Online Appendix to "Customer Acquisition, Business Dynamism and Aggregate Growth"

Marek Ignaszak and Petr Sedláček

Contents

A Benchmark model: Additional analytical results

This appendix provides a formal equilibrium definition in the benchmark model and all the proofs to Propositions 1 to 3 in the main text. It also offers additional analytical results pertaining to firm growth and Gibrat's law and planner's allocations.

For the sake of notational clarity, we drop the firm subscript j whenever it does not create confusion. As in the main text, we use primes ($'$) to denote next period values.

A.1 Equilibrium definition

We begin by providing a formal equilibrium definition for our benchmark model.

Benchmark model: Equilibrium definition. *A balanced growth path equilibrium of our benchmark model consists of the following tuple in every period with* $j \in \Omega$*:* c_j, p_j, d_j, b_j, q_j, x_j, n_j, s_j, $v(q_j, b_j)$, r, W, C, Q, g, *such that (i) output and prices,* c_j *and* p_j*, satisfy, (5) and (12), (ii) firms' demand stock,* d_j*, is given by (9) and firms' customer bases and productivity levels,* b_j *and* q_j*, evolve according to (10) and (8), (iii) optimal innovation probabilities,* x_j*, satisfy (13), (iv) labor demand,* n_j *and* s_j*, satisfy (6) and (7), (v) firm values,* $v(q_j, b_j)$*, satisfy (11), (vi) the interest and wage rates,* r *and* W*, satisfy (3) and (14), (vii) aggregate consumption,* C*, is defined by (1), and (viii) the aggregate productivity index,* Q*, is given by (15) and it's growth equals* $1 + g = Q'/Q$*.*

A.2 Proofs of Propositions 1 to 3

In this subsection, we provide proofs of Propositions 1 to 3.

Proof of Proposition 1 . Substituting the demand constraint (5) and optimal pricing (12) into the consumption aggregator (1), we can write $W = Q(\eta - 1)/\eta = Q/2$. Using this expression, together with optimal labor supply (4) and pricing conditions (12), stationarized per-period profits are given by $\hat{\pi} = b\hat{q}\left[\frac{1}{4\upsilon} - \bar{s}\frac{x^2}{2}\right]$. Firm value can, therefore, be written as

$$\hat{v}(\hat{z}) = \hat{\pi} + \beta(1-\delta)\left[x\hat{v}(\hat{z}^+) + (1-x)\hat{v}(\hat{z}^-)\right], \tag{A1}$$

where $\hat{z} = (b, \hat{q})$, $\hat{q} = q/Q$, $\hat{z}^+ = (b(1+g_b), \hat{q}(1+\lambda)/(1+g))$ and $\hat{z}^- = (b(1+g_b), \hat{q}/(1+g))$, implying $x = \frac{\beta(1-\delta)}{\bar{s}b\hat{q}}\left(\hat{v}(\hat{z}^+) - \hat{v}(\hat{z}^-)\right)$. Using our guess for the value function, $\hat{v} = \mathcal{A}b\hat{q}$, yields

$$x = \frac{\beta(1-\delta)}{\bar{s}(1+g)}(1+g_b)\lambda\mathcal{A}. \tag{A2}$$

Using (A2) in (A1), equating it with our guess that $\hat{v} = \mathcal{A}b\hat{q}$ and solving for $\mathcal{A}$ as the positive real solution to the resulting quadratic equation implies

$$\mathcal{A} = \left[\left(\frac{1}{4\upsilon} - \bar{s}x^2/2\right)\right]\left[1 - \frac{1+g_b}{1+g}\beta(1-\delta)\left(1+\lambda x\right)\right]^{-1} > 0, \tag{A3}$$

where $x = x(\mathcal{A})$ is defined in (A2). Finally, let $\widetilde{\beta} = \beta\frac{1-\delta}{1+g}$ and $\overline{\beta} = \overline{\beta}(x) = \frac{1}{1-\widetilde{\beta}(1+\lambda x)(1+g_b)}$. Then, taking g as given in partial equilibrium (PE) and using (A3), we totally differentiate (A2) – noting that $\frac{\partial\mathcal{A}}{\partial x} = 0$ by the envelope theorem. Collecting terms gives the following

$$\epsilon_{x,g_b}|_{PE} = \frac{\mathrm{d}x}{\mathrm{d}g_b}\frac{g_b}{x}|_{PE} = \frac{g_b}{1+g_b}\left(1 + (1+g_b)\overline{\beta}\widetilde{\beta}(1+\lambda x)\right) > 0.$$

Doing the same, but accounting for the fact that $g = \lambda x$ in general equilibrium (GE), we get

$$\epsilon_{x,g_b}|_{GE} = \frac{\mathrm{d}x}{\mathrm{d}g_b}\frac{g_b}{x}|_{GE} = \frac{g_b}{1+g_b}\left(\frac{1+g}{1+2g} + \frac{(1+g_b)\beta(1-\delta)}{1-(1+g_b)(1-\delta)\beta}\right) > 0.$$

□

Proof of Proposition 2. For part (a) of Proposition 2, let us first denote the mass of firms of a particular age a with μ_a. Given exogenous exit in the benchmark model, the law of motion for the mass of firms is simply given by $\mu_{a+1} = (1-\delta)\mu_a$. In addition, note that with exogenous customer accumulation and common initial values (b_e), all firms of the same age will also have the same level of customer base, b_a. Therefore, we can write the following

$$\begin{aligned}
Q' =& \sum_{a=0} \int_{j\in\mu_a} b_a q_j' dj = \sum_{a=1} \int_{\hat{q}} b_a \hat{q} Q' \mu_a(\hat{q}) + \int_{\hat{q}^e} b_e \hat{q}^e Q(1+\lambda x)\mu_0(\hat{q}^e) \\
=& \underbrace{\sum_{a=0} \int_{\hat{q}} b_a (1+g_b)\hat{q} Q(1+\lambda x)(1-\delta)\mu_a(\hat{q})}_{\text{surviving firms from previous period}} + \underbrace{\int_{\hat{q}^e} b_e \hat{q}^e Q(1+\lambda x)\mu_0(\hat{q}^e)}_{\text{startups}} \\
=& (1+\lambda x)\left[\sum_{a=1} \int_{\hat{q}} b_a \hat{q} Q \mu_a(\hat{q}) + \int_{\hat{q}^e} b_e \hat{q}^e Q_{-1}(1+\lambda x)\mu_0(\hat{q}^e)\right] = (1+\lambda x)Q,
\end{aligned}$$

where in the last equality we guess that $Q' = (1+\lambda x)Q$, which we subsequently confirm. Therefore, $1+g = Q'/Q = 1+\lambda x$. Finally, using Proposition 1, we have that $\partial g/\partial(1+g_b) = \lambda \partial x/\partial(1+g_b) > 0$.

Part (b) follows directly from the above characterization of the growth rate, we have $\mathrm{d}g/\,\mathrm{d}g_b = \lambda\,\mathrm{d}x/\,\mathrm{d}g_b$. The latter is positive by virtue of Proposition 1.

As for part (c), let us define counterfactual aggregate growth – which results from firms expecting zero demand stock growth while holding all else equal (including expected aggregate growth and firm-level profits) – as $g^c = \lambda x^c$, where x^c is given by Proposition 1 with $g_b = 0$. To consider the upper bound, define $g = \lambda x$, where x is given by Proposition 1 with $g_b = \delta/(1-\delta)$.

Then, $x^c = \frac{\beta(1-\delta)}{1-\beta(1-\delta)} \frac{\widetilde{\pi}\lambda}{\upsilon(1+g)}$ and $x = \frac{\beta}{1-\beta} \frac{\widetilde{\pi}\lambda}{\upsilon(1+g)}$, where $\widetilde{\pi} = 1/(4\upsilon) - \bar{s}x^2/2$. The share of growth accounted for by customer acquisition is then

$$\frac{g-g^c}{g} = \frac{x-x^c}{x} = 1 - \frac{(1-\delta)(1-\beta)}{1-\beta(1-\delta)}.$$

□

Proof of Proposition 3. First, observe that the impact of R&D subsidies τ is isomorphic to the effect of changes in R&D cost scaling $\bar{s}$. To economize on notation, we will derive directly the elasticity with respect to the cost – noting that it has the opposite qualitative effect, i.e. $\partial x/\partial\tau_s = -\partial x/\partial\bar{s}$. Taking g as given and differentiating (A2) with respect to $\bar{s}$ gives:

$$\epsilon_{x,\bar{s}} \equiv \frac{\mathrm{d}x}{\mathrm{d}\bar{s}}\frac{\bar{s}}{x} = -\left(1 + \underbrace{\frac{\bar{s}x^2/2}{1/(4\upsilon) - \bar{s}x^2/2}}_{\text{R\&D expenses-to-profits, } s_p}\right) < 0 \tag{A4}$$

Not surprisingly more expensive R&D (lower subsidy) leads to a lower innovation rate.

Given $\epsilon_{x,\bar{s}}$, we can compute how the magnitude of the elasticity, $|\epsilon_{x,\bar{s}}|$, changes with g_b. In other words, we ask how the responsiveness of the economy to the R&D costs varies with the rate of underlying firm-level demand stock growth. Given the above, we can write:

$$\frac{\mathrm{d}|\epsilon_{x,\bar{s}}|}{\mathrm{d}g_b} = 2\left(2\upsilon x^3\right)^{-1}\frac{\mathrm{d}x}{\mathrm{d}g_b} > 0$$

The above shows that the innovation rate is more responsive to cost changes when customer base growth is faster.

Finally, notice that the elasticity of innovation to R&D subsidies directly depends on the share of R&D expenditures in profits, $s_p = \frac{\bar{s}x^2/2}{1/(4\upsilon)-\bar{s}x^2/2}$. Therefore, the elasticity of innovation with respect to the R&D cost (subsidy) is decreasing (increasing) in the R&D expenses-to-profits ratio, $\partial\epsilon_{x,\bar{s}}/\partial s_p < 0$.

□

A.3 Gibrat's law

Proposition 4 (Firm size growth). *Expected growth of surviving incumbents is*

$$\frac{n'+s'}{n+s}=\frac{1+g_b}{1+g}\left(1+\lambda x\right).$$

Proposition 4 makes explicit that firm size growth depends both on demand stock growth, g_b, and production efficiency increases, $1+\lambda x$. As noted in Proposition 1, the optimal innovation rate x is constant and common to all firms. Therefore, our theoretical framework satisfies Gibrat's law that firm growth is independent of size.[1]

Proof. Firm size is given by $e = n + s$. Using the demand constraint, we can then write the (stationarized) firm size as

$$\hat{e}=b\hat{q}\left[\frac{1}{2\upsilon}+\bar{s}x^2\right], \tag{A5}$$

where the term in square brackets is constant. Firm size is, therefore, proportional to $b\hat{q}$. Hence, we can write the expected (with respect to the innovation probability) firm size growth rate as

$$\mathbb{E}\,\frac{\hat{e}'}{\hat{e}}=\frac{b'\hat{q}'}{b\hat{q}}=\frac{1+g_b}{1+g}\left(1+\lambda x\right) \tag{A6}$$

where the last equality follows because $b'/b = 1 + g_b$ and $\mathbb{E}\,\hat{q}'/\hat{q} = \mathbb{E}\,\frac{Q}{Q'}\frac{q'}{q} = \frac{1+\lambda x}{1+g}$. □

A.4 The planner's R&D allocation

Given that we are not interested in monopolistic distortions as such (recall that in the benchmark model markups are fixed), we assume that the planner is restricted to the decentralized production and pricing choices and only focus on R&D decisions.

Proposition 5 (Socially optimal innovation). *Let stars (*) denote the quantities in the planner's allocation. The socially optimal innovation rates satisfy*

$$\underbrace{2\upsilon\bar{s}x^*}_{\textit{marginal costs}}=\underbrace{\frac{\beta(1-\delta)}{(1+g^*)(1-\beta)}}_{\textit{discounting}}\underbrace{\lambda(1+g_b)}_{\textit{flow benefits}}. \tag{A7}$$

Known R&D distortions. For convenience, we repeat the innovation rates in the market allocation, where we explicitly write out the expression for $\mathcal{A}(x)$:

$$\underbrace{2\hat{W}\bar{s}x(\theta)}_{\textit{marginal costs}}=\underbrace{\frac{\beta(1-\delta)}{(1+g)\left(1-\beta(1-\delta)(1+g_b)\right)}}_{\textit{discounting}}\underbrace{\lambda(1+g_b)\widetilde{\pi}}_{\textit{flow benefits}}, \tag{A8}$$

where $\widetilde{\pi}=1/(4\upsilon)-\bar{s}x^2/2$.

[1] While Gibrat's law has been widely documented for large firms, young and small businesses seem to deviate from it (see e.g. Evans, 1987).

The above incorporates three well known differences between the socially optimal and market innovation rates in endogenous growth models. First, there is a difference in the (marginal) costs of innovation because the individual firm faces wage costs, while the planner takes into account the disutility of labor. This is the so called *monopolistic distortion*, which arises because wages are partly determined by production workers in monopolistically competitive firms. The wage may, therefore, be inefficiently low and thus lead to over-investment into R&D.

Second, the planner and the individual firm differ in their *discounting*. While the planner takes into account that any innovation lasts forever, the individual firm only takes into account its own lifetime. Therefore, because the individual firm discounts the future more heavily, this leads to under-investment in R&D.

Third, the individual firm only takes into account the private flow benefits (implicitly defined through $\widetilde{\pi}\lambda(1+g_b)$), while for the planner the benefit is the contribution to growth ($\lambda(1+g_b)$). This effect is referred to as *limited appropriability*. The fact that the private firm cannot appropriate the entire benefits from its own innovation typically leads to under-investment in R&D.

Note that demand stock growth $(1+g_b)$ shows up directly in the flow benefits of both the planner and the individual firm in the decentralized economy. This simply reflects the fact that firms are characterized by different weights in the aggregate productivity index Q, which the planner also takes into account. However, this direct effect does not drive a wedge between the decisions made by the planner and those carried out in the decentralized economy.

Novel R&D distortion. We now analyze how the presence of firm-level market size growth creates a difference between the privately and socially optimal level of innovation rates. Typically, endogenous growth models exhibit *under-investment* in R&D as the monopolistic distortion discussed above is not strong enough to overturn the other two effects (see e.g. Aghion and Howitt, 1994; Denicolo and Zanchettin, 2014).[2] Our model, however, puts forward a novel channel through which individual firms may over-invest in R&D.

