Appendix for *"Population Growth and Firm Dynamics"*

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

$$Y_t = N_t^{\frac{1}{\sigma-1}} Q_t \mathcal{M}_t L_t^P \qquad \text{where} \qquad \mathcal{M}_t = \frac{\left(\int \mu^{1-\sigma}(q/Q_t)^{\sigma-1} dF_t(q,\mu)\right)^{\frac{\sigma}{\sigma-1}}}{\int \mu^{-\sigma}(q/Q_t)^{\sigma-1} dF_t(q,\mu)}. \tag{A-2}$$

Similarly,

$$w_t L_P = \Lambda_t Y_t \qquad \text{where} \qquad \Lambda_t = \frac{\int \mu^{-\sigma}(q/Q_t)^{\sigma-1} dF_t(q,\mu)}{\int \mu^{1-\sigma}(q/Q_t)^{\sigma-1} dF_t(q,\mu)}. \tag{A-3}$$

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

$$py_i = \mu_i^{1-\sigma}\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\left(\frac{1}{\mathcal{M}_t\Lambda_t}\right)^{\sigma-1}\frac{Y_t}{N_t} \tag{A-4}$$

$$\pi_i = \left(1-\frac{1}{\mu_i}\right)\times\mu_i^{1-\sigma}\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\left(\frac{1}{\mathcal{M}_t\Lambda_t}\right)^{\sigma-1}\frac{Y_t}{N_t}. \tag{A-5}$$

If markups are constant, (A-4) reduces to

$$py_i = \left(\frac{q_i}{Q_t}\right)^{\sigma-1} \frac{Y_t}{N_t} \quad \text{and} \quad \pi_i = \left(\frac{\mu-1}{\mu}\right)\left(\frac{q_i}{Q_t}\right)^{\sigma-1} \frac{Y_t}{N_t}.$$

A-1.1.2 Aggregate Growth Rate

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

$$g_Q = \frac{\dot{Q}_t}{Q_t} = \left(\frac{\lambda^{\sigma-1}-1}{\sigma-1}\right)\tau_t + \frac{(\overline{\omega}^{\sigma-1}-1)}{\sigma-1}\nu_t + I. \tag{A-6}$$

The growth rate of labor productivity is given by

$$g_t^{LP} = \frac{d}{dt}\ln\left(Q_t N_t^{\frac{1}{\sigma-1}}\right) \quad = \quad g_t^Q + \frac{1}{\sigma-1}g_t^N = I + \left(\frac{\lambda^{\sigma-1}-1}{\sigma-1}\right)\tau_t + \frac{\overline{\omega}^{\sigma-1}}{\sigma-1}\nu_t - \frac{1}{\sigma-1}\delta. \tag{A-7}$$

A-1.1.3 Derivation of value function $V_t(q)$ in the model with entry (Equation (8))

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

$$U_t = \frac{(\mu-1)\left(\frac{1}{Q_t}\right)^{\sigma-1}\frac{L_t^P}{N_t}w_t}{r_t+\tau_t+\delta-g_U}. \tag{A-8}$$

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-8) yields (8).

A-1.1.4 Proof of Proposition 1

Equations (12) and (13) directly from $\nu = (1-\alpha)z - \delta$ in (11). To derive (14), consider the free entry condition in (A-9). Using the Euler equation $g_{Y/L} = r_t - \rho$ and $g_w = g_{Y/L} - g_{\ell^P}$ (see (1)). Furthermore, note that

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

The free entry condition in (A-9) thus implies that

$$\frac{1}{\varphi_E} = \frac{(\mu-1)\overline{q}^{\sigma-1}}{\rho + \frac{\overline{q}^{\sigma-1}}{1-\alpha}\delta + \left(\frac{\overline{q}^{\sigma-1}}{1-\alpha}-1\right)(g_{\mathcal{N}}+\eta)+g_{\ell^P}}\frac{\ell_t^P}{\mathcal{N}_t}. \tag{A-9}$$

Similarly, the resource constraint in (9) can be written as

$$\left(\frac{1-\ell_t^P}{\mathcal{N}_t}\right)=\frac{1}{\varphi_E}\frac{v_t}{1-\alpha}=\frac{1}{\varphi_E}\frac{g_{\mathcal{N}}+\delta}{1-\alpha}=\frac{1}{\varphi_E}\frac{g_{\mathcal{N}}+\eta+\delta}{1-\alpha}. \tag{A-10}$$

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, the determine the path $\{\mathcal{N}_t,\ell_t^P\}_t$.

Along the BGP, $g_{\mathcal{N}}=g_{\ell^P}=0$. Equations (A-9) and (A-10) can then be solved for $\mathcal{N}$ and ℓ^P given in (14). In addition, (A-9) and (A-10) also characterize the transitional dynamics depicted in Figure 2. Rearranging terms in (A-10) and substituting for $g_{\mathcal{N}}$ in (A-9) 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 $\Xi_t([q_i])$ as

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

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

$$r_tV_t(q)-\dot{V}_t(q)=\pi_t(q)+I\frac{\partial V_t(q)}{\partial q}q-(\tau+\delta)V_t(q)+\Xi_t, \tag{A-11}$$

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-11) can then be written as

$$(r_t+\tau+\delta-g_U)q^{\sigma-1}U_t+(r+\tau+\delta-g_M)M_t=\left((\mu-1)\left(\frac{1}{Q_t}\right)^{\sigma-1}\frac{L_t^P}{N_t}w_t+I(\sigma-1)U_t\right)q^{\sigma-1}+\Xi_t. \tag{A-12}$$

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

$$\Xi_t = \max_x \left\{ x \left(\alpha V_t^{CD} + (1-\alpha) V_t^{NV} \right) - \frac{1}{\varphi_x} x^{\zeta} w_t \right\} = \frac{\zeta - 1}{\varphi_x} x^{\zeta} w_t,$$

where

$$x = \left(\frac{\varphi_x}{\zeta} \right)^{\frac{1}{\zeta-1}} \left(\alpha \frac{V_t^{CD}}{w_t} + (1-\alpha) \frac{V_t^{NV}}{w_t} \right)^{\frac{1}{\zeta-1}}. \tag{A-13}$$

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

$$V_t^{CD} = \int V_t(\lambda q) \, dF_t(q) = V_t(\lambda Q_t) \text{ and } V_t^{NV} = \int V_t(\omega Q_t) \, d\Gamma(\omega) = V_t(\overline{\omega} Q_t).$$

This concludes the proof of Proposition 2.

A-1.1.6 Characterization of Equilibrium

In this section we characterize the full equilibrium of our economy. We maintain the assumption that the free entry condition is binding along the equilibrium path. The equilibrium is characterized by the following conditions:

1. The evolution of aggregate productivity is given by, see (A-6), $g_Q = \frac{\dot{Q}_t}{Q_t} = \frac{\lambda^{\sigma-1}-1}{\sigma-1}\tau_t + \frac{\overline{\omega}^{\sigma-1}-1}{\sigma-1}\nu_t + I$.

2. The rate of creative destruction is linked to the growth rate of N_t according to $\tau = \frac{\alpha}{1-\alpha}\nu_t$, where $\nu_t = (1-\alpha)(z_t + x)$. Note that x is constant because of the binding free entry condition.

3. Labor market clearing requires $L_t = L_{Pt} + L_{Rt}$, where $L_{Rt} = N_t \frac{1}{\varphi_E}\left(z_t + \frac{1}{\zeta}x\right)$. Hence, $\frac{L_t}{N_t} = \frac{L_{Pt}}{N_t} + \frac{1}{\varphi_E}\left(z_t + \frac{1}{\zeta}x\right)$.

4. 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 $w_t = \frac{1}{\mu} Y_t / L_t^P$. Note that market clearing requires $C_t = Y_t$. Hence, the growth rate of per capita consumption is given by

$$g_c = g_Y - \eta = g_w + g_{L^P} - \eta, \tag{A-14}$$

   where $g_w = \frac{1}{\sigma-1} g_N + g_Q$ (see (A-7)). The real interest rate is thus given by $r = \rho + g_w + g_{L^P} - \eta$.

