Econometrica Supplementary Material

SUPPLEMENT TO "HETEROGENEOUS MARKUPS, GROWTH, AND ENDOGENOUS MISALLOCATION"

(*Econometrica*, Vol. 88, No. 5, September 2020, 2037–2073)

MICHAEL PETERS
Department of Economics, Yale University

A-1. THEORETICAL RESULTS

A-1.1. *Cournot Competition*

IN THIS SECTION, I briefly discuss the model for the case where firms compete à la Cournot. This assumption makes the analysis much less tractable compared to the setup with Bertrand competition. To see this, consider product i and suppose that there are N_i potential producers with productivity $[q_{fi}]_{f=1}^{N_i}$. Let $[y_{fi}]_{f=1}^{N_i}$ denote the quantities these firms optimally set. Given the Cobb–Douglas demand across products, firm f solves the problem

$$\max_{y_f} p_i y_f - \frac{1}{q_f} w y_f \quad \text{s.t.} \quad p_i\left(\sum_{j=1}^{n} y_{ji}\right) = Y.$$

It is easy to verify that the equilibrium markup of firm f in product i is given by

$$\mu_{fi} = \frac{p_i}{\frac{1}{q_f} w} = \frac{q_i}{\Gamma\big([q_{fi}]\big)},$$

where $\Gamma([q_{fi}]) \equiv (\frac{1}{N-1}\sum_f \frac{1}{q_f})^{-1}$. Furthermore, the equilibrium quantities y_f and profits π_f are given by

$$y_f = \left(\frac{1-\mu_{fi}^{-1}}{\mu_{fi}}\right) q_i \frac{Y}{w} \quad \text{and} \quad \pi_f = Y\big(1-\mu_{fi}^{-1}\big)^2.$$

These equations highlight how the assumption of Bertrand competition simplifies the analysis. To compute profits of firm f for product i, one needs to know the number of competing firms in product i, N_i, and the distribution of productivity among them, $[q_{fi}]$, to compute $\Gamma([q_{fi}])$. This complicates the dynamic problem of the firm considerably. In addition, the mapping between the model and the data is more intricate. For the Bertrand case, one can treat the productivity of the most productive competitor as a latent variable as—according to the theory—only a single firm produces each individual product. For the case of Cournot competition, one has to specify which firms (in the data) are producing a particular product to compute firm-level markups.

A-1.2. *Proof of Proposition 1*

I prove Proposition 1 in two steps. I first derive the value function (6). Then I show that the conditions in Proposition 1 uniquely define the optimal choices for (I, x, z).

Michael Peters: m.peters@yale.edu

 https://doi.org/10.3982/ECTA15565

Solving for the Value Function $V_t(\cdot)$ in (6). To derive (6), conjecture first that the value function takes an additive form[1]

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

V_t^P and V_t^M are therefore defined by the differential equations

$$\begin{aligned} rV_t^M(\Delta)-\dot V_t^M(\Delta)=\pi_t(\Delta)-\pi_t(1)-\tau V_t^M(\Delta)&\\ +\max_I\big\{I\big[V_t^M(\Delta+1)-V_t^M(\Delta)\big]-c^I(I,\Delta)w_t\big\}& \end{aligned} \tag{A-2}$$

and

$$\begin{aligned} rV_t^P(n)-\dot V_t^P(n)=n\pi_t(1)+\sum_{i=1}^n\tau\big[V_t^P(n-1)-V_t^P(n)\big]&\\ +\max_X\big\{X\big[V_t^P(n+1)+V_t^M(1)-V_t^P(n)\big]-c^X(X,n)w_t\big\}.& \end{aligned}$$

Now consider a steady state where both value functions grow at rate g. Then I can write (A-2) as

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

Conjecture that

$$V_t^M(\Delta)=\kappa_t-\alpha_t\lambda^{-\Delta}. \tag{A-3}$$

Then $V_t^M(\Delta+1)-V_t^M(\Delta)=\frac{\lambda-1}{\lambda}\alpha_t\lambda^{-\Delta}$ and the optimal innovation rate I solves

$$I_t=\left(\frac{\lambda-1}{\lambda}\frac{\varphi_I}{\zeta}\frac{\alpha_t}{w_t}\right)^{\frac{1}{\zeta-1}}. \tag{A-4}$$

Suppose that $\frac{\alpha_t}{w_t}$ is constant along the BGP (which I will verify below). Equation (A-4) then implies that $I_t(\Delta)=I$. Using (A-4), (A-3), and the Euler equation $\rho=r-g$ yields

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

so that $\kappa_t=\frac{\frac{1}{\lambda}Y_t}{\rho+\tau}$ and $\alpha_t=\frac{Y_t-\frac{\zeta-1}{\varphi_I}I^\zeta w_t}{\rho+\tau}$. Hence, (A-3) yields

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

[1]The analysis in this section only contains the most important steps. A detailed derivation is contained in the Supplemental Material.

Note also that this implies that

$$V_t^M(1) = (\zeta - 1)\frac{1}{\lambda}\frac{1}{\varphi_I}\frac{I^\zeta}{\rho + \tau}w_t. \tag{A-5}$$

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

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

The optimality condition for x reads

$$v_t + V_t^M(1) = \frac{\zeta}{\varphi_x}x^{\zeta-1}w_t. \tag{A-6}$$

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

$$(\rho + \tau)v_t = \pi_t(1) + (\zeta - 1)\frac{1}{\varphi_x}x^\zeta w_t. \tag{A-7}$$

To solve for v_t, let $v_t = \overline{v}w_t$. The unknowns $(x, \overline{v})$ are then determined from (A-6), (A-7), and (A-5) as

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

$$\overline{v} = \frac{\dfrac{\lambda - 1}{\lambda}\dfrac{Y_t}{w_t} + \dfrac{\zeta - 1}{\varphi_x}x^\zeta}{\rho + \tau}.$$

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

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

Substituting into (A-1) yields (6).

Existence and Uniqueness. I now prove existence and uniqueness of the equilibrium. I need to solve for the tupel (I, x, z). Alternatively, I can solve for (I, x, τ) and then solve for $z = \tau - x$. From the static allocations, I know that $Y_t\Lambda_t = w_tL_t^P$. Note that Λ_t is a known function of τ/I (see Proposition 2) and hence I write it as $\Lambda(\frac{\tau}{I})$. To solve for L_t^P, I need the labor market clearing condition. Note that $L_t^z = \frac{1}{\varphi_I}z$, $L_t^x = \frac{1}{\varphi_x}x^\zeta$, and

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

Hence, the equilibrium is defined by the four equations

$$1 = \Lambda\left(\frac{\tau}{I}\right)\frac{Y_t}{w_t} + \frac{1}{\varphi_I}I^\zeta\Lambda\left(\frac{\tau}{I}\right) + \frac{1}{\varphi_z}(\tau - x) + \frac{1}{\varphi_x}x^\zeta, \tag{A-8}$$

$$\frac{1}{\varphi_z} = \frac{\frac{\lambda - 1}{\lambda}\frac{Y_t}{w_t} + \frac{\zeta - 1}{\varphi_x}x^{\zeta} + \frac{\zeta - 1}{\varphi_I}\frac{1}{\lambda}I^{\zeta}}{\rho + \tau}, \tag{A-9}$$

$$\frac{Y_t}{w_t} = \frac{\lambda}{\lambda - 1}\frac{\zeta}{\varphi_I}I^{\zeta-1}(\rho + \tau) + (\zeta - 1)\frac{1}{\varphi_I}I^{\zeta}, \tag{A-10}$$

$$\frac{Y_t}{w_t} = \frac{\zeta}{\varphi_x}x^{\zeta-1}(\rho + \tau)\frac{\lambda}{\lambda - 1} - \frac{\lambda}{\lambda - 1}\frac{\zeta - 1}{\varphi_x}x^{\zeta} - \frac{1}{\lambda - 1}\frac{\zeta - 1}{\varphi_I}I^{\zeta}. \tag{A-11}$$