To understand this, notice that the discount rate of individual firms in (A8) is affected by market size growth. In contrast to this, the discounting of the social planner in (A7) does not feature customer base growth. The intuition for this rests on the fact that innovation of individual firms is governed by growth in future profitability, which is partly driven by life-cycle growth unrelated to productivity enhancements. Firm-level customer base growth, therefore, provides an extra boost to the incentives of individual businesses to conduct R&D. In contrast, the social planner internalizes the fact that the aggregate customer base is fixed and focuses only on productivity growth.[3]

We dub this effect the *profitability distortion* because it comes about only if firm profitability is affected by factors other than productivity growth – in our case, by customer accumulation. In principle, the profitability distortion could lead to *over-investment* in R&D if customer base growth is strong enough.[4] But even if this channel is not strong enough, and the economy is characterized by

[2]Our model lacks another distortion, the *business stealing effect* present in models based on the tradition of Klette and Kortum (2004). However, Denicolo and Zanchettin (2014) highlight that business stealing itself cannot lead to over-investment in standard endogenous growth models.

[3]Note that firms (products) with higher customer base growth become more important for welfare as their market share increases. This is reflected in the flow benefits term, but this effect is present in both the planner's allocation (A7) as well as the decentralized economy (A8).

[4]Indeed, when $(1+g_b) > \frac{1}{1-\delta}$, individual firms will discount the future *less* than the social planner. In the benchmark model, with all businesses enjoying the same speed of exogenous customer base growth, $1+g_b$ must in fact be smaller than $1/(1-\delta)$ in order for the aggregate customer base to be stationary. However, in richer environments with (time varying) firm heterogeneity (like our generalized model), businesses may temporarily go through periods in which $1+g_b > 1/(1-\delta)$

under-investment in R&D at the aggregate level, the profitability distortion creates heterogeneity in the extent to which firms under-invest. In particular, firms with high demand stock growth are likely to under-invest less, potentially opening the door to more targeted R&D policies.

Proof. In what follows, we consider a utilitarian planner, who gives equal weight to all workers, in an economy adhering to Assumption 1.

$$\sum_{t=0}^{\infty} \beta^t \left[\ln C_t - \upsilon N_t\right].$$

We start by noting that, on the BGP, the above objective function can be simplified to[5]

$$\left[\ln C + \ln(1+g)\frac{\beta}{1-\beta} - \upsilon N\right],$$

where $C = \left(\int_j b_j^{\frac{1}{2}} c_j^{\frac{1}{2}}\, \mathrm{d}j\right)^2$ and $N = \int_j (n_j + s_j)\, \mathrm{d}j = \int_j \frac{c_j}{q_j} + \frac{b_j q_j}{Q}\bar{s} x_j^2\, \mathrm{d}j$.

Let us first discuss the static decision of allocating consumption varieties c_j in every period. The planner's allocation, denoted by a star, takes the following form

$$c_j^* = b_j \left(\frac{q_j}{\upsilon}\right)^2 \frac{1}{C^*}, \tag{A9}$$

which gives rise to the following aggregate consumption value[6]

$$C^* = \frac{Q^*}{\upsilon}, \tag{A10}$$

where $Q^* = \left(\int_j b_j q_j\, \mathrm{d}j\right)$.

Let us now substitute out the optimal allocation of consumption varieties (A9) in the planner's problem.[7] The planner is left with choosing optimal innovation rates. Given the common growth profiles of the demand stock, this amounts to choosing age- and productivity-specific innovation rates for incumbents, $x_a(\hat{q})$, and for entrants, $x_e(\hat{q}^e)$, where $\hat{q} = q/Q$. The planner solves:

$$\max_{x_a(\hat{q}), x_e(\hat{q}^e)} \left[\ln Q^* - \ln\upsilon + \ln(1+g)\frac{\beta}{1-\beta} - \upsilon N\right],$$

where

$$N = \frac{1}{\upsilon} + \sum_{a=0} \int_{\hat{q}} (1+g_b)^{a+2} \hat{q}\bar{s} x_a(\hat{q})^2 \mu_a(\hat{q}) + (1+g_b) \int_{\hat{q}^e} \hat{q}^e \bar{s} x_e(\hat{q}^e)^2 H_e(\hat{q}^e), \tag{A11}$$

$$1 + g = \frac{Q'}{Q} = \left[\begin{array}{c} \sum_{a=0}(1+g_b)^{a+3} \int_{\hat{q}} \hat{q}(1+\lambda x_a(\hat{q}))(1-\delta)\mu_a(\hat{q}) \\ \\ +(1+g_b)^2 \int_{\hat{q}^e} \hat{q}^e(1+\lambda x_e(\hat{q}^e))(1-\delta) H_e(\hat{q}^e) \end{array}\right]. \tag{A12}$$

---

and, therefore, overinvest into R&D.

[5]Note that $\sum_t \beta^t \ln C_t = \sum_t \beta^t \left(\ln C_0 t \ln(1+g)\right) = \frac{\ln C_0}{1-\beta} + \ln(1+g)\frac{\beta}{(1-\beta)^2}$. In addition, N is stationary and therefore its discounted value is equal to $N/(1-\beta)$. Finally, we denote C_0 as C since the composition of the consumption good is fixed over time.

[6]We use (A9) in the definition of the consumption bundle to express C^*.

[7]We use the fact that $C^* = Q^*/\upsilon$ to write aggregate labor supply as $N = \int_j c_j^*/q_j + s_j\, \mathrm{d}j = 1/\upsilon + \int_j s_j\, \mathrm{d}j$.

The resulting socially optimal innovation rates are given by

$$x_a(\hat{q}) = x_e(\hat{q}^e) = \frac{1}{2\bar{s}\upsilon}\frac{\beta(1-\delta)}{(1+g)(1-\beta)}(1+g_b)\lambda. \tag{A13}$$

B Generalized model: Solution and simulation

In this appendix, we provide more information on the solution and estimation of the generalized model. In addition, we also describe details of the "counterfactual" R&D choices used to isolate the influence of customer acquisition on model results.

B.1 Solution method

Recall from Section 4 that instead of parameterizing the exogenous component of firms' customer base, we parameterize the overall customer base and later back out the endogenous and exogenous contributions. Therefore, as noted in the main text, firms face four state variables, $z_j = (b_j, \hat{q}_j, \theta_j, \chi_j)$: firm-level customer base, b_j, relative productivity, $\hat{q}_j$, transitory shocks, θ_j, and the joint customer base life-cycle growth factor, χ_j.

Grid for firm-level state variables. We discretize the four dimensions of the state space using 41, 101, 21, and 31 points, respectively. The grid is equidistant in the log of variables. These relatively large grids allow us to capture the full extent of the firm life-cycle dynamics induced by the evolution of the demand stock.

The grid for $\ln b$ spans the interval $[-5, 7]$. The choice of the upper and lower bound is dictated by the desire to replicate the empirical size distribution of firms (see Section 4.3 for details).

We opt for a relatively dense productivity grid as we want to capture adequately the discrete productivity jumps by factor $1+\lambda$ while at the same time allowing for cross-sectional productivity dispersion consistent with the empirical evidence. The resulting equidistant grid for log productivity ranges from -0.5 to 0.99 and generates TFPQ and TFPR dispersion in line with existing literature.

The grid for χ spans the interval $[-3, 4]$, i.e. includes shrinking as well as growing firms, chosen to capture the rich pattern of log size autocovariance in the data, as illustrated in Figure 1.

Finally, we discretize the AR(1) process for $\ln \theta$ using the Rouwenhorst (1995) method.

Solution algorithm. The quantitative model is challenging to solve for two main reasons. Firstly, firm's pricing and R&D investment problems are highly non-linear calling for global solution methods. Second, due to the endogenous customer base accumulation, both the pricing and innovation decisions are forward looking and lack closed form solutions.

To solve the quantitative model numerically, we use value function iteration nested inside iteration on the markup and employment policy functions. The following algorithm describes our solution strategy:

1. guess optimal markups, firm values and production employment as a function of state variables, $\mu(z)$, $v(z)$ and $n(z)$, respectively

2. solve firm values, $v(z)$, using fixed point iteration

a) given $\mu(z)$ and $x(z)$ (which is implied by our guess of $v(z)$ and $n(z)$), compute the continuation value $\tilde{v}_j = \max\left[0, \mathbb{E}_\theta v_j - W\phi\right]$, where $\mathbb{E}$ is computed using our assumption on transitory shocks, θ.

b) $\mu(z)$ together with $x(z)$ can then be used to compute firm employment $n(z)$ and $s(z)$, allowing us to compute a new guess for firm value using (20)

c) using the new guess of firm value, we go back to 2a) and repeat this part of the loop until convergence

3. solve for optimal markups, $\mu(z)$, using the pricing condition (19)

4. using the new value of markups, go back to 1 and iterate until convergence.

Note that in the generalized model we normalize wage W and consumption $\hat{C}$ to unity, using the disutility of labor and mass of potential entrants as normalizing constants. The remaining equilibrium variable - the aggregate growth rate g - is one of the estimated targets described in Section B.2 below.

B.2 Simulation

To estimate moments pertaining to firm lifecycle dynamics or aggregate economic growth, we need to approximate the stationary firm distribution. Towards this end, we use on-grid stochastic simulation.

Simulating an unbalanced panel of heterogeneous firms. We draw potential entrants from the initial distribution $H_e(z_e)$ over the idiosyncratic states. We determine which entrants continue operating based on their exit decision as described Section 3.1. We draw new potential entrants up to a point in which we have 40 thousand surviving startups.

From then on, we simulate the lifecycle of each firm for 51 years. This simulation length is roughly in line with the empirical sample length in the BDS and Compustat datasets. Moreover, as the customer base exhibits decreasing marginal returns to scale and R&D investment become increasingly expensive as firms become larger, their growth rate eventually declines and there is little to be gained in terms of precision when allowing for a longer lifespan.

In each period, firms face transitory, exit and innovation shocks, and make decision regarding pricing, employment, investment into R&D, and exit. Based on these choices, firms move between grid points along three dimensions of the statespace (the firm-specific customer base growth factor χ is determined at entry and fixed over the firm's life). In addition, at each period and for every firm, we track the endogenous and exogenous components of firms' customer bases.

While the firm is alive, in each period we simulate the following steps. At the beginning of each period, before any shock is realized, firms decide whether to pay the fixed cost and continue operating or avoid paying the cost and exit. Firms that decided to continue operating receive transitory shocks following our AR(1) specification in (27) and observe the realization of the innovation shock determining their current period productivity. The probability of a successful innovation was determined by R&D expenses in the previous period. Next, firms decide on investment into R&D based on their first-order condition and set the optimal markup. These decisions imply certain employment and profits. Then, the firm transitions into the next period and the procedure repeats.

B.3 Isolating the influence of customer acquisition

Finally, let us describe the computational details of the counterfactual R&D decisions which allow us to decompose results from our generalized model into the contribution of customer acquisition. Importantly, note that when computing the counterfactual decisions, all structural parameters and equilibrium variables remain at the values estimated in the baseline economy. Therefore, our approach offers a partial equilibrium decomposition. Moreover, the solution and simulation algorithms as well as the state space discretization method are all intact.

Counterfactual firm values. In addition to firms' optimal decisions, we also solve firms' counterfactual R&D decisions assuming that firms' believe their customer bases to be fixed, i.e. $b' = b$. This means that firm value can be written as

$$v(b, \hat{q}, \theta, \pi) = \max_{p,n,s} \left\{ \begin{array}{c} pc - W(n+s) + \\ \frac{1-\delta}{1+R'} \Big[x\tilde{v}(b, \hat{q}(1+\lambda), \theta', \pi') + (1-x)\tilde{v}(b, \hat{q}, \theta', \pi') \Big] \end{array} \right\}.$$

Given the formula for the firm value above, we solve for firm values in the same way as described above. This provides us with the desired "counterfactual" firm-level R&D decisions.

Simulating using counterfactual R&D decisions. When simulating the economy we let the firm-level state variables follow their respective "true" laws of motion. In other words, at *each point* of the simulation, we compute firms' optimal decisions (as described in the previous paragraphs) as well as firms' counterfactual decisions. This allows us to compare – for the same values of the state-variables – optimal and counterfactual decisions.

Note that when computing counterfactual counterparts, we *only* use firms' counterfactual R&D decisions. In particular, we consider the same set of firms as in the baseline (i.e. we do not use the exit rule implied by firms counterfactual firm values), nor do we consider counterfactual pricing decisions.

Finally, we can use the counterfactual R&D decisions (and thus productivity values) to compute the counterfactual transition of firm-level customer base using (16). This enables us to speak of the "reallocation" of market activity across heterogeneous firms as a result of differences in firm-level R&D between the baseline and the counterfactual case.

C Generalized model: Estimation

In this section we provide more details on the procedure used to estimate the link between firm-specific demand stock growth and subsequent productivity improvements.

General strategy. As described in the main text in Section 4, we estimate estimate our model using indirect inference. Towards this end, we make use of the unbalanced panel of firms described above. In particular, we minimize the Euclidean distance between 253 moments in the model and in the data. Each moment is equally weighted, with the exception of the equilibrium growth rate, g, to which we assign a value five times as large as all other moments.

Our targets can be grouped into four sets: (i) average growth of real GDP and average R&D to sales ratio (2 moments); (ii) the regression coefficients of sales-to-cost and R&D-to-sales ratios on firm age (2 moments); (iii) the firm size life-cycle profile (20 moments); (iv) the firm exit life-cycle profile

(19 moments); and (v) the upper triangle of the autocovariance matrix of log-employment, by age and for a balanced panel of firms surviving up to at least the age of 19 years (210 moments).

C.1 Data

In this subsection, we describe the data sources used in the parameterization of our model.

Business Dynamic Statistics and other sources. Our model matches the life-cycle profiles of firm exit rates, sizes (employment) and the autocovariance structure of log-employment. While based on BDS data, the latter is not reported by the Census directly. However, we take these targeted values from Sterk ⓡ al. (2021), who used the Longitudinal Business Database – the micro-data underlying the BDS – to compute the autocovariance structure. The moments are provided on their websites. Finally, consistent with the latest BDS sample, we use real GDP per capita growth between 1979 and 2019 as a measure of aggregate growth.

Compustat data. As the main data source with firm-level information we use Compustat. In particular, we use the Compustat fundamental annual data from 1950 to 2019. All nominal variables are deflated with the GDP deflator.