5. To derive the implications for the free entry condition, note that (A-12) implies that

$$\frac{1}{\varphi_E} = \overline{q}^{\sigma-1} u_t + m_t$$

where

$$m_t = \frac{\frac{\zeta-1}{\varphi_x}x^{\zeta}}{\rho + g_{L^P} - \eta + \tau + \delta - g_m}$$

$$u_t = \frac{(\mu-1)\frac{L_t^P}{N_t}}{\rho + g_{L^P} - \eta + \tau + \delta - g_u + (\sigma-1)(g_Q - I)}.$$

Now define $\mathcal{N}_t = \frac{N_t}{L_t}$ and $\ell_t^P \equiv \frac{L_t^P}{L_t}$. Also note that $\tau = \alpha(z+x) = \frac{\alpha}{1-\alpha}\nu_t$. Then we can write the free entry condition as

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

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

$$\frac{1}{\varphi_E} = \frac{\overline{q}^{\sigma-1}(\mu-1)}{\rho + g_\ell + \delta - g_u + \left(\frac{\overline{q}^{\sigma-1}}{1-\alpha} - 1\right)\nu_t}\frac{\ell_t^P}{\mathcal{N}_t} + \frac{\frac{\zeta-1}{\varphi_x}x^{\zeta}}{\rho + g_\ell + \frac{\alpha}{1-\alpha}\nu_t + \delta - g_m} \quad \text{(A-15)}$$

$$\frac{1-\ell_t^P}{\mathcal{N}_t} = \frac{1}{\varphi_E}\left(\frac{\nu_t}{1-\alpha} - \frac{\zeta-1}{\zeta}x\right), \quad \text{(A-16)}$$

where g_u and g_m are the growth rates of u_t and m_t given in

$$m_t = \frac{\frac{\zeta-1}{\varphi_x}x^{\zeta}}{\rho + g_\ell + \frac{\alpha}{1-\alpha}\nu_t + \delta - g_m} \quad \text{and} \quad u_t = \frac{(\mu-1)\,\ell_t^P/\mathcal{N}_t}{\rho + g_\ell + \delta - g_u + \left(\frac{\overline{q}^{\sigma-1}}{1-\alpha} - 1\right)\nu_t}.$$

For a given initial condition $\mathcal{N}_0$ and the terminal condition that $\ell_t^P \to \overline{\ell}^P$ and $m_t \to m$ and $u_t \to u$ one can solve for the dynamic path $\left\{\ell_t^P, \mathcal{N}_t\right\}_t$.

A-1.1.7 Balanced Growth Path

Along a BGP, income per capita grows at a constant rate. (A-14) implies that

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

Along the BGP it also has to be the case that $\ell^P = L_t^P/L_t$ is constant. Hence, g^N is constant along a BGP. (A-16) therefore implies that $\mathcal{N}_t$ has to be constant, i.e. $g_N = \nu - \delta = \eta$. Hence, along the BGP the mass of products N_t grows at the same rate as the population. 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

$$\frac{1}{\varphi_E} = \frac{\overline{q}^{\sigma-1}(\mu-1)}{\rho + \left(\frac{\overline{q}^{\sigma-1}}{1-\alpha} - 1\right)\eta + \frac{\overline{q}^{\sigma-1}}{1-\alpha}\delta}\frac{\ell_t^P}{\mathcal{N}_t} + \frac{\frac{\zeta-1}{\varphi_x}x^{\zeta}}{\rho + \frac{\alpha\eta+\delta}{1-\alpha}} \quad \text{(A-17)}$$

$$\frac{1-\ell^P}{\mathcal{N}} = \frac{1}{\varphi_E}\left(\frac{\eta+\delta}{1-\alpha} - \frac{\zeta-1}{\zeta}x\right). \quad \text{(A-18)}$$

To characterize the solution, note that the free entry condition (A-17) defines a relationship $\mathcal{N}^{FE}\left(\ell^P\right)$ which is increasing and satisfies $\lim_{\ell^P\to 0}\mathcal{N}^{FE}\left(\ell^P\right)=0$ and $\mathcal{N}^{FE}(1)=\left(\frac{1}{\varphi_E}-\frac{\frac{\zeta-1}{\varphi_x}x^{\zeta}}{\rho+\frac{\alpha\eta+\delta}{1-\alpha}}\right)^{-1}\frac{(\mu-1)\overline{q}^{\sigma-1}}{\rho+\left(\frac{\overline{q}^{\sigma-1}}{1-\alpha}-1\right)\eta+\frac{\overline{q}^{\sigma-1}}{1-\alpha}\delta}$. Similarly, the resource constraint (A-18) defines a schedule $\mathcal{N}^{RC}\left(\ell^P\right)$, which is decreasing and satisfies $\mathcal{N}^{RC}(0)=\frac{\varphi_E}{\frac{\eta+\delta}{1-\alpha}-\frac{\zeta-1}{\zeta}x}$ and $\lim_{\ell^P\to 1}\mathcal{N}^{FE}\left(\ell^P\right)=0$. Hence, 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 different moments of the process of firm dynamics. In particular, we derive (i) the survival function $S(a)$ in (24), (ii) the average number of products by age $\overline{n}(a)$ in (24), and (iii) the pareto tail of the product distribution ζ_n in (26).

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

$$\dot{p}_n(a)=(n-1)\,x p_{n-1}(a)+(n+1)(\tau+\delta)\,p_{n+1}(a)-n(x+\tau+\delta)\,p_n(a).$$

Because exit is an absorbing state, $\dot{p}_0(a)=(\tau+\delta)\,p_1(a)$. The solution to this set of differential equations is (see Klette and Kortum (2004))

$$p_0(a)=\frac{\tau+\delta}{x}\gamma(a) \qquad \text{(A-19)}$$

$$p_1(a)=(1-p_0(a))(1-\gamma(a))$$

$$p_n(a)=p_{n-1}(a)\,\gamma(a) \qquad \text{(A-20)}$$

where

$$\gamma(a)=\frac{x\left(1-e^{-(\tau+\delta-x)a}\right)}{\tau+\delta-x\times e^{-(\tau+\delta-x)a}}. \qquad \text{(A-21)}$$

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

$$\psi \equiv x-\tau-\delta=x-\frac{\alpha}{1-\alpha}(\eta+\delta)-\delta=x-\frac{\alpha\eta+\delta}{1-\alpha}. \qquad \text{(A-22)}$$

Hence, ψ is decreasing in η. Also note that $\psi=\eta-z$. To make the firm-size distribution stationary, we need that $\eta>x-\tau-\delta$. Using equation (A-22), this implies that $z>0$, i.e. stationary requires the entry flow to be positive. From this solution for $p_n(a)$ we can calculate both the survival rate and the cross-sectional age distribution.

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

$$S(a)=1-p_0(a)=\frac{\psi e^{\psi a}}{\psi-x\left(1-e^{\psi a}\right)}, \qquad \text{(A-23)}$$

which is equation (24) in the main text. The average age is given by (again see Klette and Kortum

(2004))

$$\mathbb{E}[\text{Age}] = \int_0^\infty (1 - p_0(a))\, da = \frac{\ln\left(\frac{\tau+\delta}{\tau+\delta-x}\right)}{x} = \frac{\ln\left(\frac{\tau+\delta}{\tau+\delta-x}\right)}{x}$$

The expected number of products by age $\overline{n}(a)$. 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 $p_n(a)$ in (A-19)-(A-20), this implies that $\overline{p}_n(a) = \gamma(a)^{n-1}(1-\gamma(a))$. Then,

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

Using (A-21), this implies $\overline{n}(a) = 1 - \frac{x}{\psi}\left(1 - e^{\psi a}\right)$, which is the expression in (24).

