E}\frac{1}{\zeta}\right)^{\frac{1}{\zeta-1}}$ and $\Xi_t = \frac{\zeta-1}{\varphi_x}x^\zeta w_t$. To solve for $V_t(q,\Delta)$, conjecture that $V_t(q,\Delta)$ takes the form $V_t(q,\Delta) =$

$k(\Delta) q^{\sigma-1} U_t + M_t$. This implies that

$$\frac{\partial V_t(q,\Delta)}{\partial q} q = (\sigma-1) k(\Delta) q^{\sigma-1} U_t \quad \text{and} \quad \frac{\partial V_t(q,\Delta)}{\partial \Delta} \Delta = k'(\Delta) \Delta q^{\sigma-1} U_t.$$

Equation (A-36) thus implies that $k(\Delta)$, U_t and M_t solve the equations

$$(r_t+\tau) M_t - \dot{M}_t = \frac{\zeta-1}{\varphi_x} x^{\zeta} w_t \tag{A-37}$$

and

$$\begin{aligned} k(\Delta) q^{\sigma-1} \left(\left(r_t + \tau - I(\sigma - 1 + \varepsilon_k(\Delta)) \right) U_t - \dot{U}_t \right) &= \pi_t(q,\Delta) \\ &= h(\Delta) \left(\frac{q_i}{Q_t} \right)^{\sigma-1} \frac{Y_t/N_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1}}, \end{aligned}$$

where $h(\Delta) = \left(1 - \frac{1}{\mu(\Delta)}\right) \mu(\Delta)^{1-\sigma}$ and $\varepsilon_k(\Delta) \equiv \frac{k'(\Delta)\Delta}{k(\Delta)}$.

Along the BGP, $\mathcal{M}_t \Lambda_t$ is constant, $w_t \propto Y_t/N_t$ and U_t grows at the same rates as $\frac{Y_t}{N_t} Q_t^{1-\sigma}$. This implies that $g_U = \frac{\dot{U}_t}{U_t} = g_w - (\sigma-1) g_Q$. Hence,

$$k(\Delta) U_t = \frac{h(\Delta) \left(\frac{1}{Q_t}\right)^{\sigma-1} \frac{1}{(\mathcal{M}\Lambda)^{\sigma-1}} \frac{Y_t}{N_t}}{r + \delta + \tau - g_w + (\sigma-1)\left(g_Q - I\right) - I \varepsilon_k(\Delta)}.$$

The solution to the value function is therefore

$$U_t = \left(\frac{1}{Q_t}\right)^{\sigma-1} \frac{1}{(\mathcal{M}\Lambda)^{\sigma-1}} \frac{Y_t}{N_t} \quad \text{and} \quad k(\Delta) = \frac{h(\Delta)}{r + \delta + \tau - g_w + (\sigma-1)(g_Q - I) - I \frac{k'(\Delta)\Delta}{k(\Delta)}}.$$

Along the BGP, $r + \delta + \tau - g_w + (\sigma-1)\left(g_Q - I\right) = \rho + \delta + \left(\frac{\bar{q}^{\sigma-1}}{1-\alpha} - 1\right)(\eta+\delta)$. Hence, $k(\Delta)$ solves the differential equation

$$k(\Delta) \mathcal{C} - I k'(\Delta) \Delta = h(\Delta) = \frac{\min\left\{\frac{\sigma}{\sigma-1}, \Delta\right\} - 1}{\min\left\{\frac{\sigma}{\sigma-1}, \Delta\right\}^{\sigma}},$$

where $\mathcal{C} = \rho + \delta + \left(\frac{\bar{q}^{\sigma-1}}{1-\alpha} - 1\right)(\eta+\delta)$. For $\Delta \geq \frac{\sigma}{\sigma-1}$ we have $k(\Delta) \mathcal{C} - I k'(\Delta) \Delta =$

$\frac{1}{\sigma-1}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}$. Hence, $k'(\Delta)=0$ and $k(\Delta)=\frac{\frac{1}{\sigma-1}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}}{\mathcal{C}}$. For $\Delta<\frac{\sigma}{\sigma-1}$, we have

$$k(\Delta)\mathcal{C}-Ik'(\Delta)\Delta \;=\; \frac{\Delta-1}{\Delta^{\sigma}}.$$

We can solve this differential equation together with the terminal condition $k\left(\frac{\sigma}{\sigma-1}\right)=\frac{\frac{1}{\sigma-1}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}}{\mathcal{C}}$.

Equation (A-37) implies M_t grows at g_w along the BGP. Hence, $M_t=\frac{1}{\rho+\tau+\delta}\frac{\zeta-1}{\varphi_x}x^{\zeta}$. Together with $w_tL_P=\Lambda_tY_t$, the value function along the BGP is given by

$$V_t(q,\Delta)=\left\{k(\Delta)\left(\frac{q}{Q_t}\right)^{\sigma-1}\frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}}\frac{\ell^P}{\mathcal{N}}+\frac{1}{\rho+\tau+\delta}\frac{\zeta-1}{\varphi_x}x^{\zeta}\right\}w_t. \tag{A-38}$$

Using (A-38) we can derive the free entry condition. The value of creative destruction is given by $V_t^{CD}=V_t(\lambda Q_t,\lambda)$. The value of variety creation is $V_t^{NV}=V_t\left(\overline{\omega}Q_t,\frac{\sigma}{\sigma-1}\right)$. The free entry condition, is thus given by

$$\frac{1}{\varphi_E}=\frac{V_t^{Entry}}{w_t}=\frac{\alpha k(\lambda)\lambda^{\sigma-1}+(1-\alpha)k\left(\frac{\sigma}{\sigma-1}\right)\overline{\omega}^{\sigma-1}}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}}\frac{\ell^P}{\mathcal{N}}+\frac{1}{\rho+\tau+\delta}\frac{\zeta-1}{\varphi_x}x^{\zeta}.$$

Because $\mathcal{M}^{\sigma-1}\Lambda^{\sigma}$ can be calculated along a BGP, the free entry condition determines $\frac{\ell^P}{\mathcal{N}}$ as in the model with constant markups.

To see that Proposition 1 still applies in the model with Bertrand competition, note first that creative destruction and variety creation are still given by $\tau=\alpha(z+x)$ and $\nu=(1-\alpha)(z+x)$. Moreover, the optimality condition for incumbent expansion x is still given by (A-11) and the free entry condition still holds. Hence, $x=\left(\frac{1}{\zeta}\frac{\varphi_x}{\varphi_E}\right)^{\frac{1}{\zeta-1}}$. These three equations together with BGP condition $g_N=\eta=\nu-\delta$ are sufficient to derive Proposition 1.

A-1.3.1 The Joint Distribution of Gaps and Productivity

In the model with Bertrand competition, the joint distribution of relative quality $\hat{q}_t=\ln(q_t/Q_t)^{\sigma-1}$ and quality gaps Δ, $F_t(\Delta,\hat{q}_t)$, emerges as a key endogenous object. To solve for $F_t(\Delta,\hat{q}_t)$, it is useful to separate the problem by focusing individually on products with competitors (i.e. where creative destructions has happened at some point in the past) and products without competitors (i.e. products which are still owned by the firms that introduced the variety originally). We denote these distributions by $F_t^C(\Delta,\hat{q})$ and $F_t^{NC}(\hat{q})$.[25] They are characterized from the two differential equations $\frac{\partial F_t^C(\Delta,\hat{q})}{\partial t}$ and $\frac{\partial F_t^{NC}(\hat{q})}{\partial t}$ given in the main text. In this section we derive these expressions. We denote the mass of products with and without competitors respectively by N_t^C and N_t^{NC}. Also recall that $\hat{q}_t$ has a drift of $g_{\hat{q}} = (\sigma - 1)\left(I - g_{Q,t}\right)$.

Let $\bar{F}_t^C(\Delta,\hat{q}) = N_t^C F_t^C(\Delta,\hat{q})$ denote the *mass* of products with a gap less than Δ and relative productivity less than $\hat{q}$, for products with a direct competitor. Similarly, let $\bar{F}_t^{NC}(\hat{q}) = N_t^{NC} F_t^{NC}(\hat{q})$ denote the *mass* of the products who have no direct competitor at time t with relative productivity less than $\hat{q}$. The evolution of $\bar{F}_t^{NC}(\hat{q})$ satisfies

$$\bar{F}_t^{NC}(\hat{q}) = \bar{F}_{t-\iota}^{NC}\left(\hat{q} - g_{\hat{q}}\iota\right)\left(1 - (\tau_t + \delta)\,\iota\right) + \left(\frac{1-\alpha}{\alpha}\right)\tau_t N_t \Gamma\left(\frac{\exp(\hat{q})}{\sigma - 1}\right)\iota,$$

where we have used the fact that the new product creation rate is $\frac{1-\alpha}{\alpha}\tau = \nu_t$. As ι becomes small this leads to the differential equation

$$\frac{\partial \bar{F}_t^{NC}(\hat{q})}{\partial t} = -g_{\hat{q}}\frac{\partial \bar{F}_t^{NC}(\hat{q})}{\partial \hat{q}} - (\tau_t + \delta)\,\bar{F}_t^{NC}(\hat{q}) + \left(\frac{1-\alpha}{\alpha}\right)\tau_t N_t \Gamma\left(\frac{\exp(\hat{q})}{\sigma - 1}\right), \tag{A-39}$$

so that

$$\frac{\partial F_t^{NC}(\hat{q})}{\partial t} = -g_{\hat{q}}\frac{\partial F_t^{NC}(\hat{q})}{\partial \hat{q}} - (\tau_t + \delta + \eta)\,F_t^{NC}(\hat{q}) + \left(\frac{1-\alpha}{\alpha}\right)\tau_t \frac{N_t}{N_t^{NC}}\Gamma\left(\frac{\exp(\hat{q})}{\sigma - 1}\right).$$

This is the equation reported in Section 3.

For the mass of products with a competitor, $\bar{F}_t^C(\Delta,\hat{q})$, we not only need to keep track of the relative quality $\hat{q}$ but also of the quality gap Δ. This mass evolves according to the differential equation

$$\begin{aligned}\frac{\partial \bar{F}_t^C(\Delta,\hat{q})}{\partial t} =& -\frac{\partial \bar{F}_t^C(\Delta,\hat{q})}{\partial \Delta} I\Delta - g_{\hat{q}}\frac{\partial \bar{F}_t^C(\Delta,\hat{q})}{\partial \hat{q}} - \bar{F}_t^C\left(\Delta,\hat{q}\right)(\tau_t + \delta) \\ &+ \lim_{s\to\infty} \tau_t \bar{F}_t^C(s,\hat{q} - \hat{\lambda}) + \tau \bar{F}_t^{NC}(\hat{q} - \hat{\lambda}),\end{aligned}$$

[25]Note that we do not need to keep track of the quality gap among products without competitors because markups are always given by $\frac{\sigma}{\sigma-1}$.

where again we defined $\hat{\lambda} = \ln \lambda^{\sigma-1}$. Defining $\bar{F}_t^C(\Delta, \hat{q}) = N_t^C F_t^C(\Delta, \hat{q})$, we get

$$\begin{aligned}
\frac{\partial F_t^C(\Delta, \hat{q})}{\partial t} &= -\Delta I \frac{\partial F_t^C\left(\Delta, \hat{q}\right)}{\partial \Delta} - g_{\hat{q}} \frac{\partial F_t^C\left(\Delta, \hat{q}\right)}{\partial \hat{q}} - (\tau_t + \delta + \eta) F_t^C(\Delta, \hat{q}) \\
&\quad + \lim_{s \to \infty} \tau_t F_t^C(s, \hat{q} - \hat{\lambda}) + \tau_t \frac{N_t^{NC}}{N_t^C} F_t^{NC}(\hat{q} - \hat{\lambda}).
\end{aligned}$$

Note that both differential equations depend on N_t^{NC}/N_t^C and N_t/N_t^{NC}. N_t^{NC} and N_t^C evolve according to

$$\dot{N}_t^{NC} = N_t \tau \left(\frac{1-\alpha}{\alpha}\right) - N_t^{NC}(\delta + \tau) \quad \text{and} \quad \dot{N}_t^C = -\delta N_t^C + N_t^{NC} \tau.$$

The steady state share of NC products is therefore given by $\frac{N_t^{NC}}{N_t} = \frac{\tau\left(\frac{1-\alpha}{\alpha}\right)}{\eta+\delta+\tau} = 1 - \alpha$.

In the left panel of Figure A-2 we display the distribution $F(\Delta, \hat{q})$ in a BGP. Multiple forces shape this distribution. On the one hand, firms increase their efficiency q over their life-cycle. This tends to generate a positive correlation between relative efficiency and efficiency gaps. On the other hand, successful creative destruction events also increase relative efficiency but reduce efficiency gaps and hence markups. Moreover, new products have - in our calibration - low efficiency (because $\overline{\omega} < 1$) and high efficiency gaps.

In the right panel we look at the efficiency distributions of the different type of products more directly. We depict the overall cross-sectional distribution of competitive products in red and compare it to the efficiency of products conditional on having a quality gap of λ (blue) and to the products that just entered and are still without a competitor (orange). The overall distribution dominates the distribution of new products in a first-order stochastic dominance sense because new products have on average lower qualities. The efficiency distribution, conditional on having a quality gap of Δ, is also lower because some of these products are non-competitive products that just experienced their first creative destruction event.

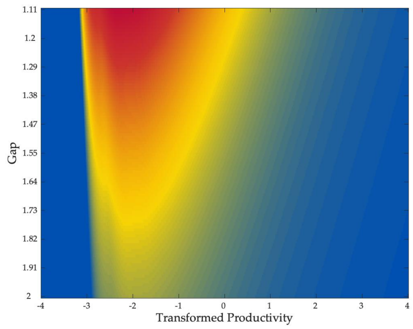
We can also derive the distribution of efficiency gaps given in (28). Let $F_t^C(\Delta)$ denote the cdf of quality gaps among products with a competitor. The distribution $F_t^C(\Delta)$ the solves the differential equation

$$\frac{\partial F_t^C(\Delta)}{\partial t} + F_t^C(\Delta) \frac{1}{N_t^C} \frac{\partial N_t^C}{\partial t} = -I\Delta \frac{\partial F_t^C(\Delta)}{\partial \Delta} - \delta F_t^C(\Delta) + (1 - F_t^C(\Delta))\tau + \frac{N_t^{NC}}{N_t^C}\tau.$$

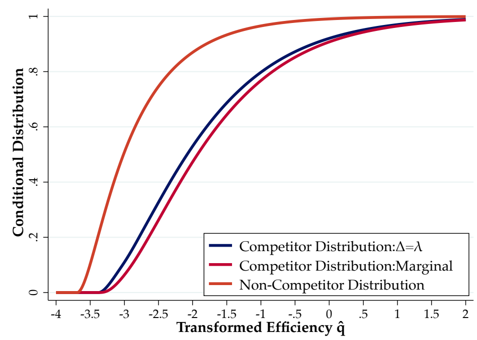
Along a BGP, this distribution is stationary, the number of competitive products grows at

Figure A-2: The Distributions of Efficiency q and Gaps Δ

(a) Joint Distribution of Efficiency and Gaps



(b) Conditional Productivity Distributions



Notes: The left panel shows the joint density of $\hat{q}$ (relative efficiency) and Δ (the gap between the leading product and the next best product) in the calibrated BGP. The right panel shows the productivity distributions in the calibrated model for three types of products: non-competitive products (orange), products which have just seen a creative destruction event and have a gap of $\Delta = \lambda$ (red) and all competitive products (blue)

rate η and $N_t^{NC}/N_t^C = \frac{1-\alpha}{\alpha}$. Hence,

$$I\Delta\frac{\partial F^C(\Delta)}{\partial\Delta} = -(\delta+\eta)F^C(\Delta) + (1-F^C(\Delta))\tau + \frac{1-\alpha}{\alpha}\tau = -(\delta+\eta+\tau)F^C(\Delta) + \frac{1}{\alpha}\tau.$$

Together with the initial condition $F^C(\lambda) = 0$ and the fact that $\frac{1-\alpha}{\alpha}\tau = \eta + \delta$, it is easy to verify that the solution to this differential equation is $F^C(\Delta) = 1 - \left(\frac{\lambda}{\Delta}\right)^{\frac{\delta+\eta+\tau}{I}}$.

