Customer Acquisition, Business Dynamism and Aggregate Growth*

Marek Ignaszak
Goethe University Frankfurt

Petr Sedláček
University of New South Wales

February 5, 2026

Abstract

Business dynamism – the process of firm entry, growth and exit – lies at the heart of modern endogenous growth models. While productivity differences have traditionally been seen as the main driving forces of business dynamism, a growing body of evidence suggests that customer acquisition is at least as important. In light of this evidence, we propose a novel endogenous growth model in which innovating firms must first acquire customers to sell their products. Estimating our model with aggregate and firm-level data, we find that expansions of firms' customer bases (market sizes) boost their incentives to innovate and shift resources towards high-growth businesses ("gazelles"). Combined, these effects explain over 40% of aggregate growth and substantially change predictions about the efficacy of growth policies. Finally, we document support for key model predictions using firm-level micro-data.

Keywords: Customer acquisition, firm heterogeneity, growth, innovation, R&D

*For useful comments furthering this project we thank Steve Bond, Vasco Carvalho, Sergio de Ferra, Maarten De Ridder, Łukasz Drozd, Georg Duernecker, Chris Edmond, Ippei Fujiwara, Pauline Grosjean, Hugo Hopenhayn, Giammario Impullitti, Leo Kaas, Greg Kaplan, Benjamin Moll, Toshihiko Mukoyama, Emi Nakamura, Bruce Preston, Łukasz Rachel, Federica Romei, Benjamin Schoefer, Mortiz Schularick, Emily Swift, Gianluca Violante, Hasin Yousaf, Francesco Zanetti and seminar participants at the Australian National University, Cambridge University, European Commission and ECFIN, Georgetown University, Goethe University Frankfurt, London Business School, Monash University, University of Alabama, University of California Berkeley, University of Copenhagen, University of Melbourne, the University of Oxford, the U.S. Census Bureau and conference participants at the Australasian Meeting of the Econometric Society, the Conference in Computing in Economics and Finance and the Macroeconomic Workshop of the Reserve Bank of Australia. Ignaszak gratefully acknowledges research support by Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy EXC 2126/1-390838866 and the Research Training Group 2281. Sedláček gratefully acknowledges the financial support of the Australian Research Council [grant number FT230100545] and, for earlier stages of the project, the financial support of the European Research Council [grant number 802145]. Ignaszak: ignaszak@uni-frankfurt.de; Sedláček: p.sedlacek@unsw.edu.au.

1 Introduction

The continuous process of firm entry, growth and exit– "business dynamism"– plays a key role in aggregate job creation and productivity growth (Haltiwanger, 2012). By now, models incorporating business dynamism have been employed to understand a range of important questions regarding e.g. international trade, long-run growth, unemployment, inequality, market power, and monetary and fiscal policy.[1] Typically, such models think of business dynamism as a purely productivity-driven phenomenon whereby less productive firms shrink and shut down, giving way to more productive and innovative businesses – so called "creative destruction."

However, innovation and creative destruction are not the sole drivers of business dynamism. In fact, a growing body of recent empirical and theoretical research stresses that a key source of business dynamism lies in customer accumulation.[2] In this paper, we argue that firm-level customer acquisition is also a quantitatively important driver of aggregate economic growth. Moreover, we show that ignoring customer accumulation leads to substantially different conclusions about the efficacy of growth policies. We then proceed to document empirical support for key theoretical predictions of our framework in firm-level micro-data.

To make our point, we propose a new model of endogenous growth which explicitly considers firm-level customer accumulation. In our framework – as in existing growth models – firms invest into improving their individual productivity levels, the ultimate source of aggregate growth. In contrast to existing growth models, however, in our framework firms must first acquire customers in order to sell their products. We consider two distinct sources of customer accumulation. First, businesses can attract customers by spending resources on marketing along the lines of e.g. Dinlersoz and Yorukoglu (2012), Sedláček and Sterk (2017) and Abaluck and Adams-Prassl (2021). Second, firms can also attract new customers by charging lower prices (markups) as in Foster et al. (2016), Gourio and Rudanko (2014) and Rudanko (2022).[3]

We first analyze a "benchmark" version of our model in which we impose specific restrictions allowing for analytical conclusions. In particular, we assume that R&D and marketing costs scale one-for-one with firm size. Note that, as we discuss in detail below, we relax these restrictions when taking the model to the data for a quantitative analysis.