We exclude observations of financial firms (SIC classification between 6000 and 6999) and utilities (SIC classification 4900-4999).[8] We consider only companies incorporated in the U.S. and traded in the U.S. stock exchange. We remove observations with clear measurement errors, such as negative sales (variable `sale` in Compustat) or negative costs of goods sold (`cogs`). Furthermore, to make sure that outliers are not driving our results, we remove observations for which values of sale to COGS ratio or the estimated markup exceed either top or bottom 1% of values. The percentiles are calculated individually for each year.

R&D intensity is defined as the ratio of the R&D expenses, `xrd` in Compustat, relative to `sales`. Again, to remove outliers, we drop top 1% percentile of annually calculated R&D intensity distribution (we remove only top percentile since in the bottom 1/5th of observations R&D intensity is zero). We then remove firms with R&D intensity above 1000% if any are left after the percentile trimming. In all Compustat variables that we use, we interpolate linearly missing values at the firm level. We do not interpolate if there are more than one consecutive missing values. After the cleaning procedure, the data sample consists of 12,864 unique firms and 119,682 firm-year observations.[9] When estimating the model parameters via the indirect inference, we apply the same sample restriction to the model-simulated data.

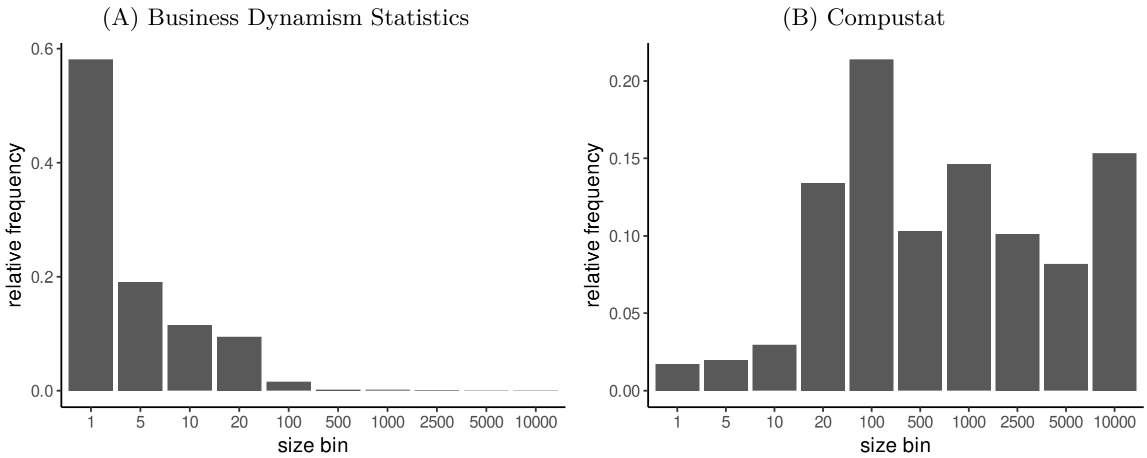
We use the following financial ratios as controls in firm level regressions. The liquidity ratio is defined as debt to current assets, (`dltt` + `dlc`)/`act`. Leverage is defined is as debt to equity, (`dltt` + `dlc`)/`ceq`. Dividend is the value of Compustat variable `dvt` deflated with the GDP deflator. The sample age in compustat is defined as the difference between the current fiscal year `fyear` and the first year of reporting balance sheet information, `year1`, which can happen prior to the IPO. For a subset of firms, we can compute the "true" age based on firm's founding date in the updated Loughran and Ritter (2004) database.[10] The advertising expenses correspond to variable `xad` in Compustat.

[8] As is standard in the literature, we remove financial firms because the balance reporting conventions are very different for those companies. We delete utilities because it is a non-representative and regulated sector.

[9] The degrees of freedom in regressions we run tend to be lower due to additional restrictions or missing data.

[10] We are able to match the founding date to 26% of firms in our sample. See Section C.3

Figure A1: Firm size distribution in BDS and Compustat



Note: The figures present relative frequency of firm employment count. BDS covers the universe of firms in the US. Compustat includes only public firms. The x-axis label for each size bin corresponds to left bin edge, i.e. the bins are 1 to 4, 5 to 9, 10 to 19, etc.

C.2 Construction of size-based Compustat weights

Compustat sample consists only of publicly listed companies that differ significantly from universe of firms in the US (i.e. the BDS data and our model, which matches the latter well, see Figure 2 in the main text) when it comes to size. To tackle this issue, we construct a set of firm-specific weights based on the firm size (employment).

Figure A1 illustrates the differences in size distribution between BDS and Compustat. In the latter, the mass of firms is shifted towards the largest firms.

Aligning model-simulated data with Compustat information. Formally, we pool all observation across all years and calculate the empirical distribution of firm size in Compustat. Let f_e denote the frequency of firm size e in Compustat.

Given that our model is parameterized to business dynamics from the BDS, we need to realign the model distribution to that of Compustat whenever comparing model moments to those which have counterparts in Compustat. Towards this end, in any "Compustat-related" regression in the model we weigh the model-simulated data with the respective empirical weights from Compustat, f_e.

C.3 Descriptive statistics

When estimating the life-cycle profiles of R&D expenses and cost shares, we rely on the number of years since IPO as a proxy for firm age. However, our results are fully robust (if not stronger) when using the actual age based on the firm founding dates. For a subset of firms, the updated Loughran and Ritter (2004) database contains information on founding dates, for a subset 3,372 out of 12,864 firms, or 26% of our sample.

In this section, we compare the characteristic of firms in the whole Compustat sample against the subset of firms which we are able to obtain the founding date. The firms matched to their founding dates tend to be smaller which suggests that the full sample consists of older firms on average. Figure A2 plots the distribution of the founding dates in Loughran and Ritter (2004) database. Indeed, most of the mass of firms is located after 1980s. This highlights the importance of controlling for size when using the 'true' age in regressions.

Table A1: Firm characteristics: full and "founding date" samples

variable	non-missing obs.	mean	st. dev.
		Full sample	
sales (mln \$)	119,682	2,832.712	14,822.799
employment (1000s)	111,623	10.47	46.189
R&D intensity	119,682	0.216	0.771
R&D expenses (mln \$)	119,682	84.685	505.375
leverage	104,815	1.676	41.448
		Founding date sample	
sales (mln \$)	23,184	1,181.597	10,444.086
employment (1000s)	20,732	5.119	26.402
R&D intensity	23,184	0.333	0.952
R&D expenses (mln \$)	23,184	56.255	438.641
leverage	20,100	1.561	45.369

Note: The Table presents sample moments for all firms in our compustat sample and subsample for which we are able to match their founding dates using Loughran and Ritter (2004) database. R&D intensity is defined as the R&D expenses to sales, $\frac{\texttt{xrd}}{\texttt{sale}}$. Leverage is defined is as debt to equity, $\frac{\texttt{dltt+dlc}}{\texttt{ceq}}$.

D Generalized model: Further details, results and extensions

In this appendix, we provide further details regarding the generalized model, including additional model results.

D.1 Equilibrium definition

We begin by providing a formal equilibrium definition of our generalized model.

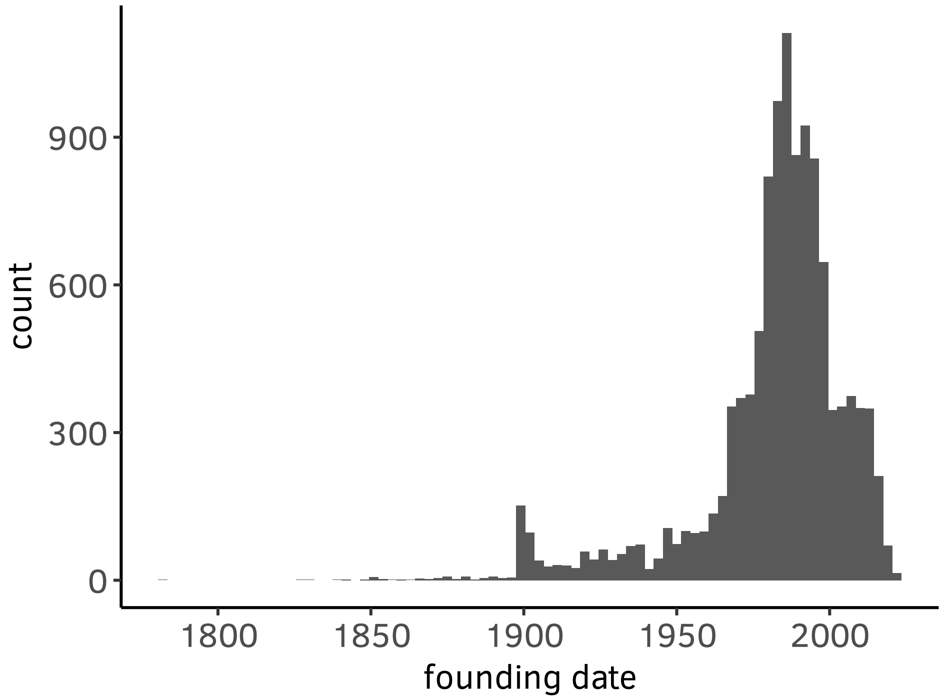
Generalized model: Equilibrium definition. *A balanced growth path equilibrium of our generalized model consists of the following tuple in every period with $j \in \Omega$: c_j, d_j, b_j, θ_j, q_j, p_j, μ_j, x_j, n_j, s_j, $v(q_j, b_j)$, z_j^*, r, W, C, Q, g, such that (i) output, c_j, satisfies (5), (ii) firms' demand stock, d_j, is given by (17) and firms' customer bases and productivity levels, b_j and q_j, evolve according to (16) and (8), (iii) prices are given by $p_j = \mu_j W/q_j(1 + \sigma s_j/n_j)$, where markups, μ_j are satisfy (19), (iv) optimal innovation probabilities, x_j, satisfy (13), (v) labor demand, n_j and s_j, satisfy (6) and (7), where $\bar{s}_j = \bar{s} n_j^\sigma$ (vi) firm values, v, satisfy (20), and the exit threshold, z_j^*, is determined by (21), (vii) the interest and wage rates, r and W, satisfy (3) and (14), (viii) aggregate consumption, C, is defined by (1), and (ix) the aggregate productivity index, Q, is given by (15) and it's growth equals $1 + g = Q'/Q$ where in all the above it is understood that firms' state variables are given by $z_j = (b_j, q_j, \theta_j, \pi_j)$.*

D.2 Impact of transitory shocks

The main text focuses on the impact of customer base accumulation. We now turn to showing how transitory shocks affect the results. We begin by inspecting their influence on firm-level innovation rates.

Table A2 reports innovation rates in the baseline and the respective contributions of customer acquisition as in the main text (Panel A). In addition, we show the same exercise when assuming that firm also believe their transitory shocks, θ, to be fixed in the future. Notice that the presence of transitory shocks *reduces* innovation rates (compare Panel A – with transitory shocks – to Panel B – with fixed transitory shocks). The effect is, however, negligible for young businesses for which customer base dynamics are the dominant force.

Figure A2: Founding dates of US public firms in Loughran and Ritter (2004) database



Note: The figure presents the frequency of founding dates in Loughran and Ritter (2004) database.

The reason for this is that positive transitory shocks are in fact associated with negative expected demand stock growth in the future. This is due to their autoregressive nature, as explained in the main text. Negative transitory shocks do not fully compensate the effect of positive ones because of endogenous firm selection. Businesses are relatively more likely to shut down when their sales are low. Therefore, firm exit endogenously skews the distribution of transitory shocks towards positive ones, driving down firm innovation rates.

D.3 Selection on "productivity" or "demand"?

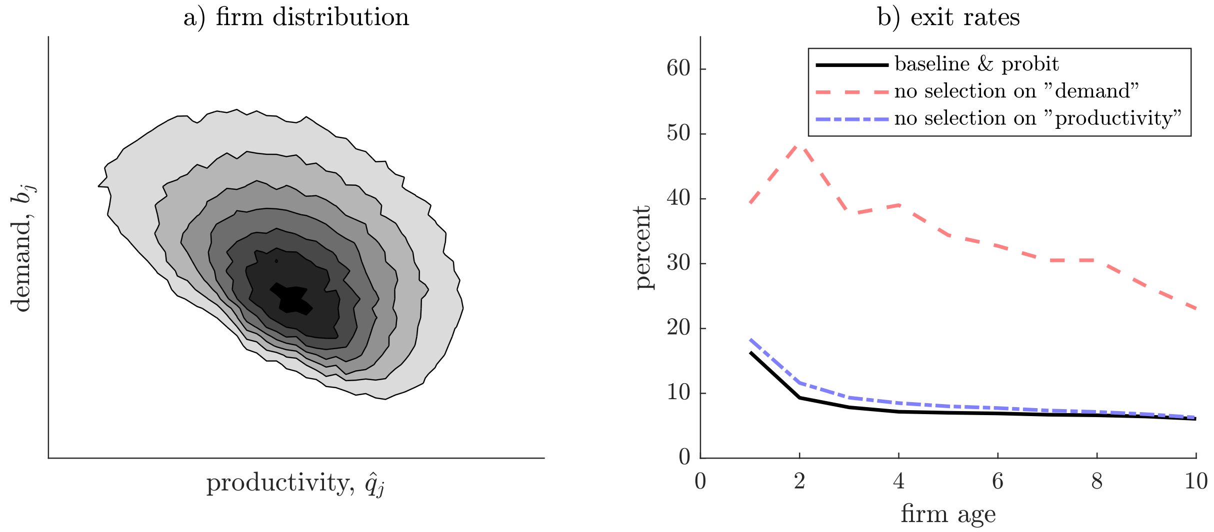
The decision to remain in operation depends on firms' expected discounted profits. Unlike in existing endogenous growth models – in which firm-level productivity is the sole source of variation in profits

Table A2: Decomposition of firm-level innovation rates

		Contribution of customer acquisition	
	Baseline	Percentage points	Share of baseline
		A: Fixed customer base only	
All firms	6.92	1.31	0.19
Young firms	7.51	3.13	0.42
		B: Fixed customer base and transitory shocks	
All firms	6.92	1.82	0.26
Young firms	7.51	3.11	0.42

Note: The first column shows "baseline" innovation rates, x, for all firms (top row) and young business not older than 5 years (bottom row). The second and third columns report the contributions of customer acquisition to innovation rates in "percentage points", i.e. $x - \underline{x}$, and as a "share of baseline" values, i.e. $(x - \underline{x})/x$. All values are in percent. Panel A shows values when firms believe their customer bases to be fixed (as in the main text) and panel B shows the same but where firms believe both their customer bases and transitory shocks, θ, to be fixed.