A-2 Quantitative Analysis

A-2.1 Data Description

Our main data is the LBD, which contains information for employment and age for the population of firms in the US. In Table A-1 we report a set of descriptive statistics from this data. The firm size distribution in the US has been changing. Between 1980 and 2010 average firm size increased from 20 employees to about 22 employees. This increase in firm size is mostly due to a change in the concentration of economic activity. As seen in Panel B, the employment share of firms with more than 10,000 employees increased and the employment share of firms with less than 20 employees declined. Finally, an important mechanisms underlying these changes in the size distribution are shifts in the age distribution. As seen in the lowest panel, young firms account for much lower share

Table A-1: Summary of Data

Aggregate Statistics				
Year	Number of Firms	Employees	Average Employment	
1980	3,606,457	73,753,303	20.04	
1995	4,613,849	99,243,906	21.20	
2010	4,953,425	111,189,088	22.15	

Size Distribution				
	Firms with <20 Employees		Firms with >10,000 Employees	
Year	Firm Share	Employment Share	Firm Share	Employment share
1980	89.38	21.58	0.0002	25.71
1995	88.95	20.74	0.0002	23.84
2010	88.88	18.8	0.0002	27.02

Age Distribution				
	Firms with <5 years		Firms with >5 years	
Year	Firm Share	Employment Share	Firm Share	Employment share
1980	13.84 %	38.50 %	86.16%	61.50 %
1995	13.12 %	35.34 %	86.88 %	64.66 %
2010	9.43%	30.02%	91.57 %	69.98 %

Notes: This table gives basic summary information about the firms in the LBD through time.

of aggregate employment then they used to in 1980.

A-2.2 Computing the sales and markups lifecycle

In this section we derive the details of our characterization of the firms' lifecycle of markup and sales that we use to calibrate the model (see Section 4.3). In particular, we show that relative sales by age of the product is given by

$$s_P(a_P) \equiv E\left[\frac{p_i y_i}{Y}\middle| a_p\right] = \mu(a_p)^{1-\sigma} e^{(\sigma-1)(I-g^Q)a_p}\overline{q}^{\sigma-1}. \tag{A-40}$$

Moreover we derive the distribution of product age a_P as a function of firm age a_f and the number of products N. Given this distribution we can then easily evaluate $s_f(a_f)$ and $\mu_f(a_f)$ computationally.

Consider a BGP where $\mathcal{M}_t$ and Λ_t are constant. Equation (A-3) then implies that sales of product i relative to average sales are

$$s_P(a_P) \equiv E\left[\frac{p_i y_i}{Y_t/N_t}\middle| a_p\right] = E\left[\mu_i^{1-\sigma}\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\middle| a_p\right]\left(\frac{1}{\mathcal{M}_t\Lambda_t}\right)^{\sigma-1}.$$

Because Δ and hence markups are a deterministic function of the age of the product, $\mu_i = \mu(a_P) = \min\left\{\lambda e^{Ia_P}, \frac{\sigma-1}{\sigma}\right\}$. Similarly, Q_t is given by $Q_t = e^{g_Q a_P} Q_{t-a_p}$.

Now consider the distribution of q_i conditional on a_P. This distribution is given by

$$P(q_i \leq q | a_P) = P(q_i \leq q | a_P, CD)\,\alpha + P(q_i \leq q | a_P, NV)\,(1-\alpha),$$

where $P(q_i \leq q | a_P, CD)$ and $P(q_i \leq q | a_P, NV)$ denotes the conditional probability, conditional on the firm having acquired product i through creative destruction or new variety creation respectively. Then

$$P(q_i \leq q | a_P, CD) = F_{t-a_P}\left(\frac{1}{\lambda} q e^{-Ia_P}\right),$$

where $F_{t-a_P}(q)$ denotes the productivity distribution at time $t - a_P$. Similarly,

$$P(q_i \leq q | a_P, NV) = \Gamma\left(q e^{-Ia_P}\frac{1}{Q_{t-a_P}}\right).$$

Hence,

$$E\left[q_i^{\sigma-1}\middle| a_P\right] = \alpha\int q^{\sigma-1}dF_{t-a_P}\left(\frac{q}{\lambda}e^{-Ia_P}\right) + (1-\alpha)\int q^{\sigma-1}d\Gamma\left(\frac{qe^{-Ia_P}}{Q_{t-a_P}}\right) = e^{(\sigma-1)Ia_P}Q_{t-a_P}^{\sigma-1}\overline{q}^{\sigma-1},$$

so that $s_P(a_P) = \mu(a_P)^{1-\sigma}e^{(\sigma-1)(I-g_Q)a_P}\overline{q}^{\sigma-1}\left(\frac{1}{\mathcal{M}_t\Lambda_t}\right)^{\sigma-1}$ (see (A-40)).

Life-Cycle Dynamics Relative sales and markups at the product level as a function of the state variables Δ and q are given by

$$\mu(\Delta) = \min\left\{\frac{\sigma}{\sigma-1},\Delta\right\} \text{ and } \frac{s_i}{Y_t/N_t} = s_P(\Delta,q) = \left(\frac{1}{\mu(\Delta)}\frac{1}{\mathcal{M}_t\Lambda_t}\frac{q}{Q_t}\right)^{\sigma-1}.$$

Relative sales and average markups of firm f as functions of the random vector $[\Delta_i, q_i]_{i=1}^{N_f}$ are then given by

$$\frac{s_{ft}}{Y_t/N_t} = \sum_{i=1}^{N_f} s_P(\Delta_i, q_i) \quad \text{and} \quad \mu_f = \frac{1}{N_f}\sum_{i=1}^{N_f}\mu(\Delta_i).$$

Expected relative sales as a function of firm age a_f are given by

$$E\left[\frac{s_{ft}}{Y_t/N_t}\middle| a_f\right] = E\left[\sum_{n=1}^{N_f} E\left[s_P(\Delta_i,q_i)|\, a_f, a_P, N_f\right]\middle| a_f\right] = E\left[\sum_{n=1}^{N_f} s_P(a_P)\middle| a_f\right],$$

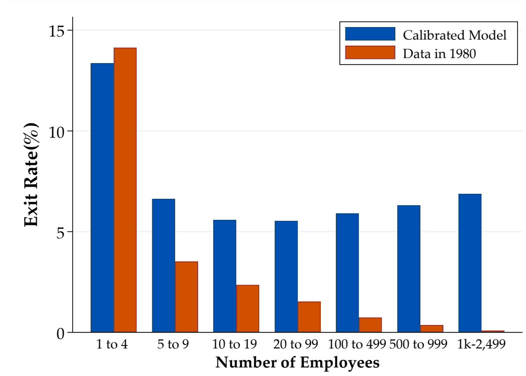
where $s_P(a_P)$ is given in (A-40). The last equality exploits the fact that conditional on product age a_P, product level sales are independent of firm age a_f and the number of products N_f. Let $f_{a_P|A_f,N}(a_P|a,n)$ denote the conditional distribution of product age a_P conditional on firm age a_f and the number of products n and recall that $\overline{p}_n(a_f) = \gamma(a_f)^{n-1}(1-\gamma(a_f))$ is the probability a firm of age a_f having n products (conditional on survival), where $\gamma(a)$ is given in (A-20). Then

$$E\left[\frac{s_{ft}}{Y_t/N_t}\middle| a_f\right] = (1-\gamma(a_f))\sum_{n=1}^{\infty} n\left(\int_{a_P} s_P(a_P) f_{a_P|A_f,N}(a_P|a_f,n)\, da_P\right)\gamma(a_f)^{n-1}.$$

Using the same logic, the average markup as a function of firm age a_f is given by

$$E\left[\mu_f|\, a_f\right] = (1-\gamma(a_f))\sum_{n=1}^{\infty}\left(\int_{a_P}\mu(a_P) f_{a_P|A_f,N}(a_P|a_f,n)\, da_P\right)\gamma(a_f)^{n-1}.$$

Figure A-3: Firm Exit Rates: Model and Data



Notes: This figure presents a comparison of the exit rates by firm size in the model (blue) and the data (orange)

Given the density $f_{a_P|A_f,N}\left(a_P|a_f,n\right)$, these expressions can be directly evaluated. In Section SM-4 in the Supplementary Material we show how to compute this density.

A-2.3 Exit Rates by Size

In Figure A-3, we depict the exit rate for different size categories. Empirically, these exit rates are declining. Our model implies that this exit rate is initially declining but essentially independent of size for firms with more than 10 employees. The reason our model has this counterfactual prediction is that (in our calibration) the thick tail of the employment distribution is driven by the distribution of product quality q and not the extensive margin of product creation. Hence, large firms are firms with a few superstar products, not those with many products. And because creative destruction is independent of product quality, such firms are as likely to exit as other firms.

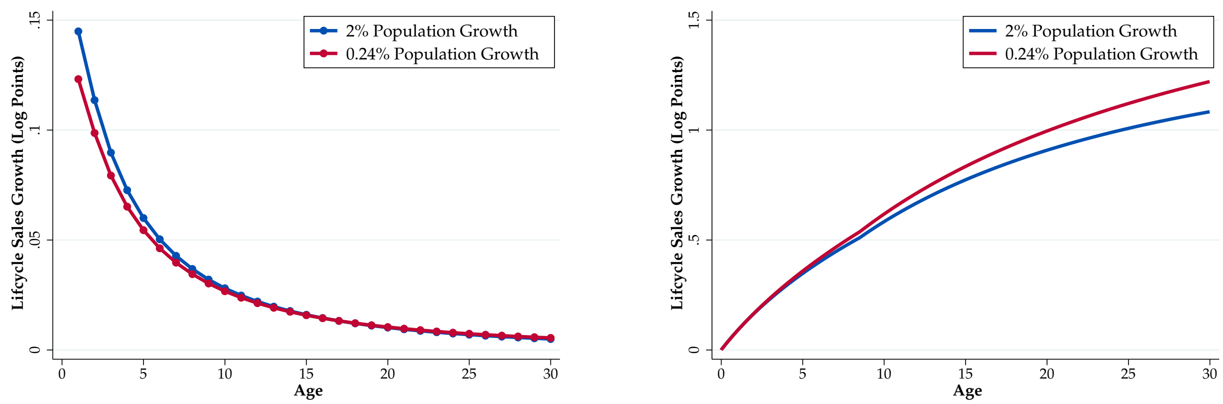
To address this counterfactual prediction, in Section SM-3 in the Supplementary Material, we extend our model to allow for type heterogeneity, whereby some young firms (sometimes described as "rockets" or "gazelles", see Sterk et al. (2021)) grow systematically at a faster rate. This extension improves the model's fit along this dimension substantially, because some large firms have many products and are thus unlikely to exit, but changes little else in the theoretical analysis.

A-2.4 Decomposing the Impact of Falling Population Growth

Our analysis in Section 5 showed that the experienced and projected decline in population growth increased firm size and markups. In principle, these patterns can be due to changes in the age distribution and changes in firms' size or markups conditional on age. In Figure A-4 we show that the lion share of these changes is due to changes in the age distribution. We show the exit rate by age (left panel) and the sales life-cycle (right panel) both in the original BGP (blue) and the new BGP when population growth is 0.24% (red line). While both the age conditional exit rate and the life-cycle do change, the changes are qualitatively small. By contrast, the age distribution, which we display in Figure A-5, shifts substantially. And because older firms are larger and exit at a lower rate, such shifts in the age distribution explain most of the observed change in concentration in our model. This result is consistent with the findings of Hopenhayn et al. (2018) and Karahan et al. (2019), who document empirically that the age-conditional allocations have been relatively constant since the 1980s.

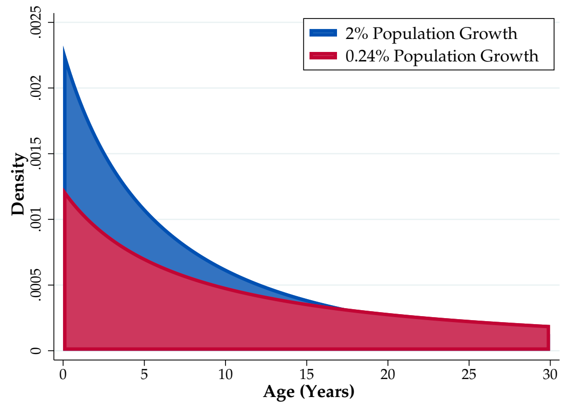
Figure A-4: Decomposing the Impact of Falling Population Growth

(a) Exit Rates (b) Lifecycle of Sales



Notes: The left panel shows the model prediction for firm exit rates by age when population growth is 2% (blue) and 0.24% (red). The right panel shows the same for sales growth.

Figure A-5: The Impact of Falling Population Growth on the Age Distribution



Notes: The figure shows the age distribution when population growth is 2% (blue) and 0.24% (red).

Supplementary Material for *"Population Growth and Firm Dynamics"*

[NOT FOR PUBLICATION]

SM-1 Deriving a Sufficient Condition for $\mathscr{N}$ to Increase in Response to Falling Population Growth

The two equations in (A-16) and (A-17) characterize $\mathscr{N}$ and ℓ^P as a function of population growth η. In this section we show that

$$\frac{\overline{q}^{\sigma-1}}{1-\alpha} > \frac{1}{\mu} \tag{SM-1}$$

is a sufficient condition for $\mathscr{N}$ to be decreasing in η. We also show that (SM-1) is the tightest condition one can derive without further restrictions on the innovation cost function.

Using (A-16) and (A-17) we can solve for $\mathscr{N}$ explicitly as

$$\frac{1}{\mathscr{N}} = \frac{1}{\varphi_E}\left[\left(1-\frac{\zeta-1}{\zeta}\frac{x}{\rho+\frac{1}{1-\alpha}\delta+\frac{\alpha}{1-\alpha}\eta}\right)\frac{\rho+\frac{\overline{q}^{\sigma-1}}{1-\alpha}\delta+\left(\frac{\overline{q}^{\sigma-1}}{1-\alpha}-1\right)\eta}{\overline{q}^{\sigma-1}(\mu-1)}+\frac{\eta+\delta}{1-\alpha}-\frac{\zeta-1}{\zeta}x\right].$$

Clearly $\mathscr{N}$ is decreasing in η if $\frac{\overline{q}^{\sigma-1}}{1-\alpha}-1>0$. Hence, consider the case $\frac{\overline{q}^{\sigma-1}}{1-\alpha}-1<0$ and define

$$\Delta \equiv -\left(\frac{\overline{q}^{\sigma-1}}{1-\alpha}-1\right) > 0.$$

Substituting Δ for $\overline{q}^{\sigma-1}$ yields

$$\frac{1}{\mathscr{N}} = \frac{1}{\varphi_E}\left[\frac{1}{1-\alpha}\left(\left(1-\frac{\zeta-1}{\zeta}\frac{x}{\rho+\frac{1}{1-\alpha}\delta+\frac{\alpha}{1-\alpha}\eta}\right)\frac{\rho+(1-\Delta)\delta-\Delta\eta}{(1-\Delta)(\mu-1)}+\eta+\delta\right)-\frac{\zeta-1}{\zeta}x\right].$$

Hence, define the function

$$
\begin{aligned}
m(\eta) &= \left(1-\frac{\zeta-1}{\zeta}\frac{x}{\rho+\frac{1}{1-\alpha}\delta+\frac{\alpha}{1-\alpha}\eta}\right)\frac{\rho+(1-\Delta)\delta-\Delta\eta}{(1-\Delta)(\mu-1)}+\eta \\
&= \frac{\rho+(1-\Delta)\delta-\Delta\eta}{(1-\Delta)(\mu-1)}-\frac{\zeta-1}{\zeta}\frac{x}{\rho+\frac{1}{1-\alpha}\delta+\frac{\alpha}{1-\alpha}\eta}\frac{\rho+(1-\Delta)\delta-\Delta\eta}{(1-\Delta)(\mu-1)}+\eta.
\end{aligned}
$$

$\mathscr{N}$ is decreasing in η if $m'(\eta)>0$. Taking the derivative yields

$$
m'(\eta)=\frac{1}{(1-\Delta)(\mu-1)}\left(\frac{\zeta-1}{\zeta}x\frac{\Delta\rho+\frac{\alpha}{1-\alpha}\rho+\frac{\alpha}{1-\alpha}\delta}{\left(\rho+\frac{1}{1-\alpha}\delta+\frac{\alpha}{1-\alpha}\eta\right)^2}+(1-\Delta)\mu-1\right).
$$

Note that $1-\Delta=\frac{\overline{q}^{\sigma-1}}{1-\alpha}\geq\frac{\alpha}{1-\alpha}>0$. Hence, we require that

$$
\frac{\zeta-1}{\zeta}x\frac{\Delta\rho+\frac{\alpha}{1-\alpha}\rho+\frac{\alpha}{1-\alpha}\delta}{\left(\rho+\frac{1}{1-\alpha}\delta+\frac{\alpha}{1-\alpha}\eta\right)^2}+(1-\Delta)\mu>1.
$$

Now note that x is only a function of the innovation technologies φ_x and φ_E. Hence, by choice of these parameters, we can make x arbitrarily small. The tightest sufficient condition is therefore given by

$$
(1-\Delta)=\frac{\overline{q}^{\sigma-1}}{1-\alpha}>\frac{1}{\mu}.
$$

SM-2 The Mechanism: Demand or Supply?

In our baseline model, population growth impacts the economy on both the labor supply side and via aggregate demand. In this section we extend to our analysis to distinguish these two forces.

