1 Overview of the Code

In this Replication Package, we explain how our code works. Essentially, it performs the following three big tasks:

- Calibration: In this part, we search for the vector of parameters that generates moments in the model along a BGP that are consistent with the data. The calibration process is discussed in Section 2 of this Replication Package.
- Computation of the Model: Once the calibrated parameters have been found, we compute the model under these parameters. Moreover, we compute it under the counterfactual population growth rate.
- Transitional Dynamics: Starting from the calibrated BGP, this part of the code uses the equilibrium conditions of the model to compute the transitional dynamics of the economy as it moves to the counterfactual BGP. This part of the code is discussed in Section 3 of this Replication Package.

2 Calibration

As pointed out in the Article, there are 11 parameters for baseline model, which are

$$\Psi = \left\{ \eta, \rho, \sigma, \zeta, \alpha, I, \varphi_e, \lambda, \delta, \varphi_x, \overline{\omega}^{\sigma-1} \right\}.$$

Three of these parameters are set exogenously, which are $\zeta = 2$, $\sigma = 4$, and $\rho = 0.01$. Moreover, the parameter η is computed from the data directly and remains invariant during the calibration process at $\eta = 0.02$. The remaining seven parameters must be calibrated internally and we discuss the process in what follows.

2.1 Preliminaries

Considering the need to calibrate seven parameters, $(\alpha, I, \varphi_e, \lambda, \varphi_x, \overline{\omega}^{\sigma-1}, \delta)$, we target seven moment conditions that the model needs to satisfy under a balanced growth path (BGP). In Table 1, we present all the target moments we consider.

Table 1: Summary of Targeted Moments

Moment	Notation	Target
Aggregate Productivity Growth	g	1.84%
Entry Rate	EntryRate	11.6%
Average Firm Size	AvgFirmSize	19.816
Markup Growth by Age 10	MarkupGrowth	10.2%
Sales Growth by Age 10	SalesGrowth	58%
Average Cost-Weighted Markup	AvgMarkup	20%
Share of Firms with + 10,000 Employees	ShareLarge	26.2%

Given such moments, the calibration algorithm works in the following steps:

1. Pick an arbitrary parameter vector $\theta \equiv (\varphi_e, \lambda, \varphi_x, \overline{\omega}^{\sigma-1}, \delta)$.
2. Given θ, we can find $\alpha(\theta)$ and $I(\theta)$ that ensure that the aggregate productivity growth rate and entry rate match their targets.
3. Based on θ, we compute the remaining five moments in the model.
4. Steps 1 to 3 characterize a function that takes $\theta \in \mathbb{R}^5$ and returns five model-implied moments. We use a solver to find the value of θ that matches these moments to their respective targets.

In the next Subsections, we discuss this algorithm in detail.

2.2 Computing Direct Parameters

Apart from $(\zeta, \eta, \rho, \sigma)$, which are calibrated exogenously, we claim that the parameters (α, I) can be directly calibrated as a function of the other parameters. The idea is that given a set of parameters $\theta = (\varphi_e, \lambda, \varphi_x, \overline{\omega}^{\sigma-1}, \delta)$, the model computes (i) the aggregate productivity growth rate and (ii) the aggregate entry rate as a function of (α, I). We can then solve a system of two equations to find (α, I) that ensure that these two objects match their data counterparts.

2.2.1 Computing Aggregate Productivity Growth

On a BGP, the rate of creative destruction is given by $\tau = \frac{\alpha}{1-\alpha}(\eta + \delta)$ as per Proposition 3 of the Article. Also according to this Proposition, the aggregate productivity growth rate is then computed as

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

where $\overline{q}^{\sigma-1} = \alpha\lambda^{\sigma-1} + (1-\alpha)\overline{\omega}^{\sigma-1}$ as per Equation 5 of the Article.

2.2.2 Computing the Aggregate Entry Rate

In any period t, the number of varieties in the economy is denoted by N_t and the number of firms by F_t. Even though both of these quantities grow on a BGP, their ratio N_t/F_t is stationary and represents the average number of varieties per firm. The aggregate entry rate in the economy is given by $\text{EntryRate}_t = z_t N_t/F_t$, where z_t is the flow entry rate. On a BGP, Proposition 3 of the Article states that $z = \frac{\eta+\delta}{1-\alpha} - x$, where $x = (\frac{\varphi_x}{\zeta\varphi_e})^{\frac{1}{\zeta-1}}$ is the rate of incumbent product creation. As a result, to compute the aggregate entry rate, we only need to characterize the average number of products per firm on a BGP, N/F. For this purpose, consider a firm of age a. In Section 1.1.8 of the Appendix, it is shown that the product distribution for such firm is given by:

$$\begin{aligned} p_a(0) &= \frac{\tau+\delta}{x} \cdot \gamma(a) \\ p_a(1) &= (1 - p_a(0)) \cdot (1 - \gamma(a)) \\ p_a(n) &= p_a(n-1) \cdot \gamma(a), \end{aligned} \tag{2}$$

where

$$\gamma(a) = \frac{x(1 - e^{\psi a})}{x(1 - e^{\psi a}) - \psi} \tag{3}$$

and $\psi \equiv x - (\tau + \delta)$ is the net rate of production accumulation. Therefore, the average number of products for a firm of age a is given by

$$\frac{N}{F}\Big|_{A=a} = \sum_{n=1}^{\infty} n \cdot \frac{p_a(n)}{1 - p_a(0)},$$

where the denominator conditions the probabilities on firm survival. As such, the (unconditional) average number of products per firm can be computed as

$$\frac{N}{F} = \int \frac{N}{F}\Big|_{A=a} \cdot \omega^f(a) \cdot da,$$

where $\omega^f(a)$ is the firm age density characterized in Section 5.2 of Supplementary Material

$$\omega^f(a) = \frac{\psi e^{(\psi-\eta)a}}{\psi - x(1 - e^{\psi a})} \cdot \left(\int \frac{\psi e^{(\psi-\eta)a'}}{\psi - x(1 - e^{\psi a'})} da' \right)^{-1}. \tag{4}$$

By manipulating such integrals and noting that $z = \eta - \psi$, one can show that

$$\frac{N}{F} = \frac{1}{z} \cdot \left(\int \frac{\psi e^{(\psi-\eta)a'}}{\psi - x(1 - e^{\psi a'})} da' \right)^{-1} \tag{5}$$

as long as $z > 0$. That is, if there is no product creation by entrants, then the number of firms cannot increase, but the number of varieties keeps expanding due to innovation by incumbents. As such, the average number of products per firm cannot be stationary. Ultimately, the aggregate entry rate on a BGP can be computed as

$$\text{EntryRate} = \left(\int \frac{\psi e^{(\psi-\eta)a'}}{\psi - x(1-e^{\psi a'})} da' \right)^{-1}. \tag{6}$$

For future reference, we note that the objects introduced in this Section allow us to compute the unconditional firm product distribution as

$$p(n) = \int \frac{p_a(n)}{1 - p_a(0)} \cdot \omega^f(a) \cdot da. \tag{7}$$

2.2.3 Constructing the System of Equations

Given the parameters $\theta \equiv (\varphi_e, \lambda, \varphi_x, \overline{\omega}^{\sigma-1}, \delta)$, we use Equations (1) and (6) to find (α, I) that match productivity growth and entry rate to the data counterparts. More specifically, we would like to find a zero for the system

$$\phi(\alpha, I; \theta) = \begin{bmatrix} \phi_1(\alpha, I; \theta) \\ \phi_2(\alpha, I; \theta) \end{bmatrix} = \begin{bmatrix} g(\alpha, I; \theta) - g_{\text{target}} \\ \text{EntryRate}(\alpha, I; \theta) - \text{EntryRate}_{\text{target}} \end{bmatrix}.$$

The solution to this system consists on $\alpha(\theta)$ and $I(\theta)$. That is, for every initial guess of parameters in the outer nest θ, we can solve for $\alpha(\theta)$ and $I(\theta)$. This method implies that α and I do not need to be included as separate variables in the outer nest of the algorithm, which enhances computational efficiency. In this Section, we argued that for every (outer) parameter vector θ, we can compute α and I to make sure that the (i) aggregate productivity growth rate and (ii) aggregate entry rate match their data counterparts. In the next Section, we discuss how to compute each of the remaining moments for a given θ.