To solve for the unknowns $(\frac{Y}{w}, I, \tau, x)$, note first that (A-11) and (A-9) imply that

$$x^{\zeta-1} = \frac{\varphi_x}{\varphi_z}\frac{1}{\zeta}. \tag{A-12}$$

This determines x in terms of parameters. I can then use (A-9), (A-10), and (A-8) to arrive at two equations in the two unknowns (τ, I):

$$1 = \Lambda\left(\frac{\tau}{I}\right)\left(\frac{\lambda}{\lambda - 1}\frac{\zeta}{\varphi_I}I^{\zeta-1}(\rho + \tau) + \zeta\frac{1}{\varphi_I}I^{\zeta}\right) + \frac{1}{\varphi_z}\tau - h(\varphi), \tag{A-13}$$

$$\frac{1}{\varphi_z} = \frac{\zeta}{\varphi_I}I^{\zeta-1} + \frac{\zeta - 1}{\varphi_I}\frac{I^{\zeta}}{\rho + \tau} + \frac{h(\varphi)}{\rho + \tau}, \tag{A-14}$$

where

$$h(\varphi) = \left(\frac{\zeta - 1}{\zeta}\right)\left(\frac{\varphi_x}{\zeta\varphi_z^{\zeta}}\right)^{\frac{1}{\zeta-1}} \geq 0. \tag{A-15}$$

Given a solution (I, τ) and x from (A-12), I can calculate $\frac{Y}{w}$ from (A-10) and $z = \tau - x$. Hence, I only have to show that (A-13) and (A-14) have a unique solution. Rewriting (A-13) and (A-14) in terms of $\tau/I = \vartheta$ yields

$$1 = \Lambda(\vartheta)\left(\frac{\lambda}{\lambda - 1}\frac{\zeta}{\varphi_I}I^{\zeta-1}(\rho + \vartheta I) + \zeta\frac{1}{\varphi_I}I^{\zeta}\right) + \frac{1}{\varphi_z}\vartheta I - h(\varphi), \tag{A-16}$$

$$\frac{1}{\varphi_z} = \frac{\zeta}{\varphi_I}I^{\zeta-1} + \frac{\zeta - 1}{\varphi_I}\frac{I^{\zeta}}{\rho + \vartheta I} + \frac{h(\varphi)}{\rho + \vartheta I}. \tag{A-17}$$

It can be shown that (A-16) and (A-17) have a unique solution for (I, ϑ).

A-1.3. *Proof of Proposition 2*

Consider first the distribution of quality gaps $\nu_t(\Delta)$. In a stationary equilibrium, $\dot{\nu}_t(\Delta) = 0$. Equation (9) implies that

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

Hence, $P(\Delta \leq d) = 1 - (\frac{1}{1+\vartheta})^d = 1 - e^{-\ln(1+\vartheta)\times d}$. This implies that log markups $\ln(\mu) = \Delta\ln(\lambda)$ are exponentially distributed with parameter θ. Similarly, $F(\mu; x) = P(\lambda^{\Delta} \leq \mu) =$

$1-\mu^{-\theta}$. To derive (10), note that[2]

$$\Lambda = \int \mu^{-1}\theta\mu^{-(\theta+1)}\,d\mu = \frac{\theta}{\theta+1},$$
$$\mathcal{M} = \exp\bigl(-E\ln(\mu)\bigr)\Lambda^{-1} = e^{-\theta^{-1}}\frac{\theta+1}{\theta}.$$

A-1.4. *Proof of Proposition 3*

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

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

The solution to this set of differential equations is given by

$$p_0(a) = 1 - e^{-\tau\times a},$$
$$p_{i+1}(a) = \left(\frac{1}{i!}\right)I^i a^i\bigl(e^{-(I+\tau)a}\bigr) \quad \text{for } i \geq 0.$$

The distribution of markups conditional on survival is then

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

This is a Poisson distribution with parameter Ia, so that $E[\Delta|a] = Ia$. Equation (11) then follows because $\ln(\mu) = \ln(\lambda)\Delta$.

The Expected Log Markup by Age: Equation (12). Firm-level markups are given by $\mu_f = \frac{py_f}{wl_f} = \frac{1}{\frac{1}{n}\sum_{j=1}^{n_i}\lambda^{-\Delta_j}}$. Hence,

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

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

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

[2]Note that these expression are derived taking the distribution of markups as continuous, even though the model implies that they are discrete. Taking this discreteness explicitly into account yields $\Lambda^{\text{Dis}} = \sum_{i=1}^{\infty}\lambda^{-i}\mu(i) = \frac{\vartheta}{\lambda(\vartheta+1)}\sum_{i=0}^{\infty}(\frac{1}{\lambda(\vartheta+1)})^i = \frac{\vartheta}{\lambda\vartheta+\lambda-1}$. Similarly, $\int_0^1\Delta(\nu)\,d\nu = \sum_{i=1}^{\infty}i\mu(i) = \vartheta\sum_{i=1}^{\infty}i(\frac{1}{1+\vartheta})^i = \frac{1+\vartheta}{\vartheta}$, so that $\mathcal{M}^{\text{Dis}} = \frac{1}{\Lambda}\lambda^{-\int_0^1\Delta(\nu)\,d\nu} = \lambda^{-\frac{1+\vartheta}{\vartheta}}\frac{\lambda-1+\lambda\vartheta}{\vartheta}$.

Then $E[\frac{1}{n}\sum_{j=1}^{n}\Delta_j|a,n] = \sum_{k=0}^{n} E[\frac{1}{n}\sum_{j=1}^{n}\Delta_j|a,n,k]P(B=k|a,n)$, where I denoted the conditioning on age a, the number of products n, and the random variable B as k. Then

$$E\left[\sum_j \frac{\Delta_j}{n}\middle|a,n\right] = E[\Delta|a,n,\mathrm{ni}] + \frac{\big(1-P(B=0|a,n)\big)\big(E[\Delta|a,n,\mathrm{i}]-E[\Delta|a,n,\mathrm{ni}]\big)}{n}, \quad \text{(A-18)}$$

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

Recovering $E[\Delta|a,n,\mathrm{ni}]$ *and* $E[\Delta|a,n,\mathrm{i}]$. Let U denote the age of the product so that

$$E[\Delta|a,n,\mathrm{ni}] = E_u\big\{E[\Delta|U=u]\big|a,n,\mathrm{ni}\big\} = E_u\left\{\sum_{i=1}^{\infty} i\times p(i,u)\middle|a,n,\mathrm{ni}\right\}, \quad \text{(A-19)}$$

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

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

Hence, $\sum_{i=1}^{\infty} i\times p(i,u) = Iu+1$. Equation (A-19) therefore implies that

$$E[\Delta|a,n,\mathrm{ni}] = 1 + I\times\int_{u=0}^{a} u f_{U|A,\mathrm{ni}}(u|a,\mathrm{ni})\,du, \quad \text{(A-20)}$$

where $f_{U|A,\mathrm{ni}}$ is the density of the conditional age distribution of a product. In the Supplemental Material, I show that this density is given by

$$f_{U|A,\mathrm{ni}}(u|a,\mathrm{ni}) = \frac{\tau e^{-\tau u} + xe^{-(x+\tau)a}e^{xu}}{1-e^{-(x+\tau)a}}. \quad \text{(A-21)}$$

From (A-20) and (A-21), one can show that

$$E[\Delta|a,n,\mathrm{ni}] = 1 + I\times\frac{\frac{1}{\tau}\big(1-e^{-\tau a}\big) - \frac{1}{x}e^{-\tau a}\big(1-e^{-xa}\big)}{1-e^{-(x+\tau)a}}. \quad \text{(A-22)}$$

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

$$E[\Delta|a,n,\mathrm{i}] = 1 + Ia. \quad \text{(A-23)}$$

Solving for $P(B=0|a,n)$. Note first that $P(B=0|a,n) = \frac{P(B=0,a,n)}{P(a,n)}$. I am going to construct $P(B=0,a,n)$. Denote this probability by $Q(n,t)$, where t is the age of the firm.