Environment

Consider the following two-sector economy. Sector 1 is exactly the same as in our baseline model. Sector 2 is comprised of a representative firm with a production technology $Y_t=\mathcal{A}_tH_t$,where H_t denotes the number of workers in sector 2. Aggregate productivity in sector 2, $\mathcal{A}_t$, grows at an exogenous rate $g_{\mathcal{A}}$. To distinguish between supply and demand, we assume that the total population consists of two separate types of people. A mass L_t of people can only work in sector 1. This mass grows at rate η. A mass H_t of people can only work in sector 2. This mass grows at rate η^H. We assume that all individuals have identical intra-temporal Cobb Douglas preferences $c_t = c_{1t}^{\vartheta} c_{2t}^{1-\vartheta}$. The L_t workers in sector 1 have the same preferences as in our baseline model and engage in borrowing and saving for inter-temporal consumption smoothing. For simplicity we assume that workers in sector 2 act as hand-to-mouth consumers and simply spend their labor income in each period.

We let k_t denote the equilibrium wage in sector 2 and S_t the price of sector 2 goods. The real price index of the consumption bundle $c_t = c_{1t}^{\vartheta} c_{2t}^{1-\vartheta}$ is thus given by

$$\mathcal{P}_t = \left(\frac{P_t}{\vartheta}\right)^{\vartheta} \left(\frac{S_t}{1-\vartheta}\right)^{\vartheta} = \left(\frac{1}{\vartheta}\right)^{\vartheta} \left(\frac{S_t}{1-\vartheta}\right)^{\vartheta}, \tag{SM-2}$$

where P_t denotes the price index in sector 1, which we for comparability with our baseline model take as the numeraire, i.e. $P_t = 1$.

The Value Function and Free Entry

We now derive the value function of firms in sector 1 and the free entry condition. Suppose aggregate demand in sector 1 was given by $\mathcal{D}_t$ with $\dot{\mathcal{D}}_t / \mathcal{D}_t = g_{\mathcal{D}}$. Total profits per product are then given by

$$\pi_{it} \;=\; \left(\frac{\mu-1}{\mu}\right) p y_{it} = \left(\frac{\mu-1}{\mu}\right) p_{it}^{1-\sigma} \mathcal{D}_t = \left(\frac{\mu-1}{\mu}\right) \mu^{1-\sigma} q^{\sigma-1} w_t^{1-\sigma} \mathcal{D}_t. \tag{SM-3}$$

The value function of an individual product is thus given by

$$r_t V_t(q) - \dot{V}_t(q) = \pi_t(q) + I \frac{\partial V_t(q)}{\partial q} q - (\tau + \delta) V_t(q) + \Xi_t.$$

Free entry still implies that $\Xi_t = \frac{\zeta-1}{\varphi_x} x^{\zeta} w_t$ and $x = \left(\frac{1}{\zeta} \frac{\varphi_x}{\varphi_E}\right)^{\frac{1}{\zeta-1}}$.

Following the same steps as before, one can show that along a BGP the value function is given by

$$V_t(q) = \frac{\left(\frac{\mu-1}{\mu}\right) \mu^{1-\sigma} w_t^{1-\sigma} \mathcal{D}_t q^{\sigma-1}}{r + \tau + \delta - g_{\mathcal{D}} - (\sigma-1)(I - g_w)} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta} w_t}{r + \tau + \delta - g_w}.$$

Hence, the expected value of product innovation is given by

$$\begin{aligned}V_t^{Entry} &= \alpha \int V_t(\lambda q)\, dF_t(q) + (1-\alpha) \int V(\omega Q_t)\, d\Gamma(\omega) \\ &= \frac{\left(\frac{\mu-1}{\mu}\right) \mu^{1-\sigma} w_t^{1-\sigma} \mathcal{D}_t (\overline{q} Q_t)^{\sigma-1}}{r+\tau+\delta-g_{\mathcal{D}}-(\sigma-1)(I-g_w)} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta} w_t}{r+\tau+\delta-g_w},\end{aligned}$$

where as before $\overline{q} = \left(\alpha\lambda^{\sigma-1} + (1-\alpha)\overline{\omega}^{\sigma-1}\right)^{\frac{1}{\sigma-1}}$. Free entry thus requires that

$$\frac{1}{\varphi_E} = \frac{V_t^{Entry}}{w_t} = \frac{\left(\frac{\mu-1}{\mu}\right) \mu^{1-\sigma} \overline{q}^{\sigma-1}}{r+\tau+\delta-g_{\mathcal{D}}-(\sigma-1)(I-g_w)} \frac{\mathcal{D}_t Q_t^{\sigma-1}}{w_t^{\sigma}} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta}}{r+\tau+\delta-g_w}.$$

The fact that the final good is the numeraire still implies that $w_t = \frac{1}{\mu} Q_t N_t^{\frac{1}{\sigma-1}}$. Given aggregate spending $\mathcal{D}_t$, it is still the case that production workers in sector 1 receive a constant share of aggregate sales. In particular

$$\Pi_t = \int_i \pi_{it} = \left(\frac{\mu-1}{\mu}\right) \mathcal{D}_t \quad \text{and} \quad w_t L_t^P = \frac{1}{\mu} \mathcal{D}_t.$$

Substituting this into the free entry condition yields

$$\frac{1}{\varphi_E} = \frac{(\mu-1)\,\overline{q}^{\sigma-1}}{r+\tau+\delta-g_{\mathcal{D}}-(\sigma-1)(I-g_w)} \times \frac{L_t^P}{N_t} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta}}{r+\tau+\delta-g_w}.$$

Note first that along a BGP it is still the case that L_t^P/N_t is constant and variety growth is tied to the rate of population growth

$$g_N = \eta.$$

This directly implies that like in our baseline model

$$\nu = \eta + \delta \quad \text{and} \quad \tau = \frac{\alpha}{1-\alpha}\nu = \frac{\alpha}{1-\alpha}(\eta+\delta).$$

And because x is still constant

$$z = \frac{\tau}{\alpha} - x = \frac{1}{1-\alpha}(\eta+\delta) - x.$$

Hence, all the results about the effect of population growth on the firm size distribution, entry and concentration are exactly as in our baseline model. This highlights that the relationship between population growth and the firm-size distribution reflects the effect of demographics on labor supply.

We can also solve for the rate of discounting along a BGP. Because total spending by sector 1 workers grows at rate g_w, the Euler equation requires that $g_w = r - \rho$. Also note that $w_t L_t^P \propto \mathcal{D}_t$ implies that $g_{\mathcal{D}} = g_w + \eta$. Finally note that $g_w = g_Q + \frac{1}{\sigma-1} g_N$ as before. This implies that the value function $V_t(q)$ is given by

$$V_t(q) = \frac{(\mu - 1)\left(\frac{q}{Q_t}\right)^{\sigma-1}}{\rho + \delta + \left(\frac{\overline{q}^{\sigma-1}}{1-\alpha} - 1\right)\nu} \frac{w_t L_t^P}{N_t} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta} w_t}{\rho + \tau + \delta}.$$

This is the same equation as in the baseline model. Given that the resource constraint is unchanged, the variety intensity $\mathcal{N}$ and the production share ℓ^P is also the same as in the baseline model.

Income per capita growth

In our baseline model, real consumption and income growth was given by

$$g_y = g_c = g_w = g_Q + \frac{1}{\sigma - 1} g_N = \left(\frac{\overline{q}^{\sigma-1} - \alpha}{\sigma - 1}\right) \frac{\eta}{1 - \alpha} + \left(\frac{\overline{q}^{\sigma-1} - 1}{\sigma - 1}\right) \frac{\delta}{1 - \alpha} + I,$$

and hence directly tied to population growth η. In this multi-sector environment, additional forces are at play. Because per-capita spending of workers in sector 1, $\mathcal{P}_t c_t$, grows at rate g_w, real consumption grows at rate

$$g_c = g_w - g_{\mathcal{P}}. \tag{SM-4}$$

Given the expression for the price index $\mathcal{P}_t$ in (SM-2), $g_{\mathcal{P}} = \vartheta g_S$,where $g_S = \frac{\dot{S}_t}{S_t}$ denotes the growth rate of prices in sector 2.

To solve for the relative price in sector 2, S_t, note that total spending by agents working sector 2 is trivially given by $\mathcal{D}_t^H = k_t H_t$. The workers in sector 1 are subject to a dynamic budget constraint. Along a BGP, total expenditure per capita grows at the rate of the wage w_t, i.e. $\mathcal{D}_t^L = \chi w_t L_t$, where χ is constant. Market clearing in sector 2 thus implies that

$$S_t \mathcal{A}_t H_t = k_t H_t = (1 - \vartheta)\left(\mathcal{D}_t^L + \mathcal{D}_t^H\right) = (1 - \vartheta)\left(\chi w_t L_t + k_t H_t\right),$$

so that the wage in sector 2 is given by $k_t = \frac{1-\vartheta}{\vartheta} \chi w_t \frac{L_t}{H_t}$. This implies that

$$S_t = \frac{1}{\mathcal{A}_t} k_t = \frac{1}{\mathcal{A}_t} \frac{1 - \vartheta}{\vartheta} \chi w_t \frac{L_t}{H_t},$$

so that

$$g_S = -g_{\mathcal{A}} + g_w + \eta - \eta^H.$$

Real consumption growth in (SM-4) is therefore given by

$$\begin{aligned}
g_c &= g_w - (1-\vartheta)\, g_S \\
&= \vartheta g_w + (1-\vartheta)\left(g_{\mathcal{A}} + \eta^H - \eta\right) \\
&= \vartheta\left(\left(\frac{\overline{q}^{\sigma-1}-\alpha}{\sigma-1}\right)\frac{\eta}{1-\alpha} + \left(\frac{\overline{q}^{\sigma-1}-1}{\sigma-1}\right)\frac{\delta}{1-\alpha} + I\right) + (1-\vartheta)\left(g_{\mathcal{A}} + \eta^H - \eta\right).
\end{aligned}$$

Hence, in this multi-sector extension of our theory, both productivity growth in sector 2, $g_{\mathcal{A}}$, and relative population growth $\eta^H - \eta$ determine the growth rate of real consumption. Both enter positively and are a source of welfare growth. Productivity $g_{\mathcal{A}}$ reduces the price of sector 2 goods and hence benefits workers in sector 1 in accordance with the expenditure share $1-\vartheta$. Similarly, relative population growth $\eta^H - \eta$ is a source of welfare growth. If labor in sector 2 becomes more abundant, relative wages in sector 2 fall. This benefits workers in sector 1 again through falling prices.

SM-3 Firm Heterogeneity: Young Firm Rockets

Very young firms tend to grow fast even conditional on survival. Luttmer (2011) discusses how a violation of Gibrat's law is needed to deliver the relatively low age of very large firms: to match the data, there must be a mechanism whereby young firms can grow quickly for a time. A similar reasoning is also discussed in Sterk et al. (2021). In our baseline model, such a mechanism is absent. Young firms do indeed violate Gibrat's law in the model, but this is only because of survival bias. Here we discuss the implications for the effects of a population growth slowdown of introducing a subset of young firms that act as "rockets", growing and innovating quickly for a time, before their growth rate slows to look like other ordinary firms (see also Acemoglu et al. (2018))

Suppose that when a firm is born, it can be a rocket (R) or slow (S) type. The only difference between the two is the speed with which a firm can invent new products, such that $x = \{x^R, x^S\}$ To highlight the central differences with the main model, we take these rates to be exogenous. In addition, assume that a rocket firm transitions into being a slow firm at rate ξ. For exposition, assume labor is perfectly substitutable between research and production, and the rate of own product improvement I is fixed.

The value of such a firm can be written

$$\begin{aligned}
r_t V_t^R\left(\left[\Delta_i, q_i\right]\right) \;=\; & \sum_{i=1}^{n} \pi_t\left(\left[\Delta_i, q_i\right]\right)+\dot{V}_t^R\left(\left[\Delta_i, q_i\right]\right) \\
& +\sum_{i=1}^{n}(\tau+\delta)\left[V_t^R\left(\left[\Delta_j, q_j\right]_{j \neq i}\right)-V^R\left(\left[\Delta_i, q_i\right]\right)\right] \\
& +\sum_{j=1}^{n} I \frac{\partial V_t^R\left(\left[\Delta_i, q_i\right]\right)}{d q_j} q_j \\
& +n \max_x\left\{x\left[\alpha \int_q V_t^R\left(\left[\Delta_i, q_i\right], 1, \lambda q\right) d F_t(q)+\right.\right. \\
& (1-\alpha) \int_\omega \int_\Delta V_t^R\left(\left[\Delta_i, q_i\right], \Delta, \omega Q_t\right) d G(\Delta) d \Gamma(\omega) \\
& \left.\left.-V_t^R\left(\left[\Delta_i, q_i\right]\right)\right]-\frac{1}{\varphi_x^R} x^\zeta w_t\right\} \\
& \xi\left(V_t^S\left(\left[\Delta_i, q_i\right]\right)-V_t^R\left(\left[\Delta_i, q_i\right]\right)\right).
\end{aligned}$$

with an analogous equation holding for V_t^S. Suppose lastly that entrants cannot choose whether they are going to be a rocket or slow, but become a rocket at entry with fixed probability κ. It can be shown that under these assumptions, the solution to the value functions are

$$V_t^R\left(\left[\Delta_i, q_i\right]\right)=\sum_{i=1}^{n} U_t\left(\Delta_i, q_i\right)+n \phi^R w_t$$

$$V_t^S\left(\left[\Delta_i, q_i\right]\right)=\sum_{i=1}^{n} U_t\left(\Delta_i, q_i\right)+n \phi^S w_t$$

Where

$$U_t\left(\Delta_i, q_i\right)=\frac{u\left(\Delta_i\right)}{g(\sigma-1)+\rho+\tau+\delta-\eta} \frac{q_i^{\sigma-1} Y_t}{\left(\mathcal{M}_t \Lambda_t\right)^{\sigma-1} N_t Q_t^{\sigma-1}}$$

and ϕ^R and ϕ^S are the solutions to

$$(\rho+\tau+\delta) \phi^R=x^R\left[\frac{1}{\varphi_E}+(1-\kappa)\left(\phi^R-\phi^S\right)\right]-\frac{1}{\varphi_x^R}\left(x^R\right)^\zeta-\xi\left(\phi^R-\phi^S\right)$$

$$(\rho+\tau+\delta) \phi^S=x^S\left[\frac{1}{\varphi_E}-\kappa\left(\phi^R-\phi^S\right)\right]-\frac{1}{\varphi_x^S}\left(x^S\right)^\zeta$$

The share of rocket firms v^R in the population depends on the entry rate, and so changes in population growth will affect the average rate of incumbent expansion. To see this, note that the share of rockets firm $Y^R_{a,t}$ of age a at time $t$denoted is given by

$$Y^R_{a,t} = \kappa e^{-\xi a}$$

where $Y^R_{a,t}$and so the share of rockets in the economy is given by integrating this object over the age distribution. The age distribution is defined by the following two pieces. First, for fast firms the fraction of firms with n products evolves with age a as

$$\dot{p}^R_n (a) = (n-1)\, x^R p^R_{n-1} (a) + (n+1) (\tau + \delta)\, p^R_{n+1} (a) - n \left(x^R + \tau + \delta \right) p^R_n (a) - \xi p^R_n (a) . \quad \text{(SM-5)}$$

Because exit is an absorbing state, $\dot{p^R}_0 (a) = (\tau + \delta)\, p^R_1 (a)$. The fraction of firms that have survived by a is $S^R (a) = \frac{1-p^R_0(a)}{\sum_{n=0}^{\infty} p^R_n(a)}$. Similarly for slow firms, we have

$$\dot{p}^S_n (a) = (n-1)\, x^S p^S_{n-1} (a) + (n+1) (\tau + \delta)\, p^S_{n+1} (a) - n \left(x^S + \tau + \delta \right) p^S_n (a) + \xi p^R_n (a) .$$

with $\dot{p^S}_0 (a) = (\tau + \delta)\, p^S_1 (a)$ and $S^S (a) = \frac{1-p^S_0(a)}{\sum_{n=0}^{\infty} p^S_n(a)}$. The total fraction of surviving firms is then given by

$$S(a) = 1 - \kappa S^R(a) - (1-\kappa) S^S(a).$$

The age distribution can be obtained from calculating the density of firms by age using

$$\omega_t (a) = \frac{(1-\alpha)\, z N_0 e^{\eta(t-a)} S (a)}{\int_{a=0}^{\infty} (1-\alpha)\, z N_0 e^{\eta(t-a)} S (a)\, da}$$

$$= \frac{e^{-\eta a} S (a)}{\int_{a=0}^{\infty} e^{-\eta a} S (a)\, da}$$

Then the share of rockets in the overall population of is

$$v^R = \int_0^{\infty} \kappa e^{-\xi a} \omega (a)\, da$$

The share of products that are owned by rockets however, is not quite the same thing. This is given by

$$\hat{v}^R = \frac{F_t^R \int_0^\infty \sum_{n=1}^\infty \frac{p_n^R(a)n}{1-p_0^R(a)} \omega^R(a)da}{F_t^S \int_0^\infty \sum_{n=1}^\infty \frac{p_n^S(a)n}{1-p_0^S(a)} \omega^S(a)da + F_t^S \int_0^\infty \sum_{n=1}^\infty \frac{p_n^R(a)n}{1-p_0^R(a)} \omega^R(a)da}$$
$$= \frac{v^R \int_0^\infty \sum_{n=1}^\infty \frac{p_n^R(a)n}{1-p_0^R(a)} \omega^R(a)da}{(1-v^R) \int_0^\infty \sum_{n=1}^\infty \frac{p_n^S(a)n}{1-p_0^S(a)} \omega^S(a)da + v^R \int_0^\infty \sum_{n=1}^\infty \frac{p_n^R(a)n}{1-p_0^R(a)} \omega^R(a)da}$$

where the numerator is the number of rocket firms times the average products of a rocket firm. Creative destruction in this economy is given by

$$\tau = \alpha \left(z + \hat{v}^R x^R + (1 - \hat{v}^R) x^S \right)$$

However, it is still the case that

$$\tau = \frac{\alpha}{1-\alpha}(\eta + \delta)$$

So while changes in the age distribution will affect the share of rockets in the population, the overall effect on growth is the same as the baseline model. To characterize the equilibrium with rockets, we close the model with the labor market clearing condition

$$L_t = L_t^P + L_t^R = L_t^P + N_t \left(\frac{1}{\varphi_E} z_t + \frac{\hat{v}^R}{\varphi_x^R} (x^R)^\zeta + \frac{(1-\hat{v}^R)}{\varphi_x^S} (x^S)^\zeta \right).$$