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. 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

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

Along the BGP the mass of firms grows at rate η. Intuitively: the distribution of firms across products is stationary and the number of products N_t is increasing at rate η. Hence, the mass of firms is increasing at rate η. Hence, along the BGP we have$\dot{\omega}_t(n) = \eta\omega_t(n)$. Denote $\chi(n) = \frac{\omega_t(n)}{N_t}$ and $z = \frac{Z_t}{N_t}$. Along the BGP, $\{\chi(n)\}_{n=1}^{\infty}$ is determined by

$$\chi(2) = \frac{\chi(1)(\tau+x+\delta+\eta) - z}{2(\tau+\delta)} \tag{A-24}$$

and

$$\chi(n+1) = \frac{\chi(n)\, n(\tau+x+\delta) + \chi(n)\eta - \chi(n-1)(n-1)x}{(n+1)(\tau+\delta)} \quad \text{for } n \geq 2 \tag{A-25}$$

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

$$\sum_{n=1}^{\infty}\chi(n)\, n = \sum_{n=1}^{\infty}\frac{\omega_t(n)}{N_t} n = \frac{\sum_{n=1}^{\infty}\omega_t(n)\, n}{N_t} = 1. \tag{A-26}$$

Hence, equations (A-24), (A-25) and (A-26) fully determine the firm-size distribution $[\chi(n)]_{n\geq 1}$. In particular, the average number of products per firm are given by $\overline{n} = \frac{1}{\sum_{n=1}^{\infty}\chi(n)}$.

The distribution described by (A-24), (A-25) and (A-26) 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[34]

[34] To map the formulation of Luttmer (2011) to our model, note that he expresses the law of motion for the number of products as $DM_1 = \lambda 2M_2 + \nu N - (\mu+\lambda)M_1$ and $DM_n = \mu(n-1)M_{n-1} + \lambda(n+1)M_{n+1} - (\mu+\lambda)nM_n$. This is the

$\varrho_n = \frac{\eta}{x-\tau-\delta}$. 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.[35]

The pareto tail of the efficiency distribution ϱ_q. In this section we derive the marginal distribution of efficiency q. In particular we derive (27), which we use to calibrate $\overline{\omega}$. Define $\hat{q}_t$ as the relative productivity of a product

$$\hat{q}_t \equiv \ln\left(q_t/Q_t\right)^{\sigma-1}. \tag{A-27}$$

The drift of $\hat{q}_t$ (conditional on survival) is given by

$$\frac{\partial \hat{q}_t}{\partial t} = (\sigma-1)\,I - (\sigma-1)\,d\ln Q_t = -\left(\frac{\alpha\left(\lambda^{\sigma-1}-1\right)}{1-\alpha} + \overline{\omega}^{\sigma-1} - 1\right)(\eta+\delta), \tag{A-28}$$

where the second equality uses (13).

Let $F_t(\hat{q})$ denote the share of products at time t with $\hat{q}_i \le \hat{q}$. 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

$$\frac{dF(\hat{q})}{dq}\frac{\partial \hat{q}_t}{\partial t} = \tau\left(F_t(\hat{q}-\hat{\lambda}) - F_t(\hat{q})\right) - (\delta+\eta)\left(F_t(\hat{q}) - \Gamma\left(\exp\left(\frac{\hat{q}}{\sigma-1}\right)\right)\right). \tag{A-29}$$

Guess that F is exponential in the tail with index ϱ_q, that is $\underset{\hat{q}\to\infty}{lim}\, e^{\varrho_q\hat{q}}(1-F(\hat{q})) = a$ for some a and ϱ_q. If we assume that Γ has a thin tail[36] then as $\hat{q}\to\infty$, (A-29) 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-28) and noting that $\tau = \frac{\alpha}{1-\alpha}(\eta+\delta)$ yields

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

---

same law of motion as ours once we chose $\nu = z$, $\mu = x$ and $\lambda = \tau+\delta$. He shows that the pareto tail is given by $\frac{\eta}{\mu-\lambda}$ or (using our notation) $\frac{\eta}{x-\tau-\delta}$.

[35]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 $x-\tau-\delta = \frac{1}{1-\alpha}(x(1-\alpha)-\alpha\eta-\delta)$. Hence, $x-\tau-\delta > 0$ implies that $x(1-\alpha)-\delta > \alpha\eta > 0$.

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

This is equation (27) in the main text. For the special case where creative destruction does not lead to any productivity advancements, i.e. $\lambda = 1$, the tail coefficient is given by $\varrho_q = \frac{1}{1-\varpi^{\sigma-1}}$.

A-1.2 Model Extensions (Section 2.7)

A-1.2.1 Endogenizing the Direction of Innovation α.

In the baseline model in we assume that innovation was undirected, i.e. the share of product innovation resulting in creative destruction (rather than new varieties) was constant and equal to α. In this section we show that we can extend our theory to a setting where the direction of innovation is a choice variable of the firm.

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

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

where

$$\Xi_t \equiv \max_{x_N} \left\{ x_N V_t^N - \frac{1}{\varphi_N} x_N^{\zeta} w_t \right\} + \max_{x_{CD}} \left\{ x_{CD} V_t^{CD} - \frac{1}{\varphi_{CD}} x_{CD}^{\zeta} w_t \right\}, \tag{A-30}$$

where 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

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

The optimal innovation rates associated with (A-30) are given by

$$x_{NV} = \left(\frac{\varphi_N}{\zeta} \frac{V_t^{NV}}{w_t}\right)^{\frac{1}{\zeta-1}} \text{ and } x_{CD} = \left(\frac{\varphi_{CD}}{\zeta} \frac{V_t^{CD}}{w_t}\right)^{\frac{1}{\zeta-1}}. \tag{A-31}$$

Note that this implies that the endogenous share of product creation directed to creative destruction is given by

$$\tilde{\alpha} = \frac{\left(\varphi_{CD} \frac{V_t^{CD}}{w_t}\right)^{\frac{1}{\zeta-1}}}{\left(\varphi_N \frac{V_t^N}{w_t}\right)^{\frac{1}{\zeta-1}} + \left(\varphi_{CD} \frac{V_t^{CD}}{w_t}\right)^{\frac{1}{\zeta-1}}},$$

i.e. the relative "bias" of innovation depends on the relative valuations. This also implies that

$$\Xi_t = \left(\frac{\zeta-1}{\varphi_{NV}} x_{NV}^{\zeta} + \frac{\zeta-1}{\varphi_{CD}} x_{CD}^{\zeta}\right) w_t, \tag{A-32}$$

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 $g_N = \eta$. Hence,

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

where Ξ_t is given in (A-32).

To solve for Ξ_t and x_{NV} and x_{CD}, we need V_t^N and V_t^{CD}. As before these are given by

$$V_t^{CD} = \int V(\lambda q)\, dF_t(q) = \left(\frac{(\mu - 1)\lambda^{\sigma - 1}}{\rho + (g_Q - I)(\sigma - 1) + \tau + \delta} \frac{L_t^P}{N_t} + \frac{\frac{\zeta - 1}{\varphi_{NV}} x_{NV}^{\zeta} + \frac{\zeta - 1}{\varphi_{CD}} x_{CD}^{\zeta}}{\rho + \tau + \delta} \right) w_t = V_t(\lambda Q_t).$$

Similarly, the value of new variety creation is given by $V_t^{NV} = V(\omega Q_t)$.

Entry 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 new 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 $\frac{1}{\varphi_E} w_t = \Xi_t$, where Ξ_t is the value of innovation given in (A-32).