This probability evolves according to the differential equation

$$\dot{Q}(n,t) = x(n-1)Q(n-1,t) + \tau(n+1)Q(n+1,t) \tag{A-24}$$
$$- n(x+\tau)Q(n,t) + \tau\big(p(n+1,t) - Q(n+1,t)\big), \tag{A-25}$$

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

$$H_Q(z,t) \equiv \sum_{n=0}^{\infty} Q(n,t)z^n. \tag{A-26}$$

Then $\frac{\partial H_Q(z,t)}{\partial z} = \sum_{n=1}^{\infty} nQ(n,t)z^{n-1}$ and $\frac{\partial H_Q(z,t)}{\partial t} = \dot{Q}(0,t) + \sum_{n=1}^{\infty} \dot{Q}(n,t)z^n$. Using (A-24), it follows that

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

where, as in (A-26), I defined $H_P(z,t) \equiv \sum_{n=0}^{\infty} p(n,t)z^n$. Now define

$$\Psi(z,t) \equiv H_P(z,t) - H_Q(z,t). \tag{A-27}$$

Then

$$\dot{\Psi}(z,t) = \big(xz^2 - (x+\tau)z + \tau\big)\frac{\partial \Psi(z,t)}{\partial z} - \frac{\tau}{z}\Psi(z,t), \tag{A-28}$$

where $\dot{H}_P(z,t) = (xz^2 - (x+\tau)z + \tau)\frac{\partial H_P(z,t)}{\partial z}$ follows the derivations in Klette and Kortum (2004). To solve for $\Psi(z,t)$, I need an initial condition. As every firm enters with a single product, $p(1,t) = 1$ and $p(n,t) = 0$ for $n \neq 1$. Similarly, $Q(n,0) = 0$ for all n. This implies that

$$\Psi(z,0) = \sum_{n=0}^{\infty} p(n,0)z^n - \sum_{n=0}^{\infty} Q(n,0)z^n = z, \tag{A-29}$$

which is the required initial condition. The solution to (A-28) with the initial condition in (A-29) is given by (see the Supplemental Material for the proof)

$$\Psi(z,t) = \frac{(\tau - x)ze^{-\tau t}}{x(z-1)e^{-(\tau-x)t} - (xz-\tau)}. \tag{A-30}$$

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

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

Equations (A-27) and (A-26) therefore imply that

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

From the definition of H_Q in (A-26), I can recover $Q(n,t)$ as the coefficients of the Taylor approximation around $z=0$. In the Supplemental Material, I show that

$$Q(n,t)=\left(1-\frac{\tau e^{-xt}-xe^{-\tau t}}{\tau-x}\right)\times p(n,t), \tag{A-31}$$

where $p(n,t)$ is described by $p(0,t)=\frac{\tau}{x}\gamma(t)$, $p(1,t)=(1-\gamma(t))(1-p(0,t))$ and $p(n,t)=\gamma(t)^{n-1}p(1,t)$ and the function $\gamma(t)$ is given in Proposition 3. Equation (A-31) has the important implication that the *conditional* probability of not having an initial product at time t is independent of n, that is,

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

Hence,

$$1-P(B=0|a,n)=\frac{\tau e^{-xa}-xe^{-\tau a}}{\tau-x}. \tag{A-32}$$

Note that $P(\text{not initial}|0,n)=0$ and $\lim_{t\to\infty}P(\text{not initial}|t,n)=1$ as required. Substituting (A-22), (A-23), and (A-32) into (A-18) yields

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

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

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

where

$$\begin{aligned}E[a_P|a_f]={}&\frac{\frac{1}{\tau}\big(1-e^{-\tau a}\big)-\frac{1}{x}e^{-\tau a}\big(1-e^{-xa}\big)}{1-e^{-(x+\tau)a}}+\left(a-\frac{\frac{1}{\tau}\big(1-e^{-\tau a}\big)-\frac{1}{x}e^{-\tau a}\big(1-e^{-xa}\big)}{1-e^{-(x+\tau)a}}\right)\\&\times\left(\frac{\tau e^{-xa}-xe^{-\tau a}}{x\big(1-e^{-(\tau-x)a}\big)}\right)\ln\left(\frac{\tau-xe^{-(\tau-x)a}}{\tau-x}\right).\end{aligned}$$

This is the required expression in (12).

A-1.5. *Proofs for Section 2.6*

Consider the distribution of firms across the number of products they produce. Let $\omega(n) \equiv F \times \tilde{\omega}(n)$, where $\tilde{\omega}(n)$ denotes the share of firms producing n products, that is, $\sum_{n=1}^{\infty} \tilde{\omega}(n) = 1$. As shown in Klette and Kortum (2004),

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

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

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

The entry rate is therefore given by

$$\text{Entry rate} = \frac{z}{F} = \frac{z}{\frac{z}{x} \times \ln\left(\frac{z+x}{z}\right)} = \frac{x}{\ln\left(\frac{z+x}{z}\right)}. \tag{A-33}$$

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

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

To derive the employment life-cycle, consider first the distribution of sales conditional on age. Note that $E[\ln l|a] = E[\ln(\frac{nY}{w\mu_f})|a] = \ln(\frac{Y}{w}) + E[\ln n|a] - E[\ln \mu_f|a]$. To calculate $E[\ln n|a]$, note that the distribution of n conditional on age is given by $f_{N|A}(n|a) = \gamma(a)^{n-1}(1-\gamma(a))$. Hence,

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

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

A-1.6. *The Model With Stochastic Step Size (Proposition 4)*

In this section, I derive the main results for the stochastic step size model. The detailed derivations are contained in the Supplemental Material. Suppose that, conditional on an innovation, the step size of the quality increase is stochastic. Let the probability of climbing k rungs of the ladder be p_k with $\sum_{k=1}^{\infty} p_k = 1$.

The Value Function and Equilibrium Conditions

As I show in the Supplemental Material, the value function is still given by

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

where

$$V_t^M(\Delta) = \frac{\pi_t(\Delta) - \pi_t(1) + (\zeta - 1)\lambda^{-\Delta}\dfrac{1}{\varphi_I} I^\zeta w_t}{\rho + \tau}$$

and $V_t^P(n) = v_t n$, where

$$v_t = \frac{\left(1 - \dfrac{1}{\lambda}\right) Y_t + (\zeta - 1)\dfrac{1}{\varphi_x} x^\zeta w_t}{\rho + \tau}.$$

The optimal innovation and expansion rates are given by

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

$$x = \left(\frac{\varphi_x}{\varphi_z} \frac{1}{\zeta} \right)^{1/(\zeta-1)}.$$

The free-entry condition is given by

$$\frac{1}{\varphi_z} = \sum_{j=1}^{\infty} \left(\frac{V_t^P(1) + V_t^M(j)}{w_t} \right) p_j = \frac{1}{\rho + \tau} \left(\frac{Y_t}{w_t} + \frac{\zeta - 1}{\varphi_x} x^\zeta + \left(\frac{\zeta - 1}{\varphi_I} I^\zeta - \frac{Y_t}{w_t} \right) E\big[\lambda^{-\Delta}\big] \right).$$

Together with the labor market condition, these equations fully determine the equilibrium.