Figure A3: Productivity and demand distribution in model



Note: Panel a) of the figure shows the distribution of demand and (relative) productivity in the simulated model. Darker shades indicate densely populated areas. White areas indicate areas of the state space which are not populated. Panel b) shows exit rates in the baseline (and the age-dependent probit which coincides with it), and two counterfactuals. One in which demand happens only on productivity ("no selection on demand") and one in which selection happens only on demand ("no selection on productivity").

– in our framework firms' profits and, in turn, their survival depend on firms' demand stocks.

In this appendix, we provide evidence that our generalized framework is consistent with empirical patterns concerning firm selection, entry and exit. Towards this end, we follow the existing literature which estimates "revenue based productivity" (TFPR) and associated "demand shocks" and analyzes their respective importance for firm churn.

Productivity and demand stocks of entering and exiting firms. The left panel of Figure A3 replicates information in the main text and shows the distribution of demand, b_j, and (relative) productivity, $\hat{q}_j$, from the simulated model. Darker shades indicate more densely populated parts of the state-space.

Since firms select on firm values and firm values are a combination of demand and productivity, there is a clear negative relationship between the two among surviving incumbents. Intuitively, if firms enjoy high demand for their product, they manage to survive despite having lower relative productivity and vice versa.

The inverse relationship between productivity and demand has profound implications for selection over firms' life-cycles. To understand this, we estimate the following regression

$$y_{j,t} = \alpha + \beta_{en} \mathbb{1}_{\text{entry}} + \beta_{ex} \mathbb{1}_{\text{exit}} + \omega_{j,t}, \tag{A14}$$

where $y_{j,t}$ is our variable of interest – either log customer base, $\ln b_{j,t}$, or relative productivity, $\ln \hat{q}_{j,t}$, of incumbent firm j in period t. The variables $\mathbb{1}_{\text{entry}}$ and $\mathbb{1}_{\text{exit}}$ are indicator functions equal to 1 if firm j in period t is an entrant or a firm that shuts down, respectively. Therefore, the estimated coefficients can be interpreted as percentage differences in demand stocks or productivity of entrants or exiting firms relative to incumbent businesses.

Table A3 shows estimates of β_{en} and β_{ex}. The results suggest that entering firms are characterized by relative productivity which is about 4 percent higher than that of incumbents. On the other hand, firms which decide to shut down have relative productivity about 4 percent below that of incumbents.

Table A3: Productivity and demand stocks: entering and exiting firms

	entry	exit
productivity	0.042	−0.029
demand stock	−1.202	−0.566

Note: The table shows estimates of coefficients on indicator variables for "entry" and "exit" in regression (A14) in the main text. Demand stock stands fo $\ln d_{j,t}$ and productivity stands for $\ln \hat{q}_{j,t}$.

In contrast, both entering and exiting firms have demand stock levels which are below the average incumbent. Entrants, however, are burdened by a substantially lower demand stock level compared to incumbents. Importantly, both the model-predicted productivity and demand stock patterns of entrants and exiters are consistent with empirical evidence (see e.g. Foster et al., 2008).

Productivity- or "demand-driven" selection? To quantify the extent to which firm selection is driven by productivity or "demand" we estimate the following probit model for firm exit

$$Pr(exit_{j,a} = 1 | b_{j,a}, \hat{q}_{j,a}) = \Phi(\beta_{0,a} + \beta_{d,a} \ln d_{j,a} + \beta_{q,a} \ln \hat{q}_{j,a}), \tag{A15}$$

where $exit_{j,t}$ is an indicator function equal to 1 if business j of age a decides to shut down (i.e. is out of operation in the next period). The right panel of Figure A3 shows that the age-dependent probit model (which mimics the exit pattern in the baseline) together with two counterfactual exit rates. "No selection on demand" is average firm exit predicted by the estimated probit model when we ignore the effect of customer acquisition on exit decisions, i.e. setting $\beta_{d,a} = 0$ for all ages. In contrast, "no selection on productivity" represents firm exit predicted by our probit model when we ignore the effect of productivity on exit decisions, i.e. setting $\beta_{q,a} = 0$ for all a.

The results show that customer acquisition is a dominant force when it comes to firm selection, consistent with empirical evidence (see e.g. Foster et al., 2008). While selection on productivity also plays a role, its quantitative effect is much weaker. Therefore, ignoring the impact of customer accumulation on exit decisions – as is implicitly assumed in existing endogenous growth models – would paint a very different picture of the process of firm selection.

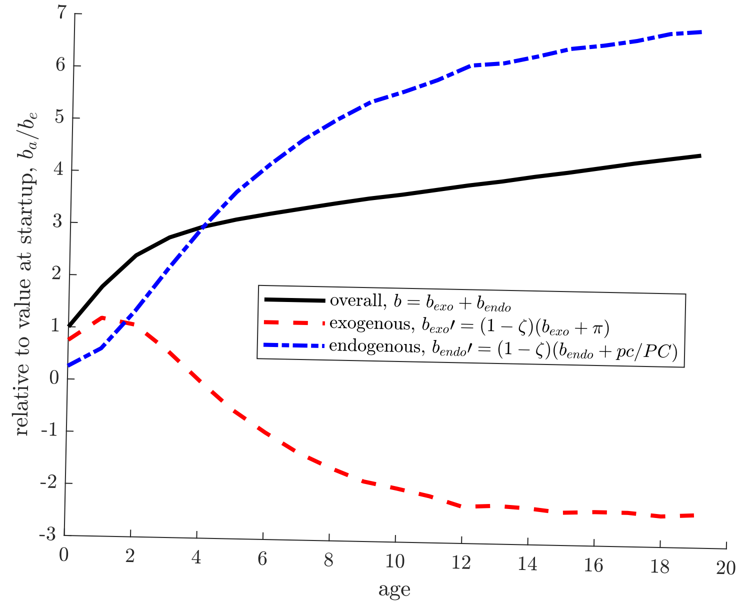
Endogenous and exogenous customer acquisition. When parameterizing our generalized model, we treat the exogenous customer accumulation factors, π_j, as non-parametric wedges between overall customer base, $b_j = b_{endo,j} + b_{exo,j}$, which we parametrize and endogenous customer base, $b'_{endo,j} = (1 - \zeta)(b_{j,endo} + p_j c_j/(PC))$, which we solve for using firms optimal pricing decision.

Figure A4 show she time-path of average customer base in our generalized model and decomposes it into the respective endogenous and exogenous contributions. As can be seen from the figure, with the exception of the first two years, the exogenous component serves only to slow down firms' customer accumulation. As noted in the main text, one interpretation of this pattern is that it captures the presence of other frictions (e.g. financial or informational) which only slow firms' growth.

D.4 Details on the restricted "productivity-only" model

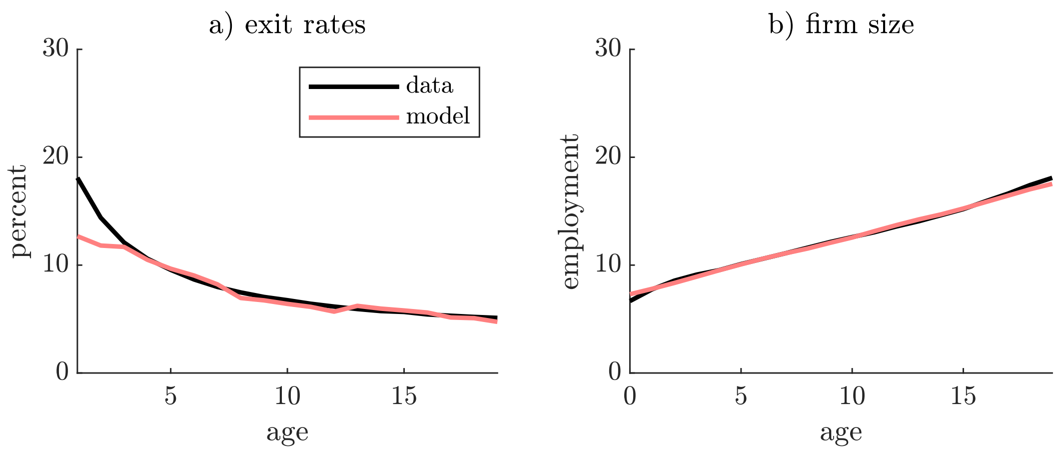
As mentioned in the main text, the restricted "productivity-only" version of our baseline model is obtained by assuming $\sigma_z = \sigma_\epsilon = \sigma_{\overline{u}} = \sigma_u = \sigma_v = 0$. The latter results in firm-level demand stocks being common and constant across all businesses.

Figure A4: Average customer base: overall, exogenous and endogenous



Note: Average customer base accumulation and the respective endogenous and exogenous components.

Figure A5: Model fit of restricted "productivity-only" model



Note: The Figure shows average firm size (employment) and exit rates by age in the restricted model and the BDS data.

We parametrize the restricted model in exactly the same way as the baseline, with the exception that we do not target the autocovariance structure of employment. This is because the productivity-only model – which lacks ex-ante heterogeneity – would not be able to match the empirical patterns as argued in Sterk ⓡ al. (2021).

Finally, in order for the model to match the steep decline in exit rates by age, we allow for the mean of the initial draws of productivity μ_q to differ from 0. Table A4 shows the model parameters and Figure A5 depicts the model fit.

D.5 Advertising expenditures: Theory and data

Our generalized model in the main text builds on empirical evidence in Foster et al. (2016) and considers the pricing theory of customer accumulation. In this appendix, we extend our generalized model and allow firms to grow their customer base not only through pricing but also via direct expenditures on advertising. Importantly, we show that this extension does not fundamentally change

Table A4: Parameter values in the restricted "productivity-only" model

parameter		value	parameter		value
A: Parameters set externally and normalizations					
β	discount factor	0.97	κ	entry cost	0.23
η	elasticity of substitution	6.00	ψ	innovation elasticity	2.00
υ	disutility of labor	1.00			
B: Parameters set via indirect inference					
λ	innovation step size	0.146	ρ_b	$\ln b$, persistence	0
$\overline{s}$	R&D efficiency	4.134	σ_b	$\ln b_e$, standard dev.	0
σ	R&D cost-size elasticity	1.205	μ_χ	χ, mean	3.608
γ	demand pass-through	0	σ_χ	χ, standard dev.	0
ϕ	fixed cost of operation	0.723	ρ_θ	$\ln \theta$, persistence	0
δ	exogenous exit rate	0.020	σ_ϵ	ϵ, standard deviation	0
σ_q	$\ln q_e$, standard dev.	0.199	μ_q	$\ln q_e$, mean	-0.506

Note: Panel A provides values for parameters set externally to our "productivity-only" model as described in the main text. Panel B reports values of parameters set via indirect inference by jointly matching the lifecycle profiles of size and exit, the R&D output share, aggregate output growth and the R&D intensity-to-size relationship.

the model-implied markup dynamics which are central to our results in the main text and would only strengthen endogenous customer accumulation – the key novel channel in our model. Finally, we also relate our model results to recent empirical studies of customer accumulation.

Firm problem. In what follows, we focus only on the firm problem. Household choices and aggregation remains the same as in the main text, even after allowing for advertising expenditures, $\widetilde{a}$. The only exception is labor market clearing which – with the addition of advertising expenditures – now also includes labor used in advertising.

Specifically, we consider a setting in which a firm's customer base evolves according to

$$b_j' = (1-\rho)\left(b_j + \frac{p_j c_j}{PC} + \alpha a_j\right), \tag{A16}$$

where α is the customer accumulation efficiency of advertising relative to the pricing channel and where a is the impact of advertising expenditures on the customer base. The latter comes at a convex cost, $\widetilde{a} = \overline{a} a^{\zeta} n^{\sigma_a}$, where $\overline{a} > 0$, $\zeta > 1$ and where we assume – analogously to R&D costs (see (18)) – that advertising costs scale with employment with an elasticity of $\sigma_a > 0$.

As in the generalized model, firms maximize current and all future discounted profits by choosing prices, production employment and now also employment in advertising. Firm value, conditional on staying in the market, is given by

$$v_j = \max_{p,n,s,a}\left[p_j c_j - W(n_j + s_j + \widetilde{a}_j) + \frac{1-\delta}{1+R}\mathbb{E}\left\{x_j \widetilde{v}_{j,+}' + (1-x_j)\widetilde{v}_{j,-}'\right\}\right],$$

where, as before, $\widetilde{v}_j = \max\left[0, \mathbb{E}_\theta v_j - W\phi\right]$.

Optimal decisions: Markups. Note that in the setting extended with advertising expenditures, marginal costs are given by $W/q_j(1 + \sigma s_j/n_j + \sigma_a \widetilde{a}_j/n_j)$. However, prices are still set as a (variable) markup over marginal costs and the optimality condition for markups remains the same as in the generalized model in (19). For convenience, we repeat it here:

$$\frac{1}{\mu_j} - \frac{1}{\overline{\mu}} = \beta(1-\delta)(1-\rho)\mathbb{E}\left[\frac{1}{\mu_j'} - \frac{1}{\overline{\mu}} + \frac{\gamma}{\eta}\frac{1}{\mu_j'}\frac{m_j'}{b_j'}\right] \tag{A17}$$

Intuitively, the firm balances the marginal benefits from additional customers against the related marginal costs. The marginal costs and benefits of expanding the customer base using pricing can be, respectively, written as follows

$$\frac{\iota_j(1-\rho)}{C} = \frac{\overline{\mu} - \mu_j}{\mu_j}, \tag{A18}$$

$$\frac{\iota_j(1-\rho)}{C} = \frac{1-\delta}{1+R}(1-\rho)\frac{C'}{C}\left[\begin{array}{c} \gamma\frac{m_j'}{b_j'}\left(1-\frac{1}{\mu_j'}\right) \\ \\ +\frac{\iota_j'(1-\rho)}{C'}\left(1+\gamma\frac{m_j'}{b_j'}\right)\end{array}\right], \tag{A19}$$

where ι_j is the lagrange multiplier on the law of motion for the customer base. From (A18) the marginal costs of expanding the customer base using pricing are related to the "markup gap", the difference between a firm's markup and the static optimum. The marginal benefits in equation (A19) are composed of the (marginal) increase in profits due to a higher customer base and the savings resulting from a reduced markup gap. These expressions will be useful in understanding optimal advertising decisions which we turn to next.