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. These objects are determined from the following conditions:

1. Because $g_N = \eta$ along the BGP, $\nu = x_{NV} + z x_{NV} = x_{NV}(1 + z) = \eta + \delta$.
2. Creative destruction τ is given by $\tau = x_{CD} + z x_{CD} = x_{CD}(1 + z)$.
3. The first order condition for x_{NV} and x_{CD} are given by (see (A-31))

$$x_N = \left(\frac{\varphi_N}{\zeta} \frac{V_t^N}{w_t}\right)^{\frac{1}{\zeta - 1}} \text{ and } x_{CD} = \left(\frac{\varphi_{CD}}{\zeta} \frac{V_t^{CD}}{w_t}\right)^{\frac{1}{\zeta - 1}}. \tag{A-33}$$

4. Equation (A-32) implies that the free entry condition is given by $\frac{1}{\varphi_E} = \frac{\Xi_t}{w_t} = \frac{\zeta - 1}{\varphi_{NV}} x_{NV}^{\zeta} + \frac{\zeta - 1}{\varphi_{CD}} x_{CD}^{\zeta}$.
5. The value functions V^{NV}/w_t and V^{CD}/w_t can be written as

$$\frac{V_t^{NV}}{w_t} = \frac{(\mu - 1)\omega^{\sigma - 1}}{\lambda^{\sigma - 1}\tau + (\omega^{\sigma - 1} - 1)\eta + \omega^{\sigma - 1}\delta} \frac{L_t^P}{N_t} + \frac{1}{\rho + \tau + \delta} \frac{1}{\varphi_E} \tag{A-34}$$

and

$$\frac{V_t^{CD}}{w_t} = \frac{(\mu-1)\,\lambda^{\sigma-1}}{\lambda^{\sigma-1}\tau + (\omega^{\sigma-1}-1)\,\eta + \omega^{\sigma-1}\delta}\frac{L_t^P}{N_t} + \frac{1}{\rho+\tau+\delta}\frac{1}{\varphi_E}. \tag{A-35}$$

These are 7 equations in 7 unknowns $(z, x_{NV}, x_{CD}, \frac{V_t^{CD}}{w_t}, \frac{V_t^{NV}}{w_t}, \tau, \frac{L_t^P}{N_t})$, which fully determine the BGP equilibrium.

We can simplify this system further and express the BGP equilibrium in terms of x_{NV} and x_{CD}. Note first that $\tau = \frac{x_{CD}}{x_{NV}}(\eta+\delta)$. From (A-33) 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 $\frac{V_t^{NV}}{w_t}$ in (A-34) and (A-35), we can solve for $\frac{L_t^P}{N_t}$ as

$$\frac{(\mu-1)}{\lambda^{\sigma-1}\frac{x_{CD}}{x_{NV}}(\eta+\delta) + (\omega^{\sigma-1}-1)\,\eta + \omega^{\sigma-1}\delta}\frac{L_t^P}{N_t} = \frac{1}{\varphi_E}\left(\frac{\frac{\zeta}{\zeta-1} - \frac{1}{\rho+\frac{x_{CD}}{x_{NV}}(\eta+\delta)+\delta}(x_N+x_{CD})}{\lambda^{\sigma-1}x_{CD}+\omega^{\sigma-1}x_{NV}}\right). \tag{A-36}$$

Solving for $\frac{V_t^{CD}}{w_t}$ and $\frac{V_t^{NV}}{w_t}$ implies that x_{NV} and x_{CD} are determined from the equations

$$x_N^{\zeta-1}\frac{\zeta\varphi_E}{\varphi_N} = \frac{\frac{\zeta}{\zeta-1}\omega^{\sigma-1}}{\lambda^{\sigma-1}x_{CD}+\omega^{\sigma-1}x_{NV}} - \frac{1}{\rho+\frac{x_{CD}}{x_{NV}}(\eta+\delta)+\delta}\frac{(\omega^{\sigma-1}-\lambda^{\sigma-1})\,x_{CD}}{\lambda^{\sigma-1}x_{CD}+\omega^{\sigma-1}x_{NV}} \tag{A-37}$$

$$x_{CD}^{\zeta-1}\frac{\zeta\varphi_E}{\varphi_{CD}} = \frac{\frac{\zeta}{\zeta-1}\lambda^{\sigma-1}}{\lambda^{\sigma-1}x_{CD}+\omega^{\sigma-1}x_{NV}} - \frac{1}{\rho+\frac{x_{CD}}{x_{NV}}(\eta+\delta)+\delta}\frac{(\lambda^{\sigma-1}-\omega^{\sigma-1})\,x_{NV}}{\lambda^{\sigma-1}x_{CD}+\omega^{\sigma-1}x_{NV}}. \tag{A-38}$$

Using that $\tau = \frac{x_{CD}}{x_{NV}}(\eta+\delta)$ we can write (A-37) and (A-38) in terms of x_{NV} and τ. With some algebra we can solve for x_{NV} explicitly as a function of τ and parameters. In particular, one can show that

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

Substituting this expression for x_{NV} into (A-37), we arrive at the equation

$$\left(\left(\frac{\lambda}{\overline{\omega}}\right)^{\sigma-1}-1\right)\frac{1}{\eta+\delta}\frac{\tau}{\rho+\delta+\tau} = \frac{\zeta}{(\zeta-1)^{\frac{\zeta-1}{\zeta}}}\left(\frac{\varphi_E}{\varphi_N}\right)^{\frac{1}{\zeta}}\frac{\left(\frac{\lambda}{\overline{\omega}}\right)^{\sigma-1}\frac{\tau}{\eta+\delta} - \left(\frac{\tau}{\eta+\delta}\right)^{\zeta}\frac{\varphi_N}{\varphi_{CD}}}{\left(1+\left(\frac{\tau}{\eta+\delta}\right)^{\zeta}\frac{\varphi_N}{\varphi_{CD}}\right)^{\frac{\zeta-1}{\zeta}}}. \tag{A-39}$$

This equation determines τ as a function of parameters. In particular, τ depends directly on η.

A Special Case Consider first a special case of this setup which is exactly isomorphic to our baseline model where the direction of innovation α is exogenous. Assume that creative destruction and new variety creation leads to the same quality improvement, i.e. $\lambda = \overline{\omega}$. This implies that the value of creative destruction and new variety creation is equalized, i.e. $V_t^{CD} = V_t^{NV} = V_t$. This in turn directly yields

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

Hence, the share of innovation activity directed to creative destruction, α, is indeed endogenous and simply determined from the relative innovation efficiencies φ_{CD} and φ_{NV}. Hence, we can write $x_{NV} = (1-\alpha)\,x$ and $x_{CD} = \alpha x$ as in the baseline model. 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.[37] Using the free entry condition yields

$$x = \left(\frac{1}{\zeta-1}\frac{1}{\varphi_E}\frac{1}{\left(\frac{(1-\alpha)^{\zeta}}{\varphi_{NV}} + \frac{\alpha^{\zeta}}{\varphi_{CD}}\right)}\right)^{1/\zeta}.$$

As in the baseline model, 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.

The General Case If $\lambda \neq \omega$, τ is determined from (A-39). We can rewrite (A-39) as

$$\left(\frac{\kappa-1}{\kappa}\right)\mathcal{A}\frac{1}{\rho+\delta+\tau}\left(1+\left(\frac{\tau}{\eta+\delta}\right)^{\zeta}\frac{\varphi_N}{\varphi_{CD}}\right)^{\frac{\zeta-1}{\zeta}} + \frac{1}{\kappa}\left(\frac{\tau}{\eta+\delta}\right)^{\zeta-1}\frac{\varphi_N}{\varphi_{CD}} = 1,$$

where $\kappa = \left(\frac{\lambda}{\omega}\right)^{\sigma-1}$ and $\mathcal{A} = \frac{(\zeta-1)^{\frac{\zeta-1}{\zeta}}}{\zeta\left(\frac{\varphi_E}{\varphi_N}\right)^{\frac{1}{\zeta}}}$. Define the function

$$h(\tau) = \left(\frac{\kappa-1}{\kappa}\right)\mathcal{A}\frac{1}{\rho+\delta+\tau}\left(1+\tau^{\zeta}\frac{1}{(\eta+\delta)^{\zeta}}\frac{\varphi_N}{\varphi_{CD}}\right)^{\frac{\zeta-1}{\zeta}} + \frac{1}{\kappa}\tau^{\zeta-1}\frac{1}{(\eta+\delta)^{\zeta-1}}\frac{\varphi_N}{\varphi_{CD}}.$$