Which we can characterize in terms of the production share ℓ^P and variety intensity $\mathscr{N}$ on the BGP as

$$\left(\frac{1-\ell_t^P}{\mathscr{N}_t} \right) = \frac{1}{\varphi_E} z + \frac{\hat{v}^R}{\varphi_x^R} (x^R)^\zeta + \frac{(1-\hat{v}^R)}{\varphi_x^S} (x^S)^\zeta.$$

Free entry requires that

$$\frac{1}{\phi_E} w_t = \kappa\alpha \int V_t^R(\lambda, q)\, dF_t(q) + \kappa(1-\alpha) \int V_t^R(\bar{\Delta}, Q_t q)\, dG(q)$$
$$+ (1-\kappa)\alpha \int V_t^S(\lambda, q)\, dF_t(q) + (1-\kappa)(1-\alpha) \int V_t^S(\bar{\Delta}, Q_t q)\, dG(q)$$

$$= \frac{\alpha u(\lambda)\lambda^{\sigma-1} + (1-\alpha)u(\bar{\omega})\bar{\omega}^{\sigma-1}}{g(\sigma-1) + \rho + \tau + \delta - \eta} \frac{L_t^P / N_t}{(\mathcal{M}_t)^{\sigma-1} \Lambda_t^\sigma} w_t + \kappa\phi_R + (1-\kappa)\phi_S$$

where in the exogenous I case, we can solve for $u(\Delta)$ from the differential equation

$$u(\Delta) = h(\Delta) + \frac{(\sigma-1)u(\Delta) + u'(\Delta)\Delta}{g(\sigma-1)+\rho+\tau+\delta-\eta} I$$

SM-3.1 Pareto Tail in the Product Distribution

Now we the following law for the evolution of the rocket distribution. Consider $n \geq 2$. Then, the number of rocket firms with each number of products n evolves according to

$$\dot{\omega}_t^R(n) = \underbrace{\omega_t^R(n-1)(n-1)x}_{\text{From } n-1 \text{ to} n \text{ products}} + \underbrace{\omega_t^R(n+1)(n+1)(\tau+\delta)}_{\text{From } n+1 \text{ to} n \text{ products}} - \underbrace{\omega_t^R(n)\,n(\tau+x+\delta)}_{\text{From } n \text{ to } n-1 \text{ or } n+1 \text{ products}} - \underbrace{\xi\omega_t^R(n)}_{\text{Transition to slow}}.$$

For $n = 1$ we have

$$\dot{\omega}_t^R(1) = \kappa Z_t + \omega_t^R(2)\,2(\tau+\delta) - \omega_t^R(1)(\tau+x+\delta).$$

Along the BGP the mass of firms grows at rate η. . Hence, the mass of firms is increasing at rate η. Hence, along the BGP we have

$$\dot{\omega^R}_t(n) = \eta\omega_t^R(n).$$

Along the BGP, $\{\nu(n)\}_{n=1}^{\infty}$ is determined by

$$\nu^R(2) = \frac{\nu^R(1)\left(\tau+x^R+\delta+\eta+\xi\right) - \frac{\kappa z}{\vartheta^R}}{2(\tau+\delta)} \tag{SM-6}$$

and

$$\nu^R(n+1) = \frac{\nu^R(n)\,n\left(\tau+x^R+\delta\right) + \nu^R(n)(\eta+\xi) - \nu^R(n-1)(n-1)x^R}{(n+1)(\tau+\delta)} \quad \text{for } n \geq 2 \tag{SM-7}$$

Again we can apply the result from Luttmer (2011) and the Pareto tail is given by

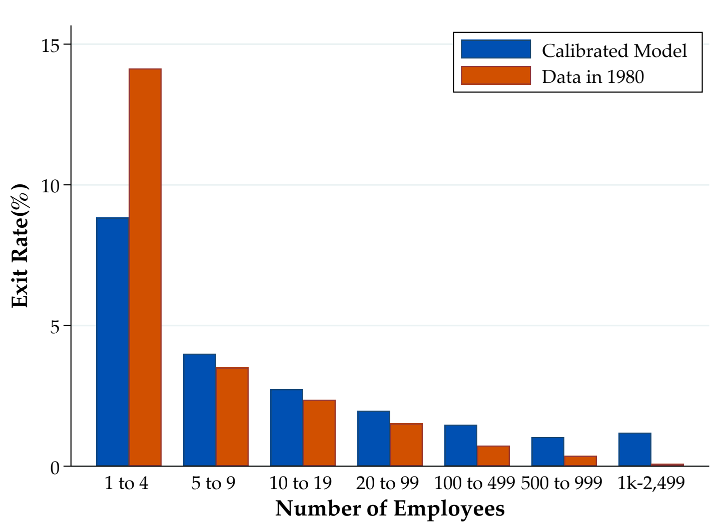
$$\begin{aligned}\zeta_n &= \frac{\eta+\xi}{x^R-\tau-\delta} \\ &= \frac{\eta+\xi}{x^R - \frac{\alpha}{1-\alpha}(\eta+\delta)-\delta}\end{aligned}$$

and we again have the result that lower population growth lowers this tail coefficient. Note also that a smaller transition rate ξ reduces the Pareto tail, that is concentration in the top rises as rockets transition into the slow types at a slower pace.

SM-3.2 Exit Rates by Size with Firm Heterogeneity

The introduction of type heterogeneity substantially improves the fit of the model against the data on exit rates by size. Simply put, this heterogeneity allows some firms to grow large by adding more products, an outcome which is relatively rare in the baseline model. Because of diversification across products, this lowers the exit rate for large firms (whereas in the baseline model, the overwhelming majority of large firms are so because they have a single, high q product). Figure SM-1 shows an illustrative calibration of the model with rockets, demonstrating a declining exit rate with size.

Figure SM-1: Exit Rate By Size with Rocket Firms



Notes: This figure shows the exit rate by size in an illustrative calibration of the model with rocket firms. The transition rate is set to $\xi = 0.25$, and the share of rockets at 0.1. The innovation rate of rockets x^R is chosen to match average sales growth by age 10 from the LBD, as in the main quantitative section, while the rate of slow firms is set to 0.1.

SM-4 Calculating the conditional density $f_{a_P|A_f,N}\left(a_P|a_f, n\right)$

We now derive the conditional density of product age a_P, $f_{a_P|A_f,N}\left(a_P|a_f, n\right)$.

For illustration, we first derive the expected age of the products in a firm's portfolio as it ages. To do so, consider the mass of firms with n products at age A. We are going to derive the law of motion for the *total number of years* the products that this mass of firms owns have been alive (think of products accumulating years for every instant they have been alive). Call this object $\Psi_A(n)$, where

$$\Psi_A(n) = \underbrace{\Lambda_A(n)n}_{\text{Total number of products by firms of age } A} \underbrace{\mathbb{E}_A[a|n]}_{\text{Average age of products of firms of age } A \text{ and } n \text{ products}} .$$

The pool of total years $\Psi_A(n)$ is equal to the number of firms of age A with n products, denoted $\Lambda_A(n)$, times the number of products they own n, times the average age of all those products $\mathbb{E}_A[a|n]$.

We are going to consider how this object evolves through a discrete time approximation. For a small time interval ι,

$$
\begin{aligned}
\mathbb{E}_A[a|n]\Lambda_A(n)n = & \underbrace{(\mathbb{E}_{A-\iota}[a|n] + \iota)\Lambda_{A-\iota}(n)n(1-(\tau+\delta+x)n\iota)}_{\text{drift from existing mass}} \\
& + \underbrace{\iota x(n-1)\Lambda_{A-\iota}(n-1)\Big((n-1)\mathbb{E}_{A-\iota}[a|n-1])\Big)}_{\text{flow in from n-1 firms}} \\
& + \underbrace{\iota(\tau+\delta)(n+1)\Lambda_{A-\iota}(n+1)\Big(n\mathbb{E}_{A-\iota}[a|n+1]\Big)}_{\text{flow in from n+1 firms}}
\end{aligned}
$$

The first term in this expression is the drift in total years from an increment of time ι, multiplied by the fraction of firms who don't drop or gain a product in this increment. Intuitively, these products age with a unit drift. The second term is the flow of total years into the pool $\Psi_A(n)$ from the mass of firms with $n-1$ products who are each gaining a product. Importantly, while they bring n products each into the year pool, only $n-1$ have a positive age, and their average age is $\mathbb{E}_{A-\iota}[a|n-1]$. Lastly, the third term is the flow from the mass of firms with $n+1$ products who are losing a product. They bring n products with average age $\mathbb{E}_A[a|n+1]$ with them.

Rewrite this as

$$
\begin{aligned}
\frac{\Psi_A(n) - \Psi_{A-\iota}(n)}{\iota} = & \Lambda_A(n)n - (\tau+\delta+x)n\mathbb{E}_{A-\iota}[a|n]\Lambda_A(n)n \\
& + x(n-1)\Lambda_A(n-1)\Big((n-1)\mathbb{E}_{A-\iota}[a|n-1])\Big) \\
& + (\tau+\delta)(n+1)\Lambda_A(n+1)\Big(n\mathbb{E}_{A-\iota}[a|n+1]\Big)
\end{aligned}
$$

so that

$$
\begin{aligned}
\frac{\Psi_A(n)}{dA} &= \Lambda_A(n)n - (\tau+\delta+x)n\mathbb{E}_A[a|n]\Lambda_A(n)n \\
&+ x(n-1)\Lambda_A(n-1)\Big((n-1)\mathbb{E}_A[a|n-1]\Big) \\
&+ (\tau+\delta)(n+1)\Lambda_A(n+1)\Big(n\mathbb{E}_A[a|n+1]\Big)
\end{aligned} \tag{SM-8}
$$

This gives us a set of equations for the evolution of $\Psi_A(n)$ for all $n > 1$ that can be solved computationally given initial conditions. We also need one for $n = 1$, which comes from

$$
\begin{aligned}
\frac{d\mathbb{E}_A[a|1]\Lambda_A(1)1}{dA} &= \Lambda_A(1) - (\tau+\delta+x)\mathbb{E}_{A-\iota}[a|1]\Lambda_A(1) \\
&+ (\tau+\delta)(2)\Lambda_A(2)\Big(\mathbb{E}_A[a|2]\Big)
\end{aligned}
$$

The initial condition is that $\mathbb{E}_0[a|n]\Lambda_0(n)n = \Psi_0(n) = 0$ for all n. The equations we solve computationally are

$$
\frac{\Psi_A(n)}{dA} = \Lambda_A(n)n - (\tau+\delta+x)n\Psi_A(n) + x(n-1)\Psi_A(n-1) + (\tau+\delta)n\Psi_A(n+1).
$$

Lastly, to recover $\mathbb{E}_A[a|n]$ after computing $\Psi_A(n)$, note that $\Lambda_A(n) = F_0 p_A(n)$, where F_0 is the initial number of firms, and $p_A(n)$ as above is the probability that a firm of age A will have n products , for which we have closed form expressions. Then $\mathbb{E}_A[a|n] = \frac{\Psi_A(n)}{\Lambda_A(n)n}$. Finally, to compute the expected age of products for surviving firms of age A, we have

$$
E_A[a] = \sum_{n=1}^{\infty} \mathbb{E}_A[a|n] \frac{p_A(n)}{1 - p_A(0)}
$$

We use this object in computing markups and sales by firm age, since product markup is a deterministic function of product age.

Full Product Age Distribution

Consider the object $X_{A,n}(a) = \Lambda_A(n)n\Phi_{A,n}(a)$, the total number of products with age less than a by firms of age A with n products. Recall that $\Lambda_A(n)$ is the total number of firms of age A with n

products. Define $\Phi_{A,n}(a)$ as the probability that a product of a firm of age A with $n$products is less than or equal to a. This evolves as

$$\begin{aligned} X_{A,n}(a) &= \Lambda_{A-\iota}(n)\Phi_{A-\iota,n}(a-\iota)n(1-(\tau+\delta+x)n\iota) \\ &+ \iota x(n-1)\Lambda_{A-\iota}(n-1)\Big((n-1)\Phi_{A-\iota,n-1}(a)+1\Big) \\ &+ \iota(\tau+\delta)(n+1)\Lambda_{A-\iota}(n+1)\Big(n\Phi_{A-\iota,n+1}(a)\Big) \end{aligned}$$

Note the difference on the second line now, because the new product has age $0 < a$.

Write this as

$$\begin{aligned} \frac{X_{A,n}(a) - X_{A-\iota,n}(a-\iota)}{\iota} =& -(\tau+\delta+x)nX_{A-\iota,n}(a-\iota) \\ &+ x(n-1)\Lambda_{A-\iota}(n-1) + x(n-1)X_{A-\iota,n-1}(a) \\ &+ (\tau+\delta)nX_{A-\iota,n+1}(a) \end{aligned} \tag{SM-9}$$

$$\begin{aligned} \frac{X_{A,n}(a) - X_{A-\iota,n}(a) + X_{A-\iota,n}(a) - X_{A-\iota,n}(a-\iota)}{\iota} =& -(\tau+\delta+x)nX_{A-\iota,n}(a-\iota) \\ &+ x(n-1)\Lambda_{A-\iota}(n-1) + x(n-1)X_{A,n-1}(a) \\ &+ (\tau+\delta)nX_{A,n+1}(a) \end{aligned}$$

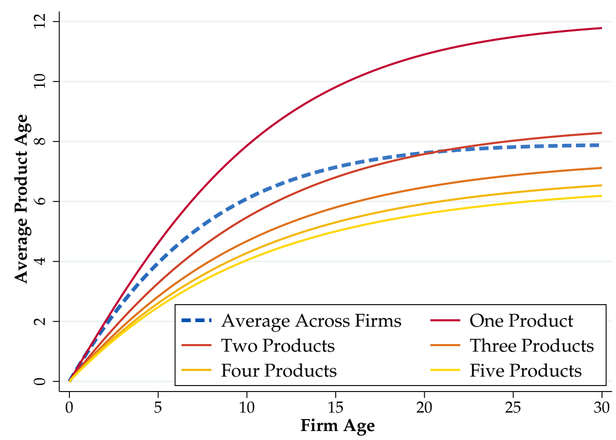
which goes to

$$\begin{aligned} \frac{\partial X_{A,n}(a)}{\partial A} + \frac{\partial X_{A,n}(a)}{\partial a} =& -(\tau+\delta+x)nX_{A,n}(a) \\ &+ x(n-1)\Lambda_{A}(n-1) + x(n-1)X_{A,n-1}(a) \\ &+ (\tau+\delta)nX_{A,n+1}(a) \end{aligned} \tag{SM-10}$$

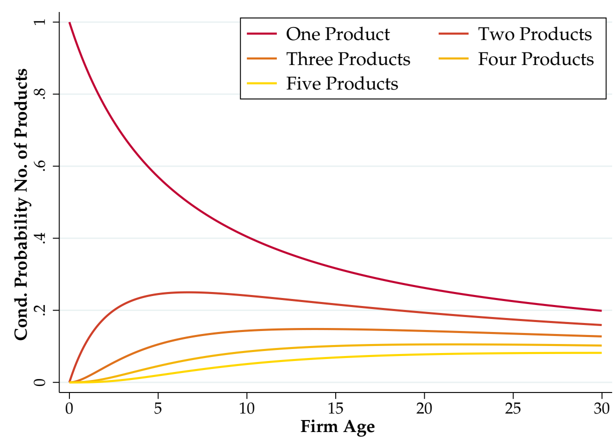
In Figure SM-2 we depict the average product by firm age (left panel) and the probability of having n products as a function of age. The left panel shows the effect of selection on the average product age of multi-product firms. Conditional on the age of the firm, the average product is declining in the number of products because newly added products are - by construction - younger. In the right panel we show five "slices" of the joint distribution of age and the number of products. One-product firms are mostly young firms as all firms enter with a single product. Old firms only rarely have a single product as they either grew or exited already. The remaining lines show that older firms are more and more likely to have many products.