2.3 Average Firm Size

Equation (5) can be used to compute the average number of products per firm on a BGP, denoted by N/F. Define $\mathscr{N}_t \equiv N_t/L_t$ as the economy variety intensity ratio. As discussed in the Article, $\mathscr{N}_t$ is stationary along the BGP and that is also true for its inverse, $\ell_t = \mathscr{N}_t^{-1}$. In this context, the average firm size measured in the number of employees can be simply computed as $\text{AvgFirmSize} = \ell N/F$. To find the inverse of the variety intensity ratio, one needs to use the two equilibrium conditions of the model. Along the BGP, the labor market clearing condition can be written as

$$\ell(1 - s^P) = \frac{1}{\varphi_e} \left(\frac{\eta + \delta}{1 - \alpha} - \frac{\zeta - 1}{\zeta} \cdot x \right), \tag{8}$$

where $s_t^P \equiv L_t^P / L_t$ is the production share of the economy. The second condition is free entry, which along the BGP can be expressed as

$$\frac{1}{\varphi_e} = \left(\alpha k(\lambda) \lambda^{\sigma-1} + (1-\alpha) k(\frac{\sigma}{\sigma-1}) \overline{\omega}^{\sigma-1} \right) \cdot \frac{\ell s^P}{\mathcal{M}^{\sigma-1} \cdot \Lambda^{\sigma}} + \frac{1}{\rho + \tau + \delta} \cdot \frac{\zeta - 1}{\zeta \varphi_e} \cdot x. \tag{9}$$

In the Equation above, $\mathcal{M}$ and Λ are the misallocation and labor wedge terms, respectively. If we know how to compute them, and we know the function $k(\cdot)$, then we can solve Equation (9) for ℓs^P and then replace this in Equation (8) to find ℓ. Finally, with $\ell s^P$and ℓ, we can back out the production share, s^P. As such, we now discuss how to obtain the function $k(\cdot)$ and also $\mathcal{M}$ and Λ.

2.3.1 Computing the function $k(\Delta)$

As per Section 1.3 of the Appendix, the function $k(\Delta)$ satisfies the differential equation

$$k(\Delta) \cdot \mathcal{C} - I\Delta \cdot k'(\Delta) = \frac{\Delta - 1}{\Delta^{\sigma}}, \tag{10}$$

where $\mathcal{C} = \rho - \eta + \frac{\eta + \delta}{1-\alpha} \cdot \overline{q}^{\sigma-1}$. The solution to this differential equation is

$$k(\Delta) = \left[\frac{\Delta}{\mathcal{C} + I(\sigma - 1)} - \frac{1}{\mathcal{C} + I\sigma} \right] \cdot \Delta^{-\sigma} + M\Delta^{\frac{\mathcal{C}}{I}}. \tag{11}$$

To obtain the constant of integration M, the Appendix sets the boundary condition $k(\frac{\sigma}{\sigma-1}) = \frac{1}{\mathcal{C}(\sigma-1)} \cdot \left(\frac{\sigma-1}{\sigma}\right)^{\sigma}$. Therefore,

$$M = \left[\frac{1}{\mathcal{C}(\sigma-1)} - \frac{\sigma}{\mathcal{C}(\sigma-1) + I(\sigma-1)^2} + \frac{1}{\mathcal{C} + I\sigma} \right] \cdot \left(\frac{\sigma}{\sigma - 1} \right)^{-\sigma - \frac{\mathcal{C}}{I}}. \tag{12}$$

2.3.2 Computing the Misallocation terms $\mathcal{M}$ and Λ

As shown in Section 3 of the Article, the misallocation terms $\mathcal{M}_t$ and Λ_t are given by

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

and

$$\Lambda_t = \frac{\int \mu^{-\sigma} \cdot (q/Q_t)^{\sigma-1} \cdot dF_t(q,\mu)}{\int \mu^{1-\sigma} \cdot (q/Q_t)^{\sigma-1} \cdot dF_t(q,\mu)}.$$

The key object for their computation is the joint distribution of efficiency and markups, $F_t(q,\mu)$. As outlined in Section 7.5 of the Supplementary Material, it is computationally useful to write

$$F_t(q,\mu) = F_t^C(q,\mu)\cdot\frac{N_t^C}{N_t} + F_t^{NC}(q,\mu)\cdot\frac{N_t^{NC}}{N_t} \equiv \widehat{F}_t^C(q,\mu) + \widehat{F}_t^{NC}(q,\mu).$$

That is, we breakdown the distribution among varieties that have a competitor (C) and varieties that do not have a competitor (NC). Finally, we perform three more transformations on these distributions. First, varieties without competitors always charge the CES markup and, therefore, the distribution $\widehat{F}_t^{NC}(q,\mu)$ is degenerate at $\mu = \frac{\sigma}{\sigma-1}$ and we can effectively supress μ as an argument of this function. Second, the markup charged by firms with competitors is a deterministic function of their productivity gap, as $\mu(\Delta) = \min\left\{\frac{\sigma}{\sigma-1}, \Delta\right\}$. As such, we can substitute μ for Δ as an argument of $F_t^C(q,\mu)$. Third, we express both distributions in terms of $\widehat{q} \equiv \log(q/Q)$, as this eliminates the need of keep track of the distinction between q_t and Q_t. As a result, we express the misallocation terms as

$$\mathcal{M}_t = \frac{\left(\int \mu(\Delta)^{1-\sigma}\cdot e^{\widehat{q}(\sigma-1)}\cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\int e^{\widehat{q}(\sigma-1)}\cdot d\widehat{F}_t^{NC}(\widehat{q})\right)^{\frac{\sigma}{\sigma-1}}}{\int \mu(\Delta)^{-\sigma}\cdot e^{\widehat{q}(\sigma-1)}\cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}\int e^{\widehat{q}(\sigma-1)}\cdot d\widehat{F}_t^{NC}(\widehat{q})} \tag{13}$$

and

$$\Lambda_t = \frac{\int \mu(\Delta)^{-\sigma}\cdot e^{\widehat{q}(\sigma-1)}\cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}\int e^{\widehat{q}(\sigma-1)}\cdot d\widehat{F}_t^{NC}(\widehat{q})}{\int \mu(\Delta)^{1-\sigma}\cdot e^{\widehat{q}(\sigma-1)}\cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\int e^{\widehat{q}(\sigma-1)}\cdot d\widehat{F}_t^{NC}(\widehat{q})}. \tag{14}$$

We now compute the distributions $\widehat{F}_t^C(\Delta,\widehat{q})$ and $\widehat{F}_t^{NC}(\widehat{q})$.

Distribution without Competitors Starting with varieties without competitors, the distribution $\widehat{F}_t^{NC}(\widehat{q})$ satisfies the differential equation[1]

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

where $g_t^Q = I + \frac{\tau_t}{\alpha}\cdot\frac{\overline{q}^{\sigma-1}-1}{\sigma-1}$. To solve for this distribution on a BGP, we use the fact that this distribution must be stationary. Setting the time derivative to zero and manipulating, we have the following differential equation

$$\frac{\partial\widehat{F}^{NC}(\widehat{q})}{\partial\widehat{q}} = d\cdot\widehat{F}^{NC}(\widehat{q}) - d(1-\alpha)\cdot\Gamma\left(\exp\left(\widehat{q}\right)\right), \tag{15}$$

where $d \equiv (\sigma-1)/(\overline{q}^{\sigma-1}-1)$. The first step towards solving this differential equation is to characterize the distribution function Γ. We assume this distribution is Pareto with scale parameter ω_m and tail parameter $\varrho = 15$. By the definition of $\overline{\omega}$, we have

$$\overline{\omega}^{\sigma-1} \equiv \mathbb{E}(\omega^{\sigma-1}) = \int \omega^{\sigma-1} d\Gamma(\omega) = \frac{\varrho\omega_m^{\sigma-1}}{\varrho-(\sigma-1)},$$

[1]Remember that g_t^N is the variety growth rate, which must equal η on the BGP.

where the last equality uses the general moment formula for the Pareto distribution. As such, we can find the scale parameter as

$$\omega_m = \overline{\omega} \cdot \left(1 - \frac{(\sigma - 1)}{\varrho}\right)^{\frac{1}{\sigma - 1}}. \tag{16}$$

Finally, the Pareto distribution is expressed as

$$\Gamma(\omega) = 1 - \left(\frac{\omega_m}{\omega}\right)^{\varrho}.$$

As such, the differential equation (15) can be written as

$$\frac{\partial \widehat{F}^{NC}(\widehat{q})}{\partial \widehat{q}} = d \cdot \widehat{F}^{NC}(\widehat{q}) - d(1-\alpha) \cdot \left[1 - \left(\frac{\omega_m}{\exp(\widehat{q})}\right)^{\varrho}\right],$$

and its solution is given by

$$\widehat{F}^{NC}(\widehat{q}) = Ce^{d\widehat{q}} + (1-\alpha) - \frac{d(1-\alpha)}{\varrho + d}\left(\frac{\omega_m}{\exp(\widehat{q})}\right)^{\varrho}. \tag{17}$$

To determine the constant of integration C, we separate the analysis in two cases: (i) if $d < 0$ (the usual case in our exercises), then we need to use $\widehat{F}^{NC}(\log(\omega_m)) = 0$, and if (ii) $d > 0$, then we use $\lim_{\widehat{q}\to\infty} \widehat{F}^{NC}(\widehat{q}) = 1 - \alpha$. In the first case, the resulting constant of integration is

$$C = -\frac{\varrho(1-\alpha)}{\varrho + d} \cdot \omega_m^{-d}, \tag{18}$$

while the second case delivers $C = 0$.