Optimal decisions: Advertising. Optimal advertising is determined by analogous conditions as pricing decisions. In particular, firms trade off marginal costs from higher advertising expenditures against the marginal benefits of expanding the customer base. Importantly, the latter is exactly the same as in the case of the pricing channel described in (A19). In contrast, marginal cost of advertising are given by

$$\frac{\iota_j(1-\rho)}{W} = \frac{\zeta}{\alpha}\frac{\widetilde{a}_j}{a_j}, \tag{A20}$$

where, given our parametrization, $W = C$ and therefore (equating (A18) and (A20)) firms optimally choose advertising consistent with the following "no-arbitrage" condition between advertising and markups

$$\frac{\zeta}{\alpha}\frac{\widetilde{a}_j}{a_j} = \frac{\overline{\mu} - \mu_j}{\mu_j}. \tag{A21}$$

The relative efficiency of advertising, α, plays a key role in determining the strength of the different channels through which to accumulate customers. If the efficiency of advertising is zero, $\alpha = 0$, then firms do not use advertising to capture customers. Conversely, with high value of α, firms use advertising relatively more compared to markups in order to attract new customers.

Implications of advertising for markup dynamics: Approach. In our model, a key parameter governing incentives to accumulate customers is γ – the pass-through from customers to demand. In the main text, we parameterize γ using the prediction that markups are increasing over firms' life-cycles (see (A17)). The strength of such an increase in markups is estimated using the sales-to-costs ratio. One may worry that not allowing firms to accumulate customers via advertising (as in the main text) skews our estimate of γ. In what follows, we show that this is in fact not the case and that the implied life-cycle profile of markups is effectively unchanged if firms also accumulate customers using advertising.

Therefore, abstracting from advertising expenditures – as we do in the main text – does not affect our estimate of γ which governs the pricing-channel of customer accumulation. Instead, allowing businesses to also accumulate customers via advertising would only strengthen endogenous customer acquisition – the key novel channel of our paper.[11]

To make our point, we proceed in three steps. First, we theoretically define model-implied markups in our two models: (i) "without advertising", as in our generalized model in the main text and (ii) "with advertising", the model extension considered in this Appendix. We do so again by focusing on observed sales to costs, though in the latter case we explicitly account for advertising expenditures in the data. Second, we estimate how advertising expenditures scale with firm size, σ_a. This is a necessary input to estimating implied markup profiles. Third, using the previous two steps, we estimate the life-cycle profile of markups in our two models.

Implications of advertising for markup dynamics: Theory. Let us now consider what the presence of firm-level advertising expenditures implies for such model-implied markup dynamics and, therefore, our parameterization of γ. Towards this end, we focus again on the sales to cost ratio which in the presence of advertising expenditures is now given by:

$$\frac{py}{W(n+s+\widetilde{a})} = \frac{p}{\frac{W}{q}\left(1+\frac{s}{n}+\frac{\widetilde{a}}{n}\right)} = \mu_{\text{ads}}\frac{1+\sigma\frac{s}{n}+\sigma_a\frac{\widetilde{a}}{n}}{1+\frac{s}{n}+\frac{\widetilde{a}}{n}}. \tag{A22}$$

where we refer to μ_{ads} as the model-implied markup "with advertising." We can now compare the life-cycle profile of μ_{ads} to that of markups implied in our generalized model "without advertising", $\mu_{\text{no ads}}$, which we repeat here for convenience:

$$\frac{py}{W(n+s)} = \frac{p}{\frac{W}{q}\left(1+\frac{s}{n}\right)} = \mu_{\text{no ads}}\frac{1+\sigma\frac{s}{n}}{1+\frac{s}{n}}. \tag{A23}$$

Implications of advertising for markup dynamics: Estimating σ_a. In the main text, we estimate $\mu_{\text{no ads}}$ using data on sales-to-costs, R&D costs to production costs and an estimate of how R&D costs scale with firm size, σ. In order to compute the life-cycle profile of μ_{ads}, we need – in addition – information on firms' advertising expenditures and an estimate of σ_a. For the former we will use "advertising expenditures" (`xad`) from Compustat. For the latter, we will follow the same procedures as when estimating σ in the main text.

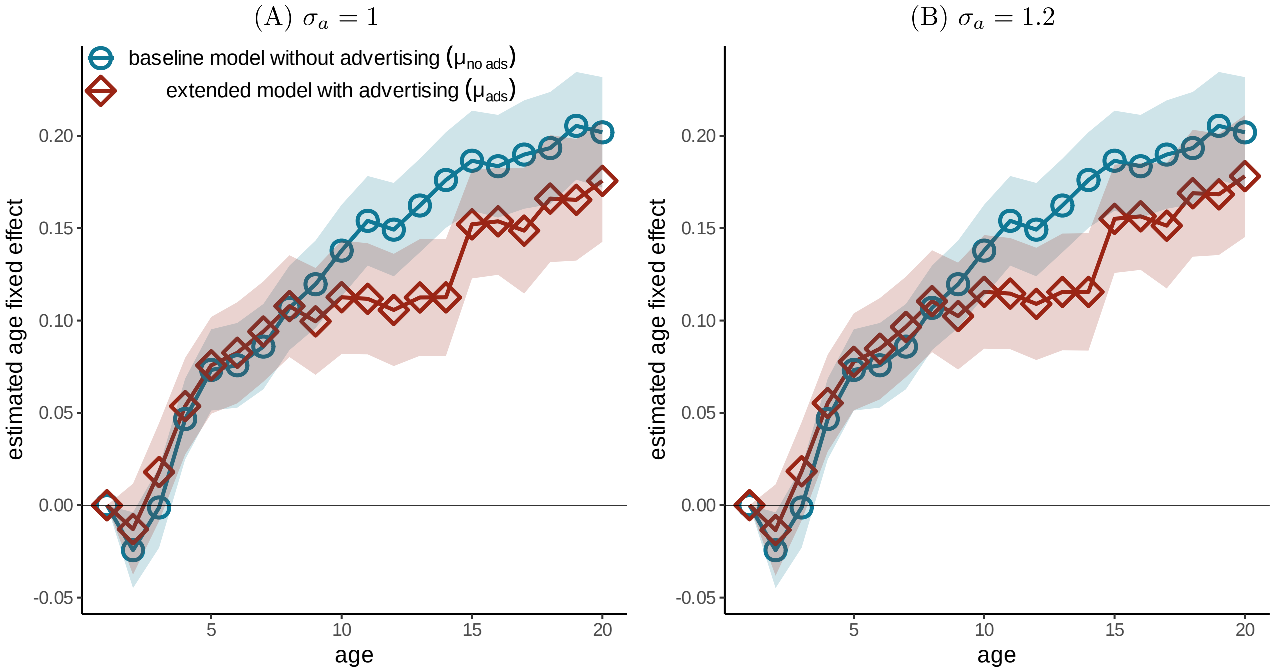
Specifically, we estimate the following using firm-level data from compustat where we use `xad` for $\widetilde{a}$:

$$\ln(1+\widetilde{a}/\text{sales})_{j,t} = \delta_{i,t} + \beta \ln \text{sales}_{j,t} + \epsilon_{j,t}, \tag{A24}$$

where $\delta_{i,t}$ are sector-time fixed effects. The estimated coefficient is -0.01 (and statistically significant) suggesting that – similar to R&D – firms' advertising intensity falls as they grow larger. The magnitude of the estimated elasticity is somewhat smaller than that for R&D (estimated at -0.04). Given that our generalized model replicates this R&D size relationship with a value of $\sigma = 1.2$, we conjecture that a reasonable value for σ_a falls into the range $\sigma_a \in [1, 1.2]$.

[11] Our indirect inference approach to estimating the strength of customer accumulation implies that the impact of advertising – which the generalized model abstracts from – is soaked up by the estimated passive life-cycle customer acquisition profiles, π, and transitory customer base shocks, θ.

Figure A6: Model-implied markups: With and without advertising (`xad`)



Note: Figure presents estimated age fixed effects from regression (A25) for our baseline model ($\mu_{\text{no ads}}$) and that extended to include advertising (μ_{ads}). In both cases, advertising expenditures are measured using `xad` in Compustat. Panel (A) estimates the latter using $\sigma_a = 1$, while Panel (B) uses a value of $\sigma_a = 1.2$. The outcome variables are expressed in logs relative to the value at age 1. Age is measured as the number of years since IPO. The shaded area represents 95% confidence intervals.

Implication of advertising for markup dynamics: Estimating μ_{ads}. With the above range of σ_a we are now able to estimate the markup profile implied by the model with advertising, μ_{ads} and compare it to that in our generalized mode, $\mu_{\text{no ads}}$. Formally, we estimate

$$y_{j,t} = \delta_{i,t} + \delta_a \text{age}_{j,t} + \epsilon_{j,t}, \tag{A25}$$

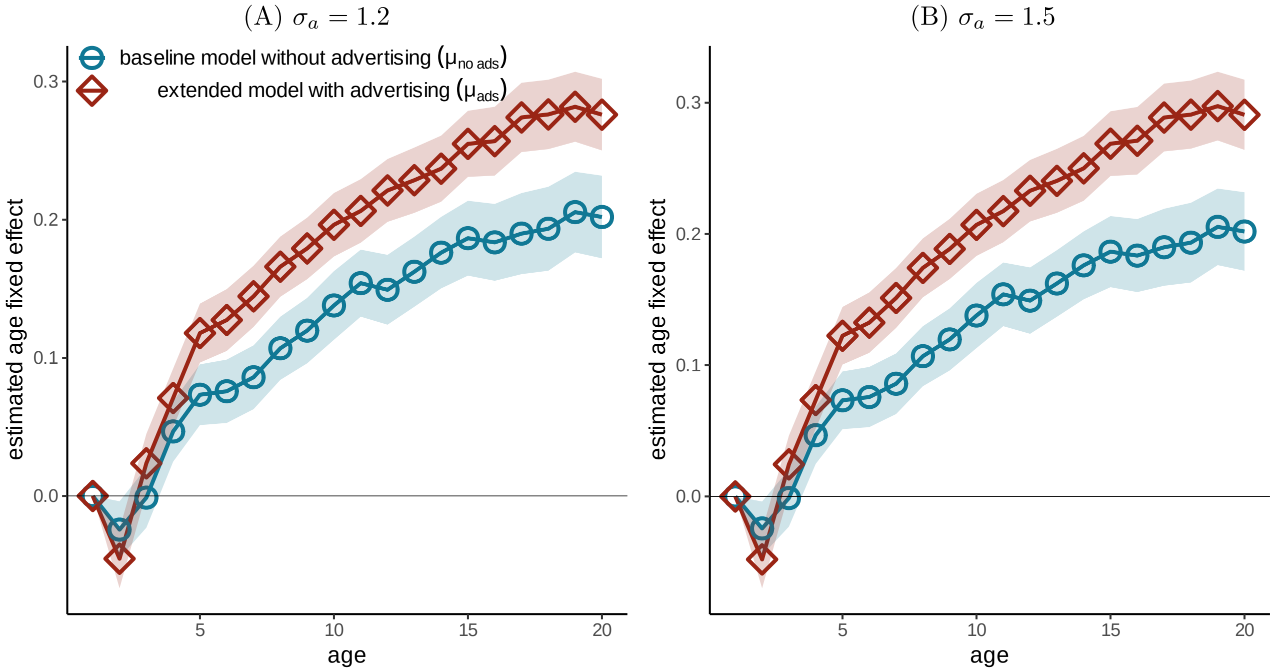
where $y_{j,t}$ is the variable of interest: either μ_{ads} or $\mu_{\text{no ads}}$. Figure A6 shows the comparison for two values of σ_a: $\sigma_a = 1$ (left panel) and $\sigma_a = 1.2$ (right panel). In both cases, however, the model-implied profiles are essentially the same as that in the baseline model in the main text.

Implication of advertising for markup dynamics: Robustness. Before moving on, we also consider a wider definition of advertising. In particular, we use the `SGA` variable in Compustat which has been used as a broader proxy of advertising (selling) expenditures in the literature (see e.g. Gourio and Rudanko, 2014). These "Selling, General and Administrative Expenses" encompasses not only advertising expenditures (`xad`), but also includes other administrative and overhead costs which cannot be assigned to a specific product but are related to business operations. The latter include also R&D expenditures which, however, we subtract in the analysis that follows.

Estimating the advertising intensity-to-size relationship in (A24) using `SGA` results in a (statistically significant) elasticity of -0.08. This is about twice as strong as the R&D intensity-to-size relationship and, therefore, we conjecture that a reasonable range for σ_a falls into the bounds of $\sigma_a \in [1.2, 1.5]$ when considering `SGA`, rather than `xad`, as our proxy for advertising expenditures.

Figure A7 replicates the results in Figure A6, but using `SGA` to measure $\widetilde{a}$. If anything, the implied markup profile is *stronger* in this case. Therefore, we conclude that allowing firms to use both pricing as well as advertising to accumulate customers does not change the strength of price-driven

Figure A7: Model-implied markups: With and without advertising (SGA)



Note: Figure presents estimated age fixed effects from regression (A25) for our baseline model ($\mu_{\text{no ads}}$) and that extended to include advertising (μ_{ads}). In both cases, advertising expenditures are measured using SGA in Compustat. Panel (A) estimates the latter using $\sigma_a = 1$, while Panel (B) uses a value of $\sigma_a = 1.2$. The outcome variables are expressed in logs relative to the value at age 1. Age is measured as the number of years since IPO. The shaded area represents 95% confidence intervals.

customer base growth in our generalized model. Extending our model to include advertising would only strengthen endogenous customer accumulation.[12]

Relation to the empirical literature on customer accumulation. Our main model specification is based on evidence from Foster et al. (2016), who use detailed information on U.S. manufacturing plants in a narrow set of commodity-like industries. Their estimates imply that markups are increasing with firm age. Similar conclusions are made by Piveteau (2021), who uses French customs data and focuses on exporting firms. Separately, Peters (2020) and Alati (2021) estimate an increasing age profile of markups – consistent with our model predictions – using Indonesian firm-level data and Compustat, respectively.

Recently, Afrouzi et al. (2021) and Fitzgerald et al. (forthcoming) provide empirical evidence in favor of the advertising channel of customer acquisition. The former uses Nielsen scanner data to measure the number of customers and sales per customer. The latter uses customs data on Irish exporters (their prices, quantities, but not advertising expenditures) and estimates a model of customer accumulation allowing for both pricing and advertising channels. Both papers conclude that the majority of sales variation is due to differences in customer bases. However, they also provide evidence that in the cross-section, markups are uncorrelated with the number of customers and that the customer base is insensitive to past sales.[13]

Consistent with Fitzgerald et al. (forthcoming) and Afrouzi et al. (2021), in our model the cross-sectional variation in sales is primarily (about 75%) driven by differences in customer bases. Despite

[12]Note that in the main text, the impact advertising (which the generalized model abstracts from) is likely absorbed by variation in the "passive" (exogenous) customer base profiles, π_j. Allowing for advertising would, therefore, likely dampen the importance of such passive life-cycle profiles.