Figure SM-2: Product Age and Firm Age

(a) Average Product Age By Firm Age

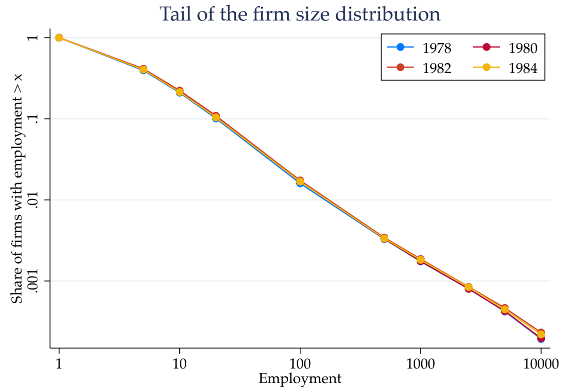


(b) Probability of Product Portfolios By Age



Notes: Panel (a) of this figure plots the average product age for a firm of n products for $n = 1, ..., 5$ as the firm ages in the calibrated model. These objects are computed using the productivity distribution $X_{A,n}(a)$ (see (SM-10)). Panel (b) plots the conditional probability of a portfolio of n products for $n = 1, ..., 5$ in the calibrated model for surviving firms by firm age.

Figure SM-3: Estimating the tail of the firm size distribution



Notes: The figure plots $\ln P_t\left(l_f > x\right)$ against $\ln x$ for different years. The data is taken from the BDS.

SM-4.1 Estimating the Pareto tail of the employment distribution

One of our target moments is the Pareto tail of the employment distribution. The distribution of firm employment at time t is - for large firms - given by $P_t\left(l_f > x\right) = \left(\bar{l}/x\right)^{\varsigma}$. Hence,

$$\ln P_t\left(l_f > x\right) = \delta - \varsigma \ln x. \tag{SM-11}$$

In Figure SM-3 we show the empirical relationship between $\ln P_t\left(l_f > x\right)$ and $\ln x$ for different years. As predicted by (SM-11), the relationship is almost perfectly linear and stable across years. When we estimate (SM-11), we find $\varsigma \approx 1$. If anything ς is slightly smaller than 1. To keep average size bounded, we require $\varsigma > 1$. We therefore opt to calibrate our model to a tail of 1.1.

SM-5 Computing the the cross-sectional size and age distribution (Section 4.4)

In Section 4.4 we reported the model-implied distribution of firm size and firm age. We now show how to derive these objects.

SM-5.1 The cross-sectional size distribution

We have expressions for the probability distribution of number of products by age p_n, as well as the age distribution of firms. So the final piece is the distribution of employment conditional on firm age and number of products. To do so, define

$$l_{A,n} = \mu_i^{-\sigma} \left(\frac{q_i}{Q_t} \right)^{\sigma-1} \frac{L_t}{N_t} \Lambda_t^{-\sigma} \mathcal{M}_t^{1-\sigma}$$

as the random variable of employment at the product level, conditional on the firm having age of A and n products. Conditional on the age of the firm and the number of products, the distribution of $l_{A,n}$ is independent across products, so we derive the distribution of the sum of these objects through a convolution. For the first product

$$Prob(l_{A,n} \le y) \equiv D^1_{A,n}(y) = Prob(\ln \left(\frac{q_i}{Q_t} \right)^{\sigma-1} \le \ln(y) + \ln(\mu_i^{\sigma}) - \ln \left(\frac{L_t}{N_t} \Lambda_t^{-\sigma} \mathcal{M}_t^{1-\sigma} \right)$$
$$=$$

Now note that $\mu = \Delta = \lambda e^{Ia}$ for a below the critical threshold. Define the joint density of (log) productivity and gaps $f^C(\tilde{q}, \Delta)$from

$$F^C(\hat{q}, \Delta) = \int_{-\infty}^{\hat{q}} \int_{\lambda}^{\Delta} f(x, y) dx dy$$

with associated conditional density $f^C_{\hat{q}|\Delta}(\tilde{q}|\Delta)$ and conditional distribution function $F^C_{\hat{q}|\Delta}(\hat{q}|\Delta)$. Lastly, denote the distribution of productivity for the non-competitor mass as $F^{NC}(\hat{q}|a)$. Given that incumbent innovation is constant for non-competitive products, the law of motion for the mass of products $\bar{F}^{NC}_a(\hat{q})$ at age a is given by

$$\frac{\partial \bar{F}^{NC}_a(\hat{q})}{\partial a} = \frac{\partial F^{NC}(\hat{q}|a)}{\partial \hat{q}} (\sigma - 1)((I(\bar{\Delta}) - \gamma) - \bar{F}^{NC}_a(\hat{q})(\tau_t + \delta)$$

with initial condition $\bar{F}^{NC}_0(\hat{q}) = \Gamma \left(\frac{\exp(\hat{q})}{\sigma - 1} \right)$. From this we can compute the conditional distribution of productivity $F^{NC}(\hat{q}|a)$. With these pieces we can compute the distribution of employment as

$$\begin{aligned} D^1_{A,n}(y) = & \int_0^{\bar{a}} d\Phi_{A,n}(a) F^C(log(y) + \log(\mu^{\sigma}) - c|\Delta(a)) \\ & + \int_{\bar{a}}^{A} d\Phi_{A,n}(a)) F^C(log(y) + \log(\bar{\mu}^{\sigma}) - c|\Delta(a)) \\ & + \int_0^{A} d\Phi_{A,n}(a)) F^{NC}(log(y) + \log(\bar{\mu}^{\sigma}) - c|a) \end{aligned}$$

where $c \equiv log(\frac{L_t}{N_t}\Lambda_t^{-\sigma}\mathcal{M}_t^{1-\sigma})$. Now we have the distribution of this object, we can define recursively the distribution of the sums of employment across products from a convolution. Define $Z^j_{A,n} = \sum_{i=1}^{j} l^i_{A,n}$ and then

$$P(Z^j_{A,n} \le y) = D^j_{A,n}(y) = \int_0^y \int_{-\infty}^{\infty} \frac{dD^{j-1}_{A,n}(x)}{dx}\frac{dD^1(z-x)}{dx}dxdz$$

for $j \ge 2$. Then, for each age of the firm we can define the conditional employment distribution as

$$Prob(E_f \le y | a_f = A) = \sum_{n=1}^{\infty} \frac{p_n(A)}{1-p_o(A)} D^n_{A,n}(y).$$

SM-5.2 The cross-sectional age distribution

Let Y_{ta} be the number of firms who are a years old at time t. The total number of firms at time t is then given by $Y_t = \int_{a=0}^{\infty} Y_{ta} da$. Let E_τ denote the number of entrants at time τ. Then

$$Y_{ta} = \underbrace{E_{t-a}}_{\text{Entrants}} \underbrace{S(a)}_{\text{Survival}}.$$

Note also that the number of entrants is given by $E_\tau = zN_\tau$. And as N_τ grows at rate η, we have $E_\tau = zN_0 e^{\eta\tau}$. Hence

$$Y_{ta} = zN_0 e^{\eta(t-a)} S(a).$$

The density of firms which are a years old is therefore given by

$$\omega^F_t(a) = \frac{Y_{ta}}{Y_t} = \frac{zN_0 e^{\eta(t-a)} S(a)}{\int_{a=0}^{\infty} zN_0 e^{\eta(t-a)} S(a)\, da} = \frac{\frac{\psi e^{-(\psi+\eta)a}}{\psi + x(1-e^{-\psi a})}}{\int_{a=0}^{\infty} \frac{\psi e^{-(\psi+\eta)a'}}{\psi + x\left(1-e^{-\psi a'}\right)} da'},$$

where the last line uses (A-22).

SM-6 Exit rates by size

To compute exit rates by size (show in Figure A-3) we can compute an exit rate per product. Then, the probability of having a number of products n by age A of the firm, conditional on being a certain size y (between l and $u$employees) is

$$Prob(A, n|l \leq y \leq u) = \frac{Prob(l \leq y \leq u|A, n) \times P(A, n)}{\sum_{A'} \sum_{n'} Prob(l \leq y \leq u|A', n') \times P(A', n')}$$

The joint probability of the age bins and number of products is $P(A, n) = p_n(A)\omega_t^F(A)$. Then we can construct the probability

$$Prob(n|l \leq y \leq u) = \sum_A \frac{Prob(l \leq y \leq u|A, n) \times P(A, n)}{\sum_{A'} \sum_{n'} Prob(l \leq y \leq u|A', n') \times P(A', n')}$$

where we are using discrete A bins to compute this object. Once we have the conditional probabilities of numbers of products by size bins, we can compute exit rates by size bins, since exit only depends on the number of products. The exit probability for each number of products n can be calculated as follows.

The probability of losing k products in an interval Δ if you lose each product at rate $\tau + \delta$ is

$$p_n(k, \tau) = e^{-\tau n \Delta} \frac{((\tau + \delta) n \Delta)^k}{k!}$$

The probability of winning m products in an interval Δ if you expand at rate x

$$g_n(m, x) = e^{-x n \Delta} \frac{(x n \Delta)^m}{m!}$$

Hence, the probability of exit when a firm has n products is

$$\begin{aligned} Prob(k - m >= n) &= E_m [Prob(k >= n + m)] = E_m \left[\sum_{k=n+m+1}^{\infty} p_n(k, (\tau + \delta)) \right] \\ &= \sum_{m=0}^{\infty} e^{-x n \Delta} \frac{(x n \Delta)^m}{m!} \sum_{k=n+m}^{\infty} e^{-\tau n \Delta} \frac{((\tau + \delta) n \Delta)^k}{k!}. \end{aligned}$$

SM-7 Computing the Transitional Dynamics

In this section we characterize the transitional dynamics of our model. In Section SM-7.1 we solve for the value function without imposing the economy to be on the BGP. In Section SM-7.2 we characterize the value of entry during the transitional dynamics. In Section SM-7.3 we use the free entry condition to characterize the differential equation for the free entry equilibrium during the transition. In Section SM-7.4 we derive the characterization of the system of equations that fully characterize the transitional dynamics. In Section (SM-7.5) we derive the differential equation for the joint distribution of quality q and quality gaps Δ, $F_t(q,\Delta)$, that we need to compute the evolution of markups along the transition.

SM-7.1 The value function

As shown in Section A-1.3, the value function is additive across products and the value of a single product with quality q and quality gap Δ is described by the HJB equation

$$r_t V_t(q,\Delta) - \dot{V}_t(q,\Delta) = \pi_t(q_i,\Delta_i) + \left(\frac{\partial V_t(q,\Delta)}{\partial q} + \frac{\partial V_t(q,\Delta)}{\partial \Delta}\frac{\partial \Delta}{\partial q}\right)\dot{q} - (\tau_t+\delta)V_t(q,\Delta) + \Xi_t,$$

where

$$\pi_t(q_i,\Delta_i) = \left(1-\frac{1}{\mu(\Delta_i)}\right)\mu(\Delta_i)^{1-\sigma}\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\frac{1}{\mathcal{M}_t^{\sigma-1}\Lambda_t^{\sigma}}\frac{L_P}{N_t}w_t.$$

Note that $\frac{\partial \Delta}{\partial q} = \frac{1}{q}\Delta$. Also note that the free entry condition still implies that $\Xi_t = \frac{\zeta-1}{\varphi_x}x^{\zeta}w_t$. Hence, the HJB equations reduces to

$$r_t V_t(q,\Delta) - \dot{V}_t(q,\Delta) = \pi_t(q_i,\Delta_i) + \left(\frac{\partial V_t(q,\Delta)}{\partial q}q + \frac{\partial V_t(q,\Delta)}{\partial \Delta}\Delta\right)I - (\tau_t+\delta)V_t(q,\Delta) + \frac{\zeta-1}{\varphi_x}x^{\zeta}w_t,$$

where $\frac{\dot{q}}{q} = I$.

Now conjecture that the value function takes the form

$$V_t(q,\Delta) = \left(\frac{q}{Q_t}\right)^{\sigma-1}U_t(\Delta) + M_t. \quad \text{(SM-12)}$$

This implies that

$$\frac{\partial V_t(q,\Delta)}{\partial q}q + \frac{\partial V_t(q,\Delta)}{\partial \Delta}\Delta = ((\sigma-1)+\varepsilon_k(\Delta))\left(\frac{q}{Q_t}\right)^{\sigma-1}U_t(\Delta),$$

where

$$\varepsilon_t(\Delta) \equiv \frac{U_t'(\Delta)\Delta}{U_t(\Delta)}. \quad \text{(SM-13)}$$

Using the conjecture in (SM-12), the HJB simplifies to the following two equations:

1. The function M_t in (SM-12) solves the differential equation

$$\left(r_t + \tau_t + \delta\right) M_t - \dot{M}_t = \frac{\zeta - 1}{\varphi_x} x^{\zeta} w_t$$

2. The function $U_t\left(\Delta\right)$ in (SM-12) solves the differential equation

$$\left(r_t + \tau_t + \delta + \left(\sigma - 1\right)\left(g_Q - I\right) - I\varepsilon_t\left(\Delta\right)\right) U_t\left(\Delta\right) - \dot{U}_t\left(\Delta\right) = h\left(\Delta\right) \frac{1}{\mathcal{M}_t^{\sigma-1}\Lambda_t^{\sigma}} \frac{L_P}{N_t} w_t$$

where

$$h\left(\Delta\right) \; = \; \left(1 - \frac{1}{\mu\left(\Delta_i\right)}\right) \mu\left(\Delta_i\right)^{1-\sigma} = \left(\frac{\mu\left(\Delta_i\right) - 1}{\mu\left(\Delta_i\right)^{\sigma}}\right) = \frac{\min\left\{\Delta, \frac{\sigma}{\sigma-1}\right\} - 1}{\left(\min\left\{\Delta, \frac{\sigma}{\sigma-1}\right\}\right)^{\sigma}}$$

and $\varepsilon_t\left(\Delta\right)$ is given in (SM-13).

SM-7.2 The value of entry

The value of entry is given by

$$V_t^{Entry} \; = \; \underbrace{\alpha \int V_t\left(\lambda q, \lambda\right) dF_t\left(q\right)}_{\text{CD with gap } \lambda \text{ and quality } \lambda q} + \underbrace{\left(1 - \alpha\right) \int V_t\left(\omega Q_t, \frac{\sigma}{\sigma - 1}\right) d\Gamma\left(\omega\right)}_{\text{New variety with gap } \frac{\sigma}{\sigma-1} \text{ and quality } \omega Q_t}$$

Using the conjecture in (SM-12), V_t^{Entry} can be written as

$$V_t^{Entry} \; = \; \alpha \lambda^{\sigma-1} U_t\left(\lambda\right) + \left(1 - \alpha\right) \overline{\omega}^{\sigma-1} U_t\left(\frac{\sigma}{\sigma - 1}\right) + M_t.$$

Upon defining $v_t^{Entry} = \frac{V_t^{Entry}}{w_t}$, $m_t = \frac{M_t}{w_t}$ and $\underline{u}_t\left(\Delta\right) = \frac{U_t(\Delta)}{w_t}$, we get

$$v_t^{Entry} = \alpha \lambda^{\sigma-1} \underline{u}_t\left(\lambda\right) + \left(1 - \alpha\right) \overline{\omega}^{\sigma-1} \underline{u}_t\left(\frac{\sigma}{\sigma - 1}\right) + m_t \qquad \text{(SM-14)}$$

where m_t solves

$$\left(r_t + \tau_t + \delta - g_w\right) m_t - \dot{m}_t = \frac{\zeta - 1}{\varphi_x} x^{\zeta} \tag{SM-15}$$

and $\underline{u}_t\left(\Delta\right)$ solves

$$\begin{aligned} \left(r_t + \tau_t + \delta - g_w + (\sigma - 1)\left(g_Q - I\right) - I\varepsilon_t\left(\Delta\right)\right)\underline{u}_t\left(\Delta\right) - \dot{\underline{u}}_t\left(\Delta\right) &= h\left(\Delta\right)\frac{1}{\mathcal{M}_t^{\sigma-1}\Lambda_t^{\sigma}}\frac{L_P}{N_t} \\ &= h\left(\Delta\right)\frac{1}{\mathcal{M}_t^{\sigma-1}\Lambda_t^{\sigma}}\ell_t s_t^P, \end{aligned} \tag{SM-16}$$

where $\ell_t = \frac{L_t}{N_t}$ and $s_t^P = \frac{L_{Pt}}{L_t}$.