Then, a solution τ^* is implicitly defined by $h(\tau^*) = 1$. Note that $h(.)$ satisfies

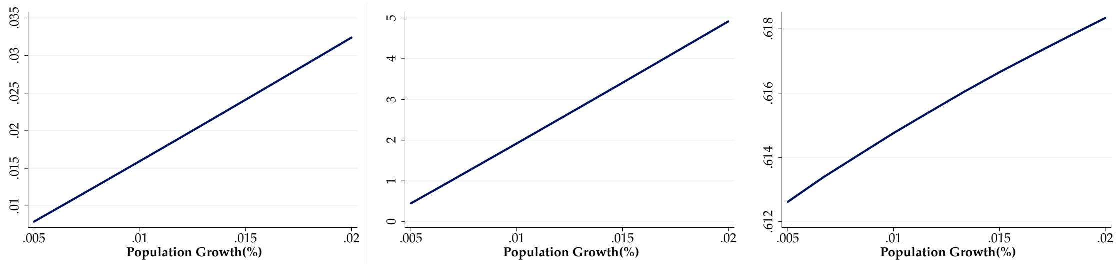
$$h(0) = 0 \quad \text{and} \quad \lim_{m\to\infty} h(m) = \infty.$$

Note also that h is continuous so that there is at least one solution $h(\tau^*) = 1$. Moreover, at least one solution satisfies $h'(\tau^*) > 0$. If there is a unique solution then $h'(\tau^*) > 0$. Hence, focus on a solution where $h'(\tau^*) > 0$. Note that an increase in η shifts the function $h(\tau)$ downwards. Hence, an

[37] Note that this solution is also implied by (A-39). If $\lambda = \varpi$, (A-39) requires that $\frac{\tau}{\eta} - \left(\frac{\tau}{\eta}\right)^{\zeta}\frac{\varphi_N}{\varphi_{CD}} = 0$. This implies that $\tau = \left(\frac{\varphi_{CD}}{\varphi_N}\right)^{\frac{1}{\zeta-1}}\eta = \frac{\alpha}{1-\alpha}\eta$.

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

(a) Creative Destruction τ (b) Rel. Importance of Entry z/x (c) Creative Destruction Share 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.

increase in η will increase τ^*. As in our baseline model, falling population growth reduces creative destruction τ.

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 $\overline{\omega} < \lambda$. To see why, consider the special case where new entrants are extremely unproductive, i.e. $\omega \approx 0$. In that case, equations (A-37) and (A-38) can be solved analytically as

$$x_{NV} = \left(\frac{\varphi_N}{\zeta \varphi_E} \frac{1}{\rho + \delta + \frac{x_{CD}}{x_{NV}} (\eta + \delta)} \right)^{\frac{1}{\zeta - 1}} \quad \text{and} \quad x_{CD}^{\zeta} = \frac{\varphi_{CD}}{\varphi_E} \frac{1}{\zeta - 1} - \frac{\varphi_{CD}}{\varphi_N} x_{NV}^{\zeta}.$$

Hence, x_{NV} is decreasing in η and x_{CD} is decreasing in x_{NV} but does not depend on η conditional on x_{NV}. Falling population growth therefore increases x_{NV} but decreases x_{CD}. Hence, falling population growth reduces α as shown in Figure A-1. To see why, define the function

$$v(m) = \frac{(\mu - 1) m^{\sigma - 1}}{\lambda^{\sigma - 1} \tau + (\omega^{\sigma - 1} - 1)(\eta + \delta) + \omega^{\sigma - 1} \delta} \frac{L_t^P}{N_t} + \frac{1}{\rho + \tau + \delta} \frac{1}{\varphi_E}.$$

Then $V_t^{NV} / w_t = v(\overline{\omega})$ and $V_t^{CD} / w_t = v(\lambda)$. Now recall that a decline in population growth reduces both τ and L_t^P / N_t. The reduction in τ increases $v(m)$ due to a decline in the discount rate. The reduction in L_t^P / N_t obviously lowers $v(m)$. The initial quality term m thus governs the weight on the present discount value of profits (the first term) as opposed to the value of innovation (the second term). If m is small, $v(m)$ is not negatively affected by the decline in L_t^P / N_t but mostly benefits from the *increase* in the innovation value. A decline in population growth therefore 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 subject to some costs. In particular, assume that the cost function (in terms of labor) of achieving a drift I of a particular product 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 (for simplicity we assume the same convexity as for firms' product creation technology). Additionally, the cost of innovation depend on firms' relative efficiency q/Q_t. Allowing for this cost-shifter is required to make the model consistent with balanced growth (see e.g. Atkeson and Burstein (2010)) 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 characterization of the value function is also strikingly similar. The value function is still additive across products and the value of a given product with efficiency q is given by

$$V_t(q) = \frac{\pi_t(q)}{\rho + \delta + \tau + \left(g^Q - \frac{\zeta-1}{\zeta} I\right)(\sigma - 1)} + \frac{\frac{\zeta-1}{\varphi_x} x^{\zeta} w_t}{\rho + \tau + \delta}. \tag{A-40}$$

Hence, the only difference to the baseline model is the term $\frac{\zeta-1}{\zeta}$ in front of I in the discount rate. Given $V_t^I(q)$ the optimal rate of own-innovation is therefore defined by

$$\max_I \left\{ I \frac{\partial V_t(q)}{\partial q} q - \left(\frac{q}{Q_t}\right)^{\sigma-1} \frac{1}{\varphi_I} I^{\zeta} w_t \right\}. \tag{A-41}$$

Using (A-40), the optimal innovation rate associated with (A-41) is given by

$$I = \left(\frac{(\sigma-1)(\mu-1) L_t^P / N_t}{\rho + \delta + \tau + \left(g^Q - \frac{\zeta-1}{\zeta} I\right)(\sigma-1)} \frac{\varphi_I}{\zeta} \right)^{\frac{1}{\zeta-1}}.$$

The free entry condition now implies

$$\frac{1}{\varphi_E} = \frac{(\mu-1)\left(\alpha\lambda^{\sigma-1} + (1-\alpha)\overline{\omega}^{\sigma-1}\right)\ell}{\rho + \delta + \tau + \left(g^Q - \frac{\zeta-1}{\zeta} I\right)(\sigma-1)} + \frac{\frac{\zeta-1}{\varphi_x}\left(\frac{1}{\zeta}\frac{\varphi_x}{\varphi_E}\right)^{\frac{\sigma}{\zeta-1}}}{\rho + \tau + \delta}.$$

Substituting this into the expression for I yields equation (32) in the main text.