Distribution with Competitors Having characterized the distribution $\widehat{F}^{NC}(\widehat{q})$, we now compute the distribution for varieties with competitors, $\widehat{F}^{C}(\Delta, \widehat{q})$. This distribution must satisfy the differential equation

$$\begin{aligned}\frac{\partial \widehat{F}^{C}_t(\Delta,\widehat{q})}{\partial t} = &-\frac{\partial \widehat{F}^{C}_t(\Delta,\widehat{q})}{\partial \Delta} \cdot I\Delta - \frac{\partial \widehat{F}^{C}_t(\Delta,\widehat{q})}{\partial \widehat{q}} \cdot (I - g^{Q}_t) \\ &- (\tau_t + \delta + g^{N}_t) \cdot \widehat{F}^{C}_t(\Delta,\widehat{q}) + \tau_t \left[\widehat{F}^{NC}_t(\widehat{q} - \widehat{\lambda}) + \lim_{s\to\infty} \widehat{F}^{C}_t(s, \widehat{q} - \widehat{\lambda})\right],\end{aligned}$$

where $\widehat{\lambda} \equiv \log(\lambda)$. When we evaluate this differential equation on a BGP, we obtain

$$\frac{\partial \widehat{F}^{C}(\Delta,\widehat{q})}{\partial \Delta} \cdot I\Delta + \frac{\partial \widehat{F}^{C}(\Delta,\widehat{q})}{\partial \widehat{q}} \cdot (I - g^{Q}) + (\tau + \delta + \eta) \cdot \widehat{F}^{C}(\Delta,\widehat{q}) = \tau \left[\widehat{F}^{NC}(\widehat{q} - \widehat{\lambda}) + \lim_{s\to\infty} \widehat{F}^{C}(s, \widehat{q} - \widehat{\lambda})\right],$$

which is a partial differential equation. Notice also that we already have the function $\widehat{F}^{NC}$. This differential equation is more complicated and we solve it in two steps.

1. Since this differential must be satisfied for every pair $(\Delta, \widehat{q})$, it must also be satisfied for Δ

sufficiently large. In other words, the marginal distribution $\widehat{F}^C(\widehat{q}) \equiv \lim_{s\to\infty} \widehat{F}^C(s,\widehat{q})$ must satisfy this differential equation. With some manipulation, we have

$$\frac{\partial \widehat{F}^C(\widehat{q})}{\partial \widehat{q}} = d \cdot \widehat{F}^C(\widehat{q}) - d\alpha \cdot \left[\widehat{F}^{NC}(\widehat{q}-\widehat{\lambda}) + \widehat{F}^C(\widehat{q}-\widehat{\lambda})\right]. \tag{19}$$

To solve this equation, we create a linear grid for efficiency $\widehat{q}$ between the points $\widehat{q}_1 \equiv -5$ and $\widehat{q}_{max} \equiv 20$. We set the boundary condition $\widehat{F}^C(\widehat{q}_1) = 0$ as we pick a negative enough value for $\widehat{q}_1$. Then, using the differential equation we compute

$$\widehat{F}^C(\widehat{q}_2) = \widehat{F}^C(\widehat{q}_1) + \left.\frac{\partial \widehat{F}^C(\widehat{q})}{\partial \widehat{q}}\right|_{\widehat{q}=\widehat{q}_1} \cdot (\widehat{q}_2 - \widehat{q}_1),$$

and proceed iteratively to obtain the marginal function $\widehat{F}^C$ over the entire $\widehat{q}$ grid. Observe that for $\widehat{q}$ such that $\widehat{q} - \widehat{\lambda} < \widehat{q}_1$, we impose $\widehat{F}^C(\widehat{q}-\widehat{\lambda}) = 0$ consistent with the boundary condition.

2. Once we solve the previous step, the original PDE can be expressed as

$$\frac{\partial \widehat{F}^C(\Delta,\widehat{q})}{\partial \Delta} = \frac{1}{I\Delta} \cdot \frac{(\eta+\delta)}{1-\alpha} \left[\frac{\partial \widehat{F}^C(\Delta,\widehat{q})/\partial \widehat{q}}{d} - \widehat{F}^C(\Delta,\widehat{q}) + \alpha \cdot \left[\widehat{F}^{NC}(\widehat{q}-\widehat{\lambda}) + \widehat{F}^C(\widehat{q}-\widehat{\lambda})\right]\right]. \tag{20}$$

From the previous step, we have a grid for efficiency $\widehat{q}$. As for productivity gaps Δ, we set a grid with $\Delta_1 = \lambda$ and $\Delta_{max} = 4$. This choice of lower bound is natural as no variety with competitors cannot have $\Delta < \lambda$. The grid is linear in logs as most of the distribution should be concentrated close to λ. To solve this PDE, we need a boundary condition and we use $\widehat{F}^C(\lambda,\widehat{q}) = 0$, for all $\widehat{q}$. To see why this is appropriate, Equation (28) of the Article states that the marginal distribution of productivity gaps has an analytical solution, which is

$$\widehat{F}^C(\Delta,\widehat{q}_{max}) = \alpha\left[1 - \left(\frac{\lambda}{\Delta}\right)^{\frac{\tau+\delta+\eta}{I}}\right].$$

Hence, we must have $\widehat{F}^C(\Delta=\lambda,\widehat{q}_{max}) = 0$. Since $\widehat{F}^C$ is non-decreasing on any of its inputs, then $\widehat{F}^C(\Delta=\lambda,\widehat{q}) = 0$ for every $\widehat{q}$. Based on such boundary condition, we compute the function for every $\widehat{q}$ as

$$\widehat{F}^C(\Delta_2,\widehat{q}) = \widehat{F}^C(\Delta_1,\widehat{q}) + \left.\frac{\partial \widehat{F}^C(\Delta,\widehat{q})}{\partial \Delta}\right|_{(\Delta,\widehat{q})=(\Delta_1,\widehat{q})} \cdot (\Delta_2 - \Delta_1),$$

and proceed iteratively to obtain the joint function $\widehat{F}$ over the entire $(\Delta,\widehat{q})$ grid. As a final comment, we note that in each step of the iteration process described above the derivative $\partial \widehat{F}^C(\Delta,\widehat{q})/\partial \widehat{q}$ enters the expression for $\partial \widehat{F}^C(\Delta,\widehat{q})/\partial \Delta$ and we compute it numerically.

Integrating for the Misallocation Terms By definition, we must have

$$\int e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}^{C}(\Delta, \widehat{q}) + \int e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}^{NC}(\widehat{q}) = 1. \tag{21}$$

To see this, notice that

$$\int e^{\widehat{q}(\sigma-1)} \cdot dF(\Delta, \widehat{q}) = \int e^{\widehat{q}(\sigma-1)} \cdot dF(\widehat{q}) = \int \left(\frac{q}{Q}\right)^{\sigma-1} \cdot dF(q) = \frac{1}{Q^{\sigma-1}} \cdot \underbrace{\int q^{\sigma-1} \cdot dF(q)}_{=Q^{\sigma-1}} = 1.$$

In practice, however, the sum of the two integrals in (21) is close but not exactly one, given numerical issues. To correct such issue, we shift our $\widehat{q}$ grid to force such integrals to add up to one. This step is fundamental for the stability of the code. Based on this shifted grid, we can compute the misallocation terms as expressed in (13) and (14) by using a standard numerical integration technique.