SM-7.3 Free Entry

Free entry requires $v_t^{Entry} = 1/\varphi_E$ so that

$$\frac{1}{\varphi_E} = \alpha\lambda^{\sigma-1}\underline{u}_t\left(\lambda\right) + \left(1 - \alpha\right)\overline{\omega}^{\sigma-1}\underline{u}_t\left(\frac{\sigma}{\sigma - 1}\right) + m_t. \tag{SM-17}$$

This also implies $\dot{v}_t^{Entry} = 0$. From (SM-14) this means $\dot{\underline{u}}_t$ and $\dot{m}_t$ satisfy the restriction.

$$0 = \alpha\lambda^{\sigma-1}\dot{\underline{u}}_t\left(\lambda\right) + \left(1 - \alpha\right)\overline{\omega}^{\sigma-1}\dot{\underline{u}}_t\left(\frac{\sigma}{\sigma - 1}\right) + \dot{m}_t.$$

Together with (SM-15), we can use this restriction to solve for m_t in terms of $\underline{u}_t$ as

$$m_t = \frac{\frac{\zeta-1}{\varphi_x}x^{\zeta} - \left(\alpha\lambda^{\sigma-1}\dot{\underline{u}}_t\left(\lambda\right) + \left(1-\alpha\right)\overline{\omega}^{\sigma-1}\dot{\underline{u}}_t\left(\frac{\sigma}{\sigma-1}\right)\right)}{r_t + \tau_t + \delta - g_w}.$$

Substituting this in (SM-17) yields

$$\begin{aligned} \frac{1}{\varphi_E} = {} & \alpha\lambda^{\sigma-1}\left(\underline{u}_t\left(\lambda\right) - \frac{\dot{\underline{u}}_t\left(\lambda\right)}{r_t + \tau_t + \delta - g_w}\right) \\ & + \left(1 - \alpha\right)\overline{\omega}^{\sigma-1}\left(\underline{u}_t\left(\frac{\sigma}{\sigma - 1}\right) - \frac{\dot{\underline{u}}_t\left(\frac{\sigma}{\sigma-1}\right)}{r_t + \tau_t + \delta - g_w}\right) + \frac{\zeta - 1}{\varphi_x}\frac{x^{\zeta}}{r_t + \tau_t + \delta - g_w}, \end{aligned} \tag{SM-18}$$

where u_t solves the differential equation (SM-16).

Note that two discount rates appear in these equations.

1. First we have $r_t + \tau_t + \delta - g_w + (\sigma - 1)\left(g_Q - I\right)$ in (SM-16). This can be written as

$$
\begin{aligned}
r_t + \tau_t + \delta - g_w + (\sigma - 1)\left(g_Q - I\right) = {} & \rho + g_{L^P} - g_\Lambda - \eta \\
& + \left(\overline{\omega}^{\sigma-1} + \lambda^{\sigma-1}\frac{\alpha}{1-\alpha}\right)\delta + \left(\overline{\omega}^{\sigma-1} - 1 + \lambda^{\sigma-1}\frac{\alpha}{1-\alpha}\right)g_t^N,
\end{aligned}
$$

where $g_\Lambda = \dot{\Lambda}_t/\Lambda_t$ and $g_{L^P} = \dot{L}_t^P/L_t^P$. Using $s_t^P = L_t^P/L_t$ we have$g_{s_t^P} = g_{L^P} - \eta$. Hence,

$$
\begin{aligned}
r_t + \tau_t + \delta - g_w + (\sigma - 1)\left(g_Q - I\right) = {} & \rho + g_{s_t^P} - g_\Lambda \\
& + \left(\overline{\omega}^{\sigma-1} + \lambda^{\sigma-1}\frac{\alpha}{1-\alpha}\right)\delta + \left(\overline{\omega}^{\sigma-1} - 1 + \lambda^{\sigma-1}\frac{\alpha}{1-\alpha}\right)g_t^N,
\end{aligned}
$$

2. Second we have the expression $r_t + \tau_t + \delta - g_w$ in (SM-18). This can be written as

$$
r_t + \tau_t + \delta - g_w \quad = \quad \rho + g_{s_t^P} - g_\Lambda + \frac{\alpha}{1-\alpha}g_N + \frac{1}{1-\alpha}\delta
$$

SM-7.4 Final Dynamic system

We now derive the final characterization equations characterizing the transitional dynamics. Note first that labor market requires

$$
\ell_t\left(1 - s_t^P\right) = \frac{1}{\varphi_E}\left(\frac{g_t^N + \delta}{1-\alpha} - \frac{\zeta - 1}{\zeta}x\right), \tag{SM-19}
$$

where $\ell_t = \frac{L_t}{N_t}$ and $s_t^P = \frac{L_{Pt}}{L_t}$. The free entry condition (SM-18) reads

$$
\begin{aligned}
\frac{1}{\varphi_E} \quad = \quad & \alpha\lambda^{\sigma-1}\left(\underline{u}_t(\lambda) - \frac{\dot{\underline{u}}_t(\lambda)}{\rho + g_{s_t^P} - g_\Lambda + \frac{\alpha}{1-\alpha}g_t^N + \frac{1}{1-\alpha}\delta}\right) \\
& + (1-\alpha)\,\overline{\omega}^{\sigma-1}\left(\underline{u}_t\left(\frac{\sigma}{\sigma-1}\right) - \frac{\dot{\underline{u}}_t\left(\frac{\sigma}{\sigma-1}\right)}{\rho + g_{s_t^P} - g_\Lambda + \frac{\alpha}{1-\alpha}g_t^N + \frac{1}{1-\alpha}\delta}\right) \qquad \text{(SM-20)} \\
& + \frac{\zeta-1}{\varphi_x}\frac{x^\zeta}{\rho + g_{s_t^P} - g_\Lambda + \frac{\alpha}{1-\alpha}g_t^N + \frac{1}{1-\alpha}\delta} \qquad \text{(SM-21)}
\end{aligned}
$$

where $\underline{u}_t(\Delta)$ solves

$$\left(\rho + g_{s_t^P} - g_\Lambda + \left(\overline{\omega}^{\sigma-1} + \lambda^{\sigma-1}\frac{\alpha}{1-\alpha}\right)\delta + \left(\overline{\omega}^{\sigma-1} - 1 + \lambda^{\sigma-1}\frac{\alpha}{1-\alpha}\right)g_t^N\right)\underline{u}_t(\Delta) \tag{SM-22}$$

$$= -I\frac{\partial u_t(\Delta)}{\partial \Delta}\Delta - \dot{\underline{u}}_t(\Delta) = h(\Delta)\frac{1}{\mathcal{M}_t^{\sigma-1}\Lambda_t^{\sigma}}\ell_t s_t^P. \tag{SM-23}$$

This is a differential equation in Δ and t. We have two terminal conditions. For $\Delta \geq \frac{\sigma}{\sigma-1}$ we have

$$h\left(\frac{\sigma}{\sigma-1}\right) = \frac{\frac{\sigma}{\sigma-1}-1}{\left(\frac{\sigma}{\sigma-1}\right)^{\sigma}} = \left(\frac{1}{\sigma-1}\right)^{1-\sigma}\frac{1}{\sigma^{\sigma}} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}}$$

and $\frac{\partial \underline{u}_t(\Delta)}{\partial \Delta} = 0$. Furthermore, we have that $\frac{1}{\mathcal{M}_t^{\sigma-1}\Lambda_t^{\sigma}}\ell_t s_t^P$ is constant in the steady state so that $\dot{\underline{u}}_t(\Delta) \to 0$.

The transitional dynamics of the system is a path of $\{\ell_t, s_t^P\}_t$ that solves the equations above. Note that given $\{\ell_t, s_t^P\}_t$, we can calculate g_t^N from SM-19. Given g_t^N we can calculate τ_t and $g_{Q,t}$ as

$$\tau_t = \frac{\alpha}{1-\alpha}\left(g_t^N + \delta\right) \tag{SM-24}$$

$$(g_Q - I)(\sigma - 1) = \left(\left(\lambda^{\sigma-1} - 1\right)\frac{\alpha}{1-\alpha} + \overline{\omega}^{\sigma-1} - 1\right)\left(g_t^N + \delta\right). \tag{SM-25}$$

As we show in Section , this is also sufficient to compute $\left\{\mathcal{M}_t^{\sigma-1}\Lambda_t^{\sigma}\right\}_t$ and g_Λ.

SM-7.5 The Evolution of $F_t(\Delta, q)$: Calculating $\{\mathcal{M}_t, \Lambda_t\}_t$

To calculate $\mathcal{M}_t$ and Λ_t we require the joint distribution of productivity q and quality gaps Δ, $F_t(\Delta, q)$. Note that $\mathcal{M}_t$ and Λ_t only depend on q via $(q/Q)^{\sigma-1}$. Hence, it is useful to characterize the distribution of $F_t(\Delta, \hat{q})$, where $\hat{q}_t = \ln(q_t/Q_t)^{\sigma-1}$. Let $F_t^C(\Delta, \hat{q})$ denote the distribution among products that have a competitor and $F_t^{NC}(\hat{q})$ denote the distribution for products without a competitor.[26] Let N_t^C and N_t^{NC} denote the mass of these products. Let $\hat{F}_t^C(\Delta, \hat{q}) = F_t^C(\Delta, \hat{q})\frac{N_t^C}{N_t}$ and $\hat{F}_t^{NC}(\hat{q}) = F_t^{NC}(\hat{q})\frac{N_t^C}{N_t}$.

[26]Recall that we do not need Δ for the non-competitor products as they all have a markup of $\frac{\sigma-1}{\sigma}$.

If we have the full evolution of $\left\{N_t, \hat{F}_t^C(\Delta,\hat{q}), \hat{F}_t^{NC}(\hat{q})\right\}_t$ we can calculate Λ_t and $\mathcal{M}_t$ as

$$
\begin{aligned}
\Lambda_t &= \frac{N_t \int \mu\left(\Delta\right)^{-\sigma} e^{\hat{q}} d\hat{F}_t^C\left(\Delta,\hat{q}\right) + N_t \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \int e^{\hat{q}} d\hat{F}_t^{NC}\left(\hat{q}\right)}{N_t \int \mu\left(\Delta\right)^{1-\sigma} e^{\hat{q}} d\hat{F}_t^C\left(\Delta,\hat{q}\right) + N_t \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \int e^{\hat{q}} d\hat{F}_t^{NC}\left(\hat{q}\right)} \\
\mathcal{M}_t &= \frac{\left(N_t \int \mu\left(\Delta\right)^{1-\sigma} e^{\hat{q}} d\hat{F}_t^C\left(\Delta,\hat{q}\right) + N_t \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \int e^{\hat{q}} d\hat{F}_t^{NC}\left(\hat{q}\right)\right)^{\frac{\sigma}{\sigma-1}}}{N_t \int \mu\left(\Delta\right)^{-\sigma} e^{\hat{q}} d\hat{F}_t^C\left(\Delta,\hat{q}\right) + N_t \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \int e^{\hat{q}} d\hat{F}_t^{NC}\left(\hat{q}\right)},
\end{aligned}
$$

where $e^{\hat{q}} = \left(q_t/Q_t\right)^{\sigma-1}$ as $\hat{q}_t = \ln\left(q_t/Q_t\right)^{\sigma-1}$ and $\mu\left(\Delta\right)$ is the markup function $\mu\left(\Delta\right) = \min\left\{\frac{\sigma}{\sigma-1}, \Delta\right\}$.

We now derive expressions to calculate the evolution of $\left\{N_t, \hat{F}_t^C(\Delta,\hat{q}), \hat{F}_t^{NC}(\hat{q})\right\}_t$. Let $\left(N_0^C, N_0^{NC}, \hat{F}_0^C(\Delta,\hat{q}), \hat{F}_0^{NC}(\right.$ be given. In practice these objects are determined in the initial BGP. This in particular implies that $N_0^C = \alpha N_t$ and $N_0^{NC} = \left(1-\alpha\right) N_t$. Suppose a path for $\left\{g_t^N\right\}$ is given. Then we can calculate and $\left\{\tau_t\right\}$ and $\left\{g_{Qt}\right\}$ from (SM-24) and (SM-25)

1. Given $\left\{\tau_t\right\}$ and $\left\{g_{Qt}\right\}$, we can calculate $\left\{\hat{F}_t^C(\Delta,\hat{q}), \hat{F}_t^{NC}(\hat{q})\right\}_t$ as follows:

   (a) The evolution of $\hat{F}_t^{NC}(\hat{q})$ is given by

   $$
   \frac{\partial \hat{F}_t^{NC}(\hat{q})}{\partial t} = -g_{\hat{q}} \frac{\partial \hat{F}_t^{NC}(\hat{q})}{\partial \hat{q}} - \left(\tau_t + \delta + g_t^N\right) \hat{F}_t^{NC}(\hat{q}) + \left(\frac{1-\alpha}{\alpha}\right) \tau_t \Gamma\left(\exp\left(\frac{\hat{q}}{\sigma-1}\right)\right),
   $$

   where $g_{\hat{q}} = \left(\sigma-1\right)\left(I - g_{Qt}\right)$ is given in (SM-25) and $\Gamma\left(\exp\left(\frac{\hat{q}}{\sigma-1}\right)\right)$ is the exogenous source distribution.

   (b) Given $\left\{\hat{F}_t^{NC}(\hat{q})\right\}_t$ we can solve for $\left\{\hat{F}_t^C(\Delta,\hat{q})\right\}_t$. In particular, $\left\{\hat{F}_t^C(\Delta,\hat{q})\right\}_t$ then solves the differential equation

   $$
   \begin{aligned}
   \frac{\partial \hat{F}_t^C(\Delta,\hat{q})}{\partial t} = &-\Delta I \frac{\partial \hat{F}_t^C\left(\Delta,\hat{q}\right)}{\partial \Delta} - g_{\hat{q}} \frac{\partial \hat{F}_t^C\left(\Delta,\hat{q}\right)}{\partial \hat{q}} \\
   &- \left(\tau + \delta + g_t^N\right) \hat{F}_t^C(\Delta,\hat{q}) + \lim_{s\to\infty} \tau \hat{F}_t^C(s, \hat{q} - \hat{\lambda}) + \tau \hat{F}_t^{NC}(\hat{q} - \hat{\lambda}),
   \end{aligned}
   $$

   where $\hat{\lambda} = \ln \lambda^{\sigma-1}$. Given that we solved for $\hat{F}_t^{NC}(\hat{q} - \hat{\lambda})$ already, this determines $\left\{\hat{F}_t^C(\Delta,\hat{q})\right\}_t$ given an initial condition $\hat{F}_0^C(\Delta,\hat{q})$

SM-7.6 Firm-level moments along the transition

Given the equilibrium path $\{g_{N,t}, z_t\}$ we can compute the time series of the entry rate and average firm size. To do so, let $\omega_t(n)$ be the mass of firms with n products at time t. Consider $n \geq 2$. Then

$$\dot{\omega}_t(n) = \underbrace{\omega_t(n-1)(n-1)x}_{\text{From } n-1 \text{ to} n \text{ products}} + \underbrace{\omega_t(n+1)(n+1)(\tau+\delta)}_{\text{From } n+1 \text{ to} n \text{ products}} - \underbrace{\omega_t(n)\,n(\tau+x+\delta)}_{\text{From } n \text{ to } n-1 \text{ or } n+1 \text{ products}} .$$

For $n = 1$ we have

$$\dot{\omega}_t(1) = Z_t + \omega_t(2)\,2(\tau+\delta) - \omega_t(1)(\tau+x+\delta).$$

Let $\nu_t(n) = \frac{\omega_t(n)}{N_t}$, which is stationary along the BGP. Then

$$\begin{aligned}
\frac{\dot{\omega}_t(n)}{N_t} &= \nu_t(n-1)(n-1)x + \nu_t(n+1)(n+1)(\tau+\delta) - \nu_t(n)\,n(\tau+x+\delta) \\
\frac{\dot{\omega}_t(1)}{N_t} &= z_t + \nu_t(2)\,2(\tau+\delta) - \nu_t(1)(\tau+x+\delta).
\end{aligned}$$

Now

$$\dot{\omega}_t(n) = \dot{\nu}_t(n) N_t + \nu_t(n)\dot{N}_t$$

so that

$$\frac{\dot{\omega}_t(n)}{N_t} = \dot{\nu}_t(n) + \nu_t(n)\,g_{N,t}$$

Hence,

$$\begin{aligned}
\dot{\nu}_t(n) &= \nu_t(n-1)(n-1)x + \nu_t(n+1)(n+1)(\tau+\delta) - \nu_t(n)\,n(\tau+x+\delta) - \nu_t(n)\,g_{N,t} \\
\dot{\nu}_t(1) &= z_t + \nu_t(2)\,2(\tau+\delta) - \nu_t(1)(\tau+x+\delta) - \nu_t(1)\,g_{N,t}.
\end{aligned}$$

Given an initial condition$\{\nu_0(n)\}_n$we can calculate the evolution of $\{\nu_t(n)\}_n$ for given $\{g_{N,t}, z_t\}$. Given $\{\nu_t(n)\}_n$ we can calculate some objects:

1. The number of firms at time t:

$$F_t = \sum_{n=1}^{\infty} \omega_t(n) = N_t \sum_{n=1}^{\infty} \nu_t(n)$$

   and hence average firm size L_t/F_t

2. The entry rate

$$Entry - rate_t = \frac{z_t N_t}{F_r} = \frac{z_t}{\sum_{n=1}^{\infty} \nu_t(n)}.$$

3. The exit rate

$$Exit - rate_t = \frac{(\tau_t + \delta) N_t \nu_t(1)}{F_t}.$$

SM-8 Characterization of the Model with Endogenous Innovation and Bertrand Competition

We now characterize the model in the case of both endogenous innovation on improving firm efficiency I and endogenous markups. For full generality we allow research labor and production labor to be imperfectly substitutable, such that the research wage v_t may not equal the production wage w_t. The state variables of the firm are efficiency q and the efficiency gap Δ for each product in the firm's portfolio, where we denote this set as $[\Delta_i, q_i]$.