The Distribution of Markups

The distribution of quality gaps $\nu(\Delta)$ solves the set of differential equations

$$\dot{\nu}_t(\Delta) = \begin{cases} -(\tau + I)\nu_t(\Delta) + I \displaystyle\sum_{j=1}^{\Delta-1} \nu_t(\Delta - j) p_j + \tau p_\Delta & \text{if } \Delta \geq 2, \\ \tau\big(p_1 - \nu_t(1)\big) - \nu_t(1) I & \text{if } \Delta = 1. \end{cases}$$

The stationary distribution is therefore given by

$$\nu(j) = \frac{1}{1 + \vartheta} \left(\sum_{m=1}^{j-1} \nu(m) p_{j-m} \right) + \frac{\vartheta}{1 + \vartheta} p_j. \tag{A-34}$$

Define the c.d.f. of quality gaps and hence markups as $\Phi(k) = \sum_{j=1}^{k} \nu(j)$. I now show that

$$\vartheta_H > \vartheta_L \to \Phi(k; \vartheta_H) > \Phi(k; \vartheta_L) \quad \text{for all } k,$$

that is an increase in ϑ reduces the distribution of markups in a first-order stochastic dominance sense. To see this, define $\alpha = \frac{\vartheta}{1+\vartheta}$. α is increasing in ϑ. Write (A-34) as

$$\nu(j) = (1-\alpha)\left(\sum_{m=1}^{j-1} \nu(m) p_{j-m}\right) + \alpha p_j.$$

The c.d.f. Φ can be written as

$$\begin{aligned}
\Phi(k) = \sum_{j=1}^{k} \nu(j) &= (1-\alpha)\sum_{j=1}^{k}\sum_{m=1}^{j-1} \nu(m) p_{j-m} + \alpha \sum_{j=1}^{k} p_j \\
&= (1-\alpha)\sum_{m=1}^{k-1} p_{k-m}\Phi(m) + \alpha \sum_{j=1}^{k} p_j.
\end{aligned}$$

Let $\Phi(k; \alpha)$ denote the c.d.f. as a function of α. Then

$$\begin{aligned}
&\Phi(k; \alpha_H) - \Phi(k; \alpha_L) \\
&\quad = (1-\alpha_H)\sum_{m=1}^{k-1} p_m \Phi(k-m; \alpha_H) + \alpha_H \sum_{j=1}^{k} p_j \\
&\qquad - (1-\alpha_L)\sum_{m=1}^{k-1} p_m \Phi(k-m; \alpha_L) - \alpha_L \sum_{j=1}^{k} p_j \\
&\quad = (1-\alpha_H)\sum_{m=1}^{k-1} p_{k-m}\big[\Phi(m; \alpha_H) - \Phi(m; \alpha_L)\big] \\
&\qquad + (\alpha_H - \alpha_L)\left[p_k + \sum_{j=1}^{k-1} p_j\big(1 - \Phi(k-j; \alpha_L)\big)\right].
\end{aligned}$$

Now note that $\Phi(1; \alpha_H) - \Phi(1; \alpha_L) = (\alpha_H - \alpha_L)p_1 > 0$. Furthermore, this implies that

$$\Phi(m; \alpha_H) - \Phi(m; \alpha_L) > 0 \quad \text{for all } m < j \to \Phi(j; \alpha_H) - \Phi(j; \alpha_L) > 0$$

as $1 - \Phi(k-j; \alpha_L) > 0$ by Φ being a c.d.f. This shows that $\Phi(m; \alpha_H) - \Phi(m; \alpha_L) > 0$ for all m.

The Case of $p_n = \frac{1-\kappa}{\kappa} \times \kappa^n$

Suppose the step size is drawn from $p_n = \frac{1-\kappa}{\kappa} \times \kappa^n$. I now show that the distribution of markups is again a Pareto distribution. The density ν_j solves the equation

$$\nu_j = (1-\alpha)\left(\sum_{m=1}^{j-1} \nu_m \frac{1-\kappa}{\kappa}\kappa^{j-m}\right) + \alpha\frac{1-\kappa}{\kappa}\kappa^j,$$

where $\alpha = \frac{\vartheta}{1+\vartheta}$. Conjecture that $\nu_j = A^{j-1}\nu_1$ for $j \geq 2$. Substituting above yields

$$\nu_j = (1-\alpha)\left(\nu_1 \frac{1-\kappa}{\kappa}\kappa^{j-1} + \sum_{m=2}^{j-1} A^{m-1}\nu_1 \frac{1-\kappa}{\kappa}\kappa^{j-m}\right) + \alpha\frac{1-\kappa}{\kappa}\kappa^j,$$

$$A^{j-1}\nu_1 = \left[(1-\alpha)\frac{1-\kappa}{\kappa}\kappa^{j-1}\left(1+\sum_{m=1}^{j-2}\left(\frac{A}{\kappa}\right)^m\right) + \kappa^{j-1}\right]\nu_1.$$

It is easy to show that $A = 1-(1-\kappa)\alpha = \frac{1+\kappa\vartheta}{1+\vartheta}$ solves this equation. Note that $\nu_1 = \frac{\vartheta}{1+\vartheta}p_1 = \frac{(1-\kappa)\vartheta}{1+\vartheta}$. Hence,

$$\nu_j = \left(\frac{1+\kappa\vartheta}{1+\vartheta}\right)^j \frac{\vartheta(1-\kappa)}{1+\kappa\vartheta}.$$

The corresponding c.d.f. is given by

$$\Phi(k) = \sum_{m=1}^{k}\nu_m = \nu_1\sum_{m=1}^{k}A^{m-1} = \nu_1\sum_{m=0}^{k-1}A^m = \nu_1\frac{1-A^k}{1-A} = 1-\left(\frac{1+\kappa\vartheta}{1+\vartheta}\right)^k.$$

Hence, $P[\Delta \leq d] = 1 - e^{k\times\ln(\frac{1+\kappa\vartheta}{1+\vartheta})}$. The distribution of markups is given by

$$P[\mu \leq m] = P[\lambda^\Delta \leq m] = P\left[\Delta \leq \frac{\ln m}{\ln\lambda}\right] = 1 - e^{\frac{\ln m}{\ln\lambda}\times\ln(\frac{1+\kappa\vartheta}{1+\vartheta})} = 1 - m^{-\frac{1}{\ln\lambda}\ln(\frac{1+\vartheta}{1+\kappa\vartheta})}.$$

Hence, the distribution is again Pareto with shape parameter $\theta(\kappa) = \frac{1}{\ln\lambda}\ln(\frac{1+\vartheta}{1+\kappa\vartheta})$. Because all aggregate wedges are expressed in terms of the Pareto tail, all other results apply directly. To derive the expression for the aggregate growth rate $g = \frac{1}{1-\kappa}(I+\tau)\ln\lambda$, note that $g = (I+\tau)\ln\lambda(\sum_{n=1}^{\infty} np_n)$ and $\sum_{n=1}^{\infty} np_n = \sum_{n=1}^{\infty} n\frac{1-\kappa}{\kappa}\times\kappa^n = \frac{1}{1-\kappa}$.

A-1.7. *The Model With CES Preferences (Proposition 5)*

In this section, I prove the main results for the model with CES preferences. For detailed derivations, I refer to the Supplemental Material. The static allocations can be derived by standard arguments. The dynamic environment for own-innovation, entry, and incumbent creative destruction is the same as in the baseline model. The only difference with respect to the baseline is that I assume that a fraction $(1-\delta)$ of creative destruction activities result in a "reset" of the quality of the destroyed product to the level λQ_t. This change is necessary to make the productivity distribution stationary. I discuss this in more detail in the Supplemental Material. All the aggregate implications are independent of the parameter δ. The need for a stationary distribution of quality only arises when taking the model to the data. For continuity with the baseline model, I still assume that the quality gap Δ after such a reset is equal to unity.