2.3.3 Computing the Average Firm Size

Now that we have the function $k(\Delta)$ and the misallocation terms $\mathcal{M}$ and Λ, we go back to the equilibrium conditions in Section 2.3 to compute the inverse of the variety intensity ratio, ℓ, and the production share, s^P. Finally, we compute

$$\text{AvgFirmSize} = \ell \cdot \frac{N}{F}. \tag{22}$$

2.4 Markup Growth, Sales Growth and Cost-Weighted Markups

Suppose we can compute (i) expected markups given product age, $\mu_P(a_P)$, (ii) expected sales given product age, $s_P(a_P)$, and (iii) expected cost-weighted markups given product age, $\mu_P^{\text{cost}}(a_P)$. To compute sales and markups at the firm-level, however, we need to aggregate over all of a firm's products. As per Section 2.2 of the Appendix, we would compute[2]

$$\begin{aligned}
\mu_f(a_f) &= \sum_{n=1}^{\infty} \left(\int \mu_P(a) \cdot f_{a|A_f,N}(a \mid a_f, n) \cdot da \right) \cdot \frac{p_{a_f}(n)}{1 - p_{a_f}(0)}, \\
s_f(a_f) &= \sum_{n=1}^{\infty} n \cdot \left(\int s_P(a) \cdot f_{a|A_f,N}(a \mid a_f, n) \cdot da \right) \cdot \frac{p_{a_f}(n)}{1 - p_{a_f}(0)}, \\
\mu_f^{\text{cost}}(a_f) &= \sum_{n=1}^{\infty} \left(\int \mu_P^{\text{cost}}(a) \cdot f_{a|A_f,N}(a \mid a_f, n) \cdot da \right) \cdot \frac{p_{a_f}(n)}{1 - p_{a_f}(0)},
\end{aligned} \tag{23}$$

where the argument a_f indicates functions defined over firm age, rather than product age a_P. In this Section, we first show how to obtain $\mu_P(a_P), s_P(a_P)$ and $\mu_P^{\text{cost}}(a_P)$. Then, we discuss the computation

[2]Here, the function $p_{a_f}(n)$ is the conditional probability of a firm having n products given its age a_f. It is the same object we introduced in Section 2.2.2.

of $f_{a|A,N}$, which is the distribution of product age a_P for a firm of age a_f and n products. Finally, we calculate the moments that are used for calibration.

2.4.1 Markups by Product Age

Consider a firm that produces a product of age a_P. If this product originally replaced an existing one, its markup is given by

$$\mu_P^C(a_P) \equiv \mathbb{E}(\mu_i \mid a_P, C) = \min\left\{\frac{\sigma}{\sigma-1}, \lambda e^{Ia_P}\right\}, \tag{24}$$

while firms without competitors charge the CES markup. Therefore, the expected markup for a product of age a_P is given by

$$\mu_P(a_P) \equiv \mathbb{E}(\mu_i \mid a_P) = \alpha\mu_P^C(a_P) + (1-\alpha)\cdot\frac{\sigma}{\sigma-1}. \tag{25}$$

2.4.2 Sales by Product Age

At the product-level, sales are given by

$$\tilde{s}_i = \mu_i^{1-\sigma}\cdot\left(\frac{q_i}{Q}\right)^{\sigma-1}\cdot\frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma-1}}\cdot\frac{Y}{N}.$$

Define normalized product-level sales as $s_i = \frac{\tilde{s}_i}{Y/N}$. As a result, the expected (normalized) sales for a product of age a_P are given by

$$s_P(a_P) \equiv \mathbb{E}(s_i \mid a_P) = \frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma-1}}\cdot\Bigg[\alpha\mu_P^C(a_P)^{1-\sigma}\cdot\mathbb{E}\Big[\Big(\frac{q_i}{Q}\Big)^{\sigma-1}\mid a_P, C\Big] + (1-\alpha)\Big(\frac{\sigma}{\sigma-1}\Big)^{1-\sigma}\cdot\mathbb{E}\Big[\Big(\frac{q_i}{Q}\Big)^{\sigma-1}\mid a_P, NC\Big]\Bigg].$$

Following Section 2.2 of the Appendix, it can be shown that

$$\mathbb{E}\Big[\Big(\frac{q_i}{Q}\Big)^{\sigma-1}\mid a_P, C\Big] = \lambda^{\sigma-1}\cdot\exp\Big((\sigma-1)\cdot(I-g^Q)\cdot a_P\Big),$$
$$\mathbb{E}\Big[\Big(\frac{q_i}{Q}\Big)^{\sigma-1}\mid a_P, NC\Big] = \overline{\omega}^{\sigma-1}\cdot\exp\Big((\sigma-1)\cdot(I-g^Q)\cdot a_P\Big).$$

Therefore, we conclude that

$$s_P(a_P) = \frac{\exp\Big((\sigma-1)\cdot(I-g^Q)\cdot a_P\Big)}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma-1}}\cdot\left[\alpha\lambda^{\sigma-1}\cdot\mu_P^C(a_P)^{1-\sigma} + (1-\alpha)\overline{\omega}^{\sigma-1}\cdot\Big(\frac{\sigma}{\sigma-1}\Big)^{1-\sigma}\right]. \tag{26}$$

2.4.3 Cost-Weighted Markups by Product Age

At the product-level, costs are given by

$$\tilde{c}_i = \mu_i^{-\sigma} \cdot \left(\frac{q_i}{Q}\right)^{\sigma-1} \cdot \frac{1}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma-1}} \cdot \frac{Y}{N}.$$

Define normalized[3] product-level sales as $c_i = \frac{\tilde{c}_i}{\Lambda Y/N}$. Following the same approach as the one to derive product level sales, we have

$$c_P(a_P) \equiv \mathbb{E}(c_i \mid a_P) = \frac{\exp\Big((\sigma-1)\cdot(I-g^Q)\cdot a_P\Big)}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}} \cdot \left[\alpha\lambda^{\sigma-1}\cdot\mu_P^C(a_P)^{-\sigma} + (1-\alpha)\overline{\omega}^{\sigma-1}\cdot\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}\right].$$

Moreover, the expected product of product-level costs and markups is computed as

$$\mathbb{E}(\mu_i c_i \mid a_P) = \frac{\exp\Big((\sigma-1)\cdot(I-g^Q)\cdot a_P\Big)}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}} \cdot \left[\alpha\lambda^{\sigma-1}\cdot\mu_P^C(a_P)^{1-\sigma} + (1-\alpha)\overline{\omega}^{\sigma-1}\cdot\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\right].$$

Therefore, the expected cost-weighted markup of a product of age a_P can be obtained as

$$\mu_P^{\text{cost}}(a_P) \equiv \frac{\mathbb{E}(\mu_i c_i \mid a_P)}{\mathbb{E}(c_i \mid a_P)} = \frac{\alpha\lambda^{\sigma-1}\cdot\mu_P^C(a_P)^{1-\sigma} + (1-\alpha)\overline{\omega}^{\sigma-1}\cdot\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}}{\alpha\lambda^{\sigma-1}\cdot\mu_P^C(a_P)^{-\sigma} + (1-\alpha)\overline{\omega}^{\sigma-1}\cdot\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}}. \tag{27}$$

2.4.4 Computing the Conditional Distribution $f_{a|A,N}(a \mid A, n)$

Let $\Phi_{A,n}(a)$ be the CDF associated with $f_{a|A,n}(a \mid A, n)$. More concretely, we have

$$\Phi_{A,n}(a) = \int_0^a f_{a|A,N}(x \mid A, n)dx.$$

As per Section 4 of the Supplementary Material, define the function $X_{A,n}(a) \equiv \Lambda_A(n)n \cdot \Phi_{A,n}(a)$, where $\Lambda_A(n)$ is the mass of firms of age A with n products. Therefore, to obtain $\Phi_{A,n}(a)$ we need to obtain $\Lambda_A(n)$ and then $X_{A,n}(a)$. Starting with the former, notice that for a small time interval ι, the function $\Lambda_A(n)$ must satisfy

$$\Lambda_A(n) = \Lambda_{A-\iota}(n)(1-(\tau+\delta+x)n\iota) + \Lambda_{A-\iota}(n+1)(\tau+\delta)(n+1)\iota + \Lambda_{A-\iota}(n-1)x(n-1)\iota,$$

[3]To understand why sales and costs have a different normalization factor, simply notice that the normalization factor is given by the respective unconditional averages. That is, $\mathbb{E}(\tilde{s}_i) = Y/N$ and $\mathbb{E}(\tilde{c}_i) = \Lambda Y/N$.

which can be rearranged to

$$\frac{\Lambda_A(n) - \Lambda_{A-\iota}(n)}{\iota} = -(\tau + \delta + x)\Lambda_{A-\iota}(n)n + \Lambda_{A-\iota}(n+1)(\tau+\delta)(n+1) + \Lambda_{A-\iota}(n-1)x(n-1).$$

By taking the limit as $\iota \to 0$, we obtain the differential equation

$$\frac{\partial \Lambda_A(n)}{\partial A} = -(\tau + \delta + x)\Lambda_A(n)n + (\tau+\delta)\Lambda_A(n+1)(n+1) + x\Lambda_A(n-1)(n-1). \tag{28}$$

To solve this differential equation, we use the following algorithm:

1. We create a discrete grid for the number of products and the maximum number of products[4] is 45. We also create a linear grid for firm age, with $A_1 = 0$ and $A_{max} = 50$. As a boundary condition, we use the fact that $\Lambda_{A=0}(n = 1) = 1$ and zero for all other values of n. This captures the idea that a firm that has just been born must have exactly one product (and we normalize the number of age zero firms to 1).