A-1.2.3 Diminishing returns to research in entry and product creation

Now we suppose that there are diminishing returns to research labor as the general state of technology improves, as suggested by the evidence in Bloom et al. (2020). In particular, suppose that the entry cost in units of labor is $\frac{1}{\varphi_E} Q_t^{\varsigma}$ with $\varsigma > 0$. Moreover, suppose that the innovation cost function for expanding the portfolio of products the firm has is

$$c_t^X(x,n) = \frac{1}{\varphi_x} x^{\zeta} n Q_t^{\varsigma}$$

In this case, the solution for the value function is given by

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

Free entry then requires that

$$\frac{1}{\varphi_E} = \frac{(\mu-1)\frac{L_t^P}{N_t}Q_t^{-\varsigma}}{(\rho+\tau+\delta+(\sigma-1)(g_Q-I)-g_{l^P})} + \frac{(\zeta-1)x}{(\rho+\tau+\delta-\varsigma g_Q)} \tag{A-42}$$

For this to hold on a BGP, it must be that $\frac{L_t^P}{N_t}Q_t^{-\varsigma}$ is constant and that L_t^P grows at the rate of population growth so that ℓ^P is constant. Hence, $\eta = g_N + \varsigma g_Q = \nu_t - \delta + \varsigma g_Q$. Note that it is still the case that $g_t^Q = I + \frac{\lambda^{\sigma-1}-1}{\sigma-1}\tau_t + \frac{\varpi^{\sigma-1}-1}{\sigma-1}\nu_t$ and that $\nu_t = (1-\alpha)(x_t+z_t) = \frac{1-\alpha}{\alpha}\tau_t$. Hence, the rate of creative destruction is given by

$$\tau = \frac{\alpha(\eta+\delta-\varsigma I)}{1-\alpha+\varsigma\left(\frac{\bar{q}^{\sigma-1}-1}{\sigma-1}\right)}.$$

As in the baseline model, τ is an increasing function of the rate of population growth η. To fully characterize the equilibrium, the research share of the economy, ℓ^P, again comes from the labor market clearing condition, which now reads $\frac{1}{\ell_t^P} = 1 + \frac{Q_t^{\varsigma}N_t}{L_t^P}\left(\frac{1}{\varphi_E}z + \frac{1}{\varphi_x}x^{\zeta}\right)$. Given that the ratio $\frac{Q_t^{\varsigma}N_t}{L_t^P}$ is fully determined from free entry and constant on the BGP (see (A-42)), and z and x are constant on the BGP, the share of production labor ℓ_t^P is constant.

The main difference between the baseline model and this modification is that the firm size distribution is not stationary along the BGP. For example, total employment of a product with average quality Q_t is given by $l_t(Q_t) = \left(\frac{Q_t}{Q_t}\right)^{\sigma-1}\frac{L_t^P}{N_t} = \frac{L_t^P}{N_t}$. Along a BGP, $\frac{\dot{l}_t(Q_t)}{l_t(Q_t)} = \eta - g_N = \varsigma g_Q$.Hence, to the extent that innovation cost rise with aggregate productivity (i.e. $\varsigma > 0$), average firm size increases along the BGP at a constant rate. Intuitively, higher profits per product are needed to offset the constantly increasing entry cost.

A-1.3 Characterization of the Model with Bertrand Competition (Section 3)

In this section we derive the results for the model with Bertrand competition described in Section 3

A-1.3.1 The Value Function

The only difference relative to the baseline case characterized in Section A-1.1.5 is that the static profit function is given by (A-4), i.e.

$$\pi\left(q_i,\Delta_i\right)=\left(1-\frac{1}{\mu\left(\Delta_i\right)}\right)\mu\left(\Delta_i\right)^{1-\sigma}\left(\frac{q_i}{Q_t}\right)^{\sigma-1}\frac{1}{\left(\mathcal{M}_t\Lambda_t\right)^{\sigma-1}}\frac{Y_t}{N_t}. \tag{A-43}$$

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

$$r_tV_t\left(q,\Delta\right)-\dot{V}_t\left(q,\Delta\right)=\pi_t\left(q,\Delta\right)+I\left\{\frac{\partial V_t\left(q,\Delta\right)}{\partial q}q+\frac{\partial V_t\left(q,\Delta\right)}{\partial\Delta}\Delta\right\}-\left(\tau+\delta\right)V_t\left(q,\Delta\right)+\Xi_t, \tag{A-44}$$

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

To characterize the value function, note first that the definition of Ξ_t still implies that

$$0=\left(\alpha V_t^{CD}+\left(1-\alpha\right)V_t^{NV}\right)-\frac{\zeta}{\varphi_x}x^{\zeta-1}w_t.$$

The free entry condition thus still requires that $w_t=\varphi_E\left(\alpha V_t^{CD}+\left(1-\alpha\right)V_t^{NV}\right)$. Hence, as in the baseline model, $x=\left(\frac{\varphi_x}{\varphi_E}\frac{1}{\zeta}\right)^{\frac{1}{\zeta-1}}$. Therefore $\Xi_t=\frac{\zeta-1}{\varphi_x}x^{\zeta}w_t\propto w_t$.

To solve for $V_t\left(q,\Delta\right)$ in (A-44) along the BGP, conjecture that $V_t\left(q,\Delta\right)$ takes the form

$$V_t\left(q,\Delta\right)=k\left(\Delta\right)q^{\sigma-1}U_t+M_t.$$

This implies that

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

Equation (A-44) thus implies that $k\left(\Delta\right)$, U_t and M_t solve the equations

$$\left(r_t+\tau\right)M_t-\dot{M}_t=\frac{\zeta-1}{\varphi_x}x^{\zeta}w_t \tag{A-45}$$

and

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

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

$$\left(r_t+\tau\right)k\left(\Delta\right)U_t-k\left(\Delta\right)\dot{U}_t=h\left(\Delta\right)\left(\frac{1}{Q_t}\right)^{\sigma-1}\frac{1}{\left(\mathcal{M}_t\Lambda_t\right)^{\sigma-1}}\frac{Y_t}{N_t}+I\left(\left(\sigma-1\right)+\varepsilon_k\left(\Delta\right)\right)k\left(\Delta\right)U_t.$$

Along the BGP, $\mathcal{M}_t\Lambda_t$ is constant and U_t grows at the same rates as $\frac{Y_t}{N_t}Q_t^{1-\sigma}$. From (A-2) and (A-3) this implies that $g_U = \frac{\dot{U}_t}{U_t} = g_w - (\sigma - 1)\, g_Q$. Hence,

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

The solution to the value function is therefore

$$U_t = \left(\frac{1}{Q_t}\right)^{\sigma-1} \frac{1}{(\mathcal{M}\Lambda)^{\sigma-1}} \frac{Y_t}{N_t}$$

and

$$k(\Delta) \quad = \quad \frac{h(\Delta)}{r + \delta + \tau - g_w + (\sigma - 1)\,(g_Q - I) - I\frac{k'(\Delta)\Delta}{k(\Delta)}}.$$

Along the BGP we have

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

Hence, the function $k(\Delta)$ solves the differential equation

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

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

$$k(\Delta)\, \mathcal{C} - Ik'(\Delta)\,\Delta \quad = \quad \frac{1}{\sigma - 1}\left(\frac{\sigma - 1}{\sigma}\right)^{\sigma}.$$

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

$$k(\Delta)\, \mathcal{C} - Ik'(\Delta)\,\Delta \quad = \quad \frac{\Delta - 1}{\Delta^{\sigma}}. \qquad \text{(A-46)}$$

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-45) implies M_t grows at g_w along the BGP. Hence,

$$M_t = \frac{1}{r + \tau + \delta - g_w} \frac{\zeta - 1}{\varphi_x} x^{\zeta} = \frac{1}{\rho + \tau + \delta} \frac{\zeta - 1}{\varphi_x} x^{\zeta},$$

because $r = \rho + g_w$. Together with the solution for $k(\Delta)$, the value function along the BGP is given by

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

Using (A-3) we get that

$$\frac{1}{(\mathcal{M}\Lambda)^{\sigma-1}}\frac{Y_t}{N_t}=\frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}}\frac{L_P}{N_t}w_t=\frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}}\frac{\ell^P}{\mathcal{N}}w_t,$$

where $\ell^P=L_t^P/L_t$ and $\mathcal{N}=N_t/L_t$ are constant along a BGP. Hence,

$$V_t(q,\Delta)=\left\{k(\Delta)\left(\frac{q}{Q_t}\right)^{\sigma-1}\frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}}\frac{\ell_t^P}{\mathcal{N}_t}+\frac{1}{\rho+\tau+\delta}\frac{\zeta-1}{\varphi_x}x^{\zeta}\right\}w_t. \tag{A-47}$$

A-1.3.2 The Free Entry Condition

Using (A-47) we can derive the free entry condition. The value of creative destruction is given by

$$V_t^{CD}=\int V_t(\lambda q,\lambda)\,dF_t(q)=\left\{k(\lambda)\lambda^{\sigma-1}\frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}}\frac{\ell^P}{\mathcal{N}}+\frac{1}{\rho+\tau+\delta}\frac{\zeta-1}{\varphi_x}x^{\zeta}\right\}w_t.$$

The value of variety creation is

$$V_t^{NV}=\int V_t\left(\omega Q_t,\frac{\sigma}{\sigma-1}\right)d\Gamma(\omega)=\left\{k\left(\frac{\sigma}{\sigma-1}\right)\overline{\omega}^{\sigma-1}\frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}}\frac{\ell^P}{\mathcal{N}}+\frac{1}{\rho+\tau+\delta}\frac{\zeta-1}{\varphi_x}x^{\zeta}\right\}w_t,$$

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. In particular, x, τ and $k(\Delta)$ can be calculated as functions of parameters and $\mathcal{M}^{\sigma-1}\Lambda^{\sigma}$ is fully determined from the joint distribution $F(\Delta,q)$, which is also only a function of parameters along the BGP.