The payoff-relevant state variable for a firm producing n products is given by $[\Delta_i, q_i]$. The value function $V_t([\Delta_i, q_i])$ therefore solves the HJB equation

$$
\begin{aligned}
& r_t V_t\big([\Delta_i, q_i]_i\big) - \dot{V}_t\big([\Delta_i, q_i]_i\big) && \text{(A-35)} \\
& \quad = \sum_{i=1}^{n} \pi_t\big([\Delta_i, q_i]_i\big) + \sum_{i=1}^{n} \tau_t\big[V_t\big([\Delta_j, q_j]_{j\neq i}\big) - V\big([\Delta_i, q_i]_i\big)\big] \\
& \quad + \max_{[I_i]_{i=1}^n}\left\{\sum_{i=1}^{n} I_i\big[V_t\big([\Delta_j, q_j]_{j\neq i}[\Delta_i+1, \lambda q_i]\big) - V_t\big([\Delta_i, q_i]_i\big)\big] - \sum_{i=1}^{n} c^I(I_i, q_i) w_t\right\} \\
& \quad + \max_X\left\{X\left[\delta \int_q V_t\big([\Delta_i, q_i]_i, 1, \lambda q\big)\, dF_t(q) + (1-\delta) V_t\big([\Delta_i, q_i]_i, 1, \lambda Q_t\big) - V_t\big([\Delta_i, q_i]_i\big)\right]\right. \\
& \quad \left. - c^X(X, n) w_t\right\}. && \text{(A-36)}
\end{aligned}
$$

As before, I continue to assume each worker employed in entry activities generates a flow of φ_z of marketable ideas. For symmetry, a fraction δ of such ideas improve the existing quality of a randomly selected product by a step size λ, and a fraction $1-\delta$ "reset" the productivity to λQ_t. The free-entry condition is therefore given by

$$
\frac{1}{\varphi_z} w_t = \delta \int_q V_t(1, \lambda q)\, dF_t(q) + (1-\delta) V_t(1, \lambda Q_t). \tag{A-37}
$$

Suppose that $c^X(X, n)$ is as in the baseline model and that $c^I(I, q)$ is given by

$$
c_t^I(I; \Delta, q) = \frac{1}{\varphi_I}\left(\frac{q}{Q_t}\right)^{\sigma-1} I^{\zeta}. \tag{A-38}
$$

In the Supplemental Material, I show that the value function $V_t([\Delta_i, q_i]_{i=1}^n)$ is given by

$$
V_t\big([\Delta_i, q_i]_{i=1}^n\big) = \sum_{i=1}^{n} \frac{\psi(\Delta_i)}{\rho+\tau+(\sigma-1)g}\left(\frac{q_i}{Q_t}\right)^{\sigma-1} \frac{Y_t}{E\big[\mu^{1-\sigma}\big]} + \frac{1}{\rho+\tau} \frac{\zeta-1}{\varphi_x}\left(\frac{\varphi_x}{\varphi_z}\frac{1}{\zeta}\right)^{\frac{\zeta}{\zeta-1}} w_t n,
$$

where $\psi(\Delta_i)$ is implicitly defined and depends only on Δ (and general-equilibrium variables). I also show that the optimal innovation rate is given by

$$
I(\Delta) = \left[\big(\psi(\Delta) - \alpha(\Delta)\big) \frac{1}{(\zeta-1)} \frac{1}{\frac{1}{\varphi_I}}\left(\frac{L^P}{E\big[\mu^{-\sigma}\big]}\right)\right]^{\frac{1}{\zeta-1}},
$$

where $\alpha(\Delta) = \big(1 - \frac{1}{\min\{\frac{\sigma}{\sigma-1}, \lambda^{\Delta}\}}\big) \min\{\frac{\sigma}{\sigma-1}, \lambda^{\Delta}\}^{1-\sigma}$. Hence, I is independent of q and constant along the BGP.

Let $\nu_t(\Delta)$ be the mass of products with quality gap Δ. This distribution satisfies the differential equation

$$
\dot{\nu}_t(\Delta) = \big(I(\Delta-1)\big)\nu_t(\Delta-1) - \big(\tau + I(\Delta)\big)\nu_t(\Delta) \quad \text{for } \Delta \geq 2.
$$

The law of motion for the mass of products with a quality gap of 1 is given by

$$\dot{\nu}_t(1) = \tau - \big(I(1) + \tau\big)\nu_t(1).$$

Along a BGP, this distribution is stationary and given by

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

Note that if $I(j) = I$, then $\nu(\Delta) = \frac{\tau}{I}(\frac{I}{\tau+I})^{\Delta} = \vartheta(\frac{1}{\vartheta+1})^{\Delta}$ as in the baseline model. Given that markups are a one-to-one function of quality gaps, the distribution of markups is also stationary and only a function of τ and $\{I(\Delta)\}_{\Delta=1}^{\infty}$.

A-2. APPENDIX: EMPIRICAL RESULTS

A-2.1. *Measuring Markups*

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

$$\ln y_{ft} = \phi(l_{ft}, k_{ft}, m_{ft}, z_{ft}) + \varepsilon_{ft}, \tag{A-39}$$

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

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

where va_{ft} is observed value added and $\hat{\varepsilon}_{ft}$ is the residual from (A-39) with va_{ft} as the dependent variable. I treat $\phi(\cdot)$ as a second-order polynomial in all (log) inputs and their interaction terms. For the specification with intermediate inputs instead of labor, the procedure is analogous.

A-2.2. *Robustness for Table II and Additional Results*

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

$$\ln(\mu_{ft}) = \delta_t + \delta_s + \beta \times \text{age}_{ft} + \alpha \times \ln(k_{ft}/l_{ft}) + \psi \times \ln(m_{ft}/l_{ft}) + x_{ft}'\gamma + u_{ft}, \tag{A-40}$$

where m/l is the observed material-labor ratio. As seen in Table A-I, the coefficient on age is always positive and mostly significant unless the dependent variable is the material share and m/l is not controlled for. The reason is that age is strongly correlated with m/l: as firms get older (and larger), they shift their resources from labor to materials.

In Table A-II, I report additional cross-sectional correlates of markups. I consider regressions akin to (20) (for the case of measuring markups by the inverse labor share) and (A-40) (for the case of using the share of material spending) and use a variety of characteristics other than age. To save space, Table A-II reports (for all the different specifications of interest) the coefficients on the respective characteristic. Each columns corresponds to a separate regression. The first four columns correspond to cross-sectional estimates for the entire sample from 1991 to 1997 (i.e., before the crisis). In particular, I confirm the result of De Loecker and Warzynski (2012) that exporters have significantly higher

TABLE A-I

THE LIFE-CYCLE OF MARKUPS: ROBUSTNESS[a]

	Measure for μ: Labor Share in Value Added, Uncorrected		
Age	0.00901	0.00454	0.00451
	(0.00130)	(0.00144)	(0.00157)
N	76,076	57,281	44,024
	Measure for μ: Material Share in Sales, Corrected		
Age	0.00715	0.00567	0.00531
	(0.000411)	(0.000313)	(0.000335)
N	55,230	55,230	42,442
	Measure for μ: Material Share in Sales, Uncorrected		
Age	0.00196	0.00120	0.000781
	(0.000718)	(0.000787)	(0.000850)
N	72,227	55,230	42,442
	Measure for μ: Labor Share in Value Added, Corrected, Controlling for $\ln(m/l)$		
Age	0.00690	0.00596	0.00614
	(0.00109)	(0.00108)	(0.00116)
N	55,212	55,212	42,434
	Measure for μ: Material Share in Sales, Corrected, Not Controlling for $\ln(m/l)$		
Age	−0.00120	−0.00146	−0.00171
	(0.000779)	(0.000779)	(0.000833)
N	55,230	55,230	42,442
	Measure for μ: Material Share in Sales, Uncorrected, Not Controlling for $\ln(m/l)$		
Age	−0.00569	−0.00568	−0.00611
	(0.000897)	(0.00104)	(0.00112)
N	72,227	55,230	42,442
Industry FE	✓	✓	✓
$\ln(k/l)$		✓	✓

[a]Robust standard errors in parentheses. I focus on the unbalanced sample of firms, who enter the market after 1990. I use the data from 1991 to 2000. All specifications include year fixed effects and industry fixed effects at the 5-digit level. Columns 2 and 3 control for $\ln(k/l)$. In column 3, I focus on the balanced panel, that is, only consider firms that survive to the end of my sample period.