2. Given the boundary condition, we use use the differential equation to compute

$$\Lambda_{A_2}(n) = \Lambda_{A_1}(n) + \left.\frac{\partial \Lambda_A(n)}{\partial A}\right|_{(A,n)=(A_1,n)} \cdot (A_2 - A_1).$$

   Then we proceed iteratively to obtain $\Lambda_A(n)$ for all values of firm age A.

Once we have the function $\Lambda_A(n)$, the next step is to obtain $X_{A,n}(a)$. As per Equation 10 of the Supplementary Material, this function satisfies the differential equation

$$\frac{\partial X_{A,n}(a)}{\partial A} + \frac{\partial X_{A,n}(a)}{\partial a} = -(\tau+\delta+x)nX_{A,n}(a) + x(n-1)\left(\Lambda_A(n-1) + X_{A,n-1}(a)\right) + (\tau+\delta)nX_{A,n+1}(a). \tag{29}$$

To solve this differential equation, we use the following algorithm:

1. For number of products and firm age, we use the same grids as we used in the previous differential equation. For product age, we use the same grid as for firm age. As for the boundary condition of $X_{A,n}(a)$, we use two arguments. First, we have $\Phi_{A,n}(a) = 1$ whenever $a \geq A$. This condition simply establishes that a firm cannot have a product that is older than the firm itself. Therefore, we have $X_{A,n}(a) = \Lambda_A(n)n$ whenever $a \geq A$. Second, notice that for every $\varepsilon > 0$, we must have $X_{A,n}(a_1 - \varepsilon) = 0$, as no product can have negative age. As such, for $a_1 = 0$ we

[4]Since we have to work with a bounded support for the number of products and the differential equation includes a term with $\Lambda_A(n+1)$, a truncation error issue arises. As such, we need to select a large upper bound for the support, to make sure that truncation errors manifest only for high values of n. Since high values of n have low probability, truncation problems will not bear consequences for the computation of the moments of sales and markup growth. As per our analysis, $n_{max} = 45$ is enough for robust moments.

can solve the simplified differential equation

$$\left(\frac{\partial X_{A,n}(a)}{\partial A}+\frac{\partial X_{A,n}(a)}{\partial a}\right)\Bigg|_{(a,A,n)=(a_1-\varepsilon,A,n)} = x(n-1)\Lambda_A(n-1).$$

With this differential equation, we can solve

$$X_{A_2,n}(a_1) = \left(\frac{\partial X_{A,n}(a)}{\partial A}+\frac{\partial X_{A,n}(a)}{\partial a}\right)\Bigg|_{(a,A,n)=(a_1-\varepsilon,A_1,n)} \cdot \varepsilon,$$

and proceed iteratively to obtain $X_{A,n}(a_1)$ for every (A,n). Remember that $X_{A_1,n}(a_1) = \Lambda_{A_1}(n)n$ as per the previous argument.

2. Finally, we need to solve $X_{A,n}(a)$ for all values of product age, a. For this purpose, we use (29) and for every $i = 1,2,3...$ we solve[5]

$$X_{A_{2+i},n}(a_2) = X_{A_{2+i-1},n}(a_1) + \left(\frac{\partial X_{A,n}(a)}{\partial A}+\frac{\partial X_{A,n}(a)}{\partial a}\right)\Bigg|_{(a,A,n)=(a_1,A_{2+i-1},n)} \cdot (a_2-a_1).$$

Then we proceed iteratively over product age a to obtain the function X over all its support.

Having $\Lambda_A(n)$ and $X_{A,n}(a)$, we compute the function $\Phi_{A,n}(a) = \frac{X_{A,n}(a)}{\Lambda_A(n)n}$. Once we have $\Phi_{A,n}(a)$, we can obtain its density $f_{a|A,n}(a \mid A,n)$ by taking the appropriate difference.

2.4.5 Computing Sales and Markup Moments

Once we have the distribution $f_{a|A,n}(a \mid A,n)$, we use Equations (23) to compute (i) expected markups for a firm of age a_f, denoted by $\mu_f(a_f)$, (ii) expected sales for a firm of age a_f, denoted by $s_f(a_f)$,and (iii) expected cost-weighted markups for a firm of age a_f, denoted by $\mu_f^{\text{cost}}(a_f)$. With the first two functions, we compute a measure of (firm) markup and sales growth by age 10. Finally, we integrate the third function for an economy wide expected cost-weighted markup measure. More specifically, we have:

$$\begin{aligned}
\text{SalesGrowth} &= \log(s_f(10)) - \log(s_f(0)) \\
\text{MarkupGrowth} &= \mu_f(10) - \mu_f(0) \\
\text{AvgMarkup} &= \int \mu_f^{\text{cost}}(a_f)\cdot\omega^f(a_f)\cdot da_f.
\end{aligned} \tag{30}$$

[5]Notice that it is fundamental that we solve this differential equation for every n and then move to the next firm age value, A. If, instead, we solve for every level of firm age A and then repeat the process for the next product n, then a problem will arise as we have dependency on $X_{A,n+1}(a)$.

2.5 Computing Employment of Large Firms

To compute the employment share of large firms, we need to obtain the cross-sectional distribution of employment. At the product-level, employment is a random variable given by

$$l_i = \mu_i^{-\sigma} \cdot \left(\frac{q_i}{Q}\right)^{\sigma-1} \cdot \frac{\ell s^P}{\mathcal{M}^{\sigma-1}\Lambda^{\sigma}}.$$

The key idea is to understand that firm age A and number of products n dictate the distribution of product age through $\Phi_{A,n}(a)$. Then, product age will condition the joint distribution of productivity gaps and efficiency, $F(\Delta, \widehat{q})$. Therefore, define $D^1_{A,n}(y)$ as the probability that a firm of age A and n products has employment lower than y for one of its products. That is,

$$D^1_{A,n}(y) \equiv \Pr(l_i \leq y \mid A, n) = \Pr(\log(l_i) \leq \log(y) \mid A, n) = \int \Pr\left(\widehat{q} \leq \frac{\log(y) + \sigma \cdot \log(\mu(a)) - \kappa}{\sigma - 1} \mid a\right) \cdot d\Phi_{A,n}(a).$$

In the notation above, we defined the constant $\kappa = \log(\ell s^P \cdot \mathcal{M}^{1-\sigma}\Lambda^{-\sigma})$ for simplicity. For the purpose of computing our calibration moment, however, we will not condition on firm age A, only on the number of products n.[6] Since the integrand in the expression above does not depend on A, we can simply compute

$$D^1_n(y) \equiv \Pr(l_i \leq y \mid n) = \Pr(\log(l_i) \leq \log(y) \mid n) = \int \Pr\left(\widehat{q} \leq \frac{\log(y) + \sigma \cdot \log(\mu(a)) - \kappa}{\sigma - 1} \mid a\right) \cdot d\Phi_n(a),$$

where $\Phi_n(a) = \int \Phi_{A,n}(a) \cdot \omega^f(A) \cdot dA$. However, it should be clear that this is the *product level* distribution of employment for a firm that has n products. To aggregate over all of this firm's products, we define[7]

$$D^n_n(y) \equiv \Pr\left(\sum_{i=1}^{n} l_i \leq y \mid n\right) = \Pr\left(l_i \leq \frac{y}{n} \mid n\right) = D^1_n\left(\frac{y}{n}\right). \tag{31}$$

We can also derive the unconditional employment distribution by summing over the number of products

$$D(y) = \sum_{n=1}^{\infty} p(n) \cdot D^n_n(y). \tag{32}$$

where $p(n)$ is defined in Equation (7). Finally, we need to notice that in practice the computation above will deliver a distribution defined over the grid of the *logarithm* of employment, not employment itself. Therefore, to obtain the density associated with $D(y)$ one needs to handle the transfor-

[6] If the employment distribution conditional on firm age was an object of particular interest, we would just proceed with $\Phi_{A,n}(a)$ rather than $\Phi_n(a)$. However, for our calibration process this would be prohibitively costly.