A-1.3.3 Proposition 1 in the model with Bertrand Competition

To see that Proposition 1 still applies in the model with Bertrand competition, note first that creative destruction and the rate of 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-13) 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-1.3.4 The Joint Distribution of Gaps and Productivity

In the model with Bertrand competition in Section 3, the joint distribution of relative quality $\hat{q}_t=\ln(q_t/Q_t)^{\sigma-1}$ (see (A-27)) 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 $F_t^{NC}(\hat{q})$.[38] 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}. Of course, $N_t = N_t^C + N_t^{NC}$. Also recall that $\hat{q}_t$ has a drift of $g_{\hat{q}} = (\sigma - 1)(I - g_{Q,t})$ (see (A-28)).

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

$$\bar{F}_t^{NC}(\hat{q}) = \underbrace{\bar{F}_{t-\iota}^{NC}\left(\hat{q} - g_{\hat{q}}\iota\right)(1 - (\tau_t + \delta)\iota)}_{\text{existing mass that survives and improves/falls}} + \underbrace{\left(\frac{1-\alpha}{\alpha}\right)\tau_t N_t \Gamma\left(\frac{\exp(\hat{q})}{\sigma - 1}\right)\iota}_{\text{new products that enter}}.$$

where we have used the fact that the new product creation rate is $\frac{1-\alpha}{\alpha}\tau = g_t^N$. Note also that $N_t(1-\alpha)$ will be the equilibrium mass of non-competitive products, which we will call N_t^{NC}. As ι becomes small this leads to the differential equation

$$\frac{\partial \bar{F}_t^{NC}(\hat{q})}{\partial t} = -g_{\hat{q}}\frac{\partial \bar{F}_t^{NC}(\hat{q})}{\partial \hat{q}} - (\tau_t + \delta)\bar{F}_t^{NC}(\hat{q}) + \left(\frac{1-\alpha}{\alpha}\right)\tau_t \frac{N_t^{NC}}{(1-\alpha)}\Gamma\left(\frac{\exp(\hat{q})}{\sigma - 1}\right). \tag{A-48}$$

Defining the distribution $F_t^{NC} \equiv \bar{F}_t^{NC}/N_t^{NC}$, A-48 implies

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

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

$$\begin{aligned}\bar{F}_t^C(\Delta, \hat{q}) = {} & \underbrace{\bar{F}_{t-\iota}^C\left(\Delta e^{-I\iota}, \hat{q} - g_{\hat{q}}\iota\right)(1 - (\tau_t + \delta)\iota)}_{\text{existing mass that survives and improves/falls}} \\ & + \underbrace{\lim_{s\to\infty} \iota\tau_t \bar{F}_{t-\iota}^C(s, \hat{q} - \hat{\lambda})}_{\text{Creative destruction of } C \text{ products below } \hat{q}\text{-}\hat{\lambda}} + \underbrace{\tau\iota\bar{F}_{t-\iota}^{NC}(\hat{q} - \hat{\lambda})}_{\text{Creative destruction of } NC \text{ products below } \hat{q}\text{-}\hat{\lambda}},\end{aligned}$$

[38] 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 $\hat{\lambda} = \ln \lambda^{\sigma-1}$. As ι becomes small this leads to the differential equation

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

Defining $\bar{F}_t^C(\Delta,\hat{q}) = N_t^C F_t^C(\Delta,\hat{q})$, we get

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

Note that the latter term depends on the relative share of products without a competitor N_t^{NC}/N_t^C. To derive N_t^{NC}/N_t^C, note that N_t^{NC} and N_t^C evolve according to

$$\dot{N}_t^{NC} = \underbrace{N_t \tau \left(\frac{1-\alpha}{\alpha}\right)}_{\text{New varieties}} - \underbrace{N_t^{NC}(\delta+\tau)}_{\text{Products that turn into C products or exit}},$$

And

$$\dot{N}_t^C = \underbrace{-\delta N_t^C}_{\text{Exiting products}} + \underbrace{N_t^{NC}\tau}_{\text{New inflows}}.$$

The steady state share of NC products is therefore given by

$$\frac{N_t^{NC}}{N_t} = \frac{\tau\left(\frac{1-\alpha}{\alpha}\right)}{\eta+\delta+\tau} = 1-\alpha, \tag{A-49}$$

i.e. the steady-state share of NC products is simply given by its share in the process of product creation.

A-1.3.5 Marginal gap distribution

We now derive the distribution of efficiency gaps given in (34). Let $F_t^C(\Delta)$ denote the cdf of quality gaps among products with a competitor. Let, as before, denote the number of competitor and non-competitor products as N_t^C and N_t^{NC}. The distribution $F_t^C(\Delta)$ the solves the differential equation

$$\frac{\partial F_t^C(\Delta)}{\partial t} + F_t^C(\Delta)\frac{1}{N_t^C}\frac{\partial N_t^C}{\partial t} = \underbrace{-I\Delta \frac{\partial F_t^C(\Delta)}{\partial \Delta}}_{\text{Upward drift of own-innovation}} - \underbrace{\delta F_t^C(\Delta)}_{\text{Exit}} + \underbrace{(1-F_t^C(\Delta))\tau}_{\text{Inflow through CD}} + \underbrace{\frac{N_t^{NC}}{N_t^C}\tau}_{\text{Inflow from } NC\text{products}}.$$

Along a BGP, this distribution is stationary (i.e. $\frac{\partial F_t^C(\Delta)}{\partial t} = 0$), the number of competitive products grows at rate η and $N_t^{NC}/N_t^C = \frac{1-\alpha}{\alpha}$ (see (A-49)). 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 of aggregate employment then they used to in 1980.

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 %

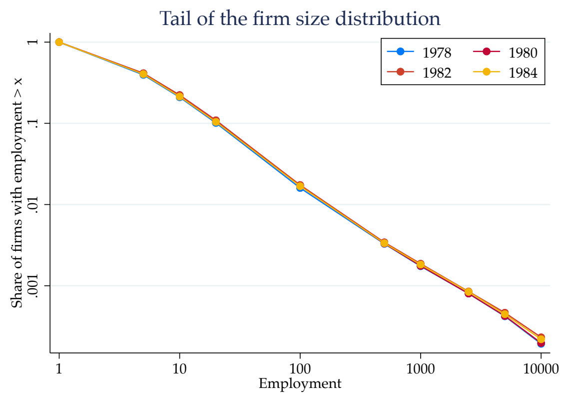
Notes: This Table gives basic summary information about the firms in the LBD through time.

A-2.2 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,

$$\ln P_t\left(l_f > x\right) = \delta - \varsigma \ln x. \quad \text{(A-50)}$$

Figure A-2: 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.

In Figure A-2 we show the empirical relationship between $\ln P_t\left(l_f > x\right)$ and $\ln x$ for different years. As predicted by (A-50), the relationship is almost perfectly linear and stable across years. When we estimate (A-50), 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.