[13]Argente et al. (2021) further document that prices decline with the age of a given *product*.

our model being based on price-driven customer accumulation, the cross-sectional correlations between the customer base and markups or (past) sales is less clear. There are at least two reasons for this. First, our model features rich firm-level heterogeneity. Specifically, firms' customer bases are driven not only by their sales, but also by exogenous life-cycle profiles, π, and transitory disturbances, θ. Second, what matters for firms' endogenous decisions are not just the current values of the different components of the demand stock, but also their *expectations*. This is because firms' pricing and investment decisions are forward-looking.

In fact, our model featuring only price-driven customer accumulation is still broadly consistent with findings in Afrouzi et al. (2021) and Fitzgerald et al. (forthcoming). Recall that the latter document, respectively, that markups are uncorrelated with the number of customers and that the customer base is insensitive to past sales. In our model, the contemporaneous, cross-sectional, correlation between markups and the customer base is -0.03. In addition, the cross-sectional correlation between the endogenous customer base and past sales is 0.107 in our framework.[14] Note further that, as explained above, allowing firms to accumulate customers with advertising – which would likely improve the model's fit to the currently untargeted moments documented in Afrouzi et al. (2021) and Fitzgerald et al. (forthcoming) – does not substantially change the model's price-driven customer base dynamics.

Finally, in our model, the customer base is a complement to R&D. This lies at the heart of the firm-level market size effect – the key channel of our model. Note that firm-level market size effects have been extensively documented in the data (see e.g. Acemoglu and Linn, 2004; Finkelstein, 2004; Lileeva and Trefler, 2010; Kyle and McGahan, 2012; Jaravel, 2019; Aghion et al., 2020). Moreover, assuming that customers and R&D are complements is also consistent with Einav et al. (2022).

That said, the contemporaneous correlation between R&D expenditures in our model is again less clear. This is both because of the rich heterogeneity mentioned above, but also because R&D costs scale with firm size as in the data. Using Compustat data, Cavenaile and Roldan-Blanco (2021) document that in the cross-section of firms, smaller businesses have higher R&D to advertising ratios. We confirm that this is the case also in our model. Proxying for advertising expenditures using the markup gap (see A21), we find that in our model sales are slightly negatively correlated (correlation coefficient of -0.06) with the R&D intensity to markup gap ratio.

E Empirical analysis: Details and additional results

In this appendix, we provide details of our empirical validation exercises in Section 6 as well as other, additional, empirical results.

E.1 Sales-to-costs and selling expenditures in the data

In the main text, we use the sales-to-cost ratio to inform us about firms' customer accumulation investments, see Section 4. In this section, we provide more details on this as well as additional results pertaining to "selling" expenditures (including e.g. advertising).

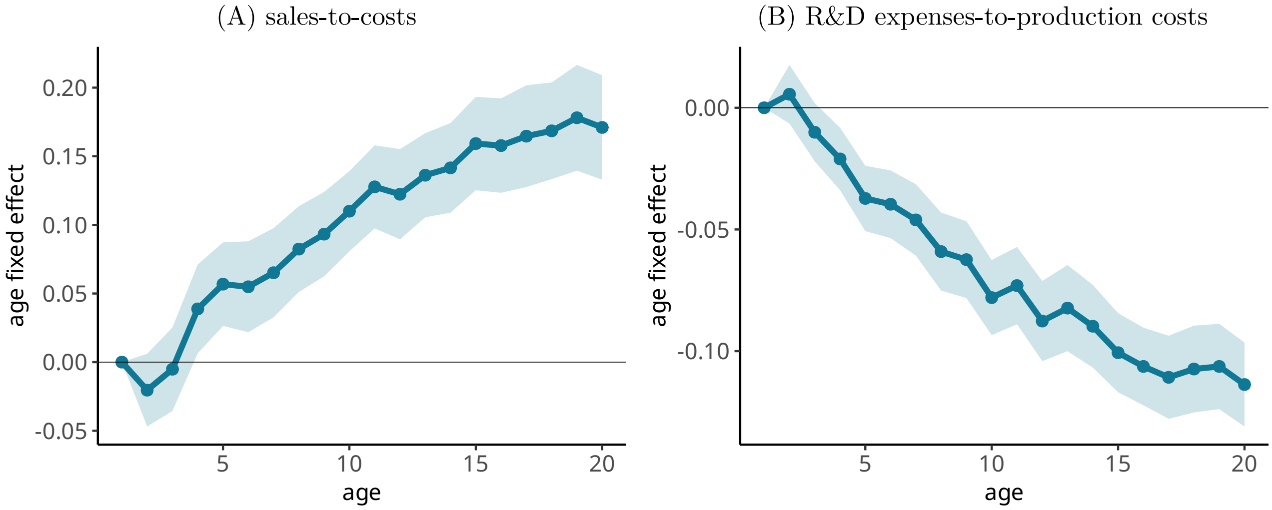
To obtain our parametrization targets, we estimate the following set of regressions

$$\ln(y_{j,t}) = \sum_a \delta_a + \delta_{i,t} + \eta_{j,t}, \tag{A26}$$

where we consider the following outcome variables, $y_{j,t}$: sales-to-costs

[14] As in Fitzgerald et al. (forthcoming), we focus on firms in the first 7 years after entering the market.

Figure A8: Life-cycle profiles of sales and costs.



Note: Figure presents estimated age fixed effects from regression (A26). Panel (A) plots revenues divided by the sum of costs of goods sold (COGS) and R&D expenses. Panel (B) presents R&D expenses divided by COGS. The outcome variables are expressed in logs relative to the value at age 1. Age is measured as the number of years since IPO. The shaded area represents 95% confidence intervals.

(`sales/(cogs+xrd)`), R&D to production costs (`xrd/cogs`), advertising expenses as a share of sales (`xad/sales`) and the SGA expenses ("selling, general and administrative expenditures) net of R&D cost as a share of sales ((`sga - xrd`)/`sales`) for firm j in year t. In the regression, δ_a is the age-a fixed effect with age defined as the number of years since IPO and $\delta_{i,t}$ marks industry-specific time fixed effects.

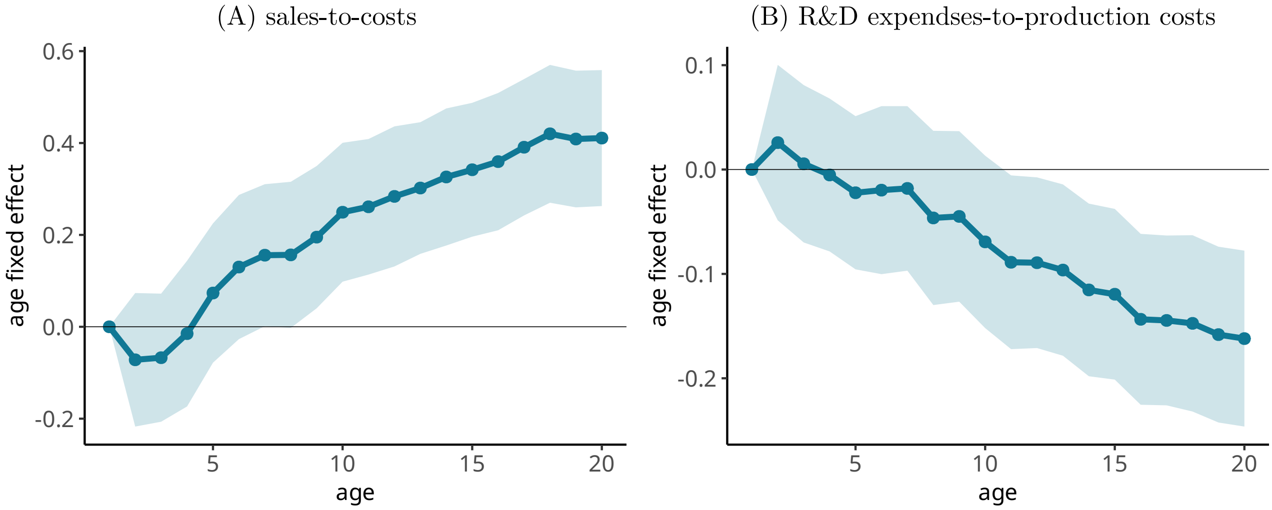
Evidence used in main text. Figure A8 presents the estimated age fixed effects – normalized such that $\delta_1 = 0$ – for the sales-to-cost and the R&D expenditures-to-production costs ratios used in the main text. As we describe in Section 4, the difference between 20 year old firms and new businesses is 0.17 log-points when it comes to sales-to-costs, but -0.11 log-points when it comes to the R&D expenditure-to-production cost ratio, s/n. Through the lens of our generalized model, these patterns are consistent only with a strongly increasing profile of markups – a form of investment into customer accumulation, see the pricing condition (19).

Robustness: Founding dates. Figure A9 presents the estimated age fixed effects – normalized such that $\delta_1 = 0$ – in which we measure the age as the number of years since firm's founding date. As described in section C.3, we are able to match 26% of firms in our data sample to their founding dates. The life-cycle profiles are, if anything, steeper than in the whole sample.

Robustness: Selling expenditures. As discussed in the previous Appendix, an alternative measure of investment into customer accumulation are firms' selling expenditures. We estimate (A26) using advertising expenses and "SGA" expenses, both expressed relative to sales.

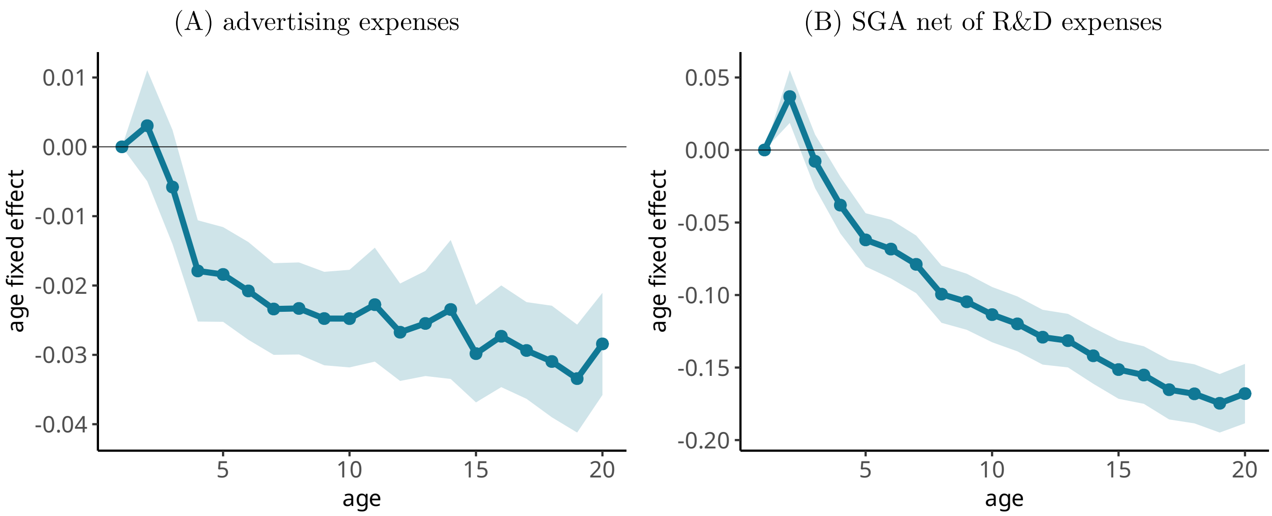
In the main text, we proxy customer accumulation investment with markups. As argued above, the empirical patterns are consistent with our model in which young firms charge lower markups in order to accumulate more customers. In other words, they invest relatively *more* at the beginning of their life-cycles and gradually taper of that investment as they age. The same patterns are observed when looking directly at advertising expenditures which too fall over firms' lifecycles, see Figure A8.

Figure A9: Life-cycle profiles of sales and costs using “true” age.



Note: Figure presents estimated age fixed effects from regression (A26). Panel (A) plots revenues divided by the sum of costs of goods sold (COGS) and R&D expenses. Panel (B) presents R&D expenses divided by COGS. The outcome variables are expressed in logs relative to the value at age 1. Age is measured as the number of years since firm’s founding date. The shaded area represents 95% confidence intervals.

Figure A10: Life-cycle profiles of advertising and SGA expenses.



Note: Figure presents estimated age fixed effects from regression (A26). Panel (A) plots advertising expenses divided by revenues. Panel (B) presents selling, general, and administrative expenses (SGA) minus R&D expenses divided by revenues. The outcome variables are expressed in logs relative to the value at age 1. Age is measured as the number of years since IPO. The shaded area represents 95% confidence intervals.

E.2 Customer accumulation and firm-level R&D: Further details

In Section 6, we document a positive relationship between R&D and customer accumulation at the 3-digit NAICS industry level. To measure the former, we use average R&D expenditures as a fraction of sales, log(1+`xrd/sales`). To measure the latter, we employ theoretical predictions of our model (also used in the parametrization of our model, in Section 4) and compute average steepness of the life-cycle sales-to-costs ratio. In this appendix, we provide further details, robustness exercises to this baseline approach as well as offer an alternative measure using “selling” expenditures (`sga`).

Details of estimation in main text. In the main text, we estimate the following regression:

$$y_i = \alpha + \beta M_i + \Gamma X_i + \varepsilon_i, \tag{A27}$$

where y_i is the average value of $\log(1 + \texttt{xrd/sales})$ in industry i, M_i is our measure of customer acquisition and X_i are additional controls.

To measure M_i, in our baseline specification we compute average life-cycle steepness in sales-to-costs. As in e.g. Peters (2020), we do so in two steps: (i) compute industry-cohort averages of the log-difference between sales-to-cost among firms of age 20 and those of age 0 and (ii) average across cohorts within an industry. Age is measured as the number of years since the IPO. When controlling for other factors, we first remove one outlier industry – "chemical manufacturing" – which is characterized by an R&D intensity almost ten times as high as the mean across all remaining industries. Including this industry only makes our results stronger.

As additional controls, X_i, we consider proxies for financial frictions, market concentration and fixed costs. Specifically, we use average leverage (debt-to-equity), the Herfindahl index and average size (employment).

The first column of Table A6 presents the results. These values underlie the right panel of Figure 5 in the main text which plots the residualized results.

Industry classification into high and low pass-through To illustrate which markets are estimated to have steep sales-to-costs profiles (i.e. a higher pass-through γ indicating industries in which customer accumulation is relatively more important), we group sectors into high and low "customer base pass-through" sectors. Specifically, we classify an industry as high pass-through if its value of life-cycle steepness of sales-to-cost exceeds the top 75th percentile of across all industries. In contrast, the industries with steepness below the 25th percentile are classified as low customer base pass-through. Table A5 presents the results.