TABLE A-II

DETERMINANTS OF LABOR PRODUCTIVITY: IMPERFECT INPUT MARKETS AND BORROWING CONSTRAINTS[a]

	Cross-Sectional Markup Heterogeneity: $\ln \mu_{ft} = \delta_t + \delta_s + \beta \times \text{FirmCharacteristic}_{ft} + \alpha \times X_{ft} + u_{ft}$						
	Exporter	FDI	Capital Market	Foreign Loan	Contrained?	Capital Constrained?	Financial Growth Barrier?
	Measure for μ: Labor Share in Value Added, Corrected						
β	0.088	0.012	0.050	0.058	−0.016	−0.011	−0.036
	(0.004)	(0.012)	(0.027)	(0.011)	(0.008)	(0.010)	(0.017)
N	134,126	134,126	134,126	134,126	16,375	16,375	16,375
R^2	0.33	0.33	0.33	0.33	0.34	0.34	0.34
	Measure for μ: Labor Share in Value Added, Uncorrected						
β	0.167	0.174	0.065	0.175	−0.055	−0.024	−0.020
	(0.005)	(0.017)	(0.033)	(0.014)	(0.009)	(0.012)	(0.023)
N	138,953	138,953	138,953	138,953	16,375	16,375	16,375
R^2	0.18	0.17	0.17	0.17	0.20	0.19	0.19
	Measure for μ: Material Share in Sales, Uncorrected						
β	0.084	0.101	0.008	0.091	−0.020	−0.015	−0.027
	(0.001)	(0.003)	(0.008)	(0.003)	(0.002)	(0.003)	(0.005)
N	134,235	134,235	134,235	134,235	16,375	16,375	16,375
R^2	0.89	0.88	0.88	0.88	0.89	0.89	0.89
	Measure for μ: Material Share in Sales, Corrected						
β	0.108	0.198	0.027	0.150	−0.037	−0.013	−0.008
	(0.003)	(0.009)	(0.020)	(0.007)	(0.005)	(0.007)	(0.012)
N	134,235	134,235	134,235	134,235	16,375	16,375	16,375
R^2	0.61	0.60	0.60	0.60	0.62	0.62	0.62

[a]Robust standard errors are shown in parentheses. The table reports the results of regressing log markups on various firm characteristics. Each column represents a separate regression and I report the coefficient on the respective firm characteristic. All regressions include a full set of 5-digit product fixed effects, a set of year fixed effects, and $\ln(\frac{k}{l})$, that is, the firm's (log) capital-labor ratio. Capital is measured as the total value of assets reported in the industrial census. The first two panels measure markups from the inverse labor share in value added. The last two panels measure markups from the inverse material share in sales. In the last two panels, I also control for the material-labor ratio $\ln(m/l)$ as in (A-40). "Exporter," "FDI," "Capital market," and "Foreign loans" are dummy variables indicating whether the firm exports and finances its investment through FDI, foreign loans, or funds from the Indonesian capital market. The last three columns use the special census supplement of the census in 1996. The survey asked whether firms were facing any constraints, whether this constraint was related to missing capital, and whether missing capital was the main obstacle in firm expansion.

markups.[3] I also explore the information in firms' financial balance sheets and show that firms who receive some part of their capital from FDI (column 2), the capital market (column 3), and from foreign loans (column 4) have higher markups. These patterns are consistent with these firms being technologically more advanced and hence having more market power. In contrast, it is harder to reconcile with markups reflecting binding financial constraints as these firms are arguably less constrained and hence should have low marginal products.

In the last three columns, I focus on the existence of constraints more directly. In 1996, the census contained direct questions on the type of constraints different firms are facing. Columns 5–7 therefore estimate the relative markup of firms who report that they face a constraint they could not overcome (column 5), that this constraint is particularly re-

[3]The results reported in Table A-II are also quantitatively consistent with De Loecker and Warzynski (2012). When they measured markups from the labor elasticity and corrected the output data for measurement error, they estimated an exporter premium of 0.078.

lated to capital (column 6), and that they did not plan to expand because of a scarcity of capital (column 7). While some coefficients are not statistically significant depending on the specification, Table A-II shows that these firms are estimated to have *lower* marginal products. This is consistent with these firms being low productivity producers who post low markups, but is harder to reconcile with a model of financial constraints.

A-2.3. *The Employment Life-Cycle in Indonesia*

In the calibration, I use the employment life-cycle in Indonesia as a moment. Focusing on the unbalanced panel of firms entering the economy after 1991, firms grow by about 8% a year. For the calibration, I therefore use the estimated profile depicted in Figure 4 and target the log difference in employment for 7-year-old firms. In Table A-III, I report additional regression results predicting firm employment from firm age. The specification is exactly the same as (20), except that I do not control for firms' capital intensity. Column 3 shows that entrants and exiting firms are substantially smaller than the average firm. Column 4 shows that entrants are not smaller as predicted by their age (in fact, they are slightly bigger) but that exiting firms are much smaller. This is exactly what the model predicts, because exiting firms are selected on n conditional on age, while entrants are not. Column 5 shows that the estimated age profile is quite similar once I condition on survival until the end of the sample.

A-2.4. *The Markup-Size Relationship*

To derive the relationship between firm-level markups and size, note that

$$E\big[\ln(\mu_f)|N=n\big]=\ln(\lambda)\times E_a\left[E\left[\frac{1}{n}\sum_{j=1}^{n}\Delta_j\middle|\text{Age}=a,N=n\right]\middle|N=n\right].$$

TABLE A-III

THE EMPLOYMENT LIFE-CYCLE IN INDONESIA[a]

	Log Employment				
Age	0.0930	0.0814		0.0923	0.0811
	(0.00193)	(0.00178)		(0.00248)	(0.00192)
Entry			−0.187	0.0466	
			(0.00821)	(0.0101)	
Exit			−0.359	−0.337	
			(0.0106)	(0.0105)	
Industry FE		✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Balanced panel					✓
N	76,106	76,106	64,958	64,958	59,602
R^2	0.035	0.218	0.212	0.231	0.239

[a] Robust standard errors in parentheses. I focus on the unbalanced panel of firms, who enter the market after 1990. I use the data from 1991 to 2000. All specifications include year fixed effects and industry fixed effects at the 5-digit level. 'Entry" and "Exit" are indicator variables for whether the firm enters (exit) the market in a given year. In column 5, I focus on the balanced panel, that is, only consider firms that survive to the end of my sample period.