[7] Here, the middle equality would require that, conditional on the number of products, product level employment is independent and identically distributed across each product. However, for this property to hold, we would need to also condition on firm age. As we mentioned, working with the distribution conditioned on both firm age and products was prohibitively costly. This abstraction did not have a substantial impact on our final employment share moment.

mation accordingly. Once this is taken care of, we have

$$\text{ShareLarge} = \left(\int_0^\infty y \cdot dD(y) \right)^{-1} \cdot \int_{10^4}^\infty y \cdot dD(y) \tag{33}$$

as the share of employment attributed to firms with at least $10,000$ employees. To conclude the derivation of this moment, we discuss the productivity distribution conditional on product age.

2.5.1 Distribution of Efficiency Conditional on Product Age

In the derivation above, we stated that

$$D_n^1(y) = \int \Pr\left(\widehat{q} \leq \frac{\log(y) + \sigma \cdot \log(\mu(a)) - \kappa}{\sigma - 1} \mid a \right) \cdot d\Phi_n(a). \tag{34}$$

Separating the varieties with and without competitors, we have

$$\begin{aligned} \Pr\left(\widehat{q} \leq \frac{\log(y) + \sigma \cdot \log(\mu(a)) - \kappa}{\sigma - 1} \mid a \right) =& \alpha \cdot F^C_{\widehat{q}|a} \left(\frac{\log(y) + \sigma \cdot \log\left(\min\left\{ \lambda e^{Ia}, \frac{\sigma}{\sigma - 1} \right\} \right) - \kappa}{\sigma - 1} \mid a \right) \\ &+ (1 - \alpha) \cdot F^{NC}_{\widehat{q}|a} \left(\frac{\log(y) + \sigma \cdot \log(\frac{\sigma}{\sigma - 1}) - \kappa}{\sigma - 1} \mid a \right). \end{aligned} \tag{35}$$

Therefore, we need to characterize the functions $F^{NC}_{\widehat{q}|a}(\widehat{q} \mid a)$ and $F^{C}_{\widehat{q}|a}(\widehat{q} \mid a)$.

Distribution Without Competitors If we consider products with age $a = 0$, their efficiency distribution is fully determined by the (exogenous) distribution of new product efficiency. That is,

$$\overline{F}^{NC}_{\widehat{q}|a}(\widehat{q} \mid a = 0) = \Gamma\left(\exp\left(\widehat{q} \right) \right). \tag{36}$$

In the formulation above, we can think of $\overline{F}^{NC}_{\widehat{q}|a}(\cdot \mid a = 0)$ as the *mass* of varieties without competitors that have age $a = 0$ (this is simply a normalization). As a result, this object must satisfy the differential equation

$$\frac{\partial \overline{F}^{NC}_{\widehat{q}|a}(\widehat{q} \mid a)}{\partial a} = -(I - g^Q) \cdot \frac{\partial \overline{F}^{NC}_{\widehat{q}|a}(\widehat{q} \mid a)}{\partial \widehat{q}} - (\tau + \delta) \cdot \overline{F}^{NC}_{\widehat{q}|a}(\widehat{q} \mid a). \tag{37}$$

Using $\overline{F}^{NC}_{\widehat{q}|a}(\cdot \mid a = 0)$ as our initial condition, we solve this differential equation in the same way we solved the differential equations in the previous sections. Notice that in the resulting function, the mass of products with age a will be scaled in proportion to their density relative to the density of

products with age $a = 0$. As such, to obtain the conditional distribution we compute

$$F^{NC}_{\widehat{q}|a}(\widehat{q} \mid a) = \frac{\overline{F}^{NC}_{\widehat{q}|a}(\widehat{q} \mid a)}{\overline{F}^{NC}_{\widehat{q}|a}(\widehat{q}_{max} \mid a)}. \tag{38}$$

Distribution With Competitors For varities with competitors, product age translates directly into productivity gaps as $\Delta = \lambda e^{Ia}$. As such, we have

$$F^{C}_{\widehat{q}|a}(\widehat{q} \mid a) = \frac{\widehat{F}^{C}(\lambda e^{Ia}, \widehat{q})}{\widehat{F}^{C}(\lambda e^{Ia}, \widehat{q}_{max})}. \tag{39}$$

In the formula above, the denominator is just the marginal distribution over productivity gaps evaluated at $\Delta = \lambda e^{Ia}$.

Putting the Distributions Together With the distributions derived above, we go back to Equation (35). However, our final concern is that the distributions are defined over the same grid for efficiency $\widehat{q}$, but this sets an implied grid in terms of y for Equation (35) that is different (i) between varieties with and without competitors, and (ii) between varities with competitors with different ages. For this reason, we interpolate these distributions to a common and exogenous grid for y and only then compute $D^1_n(y)$ with (34).

2.5.2 Computing the Entrant Employment Distribution

Based on the Section above, one could compute the employment distribution for entrant firms, which would just be $D^n_{A=0,n}(y)$. However, obtaining this distribution is particularly easy because we do not need to solve the differential equation for $\overline{F}^{NC}_{\widehat{q}|a}(\cdot \mid a = 0)$ as the boundary condition is everything we need. Also, if one wants $D^n_{A,n}(y)$ for a single level of age (in this case, $A = 0$), this would not be more demanding from a computational standpoint. In the Article, the $90 - 10$ ratio for the entrants employment distribution is mentioned and it can be computed based on the appropriate quantiles of the distribution $\sum_{n=1}^{\infty} D^n_{A=0,n}(y) \cdot p(n)$.

2.6 Matching the Model with the Data

Over the previous Subsections, we discussed how to compute the moments of interest with the model *for a given vector of parameters*. The final step is to search for the vector of parameters that minimize the distance of the model moments to the data moments. For this purpose, we need to find $\theta \in \mathbb{R}^5$such

that $H(\theta) = 0$, where

$$H(\theta) = \begin{bmatrix} \text{AvgFirmSize}(\theta) - \text{AvgFirmSize}_{\text{target}} \\ \text{MarkupGrowth}(\theta) - \text{MarkupGrowth}_{\text{target}} \\ \text{SalesGrowth}(\theta) - \text{SalesGrowth}_{\text{target}} \\ \text{AvgMarkup}(\theta) - \text{AvgMarkup}_{\text{target}} \\ \text{ShareLarge}(\theta) - \text{ShareLarge}_{\text{target}} \end{bmatrix}.$$

To find such θ, we use a minimization package.[8] With this procedure, we find a parameter vector that matches the data perfectly. In Table 2, we present a summary of the calibration process.

Table 2: Summary of Calibration

Parameter		Moment	Target	Model
φ_e	7.13%	Aggregate Productivity Growth	1.84%	1.84%
λ	1.08	Entry Rate	11.6%	11.6%
I	2.29%	Average Firm Size	19.816	19.815
δ	5.40%	Markup Growth by Age 10	10.2%	10.2%
φ_x	1.86%	Sales Growth by Age 10	58%	58%
α	60.5%	Average Cost-Weighted Markup	20%	20%
$\overline{\omega}^{\sigma-1}$	0.1502	Share of Firms with + 10,000 Employees	26.2%	26.2%

2.7 Computing Exit Rates by Firm Age

Even though exit rates by age are not used as part of the calibration process, they are depicted on Figure 7b of the Article. Therefore, we explain their computation in this Section. As shown in Section 6 of the Supplementary Material, the probability of exit for a firm with n products is given by

$$\text{Exit}_n = \Pr(k - m \geq n) = \sum_{m=0}^{\infty} e^{-xn} \cdot \frac{(xn)^m}{m!} \cdot \sum_{k=n+m}^{\infty} e^{-(\tau+\delta)n} \cdot \frac{((\tau+\delta)n)^k}{k!}. \tag{40}$$

In the notation above, k represents the number of products that could be destroyed in a period (either naturally via δ or by creative destruction via τ) and m represents the number of products that could be created, via x. Based on the expression above, we can calculate the exit rate conditional on firm age as

$$\text{Exit}_A = \sum_{n=1}^{\infty} p_A(n) \cdot P(k - m \geq n). \tag{41}$$

[8] John D'Errico (2024). fminsearchbnd, fminsearchcon (https://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd-fminsearchcon), MATLAB Central File Exchange. Retrieved June 25, 2024.

3 Transitional Dynamics

Once we calibrate the model, we can compute it under the resulting θ. By doing that, we obtain the objects of interest under the calibrated BGP. Our next step is to compute the counterfactual BGP. For this purpose, we simply change $\eta = 2\%$ to the counterfactual $\eta = 0.24\%$ and compute the model once again. Having the original and the counterfactual BGP, in this Section we discuss how the economy transitions from the former to the latter.