Table A5: 3-digit NAICS industry groups: High and low steepness of sales-to-costs

High, above $75th$ percentile		Low, below $25th$ percentile	
ID	Description	ID	Description
111	Crop Production	211	Oil and Gas Extraction
113	Forestry and Logging	213	Support Activities for Mining
212	Mining (except Oil and Gas)	236	Construction of Buildings
235	Special Trade Contractors	313	Textile Mills
237	Heavy and Civil Engineering Construction	314	Textile Product Mills
325	Chemical Manufacturing	323	Printing and Related Support Activities
339	Miscellaneous Manufacturing	324	Petroleum and Coal Products Manuf.
441	Motor Vehicle and Parts Dealers	326	Plastics and Rubber Products Manuf.
444	Building Material, Garden Equip. and (...)	327	Nonmetallic Mineral Product Manuf.
447	Gasoline Stations	332	Fabricated Metal Product Manuf.
486	Pipeline Transportation	424	Merchant Wholesalers, Nondurable Goods
515	Broadcasting (except Internet)	425	Wholesale Electronic Mkts., Agents (...)
517	Telecommunications	484	Truck Transportation
519	Other Information Services	488	Support Activities for Transportation
533	Lessors of Nonfinancial Intangible Assets (...)	492	Couriers and Messengers
611	Educational Services	524	Insurance Carriers and Related Activities
621	Ambulatory Health Care Services	532	Rental and Leasing Services
622	Hospitals	711	Performing Arts, Spectator Sports, and (...)
623	Nursing and Residential Care Facilities	722	Food Services and Drinking Places
811	Repair and Maintenance	812	Personal and Laundry Services

Robustness: Using "true" age to compute life-cycle profiles. The main text defines firm age as the number of years since a given firm appears in the Compustat sample, i.e. since its IPO. We now move to measuring steepness based on the "true" age, that is calculated relative to firms' actual founding dates. This setting, however, comes with a smaller sample of firms complicating our analysis. This happens for two reasons. First, the sub-sample of firms for which we have founding dates is smaller compared to the entire set of firms in Compustat. Second, firms enter the Compustat

Table A6: Estimation results for regression (A27).

	baseline	true age	selection	SGA
sales-to-cost steepness	0.09**	0.06***	0.03***	
	(0.04)	(0.02)	(0.01)	
SGA ex R&D to sales				0.29***
				(0.05)
size	-0.02***	-0.02***	-0.03***	-0.003
	(0.004)	(0.007)	(0.006)	(0.004)
leverage	-0.10***	-0.13***	-0.09***	-0.08***
	(0.03)	(0.04)	(0.03)	(0.02)
HHI	-0.04**	-0.05	-0.07*	-0.004
	(0.02)	(0.05)	(0.04)	(0.01)
intercept	0.21***	0.26***	0.29***	0.03
	(0.03)	(0.05)	(0.04)	(0.04)
Observations	79	59	60	79
R^2	0.38	0.45	0.49	0.55

Note: In all columns, the dependent variable is average value of $\log(1+R\&D/sales)$. The first column ("baseline") reports regression estimates from the specification in the main text. The second column ("true age") does the same but uses a subsample of firms for which we observe age since founding. The third column corresponds to the life-cycle steepness computed at the firm-level and only then calculating the respective industry-level means. The fourth column ('SGA") reports the results when customer acquisition incentives are measured by selling expenses. In all cases, we consider specifications which also control for mean leverage (debt-to-equity), the Herfindahl index of industry concentration and mean size (employment). Standard errors in parentheses. P-values: $* = 0.1$, $** = 0.05$, $*** = 0.01$.

sample at different "true" ages, limiting further the set of young firms. These two issues combined result in a small set of young firms at the 3-digit industry level.

Therefore, we adopt a somewhat broader approach in defining young and old firms. Specifically, we classify a given firm-year observation as a "young" firm if firm's age lies in the bottom 25% of cohort-specific true age distribution and as "old" if it is among the top 75% largest values in the cohort.[15] We then compute the log-difference in median sales-to-costs between old and young firms. The remaining details are the same as in the baseline exercise. The second column of Table A6 shows that the results are similar to those in the main text (first column).

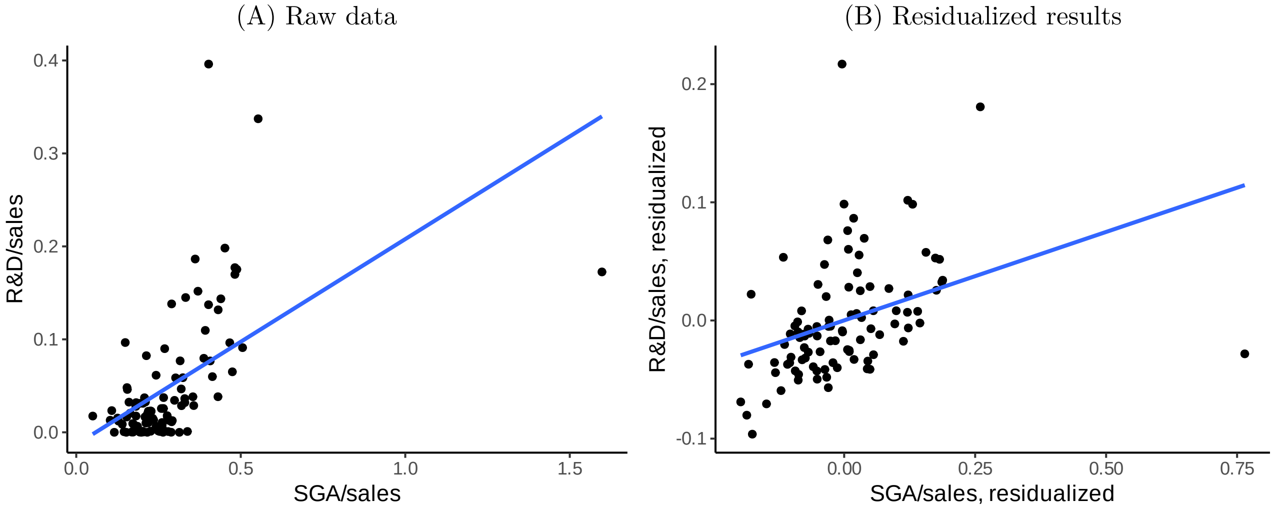
Robustness: Firm selection. Next, we consider our baseline specification, but account for firm selection. We do so by first computing life-cycle steepness of sales-to-costs at the firm-level and only then calculating the respective industry-level means. This, in contrast to the baseline specification which computes sales-to-cost steepness as the difference between average values of old and young firms, controls for firm selection. The third column in Table A6 shows that our results remain robust also to this alternative.

Robustness: Selling expenditures. Finally, as mentioned in the main text, modelling customer accumulation via pricing (markups) is only one possibility. While not considered in our model, a natural alternative is to consider customer acquisition via direct "selling" expenditures (including e.g. advertising).

Therefore, in this subsection, we consider whether industries which invest relatively more into selling expenditures – i.e. those with more active customer accumulation and, therefore, faster customer

[15]For reference, in the pooled data, the 25th percentile corresponds to 10 years since founding and 75th to 31 years since founding.

Figure A11: Industry variation in R&D intensity and selling expenditures



Note: The figure shows the relationship between average R&D intensity, log(1+`xrd`/`sales`), and selling expenditures-to-sales, log(1+`sga`-`xrd`/`sales`), across 3-digit NAICS industries. Panel A does so using "raw data". Panel B, "residualized results", first regresses both variables on mean leverage (debt-to-equity), the Herfindahl index of market concentration and mean firm size (employment) and plots the respective residuals.

base growth – also display higher R&D expenditures. Selling expenditures are conveniently measured in Compustat as "selling, general and administrative" expenditures (`sga`).

Therefore, we estimate equation (A27) but with the measure of customer accumulation, M_i, being the average selling expenditure-to-sales ratio. Because selling expenditures also include R&D, we explicitly account for those: $M_i = (\texttt{sga}_i\text{-}\texttt{xrd}_i)/\texttt{sales}_i$. All other details are the same as above.

Figure A11 visualizes the relationship, as in the main text. The left panel shows the raw data, while the right panel shows the residualized results after controlling for industry-level size and age. As in the main text when using our model-based approach, also in both of the cases presented here, the relationship is strongly positive (and statistically significant).

E.3 Firm-level impact of R&D subsidies: Further details

R&D tax credit policies are widespread in the US. Since Minnesota first introduced this policy in 1982, 32 states have legislated a tax credit for firms' R&D expenses (as of 2006; see Wilson, 2009). For the purpose of R&D tax credit, US tax authorities use the following definition of R&D expenses: "the wages, materials expenses, and rental costs of certain property and equipment incurred in performing research undertaken to discover information that is technological in nature for a new or improved business purpose." (Wilson, 2009, p. 431-432).

The generosity of these policies as well as the time of rollout differs across the US states. We use these cross-states differences to directly test the predictions of our model. In particular, Proposition 3 suggests that firms that expect a higher customer base growth g_b are *more sensitive* to the marginal cost of R&D. To test this prediction we interact the impact of R&D tax credit with firm age. Younger firms on average exhibit a steeper customer base accumulation rate. Furthermore, Proposition 3 predicts that the elasticity of R&D expenses with respect to subsidies is increasing in the R&D cost share in profits, $\frac{W s_j}{\pi_j}$. Next, we bring both predictions to the data.

Estimation strategy. Formally, we estimate

$$\begin{aligned}\log(\text{R\&D})_{j,t} =& \alpha\tau_{s,t} + \beta_a \mathbf{1}(\text{age}_{j,t} \leq 5) \times \tau_{s,t} + \beta_b \mathbf{1}(\bar{s}_{p,j} > \bar{s}_p^{\text{med}}) \times \tau_{s,t} \\ &+ \Gamma X_{j,t} + \delta_j + \delta_s + \delta_t + \epsilon_{j,t},\end{aligned} \tag{A28}$$

where $R\&D_{j,t}$ denotes R&D expenses, δ_j, $\delta_{i,t}$, and δ_s denote firm, industry-specific time, and state fixed effects, respectively, $\tau_{s,t}$ denotes R&D tax credit, $\mathbf{1}(\text{age}_{j,t} \leq 5)$ is an indicator variable for firms younger than 6 years since IPO, $\mathbf{1}(\bar{s}_{p,j} > \bar{s}_p^{\text{med}})$ denotes the indicator variable equal to one if the firm-specific average R&D expenditures-to-profits ratio, $\bar{s}_{p,j}$, exceeds its median value over all firms, $\bar{s}_p^{\text{med}}$, which we describe in greater detail below, and $X_{j,s,t}$ mark firm and state level controls.[16,17] The parameters of interest are α, measuring the association between the level of the R&D tax credit $\tau_{s,t}$ and the R&D expenses and β indicating whether firms with characteristic $\tilde{x}_{j,t}$ are more or less sensitive to the R&D policy.

As for our policy variable $\tau_{s,t}$, we explore four alternative measures of the effective cost of R&D which we describe in greater detail below. These measures use as a starting point the statutory R&D tax credit rate in a given year in a given state and adjust it with other state and federal corporate taxes as well as correct for the eligibility criteria.

Data: R&D subsidies. In the main text, we use information on the user cost of R&D as collected by Wilson (2009). We now briefly describe alternative measures of the generosity R&D subsidies considered in Wilson (2009) and refer an interested reader to this article for more details. As in Cavenaile and Roldan-Blanco (2021) we merge the original Wilson (2009) dataset with information on R&D tax credit changes between 2007 and 2009 collected in Falato and Sim (2014).

The most direct measure of the generosity of the R&D subsidies in a given state is the statutory tax credit rate. The tax credit allows firms to deduct from the income tax liabilities cost of workers, materials, and capital used in the process of discovering new technologies. However, the tax credit itself constitutes a taxable income in some states (so called credit "recapture"). Therefore, Wilson (2009) computes for each state in each year a tax-adjusted credit rate according to the formula,

$$\text{tax-adjusted rate} = \text{statutory credit rate} \times (1 - \sigma \times \tau^e),$$

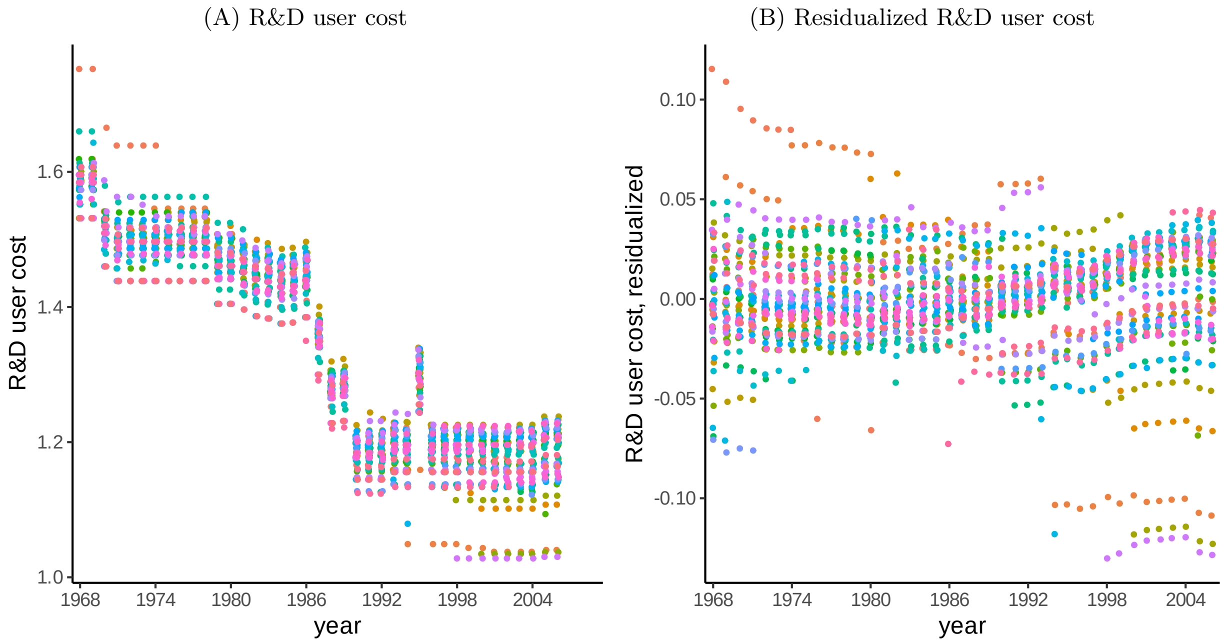
where σ marks the share of tax credit subject to corporate taxation and τ^e is the state-level effective corporate income tax rate. The tax-adjusted rate may be a more adequate measure of the tax treatment of the R&D expenses.