The value function $V_t([\Delta_i, q_i])$ solves the HJB equation

$$\begin{aligned} r_t V_t([\Delta_i, q_i]) \quad = \quad & \sum_{i=1}^{n} \pi_t([\Delta_i, q_i]) + \dot{V}_t([\Delta_i, q_i]) + \sum_{i=1}^{n} (\tau + \delta) \left[V_t\left([\Delta_j, q_j]_{j \neq i} \right) - V([\Delta_i, q_i]) \right] \\ & + \max_{\{I_t(j)\}_{j=1}^{n}} \sum_{j=1}^{n} \left\{ I_t(j) \frac{\partial V_t([\Delta_i, q_i])}{\partial q_j} q_j - c^I(I_t(j), \Delta_j, q_j) v_t \right\} \\ & + \max_X \left\{ X \left[\alpha \int_q V_t([\Delta_i, q_i], 1, \lambda q) dF_t(q) \right. \right. \\ & \left. + (1-\alpha) \int_\omega \int_\Delta V_t([\Delta_i, q_i], \Delta, \omega Q_t) dG(\Delta) d\Gamma(\omega) - V_t([\Delta_i, q_i]) \right] \\ & \left. - c^X(X, n) v_t \right\}. \end{aligned}$$

Conjecture first that the value function is additive, i.e.

$$V_t([\Delta_i, q_i]) = \sum_{i=1}^{n} V_t(\Delta_i, q_i). \qquad \text{(SM-26)}$$

We then have:

1. Loss through creative destruction:

$$V_t\left([\Delta_j, q_j]_{j \neq i} \right) - V_t([\Delta_i, q_i]) = -V_t(\Delta_i, q_i)$$

2. Value of own-innovation:
$$\frac{\partial V_t\left([\Delta_i, q_i]\right)}{\partial q_i} q_i = \frac{dV\left(\Delta_i, q_i\right)}{dq_i} q_i$$

3. Value of product creation:
$$\int_q V_t\left([\Delta_i, q_i]_{i=1}^n, 1, \lambda q\right) dF_t\left(q\right) - V_t\left([\Delta_i, q_i]_{i=1}^n\right) = \int_q V_t\left(1, \lambda q\right) dF_t\left(q\right)$$

   and
$$\int_\omega \int_\Delta V_t\left([\Delta_i, q_i], \Delta, \omega Q_t\right) dG(\Delta) d\Gamma\left(\omega\right) - V_t\left([\Delta_i, q_i]_{i=1}^n\right) = \int_\omega \int_\Delta V_t\left(\Delta, \omega Q_t\right) dG(\Delta) d\Gamma\left(\omega\right).$$

   Hence,
$$\begin{aligned} & X\Big[\alpha \textstyle\int_q V_t\left([\Delta_i, q_i], 1, \lambda q\right) dF_t\left(q\right) \\ & + \left(1-\alpha\right) \textstyle\int_\omega \int_\Delta V_t\left([\Delta_i, q_i], \Delta, \omega Q_t\right) dG(\Delta) d\Gamma\left(\omega\right) - V_t\left([\Delta_i, q_i]\right)\Big] \\ = \quad & X\left[\alpha \textstyle\int_q V_t\left(1, \lambda q\right) dF_t\left(q\right) + \int_\omega \int_\Delta V_t\left(\Delta, \omega Q_t\right) dG(\Delta) d\Gamma\left(\omega\right)\right]. \end{aligned}$$

Hence we can write the value function as

$$\begin{aligned} r_t \sum_{i=1}^n V_t\left(\Delta_i, q_i\right) &= \sum_{i=1}^n \pi_t\left(\Delta_i, q_i\right) + \sum_{i=1}^n \dot{V}_t\left(\Delta_i, q_i\right) - \left(\tau + \delta\right) \sum_{i=1}^n V_t\left(\Delta_i, q_i\right) + \\ & \max_{\{I_t(i)\}_{i=1}^n} \sum_{i=1}^n \left\{ I_t\left(i\right) \frac{dV_t\left(\Delta_i, q_i\right)}{dq_i} q_i - c^I\left(I_t(i), \Delta_i, q_i\right) v_t \right\} + \\ & n \max_x \left\{ x\left[\alpha \int_q V_t\left(1, \lambda q\right) dF_t\left(q\right) + \left(1-\alpha\right) \int_\omega \int_\Delta V_t\left(\Delta, \omega Q_t\right) dG(\Delta) d\Gamma\left(\omega\right)\right] - c^X\left(x, 1\right) v_t \right\}. \end{aligned}$$

Hence, the value function is indeed additive and we can focus on a single product with state (Δ_i, q_i), which solves the HJB equation

$$\begin{aligned} \left(r_t + \tau + \delta\right) V_t\left(\Delta_i, q_i\right) &= \pi_t\left(\Delta_i, q_i\right) + \dot{V}_t\left(\Delta_i, q_i\right) + \max_I \left[I \frac{dV\left(\Delta_i, q_i\right)}{dq_i} q_i - c^I\left(I, \Delta_i, q_i\right) v_t\right] + \\ & \max_x \left\{ x\left[\alpha \int_q V_t\left(1, \lambda q\right) dF_t\left(q\right) + \left(1-\alpha\right) \int_\omega \int_\Delta V_t\left(d, \omega Q_t\right) dG(\Delta) d\Gamma\left(\omega\right)\right] - c^X\left(x, 1\right) v_t \right\}. \end{aligned}$$

Note that the firm can choose a separate rate I_i for each product i, subject to a paying a cost $c_t^I\left(I_i, \Delta_i, q_i\right)$

which depends on that rate. We work with a particular convenient cost function, such that

$$c_t^I(I, \Delta, q) = \varphi_I \frac{q^{\sigma-1}}{Q_t^{\sigma-1}} I^{\zeta},$$

so that cost is increasing in a scaled-version of relative productivity.[27]

The value function above can be written as

$$V_t(\Delta_i, q_i) = U_t(\Delta_i, q_i) + \phi_t, \tag{SM-27}$$

where $U_t(\Delta_i, q_i)$ captures the rents from the existing products and ϕ_t describes the expansion value. From above, these functions solve the (partial) differential equations

$$\begin{aligned}
(r_t + \tau + \delta) U_t(\Delta_i, q_i) &= -\dot{U}_t(\Delta_i, q_i) && \text{(SM-28)}\\
&+ \pi_t(\Delta_i, q_i) + \max_I \left[I \frac{dU(\Delta_i, q_i)}{dq_i} q_i - c^I(I, \Delta_i, q_i) v_t \right], && \text{(SM-29)}
\end{aligned}$$

and

$$\begin{aligned}
(r + \tau + \delta)\phi_t - \dot{\phi}_t = \max_x \Bigg\{ x \Bigg[\alpha \int_q V_t(1, \lambda q) \, dF_t(q) && \text{(SM-30)}\\
+ (1-\alpha) \int_\omega \int_\Delta V_t(\Delta, \omega Q_t) \, dG(\Delta) d\Gamma(\omega) \Bigg] && \text{(SM-31)}\\
- c^X(x, 1) v_t \Bigg\}. && \text{(SM-32)}
\end{aligned}$$

Solving for $U_t(\Delta_i, q_i)$. Conjecture that the value function of a particular product takes the following form

$$U_t(\Delta_i, q_i) = \frac{u(\Delta_i)}{g(\sigma-1) + \rho + \tau + \delta - \eta} \frac{q_i^{\sigma-1} Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t Q_t^{\sigma-1}}. \tag{SM-33}$$

We need to determine the function $u(.)$ Note that Y_t is growing at rate $g + \eta$, that N_t is growing at rate η, that $g = g_Q + \frac{1}{\sigma-1}\eta$ and that $Q_t^{\sigma-1}$ is growing at rate $g_Q(\sigma-1)$. Hence, (SM-33) implies that - along the BGP - we have

[27] It can be show that this cost function is consistent with balanced growth.

$$\frac{\dot{U}_t(\Delta_i, q_i)}{U_t(\Delta_i, q_i)} = \frac{\dot{Y}_t}{Y_t} - \frac{\dot{N}_t}{N_t} - g_Q(\sigma - 1) \tag{SM-34}$$
$$= g_Q + \frac{1}{\sigma - 1}\eta - g_Q(\sigma - 1)$$
$$= g(2 - \sigma) + \eta.$$

So for a fixed q, whether profits are shrinking or rising with g depends on whether $\sigma \lesseqgtr 2$. Note also that because of log preferences we have $r = g + \rho$. Combing this with (SM-28) yields

$$(g(\sigma - 1) + \rho + \tau + \delta - \eta)\, U_t(\Delta_i, q_i) = \pi_t(\Delta_i, q_i) + \max_I \left[I \frac{dU(\Delta_i, q_i)}{dq_i} q_i - \varphi_I \frac{q_i^{\sigma-1}}{Q_t^{\sigma-1}} I^{\zeta} v_t \right].$$

Now, the markup is given by

$$\Delta_i = \frac{q_i}{q_i^C}$$

for $\Delta_i \leq \frac{\sigma}{\sigma-1}$, so that

$$\frac{dU(\Delta_i, q_i)}{dq_i} q_i = \frac{(\sigma - 1)u(\Delta_i)}{g(\sigma - 1) + \rho + \tau + \delta - \eta} \frac{q_i^{\sigma-1} Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t Q_t^{\sigma-1}} + \frac{u'(\Delta_i)\frac{q_i}{q_i^C}}{g(\sigma - 1) + \rho + \tau + \delta - \eta} \frac{q_i^{\sigma-1} Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t Q_t^{\sigma-1}}$$
$$= \frac{((\sigma - 1)u(\Delta_i) + u'(\Delta_i)\Delta_i)}{g(\sigma - 1) + \rho + \tau + \delta - \eta} \frac{q_i^{\sigma-1} Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t Q_t^{\sigma-1}}$$

Optimality for I then requires

$$\frac{(\sigma - 1)u(\Delta_i) + u'(\Delta_i)\Delta_i}{g(\sigma - 1) + \rho + \tau + \delta - \eta} \frac{q_i^{\sigma-1} Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t Q_t^{\sigma-1}} = \zeta \varphi_I \frac{q_i^{\sigma-1}}{Q_t^{\sigma-1}} I^{\zeta-1} v_t$$

Solving for I yields

$$I(\Delta) = \left(\frac{(\sigma - 1)u(\Delta) + u'(\Delta)\Delta}{g(\sigma - 1) + \rho + \tau + \delta - \eta} \frac{Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t v_t} \frac{1}{\zeta \varphi_I} \right)^{\frac{1}{\zeta-1}} \tag{SM-35}$$
$$= \left(\frac{(\sigma - 1)u(\Delta) + u'(\Delta)\Delta}{g(\sigma - 1) + \rho + \tau + \delta - \eta} \frac{L_t^P / N_t}{\mathcal{M}_t{}^{\sigma-1} \Lambda_t^{\sigma} \varphi_I \zeta} \frac{w_t}{v_t} \right)^{\frac{1}{\zeta-1}} \tag{SM-36}$$

where the second equality uses $Y_t \Lambda_t = L_t^P w_t$. Hence,

$$\max_I \left\{ I \frac{(\sigma-1)u(\Delta) + u'(\Delta)\Delta}{g(\sigma-1)+\rho+\tau+\delta-\eta} \frac{q^{\sigma-1} Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t Q_t^{\sigma-1}} - \varphi_I \frac{q^{\sigma-1}}{Q_t^{\sigma-1}} I^{\zeta} v_t \right\} = (\zeta-1)\,\varphi_I I(\Delta)^{\zeta} v_t \frac{q^{\sigma-1}}{Q_t^{\sigma-1}}.$$

Along with the expression for profits this yields

$$\begin{aligned}(g(\sigma-1)+\rho+\tau+\delta-\eta)\, U_t(\Delta_i, q_i) = h(\Delta_i) \left(\frac{q_i}{Q_t}\right)^{\sigma-1} \frac{Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t} \\ + (\zeta-1)\varphi_I \frac{q_i^{\sigma-1}}{Q_t^{\sigma-1}} I(\Delta_i)^{\zeta} v_t,\end{aligned}$$

or

$$u(\Delta_i) \frac{q_i^{\sigma-1} Y_t}{(\mathcal{M}_t \Lambda_t)^{\sigma-1} N_t Q_t^{\sigma-1}} = \frac{h(\Delta_i)}{(\mathcal{M}_t \Lambda_t)^{\sigma-1}} \frac{q_i^{\sigma-1}}{Q_t^{\sigma-1}} \frac{Y_t}{N_t} + (\zeta-1)\,\varphi_I I(\Delta_i)^{\zeta} \frac{q_i^{\sigma-1}}{Q_t^{\sigma-1}} \frac{\Lambda_t Y_t}{N_t} \frac{N_t}{L_t^P} \frac{v_t}{w_t},$$

So if $u(\Delta)$ solves

$$u(\Delta) = h(\Delta) + (\zeta-1)\,\varphi_I I(\Delta)^{\zeta} \frac{N_t}{L_t^P} (\mathcal{M}_t)^{\sigma-1} \Lambda_t^{\sigma} \frac{v_t}{w_t} \tag{SM-37}$$

then our guess is verified .Substituting the optimal solution for I in (SM-35) yields

$$u(\Delta) = h(\Delta) + \frac{(\zeta-1)\,\varphi_I}{((g(\sigma-1)+\rho+\tau+\delta-\eta)\,\varphi_I \zeta)^{\frac{\zeta}{\zeta-1}}} \left((\sigma-1)u(\Delta) + u'(\Delta)\Delta\right)^{\frac{\zeta}{\zeta-1}} \left(\frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^{\sigma}} \frac{w_t}{v_t}\right)^{\frac{1}{\zeta-1}}. \tag{SM-38}$$

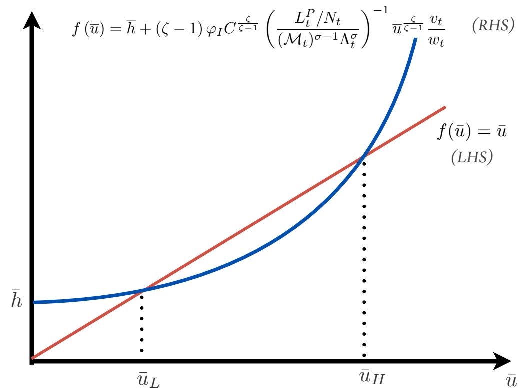
This is a non-linear differential equation in $u(\Delta)$ given parameters and the general equilibrium statistic $\frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^{\sigma}} \frac{w_t}{v_t}$, which is constant on the BGP. Given $\{u(\Delta)\}_{\Delta}$, firms' optimal innovation rate is given by

$$I(\Delta) = \left(\frac{u(\Delta) - h(\Delta)}{(\zeta-1)\,\varphi_I} \frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^{\sigma}} \frac{w_t}{v_t}\right)^{1/\zeta}. \tag{SM-39}$$

Note crucially that there is no dependence on a product's efficiency; innovation only depends on the gap between the efficiency of the firm and the next best producer.