Above, I already showed that

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

where $E[\Delta_j|a,n,\text{ni}]$ is given in (A-22), $E[\Delta_j|a,n,\text{i}]$ is given in (A-23), and $1-P(B=0|a,n)$ is given in (A-32). In particular, none of these objects depend on n. Hence,

$$\begin{aligned}E[\ln\mu_f|n]&=\ln(\lambda)\int_a\left[E[\Delta|a,\text{ni}]+\frac{\big(1-P(B=0|a)\big)}{n}\big(E[\Delta|a,\text{i}]-E[\Delta|a,\text{ni}]\big)\right]f(a|n)\,da\\&=\ln(\lambda)\left(1+I\int_a\left[g(a,x,\tau)+\frac{\big(1-P(B=0|a)\big)}{n}\big(a-g(a,x,\tau)\big)\right]f(a|n)\,da\right),\end{aligned}$$

where

$$g(a,x,\tau)=\frac{\frac{1}{\tau}\big(1-e^{-\tau a}\big)-\frac{1}{x}e^{-\tau a}\big(1-\exp(-xa)\big)}{1-\exp\big(-(x+\tau)a\big)},$$

and $f(a|n)$ is the distribution of age conditional on size, which is given by

$$f(a|n)=\frac{\big(1-\gamma(a)\big)\gamma(a)^{n-1}\left(1-\frac{\tau}{x}\gamma(a)\right)}{\frac{1}{x}\frac{1}{n}\left(\frac{x}{\tau}\right)^{n}},\tag{A-41}$$

where $\gamma(a)=\gamma(a)=\frac{x(1-e^{-(\tau-x)a})}{\tau-xe^{-(\tau-x)a}}$.[4]

The expressions above fully determine the average log markup as a function of n. In Figure A-1, I show the results for the calibrated model. The left panel shows the average markup as a function of size, the right panel shows the stationary firm size distribution. Average markup is increasing in size—at least for the vast majority of firms. The very top part of the sales distribution where markups are declining in size is of course closely

[4]To derive (A-41), note first that the mass of firms with n products is given by

$$M_n=\frac{\theta}{n}\left(\frac{1}{1+\theta}\right)^n=\frac{\frac{z}{x}}{n}\left(\frac{1}{1+\frac{z}{x}}\right)^n=\frac{z}{x}\frac{1}{n}\left(\frac{x}{\tau}\right)^n.$$

The probability of having n products at time t when born at time $t-a$ is given by $p_n(a)$, where

$$p_n(a)=\big(1-\gamma(a)\big)\gamma(a)^{n-1}\left(1-\frac{\tau}{x}\gamma(a)\right).$$

Each period, z firms enter. Hence, $M_n=\int_{a=0}^{\infty}zp_n(a)\,da$. Then conditional distribution is therefore given by

$$f(a|n)=\frac{zp_n(a)}{M_n}=\frac{\big(1-\gamma(a)\big)\gamma(a)^{n-1}\left(1-\frac{\tau}{x}\gamma(a)\right)}{\frac{1}{x}\frac{1}{n}\left(\frac{x}{\tau}\right)^n},$$

which is the expression in (A-41).

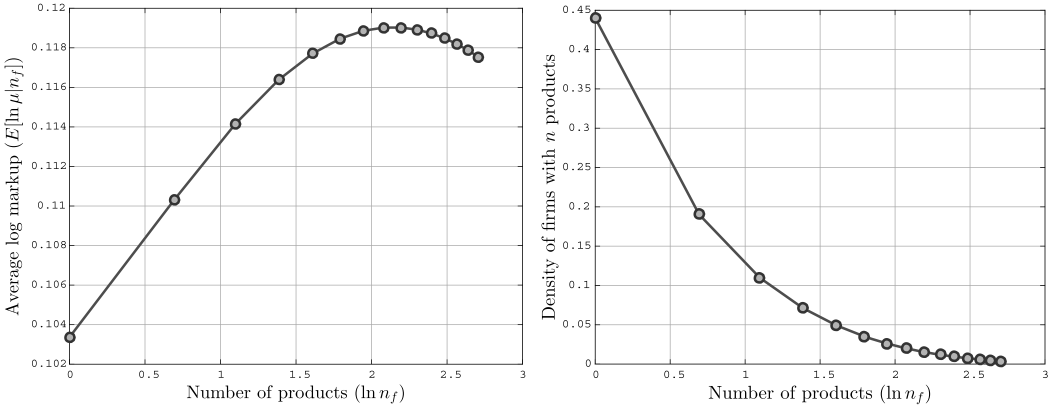

FIGURE A-1.—MARKUPS BY SIZE. *Notes*: In the left panel, I depict the average (log) markup as a function of $\ln n$. In the right panel, I depict the stationary distribution of sales. The endogenous flow rates (I, x, z) and the parameter λ correspond to the calibration of the baseline model.

related to the top of the age distribution where markups are declining in age; see Figure 1. Figure A-1 also shows that the quantitative effect of firm size on the average markup is very small. Increasing size by one log point (say from 1 to 2) increases the average markup by 0.6%. This is the elasticity reported in the main text. The firm-level data show a stronger relationship between markups and size. In Table A-IV, I report the results of regressions of log markups (columns 1 and 2) and log TFPR (columns 3 and 4) on log sales. The estimated elasticity is consistently estimated to be around 0.2, that is, much larger than in the model.

A-2.5. *Robustness of the Calibration Results*

Sensitivity With Respect to ζ. In Figure A-2, I study the sensitivity of the results with respect to the elasticity ζ. More specifically, I consider a range of ζ between 1.5 and 3 and redo the analysis around Table VI. I report three moments: the change in the number of firms, the change in the share of small firms, and the change in the growth rate. For the baseline, I assumed a value of $\zeta = 2$. In the left panel of Figure A-2, I report the change

TABLE A-IV

MARKUPS, TFPR AND SIZE[a]

	Dependent Variable			
	Log Markups ($\ln \mu_f$)		Log TFPR ($\ln py/(k^{\alpha}l^{1-\alpha})$)	
Log sales	0.192	0.168	0.163	0.229
	(0.000925)	(0.00109)	(0.00119)	(0.00112)
$\ln k/l$		0.0547		−0.242
		(0.00134)		(0.00137)
N	176,958	138,953	122,578	122,578
R^2	0.306	0.311	0.153	0.354

[a]Robust standard errors in parentheses. All specifications include year fixed effects and 5-digit industry fixed effects. $\ln k/l$ denotes the (log) capital-labor ratio at the firm level.

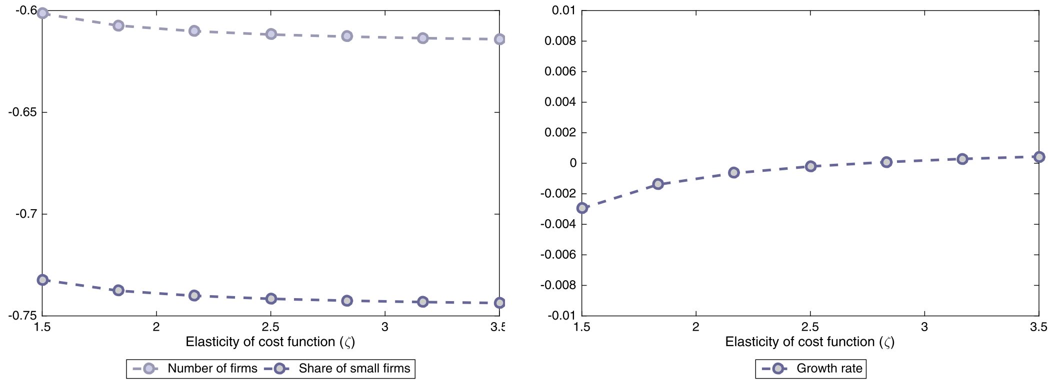

FIGURE A-2.—SENSITIVITY WITH RESPECT TO ζ. *Notes*: The figure displays the change in the number of firms and the share of small firms (left panel) and the change in the growth rate (right panel) for different values of ζ. Specifically, for each ζ, I recalibrate the model to the U.S. and Indonesian moments contained in Table VI and calculate the change in the equilibrium outcomes. The baseline specification is $\zeta = 2$.

in the number of firms and the share of small firms. It is seen that this moment is not substantially affected by the particular value of ζ: as for the baseline case, the number of firms declines by 74%, the share of small firms by 85%. In the right panel, I depict the change in the growth rate. Two results stand out: First of all, depending on the value of ζ, the growth rate can either increase or decrease in response to a reduction of entry and expansion barriers. Second, the implied changes in the growth rate are relatively small. Hence, the result that large changes in the firm size distribution are consistent with an essentially stable distribution of income is robust.