Another measure of R&D subsidies takes into account the differences in the design of the tax credit across states. In particular, the value of expenses eligible for tax credit can be determined in a non-incremental way in which all qualified R&D is eligible for the credit. Alternatively, in an incremental design only R&D above a certain base level are subject to tax credit. This base level can be calculated based on firm's activity in a fixed past period, or as a moving-average base over a given number of most recent periods. Therefore, for a state in which the tax credit eligibility threshold is

[16] When states have several credit brackets, we use the top marginal rate. However, this is the case only in 75 out of 2,244 state-year observations.

[17] At the firm-level, the control variables are: financial leverage, liquidity, dividend paid, log level of sales and its growth rate, all lagged by one period to make sure they are not affected by the time-t R&D tax credit. At the state-level, the control variables are: the corporate tax rate, employment growth and GDP growth. Note that including only state-level controls delivers similar results.

Figure A12: Variation in R&D user cost across US states and over time



Note: Each dot in the figure represents one state-year observation. Each of the 50 US states is consistently represented by a unique color throughout the entire period. The left panel (A) present the values of R&D user cost while the right panel (B) presents the residuals from regression (user cost)$_{s,t} = \delta_t + \delta_s + \varepsilon_{s,t}$.

based on the n-period moving-average of past R&D activity, the marginal effective tax credit reads

$$m = \text{statutory credit rate} \times (1 - \sigma \times \tau^e) \times \left(1 - \frac{1}{n}\sum_{k=1}^{n}(1 + r_{t+k})^{-k}\right), \tag{A29}$$

where r is the real interest rate. For states with a non-incremental tax credit design, the marginal credit rate is equal to the tax-adjusted rate defined above.

Given the intermediate components above, we can compute the state level R&D user cost as in Wilson (2009). The user cost is computed as

$$\text{R\&D user cost} = \frac{1 - v\,(m_s - m_f) - \left(\tau_f^e + \tau_f^e\right)}{1 - \left(\tau_f^e + \tau_f^e\right)}\,(r + \delta)\,,$$

where m_s is the state level marginal effective R&D tax credit and m_f its *federal* equivalent (see equation (A29)), τ_s^e and τ_f^e are the state and federal, respectively, effective corporate income tax rates, and v is the share of measured R&D expenditures considered qualified R&D in the tax code.[18] Based on the IRS Statistics on Income data, Wilson (2009) calibrates $v = 0.5$. The R&D capital depreciation rate is calibrated to $\delta = 0.15$. Note that the interpretation of the sign is reversed in the case of R&D user cost: a high value means low generosity of R&D tax credit.

Finally, Figure A12 provides evidence of the geographical and time variation in the generosity of R&D subsidies in the U.S. which we use in our estimation. Each dot in the figure represents the R&D user cost value in a given state-year pair. The left panel (A) presents the R&D user cost values while the right panel (B) shows the user cost residualized with state and time fixed effects.

[18]The notion of measured R&D concerns the definition of R&D expenses employed in the National Science Foundation's Industrial Research and Development database. This may be different to the definition of R&D expenses in the tax code.

Data: Other variables. In order to estimate (A28), we merge firm-level data on public firms in the US with the state level data R&D tax credit between 1982 and 2009 collected by Wilson (2009); Falato and Sim (2014).[19] On the firm side, we use the same Compustat sample as described in Section C.1 above. We impose one additional restriction, namely that we include only observations with positive R&D expenses. We drop firms with highest and lowest 1% of their revenue growth where the percentiles are computed annually. We measure age as the number of years since the IPO and compute the R&D share in profit as $\frac{W s_j}{\pi_j} = \frac{\texttt{xrd}}{\texttt{sale}-\texttt{xrd}-\texttt{cogs}}$.[20]

To avoid a mechanical relationship between the R&D share in profits on the right-hand side of the estimated equation and the R&D expenses on the left-hand side, we classify firms into "high R&D share" and "low R&D share" based on its long-run average. Concretely, for each firm we compute firm specific mean R&D share $\frac{W s_j}{\pi_j}$ after residualizing industry-specific time trends.[21] Next, we compute median of these firm-specific shares and classify a firm as a "high share" if it lies strictly above the median.

The identification assumption is that conditional on the aggregate and firm level fixed effects and observable characteristics, had there been no tax credit changes in a given state s, the firm-level R&D expenses would have evolved identically to the R&D expenses of firms in the remaining states s'.

Results. Results for our baseline specification using the R&D user cost are in Table A7. These are identical to the results in the main text, see Section 6. The regression estimates show that R&D subsidies increase firms' innovation expenses and that this effect is stronger for younger firms, but weaker for businesses which simultaneously invest more into customer accumulation. These results are consistent with predictions from our generalized model. Note in particular that, as in Proposition 3, firms with faster customer base growth (proxied here by age) respond more strongly. Moreover, firms with higher share of R&D costs in profits respond more.

Robustness: Alternative measures of R&D tax credits. Tables (A8) to (A10) provide results for the three alternative measures of tax credits. In all cases, R&D subsidies increase firm-level innovation expenditures, they do so more for younger firms and less so for businesses which also simultaneously invest into customer accumulation.

[19] This data has also been used in Cavenaile and Roldan-Blanco (2021) to estimate the elasticity of advertising expenses with respect to the R&D cost. Here, we explore the heterogenous impact of R&D subsidies on R&D expenses itself.

[20] Our results are robust to using net operating income (`oibdp`) are a measure of profits.

[21] Formally, we estimate $\log \frac{W_t s_{j,t}}{\pi_{j,t}} = \delta_j + \delta_{i,t}$, where δ_j and $\delta_{i,t}$ are firm and industry-time fixed effects. W group firms by median split on δ_j.

Table A7: R&D tax credit and firm-level R&D expenses

	R&D expenses		
	(i)	(ii)	(iii)
R&D user cost	-1.94**	-1.66*	-0.683
	(0.949)	(0.897)	(0.694)
young firm	-0.075**	0.913***	0.781***
	(0.034)	(0.232)	(0.197)
R&D user cost × young firm		-0.740***	-0.633***
		(0.170)	(0.143)
R&D user cost × high R&D share			-1.19***
			(0.178)
additional controls	✓	✓	✓
firm fixed effects	✓	✓	✓
time × industry fixed effects	✓	✓	✓
state fixed effects	✓	✓	✓
Observations	60,371	60,371	57,483
R^2	0.919	0.920	0.921
Within R^2	0.008	0.013	0.025

Note: The table reports coefficient estimates from (33). R&D subsidies measured in terms of R&D user cost, hence low values mean a high R&D subsidy. Additional controls include log revenues and financial variables: liquidity ratio (debt to current assets), leverage (debt to equity), and the log of dividend paid. Moreover, we control for overall state level corporate income tax, employment growth, and output growth at the state level. Stars indicate p-value, "$***$"= 0.01, "$**$"= 0.05, "$*$"= 0.10. Standard errors clustered at the state level in parentheses

Table A8: Marginal effective R&D tax rates and firm-level R&D.

	R&D expenses		
	(i)	(ii)	(iii)
marginal effective tax credit	1.62*	1.30	-0.306
	(0.869)	(0.797)	(0.502)
young firm	-0.076**	-0.107***	-0.097***
	(0.034)	(0.027)	(0.028)
marginal effective tax credit × young firm		1.40*	1.15*
		(0.698)	(0.678)
marginal effective tax credit × high R&D share			2.68***
			(0.955)
additional controls	✓	✓	✓
firm fixed effects	✓	✓	✓
time × industry fixed effects	✓	✓	✓
state fixed effects	✓	✓	✓
Observations	60,371	60,371	57,483
R^2	0.919	0.919	0.920
Within R^2	0.007	0.009	0.013

Note: The table reports coefficient estimates from (33) using marginal effective tax rate as a measure of R&D subsidy. Additional controls include log revenues and financial variables: liquidity ratio (debt to current assets), leverage (debt to equity), and the log of dividend paid. Moreover, we control for overall state level corporate income tax, employment growth, and output growth at the state level. Stars indicate p-value, "$***$"= 0.01, "$**$"= 0.05, "$*$"= 0.10. Standard errors clustered at the state level in parentheses

Table A9: Results using the statutory tax credit rate.

	R&D expenses		
	(i)	(ii)	(iii)
statutory tax credit	1.01	0.664	-1.02**
	(0.653)	(0.626)	(0.461)
young firm	-0.080**	-0.113***	-0.098***
	(0.032)	(0.030)	(0.029)
statutory tax credit × young firm		1.19**	0.898*
		(0.508)	(0.511)
statutory tax credit × high R&D share			3.06***
			(0.817)
additional controls	✓	✓	✓
firm fixed effects	✓	✓	✓
time × industry fixed effects	✓	✓	✓
state fixed effects	✓	✓	✓
Observations	60,371	60,371	57,483
R^2	0.919	0.919	0.920
Within R^2	0.005	0.007	0.013

Note: The table reports coefficient estimates from (33) using statutory tax credit rate as a measure of R&D subsidy. Additional controls include log revenues and financial variables: liquidity ratio (debt to current assets), leverage (debt to equity), and the log of dividend paid. Moreover, we control for overall state level corporate income tax, employment growth, and output growth at the state level. Stars indicate p-value, "∗ ∗ ∗"= 0.01, "∗∗"= 0.05, "∗"= 0.10. Standard errors clustered at the state level in parentheses

Table A10: Results using the effective tax credit rate.

	R&D expenses		
	(i)	(ii)	(iii)
effective tax credit	0.943	0.565	-1.09**
	(0.641)	(0.610)	(0.480)
young firm	-0.081**	-0.114***	-0.099***
	(0.031)	(0.030)	(0.029)
effective tax credit × young firm		1.26**	0.935*
		(0.562)	(0.555)
effective tax credit × high R&D share			3.08***
			(0.897)
additional controls	✓	✓	✓
firm fixed effects	✓	✓	✓
time × industry fixed effects	✓	✓	✓
state fixed effects	✓	✓	✓
Observations	60,371	60,371	57,483
R^2	0.919	0.919	0.920
Within R^2	0.005	0.007	0.013

Note: The table reports coefficient estimates from (33) using the effective tax credit rate as a measure of R&D subsidy. Additional controls include log revenues and financial variables: liquidity ratio (debt to current assets), leverage (debt to equity), and the log of dividend paid. Moreover, we control for overall state level corporate income tax, employment growth, and output growth at the state level. Stars indicate p-value, "∗ ∗ ∗"= 0.01, "∗∗"= 0.05, "∗"= 0.10. Standard errors clustered at the state level in parentheses

References

Acemoglu, Daron and Joshua Linn (2004) “Market Size in Innovation: Theory And Evidence from the Pharmaceutical Industry,” *Quarterly Journal of Economics*, 119 (3), 1049–1090.

Afrouzi, Hassan, Adres Drenik, and Ryan Kim (2021) “Growing by the Masses - Revisiting the Link between Firm Size and Market Power.”

Aghion, Philipe, Antonin Bergeaud, Matthieu Lequien, and Marc Melitz (2020) “The Heterogenous Impact of Market Size on Innovation: Evidence from French Firm-Level Exports,” mimeo.

Aghion, Philippe and Peter Howitt (1994) “Growth and Unemployment,” *Review of Economic Studies*, 61, 477–494.

Alati, Andrea (2021) “Initial Aggregate Conditions and Heterogeneity in Firm-Level Markups.”

Argente, David, Doireann Fitzgerald, Sara Moreira, and Anthony Priolo (2021) “How do Entrants Build Market Share? The Role of Demand Frictions.”

Cavenaile, Laurent and Pau Roldan-Blanco (2021) “Advertising, Innovation and Economic Growth,” *American Economic Journal: Macroeconomics*, 13 (3).

Denicolo, Vincenzo and Piercarlo Zanchettin (2014) “What Causes Over-Investment in R&D In Endogenous Groth Models,” *Economic Journal*, 124, 1192–1212.

Einav, Liran, Peter Klenow, Jonathan Levin, and Raviv Murciano-Goroff (2022) “Customers and Retail Growth.”

Evans, David (1987) “Tests of Alternative Theories of Firm Growth,” *Journal of Political Economy*, 95, 657–674.

Falato, Antonio and Jae Sim (2014) “Why Do Innovative Firms Hold So Much Cash? Evidence from Changes in State R&D Tax Credits,” *FEDS Working Paper*.

Finkelstein, Amy (2004) “Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry,” *Quarterly Journal of Economics*, 119 (2), 527–564.

Fitzgerald, Doireann, Stefanie Haller, and Yaniv Yedid-Levi (forthcoming) “How Exporters Grow,” *Review of Economic Studies*.

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

——— (2016) “The Slow Growth of New Plants: Learning about Demand?” *Economica*, 83 (329), 91–129.

Gourio, François and Leena Rudanko (2014) “Customer Capital,” *The Review of Economic Studies*, 81 (3), 1102–1136.

Jaravel, Xavier (2019) “The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector,” *Quarterly Journal of Economics*, 134 (2), 715–783.

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

Kyle, Margaret and Anita McGahan (2012) "Investments in Pharmaceuticals Before and After TRIPS," *The Review of Economics and Statistics*, 94 (4), 1157–1172.

Lileeva, Alla and Daniel Trefler (2010) "Improved Access to Foreign Markets Raises Plant-Level Productivity... For Some Plants," *The Quarterly Journal of Economics*, 125 (3), 1051–1099.

Loughran, Tim and Jay Ritter (2004) "Why Has IPO Underpricing Changed over Time?" *Financial Management*, 33 (3), 5–37, `http://www.jstor.org/stable/3666262`.

Peters, Michael (2020) "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 88 (5), 2037–2073, https://doi.org/10.3982/ECTA15565.

Piveteau, Paul (2021) "An Empirical Dynamic Model of Trade with Customer Accumulation," *American Economic Journal: Microeconomics*, 34 (4), 23–63.

Rouwenhorst, K Geert (1995) "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Cooley, T.F. (Ed.), Frontiers of Business Cycle Research*: Princeton University Press.

Sterk, Vincent ⓡ Petr Sedláček ⓡ Benjamin Pugsley (2021) "The Nature of Firm Growth," *American Economic Review*, 111 (2).

Wilson, Daniel J (2009) "Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits," *The Review of Economics and Statistics*, 91 (2), 431–436.