To solve this differential equation, we require a boundary condition. The boundary condition comes

Figure SM-4: Two Potential Solutions to Equation (SM-41)



Notes: The figure displays the left hand side (LHS) and right hand side (RHS) of equation (SM-40).

from considering a firm with efficiency gap $\Delta \geq \overline{\Delta} = \frac{\sigma}{\sigma-1}$. This firm will set a markup of $\frac{\sigma}{\sigma-1}$. Hence,

$$h(\Delta) = \left(\frac{1}{\sigma}\right)\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} = \overline{h} \text{ for } \Delta \geq \overline{\Delta}.$$

Moreover, conditional on survival, it will never set a different markup. Because we restrict our analysis to value functions which only depend on pay-off relevant state variables, we know that $u(\Delta) = \overline{u}$ for $\Delta \geq \overline{\Delta}$, and $u'(\Delta) = 0$. Hence, (SM-35) implies that (for $\Delta \geq \overline{\Delta}$)

$$I = (C\overline{u})^{\frac{1}{\zeta-1}}$$

and

$$\overline{u} = \overline{h} + (\zeta-1)\varphi_I C^{\frac{\zeta}{\zeta-1}} \left(\frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^\sigma}\right)^{-1} \overline{u}^{\frac{\zeta}{\zeta-1}} \frac{v_t}{w_t}, \tag{SM-40}$$

where

$$C \equiv \frac{(\sigma-1)}{(g(\sigma-1)+\rho+\tau+\delta-\eta)\,\varphi_I\zeta} \frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^\sigma} \frac{w_t}{v_t}.$$

To see that there is a unique solution $I(\Delta)$, note that (generically) (SM-40) has two solutions in the positive orthant- $\overline{u}_L$ and $\overline{u}_H$. These are depicted in Figure (SM-4) below. We will now show that $\overline{u}_H > \overline{u}_L$ is in fact not a solution for the value function. To see this define the function

$$f(\overline{u}) = \overline{h} + (\zeta-1)\varphi_I C^{\frac{\zeta}{\zeta-1}} \left(\frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^\sigma}\right)^{-1} \overline{u}^{\frac{\zeta}{\zeta-1}} \frac{v_t}{w_t}. \tag{SM-41}$$

Note that

$$f'(\overline{u}_H) = \zeta\varphi_I C^{\frac{\zeta}{\zeta-1}} \left(\frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^\sigma}\right)^{-1} \overline{u}_H^{\frac{1}{\zeta-1}} \frac{v_t}{w_t}.$$

Also note that $f'(\overline{u}_H) > 1$ as the curve intersects from below at $\overline{u}_H$. This implies that

$$\zeta\varphi_I C^{\frac{\zeta}{\zeta-1}} \left(\frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^\sigma}\right)^{-1} \overline{u}_H^{\frac{1}{\zeta-1}} \frac{v_t}{w_t} > 1.$$

Using that $I = (C\overline{u}_H)^{\frac{1}{\zeta-1}}$, this inequality implies that

$$\begin{aligned} 1 &< \zeta\varphi_I C \left(\frac{L_t^P/N_t}{(\mathcal{M}_t)^{\sigma-1}\Lambda_t^\sigma}\right)^{-1} (C\overline{u}_H)^{\frac{1}{\zeta-1}} \frac{v_t}{w_t} \\ &= \frac{\sigma-1}{(g(\sigma-1)+\rho+\tau+\delta-\eta)} I. \end{aligned} \tag{SM-42}$$

But now consider a firm with efficiency q and $\Delta \geq \overline{\Delta}$. Using the solution for the value function in (SM-40)

$$\bar{u}_H = h(\bar{\Delta}) + I\left(\frac{(\sigma-1)\overline{u}_H}{g(\sigma-1)+\rho+\tau+\delta-\eta}\right) - \varphi_I I^\zeta v_t \frac{N_t(\mathcal{M}_t\Lambda_t)^{\sigma-1}}{Y_t}$$

Rearranging terms and using the expression for profits (A-3)

$$\overline{u}_H = \frac{\left(\pi_i(\bar{\Delta}) - \varphi_I I^\zeta \left(\frac{q_i}{Q_t}\right)^{\sigma-1} v_t\right) \frac{N_t(\mathcal{M}_t\Lambda_t)^{\sigma-1}}{Y_t\left(\frac{q_i}{Q_t}\right)^{\sigma-1}}}{1 - \frac{I(\sigma-1)}{g(\sigma-1)+\rho+\tau+\delta-\eta}}.$$

Note that $\pi_i(\bar{\Delta}) - \varphi_I I^\zeta \left(\frac{q_i}{Q_t}\right)^{\sigma-1} v_t$ is simply the per period cash flow of the firm. Hence, for $\overline{u}_H$ to be positive, it has to be the case that

$$1 > \frac{\sigma-1}{g(\sigma-1)+\rho+\tau-\eta} I.$$

This contradicts (SM-42).

SM-8.0.1 Solving for ϕ_t from (SM-30)

To solve for ϕ_t in (SM-30), by free entry it must be that

$$\alpha \int_q V_t(\lambda, \lambda q)\, dF_t(q) + (1-\alpha) \int_\omega \int_\Delta V_t(\Delta, \omega Q_t)\, dG(\Delta) d\Gamma(\omega) = \frac{v_t}{\varphi_E}.$$

Substituting this in (SM-30) yields

$$(r+\tau+\delta)\,\phi_t - \dot{\phi}_t = \max_x \left\{ x\frac{1}{\varphi_e} - c^X(x,1) \right\} v_t.$$

Along the BGP, research wages grow at rate g. Hence, $\dot{\phi}_t = g\phi_t$ so that

$$\phi_t = \frac{\max_x \left\{ x\frac{1}{\varphi_E} - c^X(x,1) \right\}}{r+\tau+\delta-g} v_t = \frac{\max_x \left\{ x\frac{1}{\varphi_e} - c^X(x,1) \right\}}{\rho+\tau+\delta} v_t.$$

For the case where $c^X(x,1) = \frac{1}{\varphi_x} x^\zeta$ we get that

$$\max_x \left\{ x\frac{1}{\varphi_E} - c^X(x,1) \right\} = \frac{\zeta-1}{\varphi_x} \left(\frac{1}{\zeta}\frac{\varphi_x}{\varphi_E} \right)^{\frac{\zeta}{\zeta-1}}$$

and

$$\left(\frac{1}{\zeta}\frac{\varphi_x}{\varphi_e} \right)^{\frac{1}{\zeta-1}} = x^*$$

so that

$$\phi_t = \frac{\frac{\zeta-1}{\varphi_x}\left(\frac{1}{\zeta}\frac{\varphi_x}{\varphi_E} \right)^{\frac{\zeta}{\zeta-1}}}{\rho+\tau+\delta} v_t. \tag{SM-43}$$

SM-8.1 The Joint Distribution of Gaps and Productivity

As in the baseline model we can compute the joint distribution of scaled efficiency and efficiency gaps. Again we define the objects $\bar{F}_t^C(\Delta, \hat{q})$ and $F_t^{NC}(\hat{q})$ as the mass of products with a gap less than Δ and relative productivity less than $\hat{q}$, for products with (without) a direct competitor. Again, we define $\hat{q}_t \equiv \ln(q_t/Q_t)^{\sigma-1}$. Hence, for a product of efficiency and gap $(\Delta, \hat{q}_t)$ we have

$$\hat{q}_{t+\iota} = (\sigma-1)((I(\Delta) - g_Q)\iota + \hat{q}_t$$

where $g_Q = \frac{dlog(Q_t)}{dt}$.

Now, the evolution of the non competitor mass satisfies

$$\bar{F}_t^{NC}(\hat{q}) = \underbrace{\bar{F}_{t-\iota}^{NC}\left(\hat{q} - (\sigma-1)((I(\bar{\Delta}) - g_Q)\iota\right)(1 - (\tau_t + \delta)\iota)}_{\text{existing mass that survives and improves/falls}} + \underbrace{(1-\alpha)(z_{t-\iota} + x_{t-\iota})N_t^{NC}\iota\Gamma\left(\frac{\exp(\hat{q})}{\sigma-1}\right)}_{\text{new products}}$$

and then using the fact that $z + x = \tau/\alpha$ we have

$$\bar{F}_t^{NC}(\hat{q}) = \bar{F}_{t-\iota}^{NC}\left(\hat{q} - (\sigma-1)((I(\bar{\Delta}) - g_Q)\iota\right)(1 - (\tau_t + \delta)\iota) + \frac{(1-\alpha)}{\alpha}\tau\iota N_t^{NC}\Gamma\left(\frac{\exp(\hat{q})}{\sigma-1}\right)$$

Write this as

$$\frac{\bar{F}_t^{NC}(\hat{q}) - \bar{F}_{t-\iota}(\hat{q})}{\iota} + \frac{\bar{F}_{t-\iota}(\hat{q}) - \bar{F}_{t-\iota}^{NC}\left(\hat{q} - (\sigma-1)((I(\Delta) - g_Q)\iota\right)}{(\sigma-1)((I(\Delta) - g_Q)\iota}(\sigma-1)((I(\Delta) - g_Q) = -(\tau_t + \delta)\bar{F}_{t-\iota}^{NC}\left(\hat{q} - (\sigma-1)((I(\Delta) - g_Q)\iota\right) + + \frac{(1-\alpha)}{\alpha}\tau N_t^{NC}\Gamma\left(\frac{\exp(\hat{q})}{\sigma-1}\right)$$

and taking the limit

$$\frac{\partial \bar{F}_t^{NC}(\hat{q})}{\partial t} = -\frac{\partial \bar{F}_t^{NC}(\hat{q})}{\partial \hat{q}}(\sigma-1)((I(\bar{\Delta}) - g_Q) - (\tau_t + \delta)\bar{F}_t^{NC} + \frac{(1-\alpha)}{\alpha}\tau N_t^{NC}\Gamma\left(\frac{\exp(\hat{q})}{\sigma-1}\right) \quad \text{(SM-44)}$$

Then define the distribution $F_t^{NC} \equiv \bar{F}_t^{NC}/N_t^{NC}$ so that

$$\frac{\partial \bar{F}_t^{NC}(\hat{q})}{\partial t} = \dot{N}_t^{NC}F_t^{NC}(\hat{q}) + N_t\frac{\partial F_t^{NC}(\hat{q})}{\partial t}$$

So rewrite SM-44 as

$$\frac{\partial F_t^{NC}(\hat{q})}{\partial t} = -\frac{\partial F_t^{NC}(\hat{q})}{\partial \hat{q}}(\sigma-1)((I(\bar{\Delta}) - g_Q) - (\tau_t + \delta + \eta)F_t^{NC} + \frac{(1-\alpha)}{\alpha}\tau\Gamma\left(\frac{\exp(\hat{q})}{\sigma-1}\right)$$

Similarly, the competitor mass evolves as

$$\begin{aligned}
\bar{F}_t^C(\Delta,\hat{q}) = & \underbrace{\bar{F}_{t-\iota}^C\left(\Delta e^{-I(\Delta)\iota}, \hat{q} - (\sigma-1)((I(\Delta) - g_Q)\iota\right)(1-(\tau_t+\delta)\iota)}_{\text{existing mass that survives and improves/falls}} \\
+ & \underbrace{\lim_{s\to\infty} \iota\tau_t\bar{F}_t^C(s, \hat{q} - (\sigma-1)log(\lambda))}_{\text{CD returns gap to } \lambda\text{,increases productivity by } \lambda} \\
+ & \underbrace{\tau\iota\bar{F}_t^{NC}(\hat{q} - (\sigma-1)log(\lambda))}_{\text{NC products get improved by CD}}
\end{aligned}$$

Rewriting and taking the limit as $\iota \to 0$

$$\begin{aligned}
\frac{\partial \bar{F}^C(\Delta,\hat{q})}{\partial t} = & -\left(\frac{\partial \bar{F}^C(\Delta,\hat{q})}{\partial \Delta}I(\Delta)\Delta + (\sigma-1)((I(\Delta) - g_Q)\frac{\partial \bar{F}^C(\Delta,\hat{q})}{\partial \hat{q}}\right) \\
& - \bar{F}_t^C\left(\Delta,\hat{q}\right)(\tau_t+\delta) \\
& + \lim_{s\to\infty} \tau_t\bar{F}_t^C(s, \hat{q} - (\sigma-1)log(\lambda)) \\
& + \tau\bar{F}_t^{NC}(\hat{q} - (\sigma-1)log(\lambda))
\end{aligned}$$

Now$\bar{F}_t^C(\Delta,\hat{q}) = N_tF_t^C(\Delta,\hat{q})$, so that

$$\frac{\partial \bar{F}_t^C(\Delta,\hat{q})}{\partial t} = \dot{N}_tF_t(\Delta,\hat{q}) + N_t\frac{\partial F_t(\Delta,\hat{q})}{\partial t}$$

So dividing through by N_t

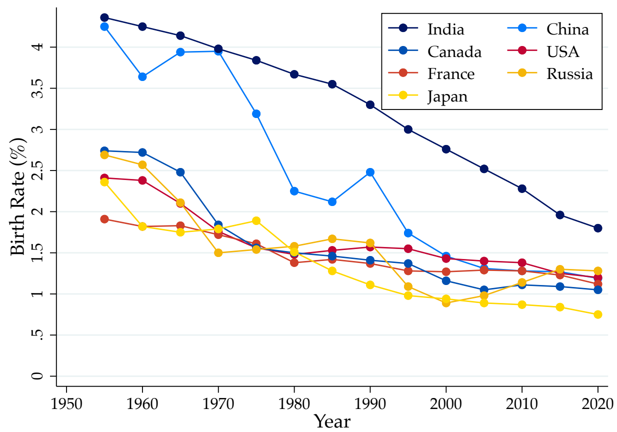
$$\frac{\partial F_t^C(\Delta,\hat{q})}{\partial t} = -\left(\frac{\partial F_t^C(\Delta,\hat{q})}{\partial \Delta} I(\Delta)\Delta + (\sigma-1)((I(\Delta)-g_Q)\frac{\partial F_t^C(\Delta,\hat{q})}{\partial \hat{q}}\right)$$
$$- F_t^C\left(\Delta,\hat{q}\right)(\tau_t+\delta+\eta_t) + \lim_{s\to\infty} \tau_t F_t^C(s,\hat{q}-(\sigma-1)log(\lambda))$$
$$+ \tau F_t^{NC}(\hat{q}-(\sigma-1)log(\lambda)).$$

SM-9 Fertility and Population Growth

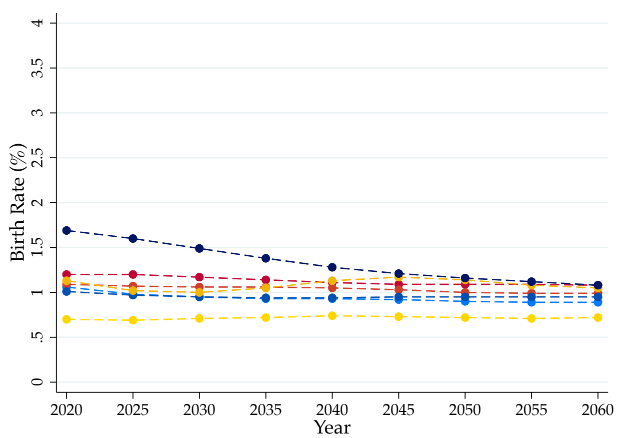
In this section we show the evolution of two major determinants of population growth in the major economies displayed in Figure 1. Figure SM-5 plots historical and projected birthrates from UN data. In all cases, these birth rates have fallen and are projected to remain low. Figure SM-6 shows the net migration rate for these economies, which is almost uniformly low and projected to remain as such.

Figure SM-5: Birth Rate Across Major Economies

(a) Historical Birth Rate



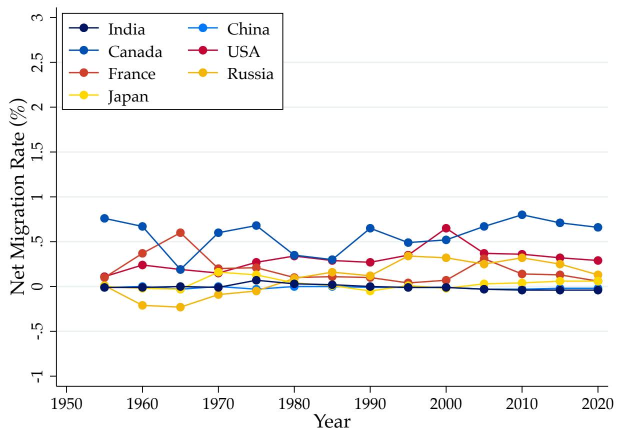
(b) Projected Birth Rate



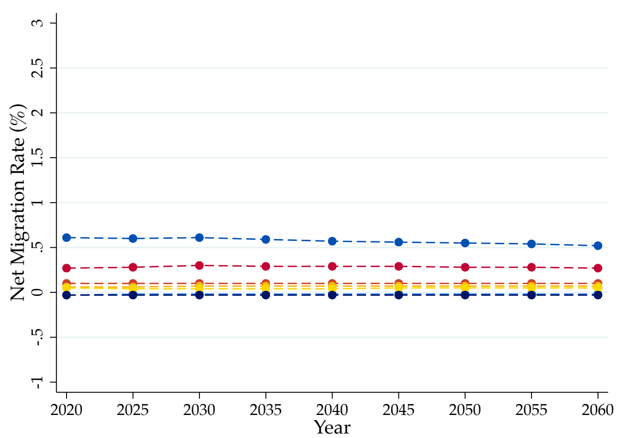
Notes: Panel (a) of this figure plots the historical "crude" birth rate from the UN World Population Prospects 2019 for several major economies. The crude birth rate is total births over total population in a given year. Panel (b) plots the UN projections for population growth in the "Medium" scenario out to to 2060.

Figure SM-6: Net Migration Across Major Economies

(a) Historical Net Migration



(b) Projected Net Migration



Notes: Panel (a) of this figure plots historical net migration growth from the UN World Population Prospects 2019 for several major economies. Panel (b) plots the UN projections for population growth in the "Medium" scenario out to to 2060.