Sensitivity With Respect to the Calibrated Moments. In Figure A-3, I study the sensitivity of these implications with respect to the underlying calibration moments. In the left panel, I depict the change in the number of firms as a function of employment growth for different values of the entry rate. The baseline calibration for the United States assumed an entry rate of 8% and that employment grows by a factor of 2 during the first 10 years of a firm's life-cycle. The effect of the employment life-cycle is quite sizable. If one were to assume that the extent of life-cycle growth in the U.S. was 2.5 instead of 2, the model would predict that the number of firms would fall by almost 90% (relative to around 60–70% in the baseline calibration). The right panel shows the implications for the aggregate growth rate. In contrast to the firm-level outcomes, the model does not predict sizable growth implications. Not only is the effect on the growth rate ambiguous in that, for example, a rate of life-cycle growth of 2.6 instead of 2 would increase U.S. growth relative to Indonesia, but even for extremely large differences in firm-dynamics, the differences in aggregate productivity growth do not exceed 0.2%.

A-2.6. *Expansion and Entry Costs in Indonesia: Empirical Evidence*

The analysis above suggests that differences in expansion costs could be an important determinant of firm size, firm growth, and misallocation. In this last section, I provide suggestive evidence for such frictions. To do so, I exploit regional variation across product markets in Indonesia. As I do not have direct information on the type of barriers different firms might face, I use the theory to suggest an empirical strategy based on the joint patterns of various firm-level outcomes.

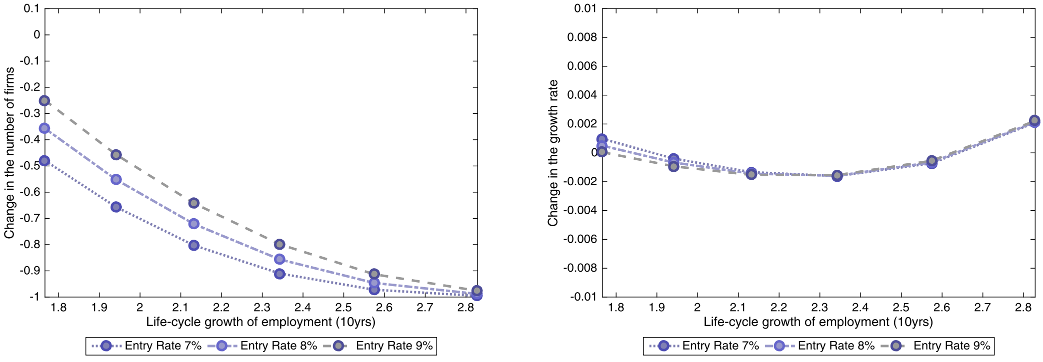

FIGURE A-3.—SENSITIVITY OF RESULTS TO UNDERLYING MOMENTS. *Notes*: This figure displays the model's predictions for the change in the number of firms (left panel) and the equilibrium growth rate (right panel) as a function of the life-cycle employment growth over 10 years and the equilibrium entry rate. The baseline results in Table VI refer to the case of a life-cycle growth of 2 and an entry rate of 8%.

The basic intuition is simple. If different regions in Indonesia differed only in their expansion costs, locations with low frictions should see fewer and bigger firms, lower entry rates, and a steeper schedule of life-cycle employment growth. Additionally, product markets in such regions should be characterized by lower markups. If, in contrast, entry costs were the dominant source of variation, it would also be the case that firm size should be negatively correlated with regional entry rates and positively correlated with the slope of life-cycle growth. However, now large firms should reside in regions with high entry costs and one would expect a positive correlation between firm size and the prevailing markups.

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

$$y_{rst} = \delta_s + \delta_t + \beta \times \text{AvgSize}_{rst} + \gamma \times \ln(\text{pop}_r) + \alpha \times Ag_r + u_{rst},$$

where δ_s and δ_t are industry and time fixed effects, AvgSize_{rst} is the average size of producers active in region r, in sector s in time t, and y_{rst} are the different outcome variables mentioned above. Moreover, I also control for the size of the population in region r and the regional agricultural share to account for the effects of market size.[5] Given the focus on the regional variation, I cluster all standard errors at the province level, to allow for correlation in the error term across industries within a province.

[5]To measure geographical characteristics, I exploit information from the Village Potential Statistics (PODES) data set in 1996. The PODES data set contains detailed information on all of Indonesia's 65,000 villages. Using the village-level data, I then aggregate this information to the province level and match these to the firm-level data. In particular, I measure the size of the population and the share of the population living in villages, which are predominantly agricultural.

TABLE A-V

FIRM SIZE, ENTRY, AND MARKUPS ACROSS PRODUCT MARKETS IN INDONESIA[a]

	Entry Rate	Exit Rate	LC Empl. Growth	Markups	
				Avg.	q^{90}
Average firm size	−0.019	−0.011	0.132	−0.054	−0.049
	(0.004)	(0.003)	(0.034)	(0.014)	(0.010)
Controls	Local population; Local agricultural share; Industry FE; Year FE				
N	455	463	463	462	462
R^2	0.369	0.320	0.640	0.378	0.369

[a] Standard errors are clustered at the level of a province and contained in parentheses. Regressions are run at the province-industry level, where industries are measured at the 3-digit level. The variables are all measured within these province-industry cells. The entry and exit rates are measured as the share of entering and exiting firms. The employment life-cycle is measured as the growth of cohort employment over the 3-year horizon. Log markups are measured as the residual from a regression of log inverse labor shares on a set of year and 5-digit industry fixed effects, "Avg." is the mean log markup, and q^{90} is the 90% quantile. "Average firm size" is the average log value added within an industry-region-year cell. All regressions contain a full set of industry and year fixed effects and control for the log of the province population and the share of villages within the province, which are agricultural. I only consider province-industry cells with at least 50 observations, and all regressions are weighted using the number of observations within each cell.

My preferred measure of size is firm sales (rather then employment), as the theory predicts that average sales only depend on x/τ. To calculate the employment life-cycle, I again rely on the panel dimension and adopt the same methodology as for Figure 4.[6] To avoid identifying the parameters from sparsely populated region-industry-year cells, I only consider cells which contain at least 50 observations, and I weight the regression by the number of observations in each bin.

Table A-V contains the results. In the first three columns, I show that average firm size in a region is negatively correlated with entry and exit rates and positively correlated with the extent of life-cycle growth. These correlations hold regardless of whether the source of variation across regions stems from entry or expansion costs. Columns 4 and 5 show that average size is negatively correlated with markups as measured by either the average markup or the 90%-quantile. This is consistent with the model if regional firm size is driven by differences in expansion costs, but not consistent with an explanation based on entry costs. While these patterns are consistent with expansion costs potentially playing an important role, they are of course only suggestive. However, they highlight the need to directly measure why the costs of entering new product markets might be systematically related to the level of development and whether these are amenable by policies.

REFERENCES

DE LOECKER, J., AND F. WARZYNSKI (2012): "Markups and Firm-Level Export Status," *American Economic Review*, 102, 2437–2471. [14-16]

KLETTE, T. J., AND S. KORTUM (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112, 986–1018. [7,9]

Co-editor Giovanni L. Violante handled this manuscript.

Manuscript received 3 August, 2017; final version accepted 20 March, 2020; available online 9 April, 2020.

[6] In particular, for each cohort, I calculate employment growth at the 3-year horizon after controlling for a full set of 5-digit product and year fixed effects and then average these growth rates at the industry-province level. I opt for a somewhat short horizon of three years to avoid losing too many observations given that I only have data for the years 1991 to 1998.