3.1 Preliminaries

As outlined in Section 7.4 of the Supplementary Material, the final dynamic system that the model satisfies in equilibrium consists on two equations. First, we have labor market clearing

$$\ell_t(1 - s_t^P) = \frac{1}{\varphi_e}\Big(\frac{g_t^N + \delta}{1-\alpha} - \frac{\zeta - 1}{\zeta} \cdot x\Big), \tag{42}$$

where $\ell_t = L_t/N_t$ and $s_t^p = L_t^P/L_t$. The second equation is free entry,[9] which reads

$$\begin{aligned}\frac{\rho + g_{s_t^P} - g_{\Lambda_t} + \frac{\alpha g_t^N + \delta}{1-\alpha}}{\varphi_e} = \frac{\zeta - 1}{\zeta} \cdot \frac{x}{\varphi_e} \\ + \ell_t s_t^P \cdot \alpha \lambda^{\sigma-1}\left(b_t(\lambda) \cdot \Big(\rho - g_{\Lambda_t} - \eta + \frac{g_t^N + \delta}{1-\alpha}\Big) - \dot{b}_t(\lambda)\right) \\ + \ell_t s_t^P \cdot (1-\alpha)\overline{\omega}^{\sigma-1}\left(b_t\Big(\frac{\sigma}{\sigma-1}\Big) \cdot \Big(\rho - g_{\Lambda_t} - \eta + \frac{g_t^N + \delta}{1-\alpha}\Big) - \dot{b}_t\Big(\frac{\sigma}{\sigma-1}\Big)\right),\end{aligned} \tag{43}$$

where $b_t(\Delta)$ satisfies the differential equation

$$\dot{b}_t(\Delta) = \frac{h(\Delta)}{\mathcal{M}_t^{\sigma-1} \cdot \Lambda_t^{\sigma}} + \Delta I \cdot \frac{\partial b_t(\Delta)}{\partial \Delta} - \left(\rho - g_{\Lambda_t} - \eta + \frac{(g_t^N + \delta) \cdot \overline{q}^{\sigma-1}}{1-\alpha}\right) \cdot b_t(\Delta), \tag{44}$$

and $h(\Delta) = (\min\{\Delta, \frac{\sigma}{\sigma-1}\} - 1)/(\min\{\Delta, \frac{\sigma}{\sigma-1}\}^{\sigma})$.

3.2 Overview of the Algorithm

Consider that at $t = 0$ the economy is on the calibrated BGP in 1980, where $\eta = \eta_0 = 2\%$. Our objective is to find the equilibrium dynamic response of the model to an exogenous path of population growth. That is, we have a vector $\{\eta_t\}_{t=0}^T$ where for $t > 0$, η_t is the "policy" variable that we change

[9] In the Supplementary Material, the free entry conditions is expressed in terms of the function $u_t(\Delta)$ rather than $b_t(\Delta)$. These functions are related and satisfy $b_t(\Delta) \equiv u_t(\Delta)/\ell_t s_t^P$.

to see how the system responds. We track the dynamic response of the system for $T = 320$ years (until 2300) and the time period is a month. Our algorithm consists on the following steps:

1. We guess a path for the production share, $\{s_t^P\}_{t=0}^T$.
2. Given the current path for the production share, we use the labor market clearing condition (42) to obtain a path for the growth rate of varieties, $\{g_t^N\}_{t=0}^T$.
3. With the path for the growth rate of varieties, $\{g_t^N\}_{t=0}^T$, we compute the variables $\left\{\tau_t, g_t^Q\right\}_{t=0}^T$ and with those we solve for the distributions, $\left\{\widehat{F}_t^{NC}(\widehat{q}), \widehat{F}_t^{C}(\Delta, \widehat{q})\right\}_{t=0}^T$. Then, we compute the path of misallocation terms, $\{\mathcal{M}_t, \Lambda_t\}_{t=0}^T$. The sequence for the labor wedge also allows us to compute $\{g_{\Lambda_t}\}_{t=0}^T$.
4. Using the sequences obtained in the previous steps, we use the differential equation (44) to compute the sequence $\{b_t(\Delta)\}_{t=0}^T$ for every Δ.
5. Finally, we obtain an updated path for the production share, $\{\widehat{s}_t^P\}_{t=0}^T$ with the free entry condition (43). While the difference between $\{s_t^P\}_{t=0}^T$ and $\{\widehat{s}_t^P\}_{t=0}^T$ is greater than some tolerance, we form a new guess for the production share as (conservative) convex combination of the original path and the updated path. Then, we take this new guess for the production share sequence and return to Step 2. We repeat the process until we find the path $\{s_t^P\}_{t=0}^T$ that is the fixed point of the routine described above.

3.2.1 Model on the Balanced Growth Path

The BGP counterparts of most objects required for the computation of this algorithm have already been introduced in the previous sections. In this section, we show how the function $b_t(\Delta)$ can be obtained on a BGP. We can evaluate the differential equation (44) on the BGP to obtain

$$\left(\rho - \eta + \frac{(\delta+\eta)\cdot \overline{q}^{\sigma-1}}{1-\alpha}\right)\cdot b(\Delta) - \Delta I \cdot \frac{\partial b(\Delta)}{\partial \Delta} = \frac{h(\Delta)}{\mathcal{M}^{\sigma-1}\cdot\Lambda^{\sigma}}.$$

Since $\partial b(\Delta)/\partial\Delta = 0$ when evaluated at $\Delta = \frac{\sigma}{\sigma-1}$, we have the boundary condition

$$b\Big(\frac{\sigma}{\sigma-1}\Big) = \left(\rho - \eta + \frac{(\delta+\eta)\cdot \overline{q}^{\sigma-1}}{1-\alpha}\right)^{-1}\cdot \frac{h(\frac{\sigma}{\sigma-1})}{\mathcal{M}^{\sigma-1}\cdot\Lambda^{\sigma}}.$$

Up to scale, this is the same differential equation that characterizes $k(\Delta)$ in (10). As such, we have a closed form solution for $b(\Delta)$ given by $b(\Delta) = k(\Delta)/\mathcal{M}^{\sigma-1}\cdot\Lambda^{\sigma}$. We evaluate this function on a grid between λ and $\frac{\sigma}{\sigma-1}$.

3.2.2 Computing the Variety Growth Rate

Suppose we are given a path for the production share, $\{s_t^P\}_{t=0}^T$. For $t = 0$, we have $g_0^N = \eta_0$ as the economy is originally on the calibrated BGP. Then, for $t = 1$ we can write (42) as

$$\ell_0 e^{(\eta_1 - g_1^N)\Delta t} \cdot (1 - s_t^P) = \frac{1}{\varphi_e}\left(\frac{g_1^N + \delta}{1-\alpha} - \frac{\zeta - 1}{\zeta} \cdot x\right),$$

where $\Delta t = 1/12$ is the factor that converts the rates from annual terms to our unit of analysis. This expression can be solved numerically for g_1^N and this allows us to compute $\ell_1 = \ell_0 \cdot e^{(\eta_1 - g_1^N)\Delta t}$. Then, we can use

$$\ell_{t-1} e^{(\eta_t - g_t^N)\Delta t} \cdot (1 - s_t^P) = \frac{1}{\varphi_e}\left(\frac{g_t^N + \delta}{1-\alpha} - \frac{\zeta - 1}{\zeta} \cdot x\right), \tag{45}$$

to solve for ℓ_t and g_t^N in the remaining periods.