A-2.3 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 4.2). In particular, we show that relative sales by age of the product is given by

$$s_P\left(a_P\right) \equiv E\left[\frac{p_i y_i}{Y}\middle| a_p\right] = \mu\left(a_p\right)^{1-\sigma} e^{(\sigma-1)\left(I-g^Q\right)a_p}\overline{q}^{\sigma-1}. \tag{A-51}$$

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 $s_f\left(a_f\right)$ and $\mu_f\left(a_f\right)$ computationally.

Consider a BGP where $\mathcal{M}_t$ and Λ_t are constant. (A-4) then implies that sales of product i relative to average sales are

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

Note first that markups are a deterministic function of Δ and Δ is a deterministic function of the age of the product. In particular, $\mu_i = \mu\left(a_P\right) = \min\left\{\lambda e^{I a_P}, \frac{\sigma-1}{\sigma}\right\}$. 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

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

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

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

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

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

Hence.

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

so that $s_P(a_P) = \mu(a_P)^{1-\sigma} e^{(\sigma-1)(I-g_Q)a_P} \overline{q}^{\sigma-1}\left(\frac{1}{\mathcal{M}_t \Lambda_t}\right)^{\sigma-1}$, which is the expression in (A-51).

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

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

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

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

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

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

where $s_P(a_P)$ is given in (A-51). 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. Letting $f_{a_P|A_f,N}(a_P|a,n)$ denote the conditional distribution of product age a_P conditional on firm age a_f and the number of products n and $\overline{p}_n(a_f)$ the probability a firm of age a_f having n products (conditional on survival). Then

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

Using that $\overline{p}_n(a_f) = \gamma(a_f)^{n-1}(1-\gamma(a_f))$ yields

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

where $\gamma(a)$ is given in (A-21). Using the same logic, the average markup as a function of firm age a_f

is given by

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

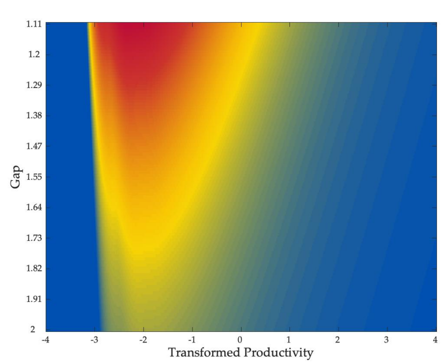
Given the density $f_{a_P|A_f,N}\left(a_P|a_f,n\right)$, these expressions can be directly evaluated. In Section SM-4 in the Supplementary Material we show how to compute this density.

A-2.4 The Joint Distribution of Efficiency and Markups

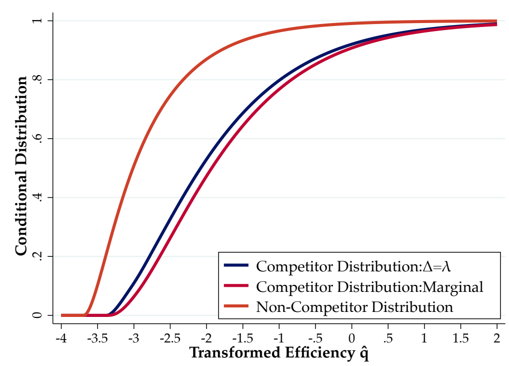
All aggregate allocations in our model depend on the misallocation wedge $\mathcal{M}$ and the labor share Λ. These aggregate wedges in turn depend on the joint distribution of relative efficiency q/Q and efficiency gaps Δ. In the left panel of Figure A-3 we display this distribution. Multiple forces 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 conditional 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.

Figure A-3: The Distributions of Efficiency q and Gaps Δ

(a) Joint Distribution of Efficiency and Gaps



(b) Conditional Productivity Distributions

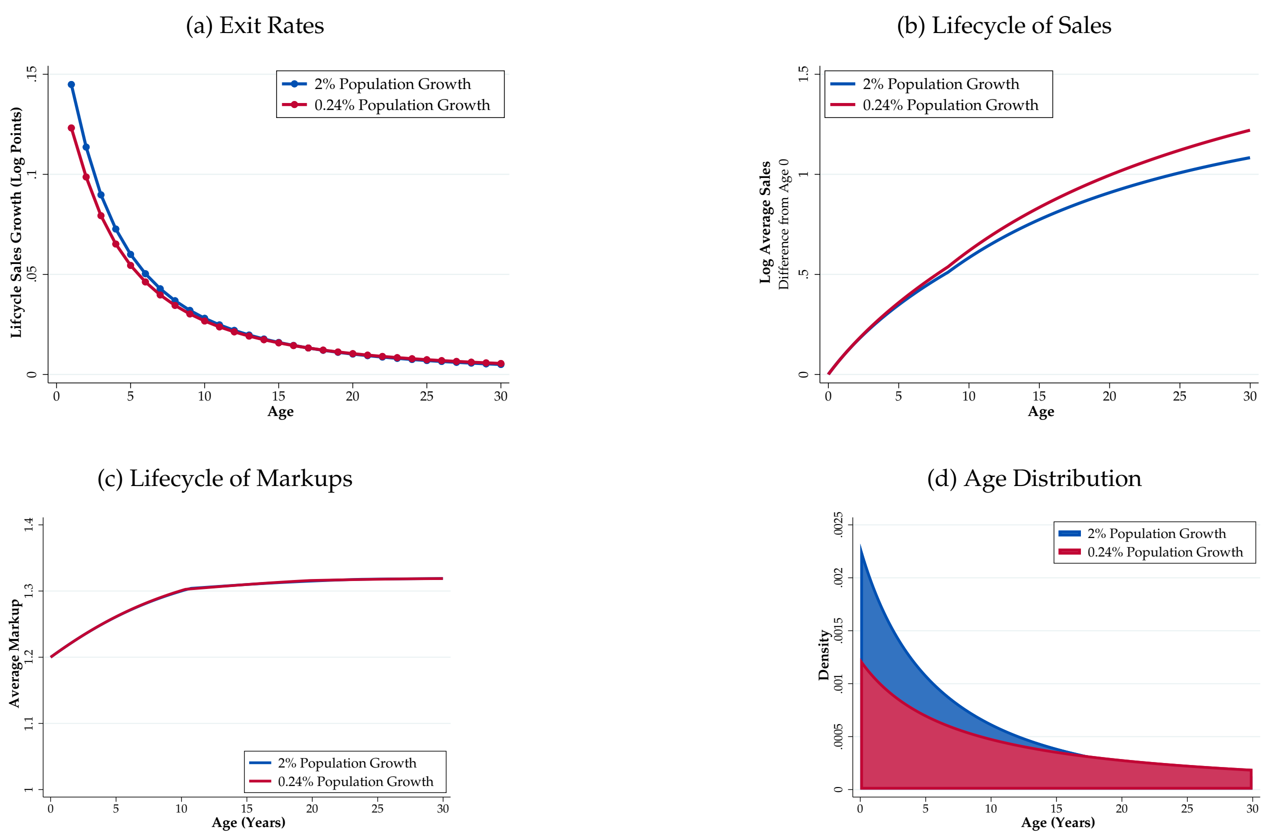


Notes: The left panel shows the joint density of $\hat{q}$ (relative efficiency) and Δ (the gap between the leading product and the next best product) in the calibrated BGP. The right panel shows the productivity distributions in the calibrated model for three types of products: non-competitive products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

A-2.5 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 substantially. In addition, the average markup also increased. In principle this rise in firm size and markups can be due to both changes in the age distribution and changes in firms' size conditional on age. In Figure A-4 we show that the lion share of these changes is due to changes in the age distribution. Consider first the top row, where 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. In the bottom row we report the life-cycle of markups (left panel) and the age distribution (right panel) again for both BGPs. The life-cycle of markups is essentially unchanged. By contrast, the age distribution shifts substantially: declining population growth causes firms to become older. And because older firms are larger, charge higher markups 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. (2016), 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



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