3.2.3 Computing the Misallocation Terms

With the path for the variety growth rate g_t^Ncomputed in the previous section, we obtain

$$\begin{aligned}
\tau_t &= \frac{\alpha}{1-\alpha} \cdot (g_t^N + \delta), \\
(I - g_t^Q) &= -\frac{(\overline{q}^{\sigma-1} - 1)}{\sigma - 1} \cdot \frac{g_t^N + \delta}{1-\alpha}.
\end{aligned} \tag{46}$$

Therefore, we can solve the differential equations presented in Section 2.3.2 for $\widehat{F}_t^{NC}(\widehat{q})$ and $\widehat{F}_t^{C}(\Delta, \widehat{q})$ respectively using the BGP distributions as the initial conditions. For convenience, we repeat these differential equations here.

$$\frac{\partial \widehat{F}_t^{NC}(\widehat{q})}{\partial t} = -\frac{\partial \widehat{F}_t^{NC}(\widehat{q})}{\partial \widehat{q}} \cdot (I - g_t^Q) - (\tau_t + \delta + g_t^N) \cdot \widehat{F}_t^{NC}(\widehat{q}) + \frac{(1-\alpha)\tau_t}{\alpha} \cdot \Gamma\left(\exp\left(\widehat{q}\right)\right), \tag{47}$$

$$\begin{aligned}
\frac{\partial \widehat{F}_t^{C}(\Delta, \widehat{q})}{\partial t} = &- \frac{\partial \widehat{F}_t^{C}(\Delta, \widehat{q})}{\partial \Delta} \cdot I\Delta - \frac{\partial \widehat{F}_t^{C}(\Delta, \widehat{q})}{\partial \widehat{q}} \cdot (I - g_t^Q) \\
&- (\tau_t + \delta + g_t^N) \cdot \widehat{F}_t^{C}(\Delta, \widehat{q}) + \tau_t \left[\widehat{F}_t^{NC}(\widehat{q} - \widehat{\lambda}) + \lim_{s \to \infty} \widehat{F}_t^{C}(s, \widehat{q} - \widehat{\lambda})\right].
\end{aligned} \tag{48}$$

The method for solving these differential equations is analogous to the one we had on the BGP. Then, we use the distributions at time t to compute the misallocation terms $\mathcal{M}_t$ and Λ_t, which are given

by[10]

$$\mathcal{M}_t = \frac{\left(\int \mu(\Delta)^{1-\sigma} \cdot e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \int e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^{NC}(\widehat{q})\right)^{\frac{\sigma}{\sigma-1}}}{\int \mu(\Delta)^{-\sigma} \cdot e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \int e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^{NC}(\widehat{q})}$$

$$\Lambda_t = \frac{\int \mu(\Delta)^{-\sigma} \cdot e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \int e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^{NC}(\widehat{q})}{\int \mu(\Delta)^{1-\sigma} \cdot e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \int e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^{NC}(\widehat{q})}.$$

Similar to what we did in Section 2.3.2, we shift the exogenous efficiency grid at every period t to ensure that

$$\int e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^C(\Delta,\widehat{q}) + \int e^{\widehat{q}(\sigma-1)} \cdot d\widehat{F}_t^{NC}(\widehat{q}) = 1.$$

holds at every t - this is key to generate a sensible transition path for the misallocation terms. Finally, we compute the growth rate of Λ_t as[11]

$$g_{\Lambda_t} = \log\left(\frac{\Lambda_t}{\Lambda_{t-12}}\right). \tag{49}$$

To conclude, we note that the misallocation terms are not updated in every iteration of our algorithm as this is the most costly part of the code. For example, we update them in the first iteration and then keep them fixed through some iterations to speed up the code before updating them again. Of course, aside from a small distance between the production shares, the code will only declare convergence on iterations for which the misallocation terms have been updated.

3.2.4 Computing the Function $b_t(\Delta)$ and Updating the Production Share

With the objects derived in the previous sections, it is straightforward to solve the differential equation (44) with the BGP function $b(\Delta)$ as an initial guess to obtain $b_t(\Delta)$. We can differentiate this function over time to get $\dot{b}_t(\Delta)$. Using the growth rate $g_{s_t^P}$ implied by our initial guess,[12] we can solve (43) for s_t^P, which will generate an updated version of the production share. We then compute the difference between the guessed production share and the updated production share. While this difference is greater than some small tolerance, we form a new guess for the production share as a convex combination of both paths. To ensure convergence, a very large weight on the old production share ($\sim$ 99.5%) is needed. Also, the presence of g_{Λ_t}makes the dynamic system quite sensitive - as

[10]Our data for the population growth path η_t ends in 2020, year after which η_t reflects forecasts. Since these forecasts are basically constant, numerical imprecisions in the differential equations generate a jagged path for the misallocation terms. While in levels this would not be a problem, the growth rate of Λ_t shows up in the dynamic system. Therefore, after 2020 we smooth the path for the misallocation terms, requiring convergence to the counterfactual BGP by 2100.

[11]While the use of backward (rather than forward) growth rates is not very important here, the use of an annual rate (rather than an annualized monthly rate) is important to smooth out numerical imprecisions associated with the solution of the differential equations.

[12]For the production share s_t^P, the use of forward growth rates, $g_{s_t^P} = \log(s_{t+12}^P / s_t^P)$, is critical for the convergence of the algorithm. With backward growth rates, if an iteration requires the production share to increase at a point, this will raise the implied production share growth rate at that point, which will then require an even higher production share increase through the system.

such, the convergence process is not monotonic and, in practice, the convergence tolerance cannot be arbitrarily small.

3.3 Firm Moments Along the Transition

Let $\omega_t(n)$ be the mass of firms with n products at time t. Denote by $p_t(n)$ the probability that any firm has n products at time t. As such, we have $\omega_t(n) = F_t p_t(n)$, where F_t is the total number of firms in period t. As such, we can define

$$\nu_t(n) \equiv \frac{\omega_t(n)}{N_t} = \frac{p_t(n)}{N_t/F_t},$$

which is stationary along a BGP. As a result, at any time period t we can compute:

1. Average Firm Size:
$$\text{AvgFirmSize}_t = \ell_t \cdot \frac{N_t}{F_t} = \ell_t \cdot \left(\sum_{n=1}^{\infty} \nu_t(n) \right)^{-1}. \tag{50}$$

2. Entry Rate:
$$\text{EntryRate}_t = z_t \cdot \frac{N_t}{F_t} = z_t \cdot \left(\sum_{n=1}^{\infty} \nu_t(n) \right)^{-1}. \tag{51}$$

3. Exit Rate:
$$\text{ExitRate}_t = (\tau_t + \delta) \cdot p_t(1) = (\tau_t + \delta) \cdot \nu_t(1) \cdot \left(\sum_{n=1}^{\infty} \nu_t(n) \right)^{-1}. \tag{52}$$

Given an equilibrium path, the sequence $\{\ell_t, z_t, \tau_t\}_{t=1}^{\infty}$ is known. Therefore, the only missing ingredient is the path $\{v_t(1), ..., v_t(n_{max})\}_{t=1}^{\infty}$. As per Section 7.6 of the Supplementary Material, the function $\nu_t(n)$ satisfies

$$\begin{aligned} \dot{\nu}_t(n) &= (n-1)\nu_t(n-1) \cdot x + (n+1)\nu_t(n+1)(\tau+\delta) - n\nu_t(n) \cdot (\tau+\delta+x) - \nu_t(n) \cdot g_t^N, \\ \dot{\nu}_t(1) &= z_t + 2\nu_t(2) \cdot (\tau+\delta) - \nu_t(1) \cdot (\tau+\delta+x) - \nu_t(1) \cdot g_t^N. \end{aligned} \tag{53}$$

To solve this set of differential equations, we only need an initial condition. On a BGP, we have

$$\nu(n) = \frac{p(n)}{N/F}, \tag{54}$$

where $p(n)$ is defined in Equation (7) and N/F is defined in Equation (5).

3.3.1 Other Moments Along the Transition

In this Section, we explain the computation of the other moments that are presented in the Article. At every period t, the firm age density is given by

$$\omega_t^f(a) = \frac{\psi_t e^{(\psi_t - g_t^N)a}}{\psi_t - x(1 - e^{\psi_t a})} \cdot \left(\int \frac{\psi_t e^{(\psi_t - g_t^N)a'}}{\psi_t - x(1 - e^{\psi_t a'})} da' \right)^{-1}, \tag{55}$$

where $\psi_t = x - (\tau_t + \delta)$. Therefore, the share of firms with age greater than 10 is given by

$$\text{ShareFirms10y}_t = \int_{10}^{\infty} \omega_t^f(a) \cdot da. \tag{56}$$

At every period t, the unconditional product distribution can be obtained as

$$p_t(k) = \nu_t(k) \cdot \left(\sum_{n=1}^{\infty} \nu_t(n) \right)^{-1}. \tag{57}$$

As such, the share of firms with k or more products can be obtained as

$$\text{ShareFirmsProds}_{k,t} = \sum_{n=k}^{\infty} p_t(n). \tag{58}$$

Finally, there is the transition of cost-weighted average markups. In our calibration, it turns out that the function $\mu_f^C(a_f)$ on the calibrated BGP is quite close to its version in the counterfactual BGP. As such, for every a_f, we compute $\mu_{f,t}^C(a_f)$ by interpolating between the functions on the calibrated and counterfactual BGPs. This simplication allows us to avoid computing the distribution $f_{a|A,n}$ along the transition, which is a complicated and large object. In this context, the average cost-weighted markup at period t is given by

$$\text{AvgMarkup}_t = \int \mu_{f,t}^C(a_f) \cdot \omega_t^f(a_f) \cdot da_f. \tag{59}$$