With our benchmark model, we can show four central theoretical results. First, higher

[1]See e.g. Melitz (2003), Klette and Kortum (2004), Kaas and Kircher (2015), Jones and Kim (2018), Peters (2020), Ottonello and Winberry (2020), Spencer (2022).

[2]See e.g. Foster et al. (2008), Dinlersoz and Yorukoglu (2012), Foster et al. (2016), Hottman et al. (2016) ,Kaas and Kimasa (2021), Bernard et al. (2022), Eslava and Haltiwanger (2021), Einav et al. (2022), Rudanko (2022).

[3]Examples price-driven customer accumulation include not only initial discount periods on products such as newspaper subscriptions, mobile phone plans or gym memberships– but also longer-term pricing strategies to build sufficiently large customer bases– including those applied by service and technology firms such as AirBnB, DoorDash, Netflix or Uber.

customer retention strengthens firms' incentives to conduct R&D through a "firm-level market size effect." To understand this, note that higher firm-level productivity allows businesses to produce at lower costs and earn higher profits from their existing customers. Being able to hang on to more customers, therefore, raises the benefits of innovation because lower production costs can apply to a larger market.

Second, the link between innovation and customer accumulation also runs in the opposite direction, creating a new endogenous feedback loop. Specifically, since more productive businesses can produce at lower costs, they can afford to invest more resources into customer acquisition. The associated increase in firms' customer bases (market shares) boosts incentives to conduct R&D and so on.

Third, aggregating firm-level innovation decisions leads to our next result: higher firm-level customer retention rates increase aggregate economic growth. However, it is important to note that this happens "only" because higher customer base levels boost firms' incentives to innovate – our first theoretical result.

Fourth, changes in firms' customer base levels alter their responsiveness to R&D subsidies. Generally speaking, the prospect of larger market shares associated with an expanded customer base tilts firms' discounting towards future periods. This makes businesses more sensitive to any changes in future profits, including those brought about by R&D subsidies and associated lower innovation costs.

To gauge the quantitative importance of our theoretical results, we generalize the benchmark model in two crucial dimensions. First, we relax the restriction that marketing and R&D costs increase one-for-one with firm size. Instead, we use micro-data to directly estimate these relationships. Second, we introduce endogenous entry and exit in order to be able to speak to the process of creative destruction. We then discipline this "full" model using a combination of aggregate moments, data on business dynamism and firm-level information on R&D, marketing and markups – the key features of our framework.

To isolate the effect of customer acquisition on firm-level outcomes, we conduct a series of decomposition exercises. First, within our full model, we consider a counterfactual scenario in which firms expect lower rates of customer retention – such that average customer base growth is zero among continuing businesses. Solving for optimal R&D decisions under these counterfactual expectations, we then compare the resulting innovation rates to those in the full model. The difference quantifies the role of customer accumulation in our full model.

As a result of lower expected customer retention, firms optimally scale back on R&D investment. This is precisely a reflection of the firm-level market size effect – our first theoretical result. Quantitatively, the counterfactual features innovation rates which are about 40% lower compared to our full model. Importantly, the vast majority of this effect is driven by the feedback loop between innovation and customer acquisition. Considering only the "direct" effect results in counterfactual firm-level innovation rates being only about 6% lower compared to the full model.

As a second decomposition exercise, we turn to analyzing the impact of customer acquisition on aggregate growth. Changes in firms' customer base levels affect aggregate growth in two distinct ways. First, the firm-level market size effect described above directly influences aggregate growth. In particular, lower expected customer accumulation weakens firm-level market size effects. In response, businesses scale down on R&D and aggregate growth is about 1/5 lower than in our full model. In addition, there is a second powerful channel through which changes in firms' customer base levels affect aggregate growth – through the reallocation of customers across firms.

In particular, the feedback loop between innovation and customer acquisition leads to an endogenous reallocation of market shares (customers) towards firms with high productivity levels. This happens because highly productive firms can afford to invest more into customer accumulation. Combined, the firm-level market size effect and the reallocation effect account for 44% of aggregate growth. Therefore, while aggregate growth has typically been considered a purely supply-side phenomenon, our framework shows that customer acquisition is a quantitatively important driver of aggregate economic growth.

As a final step in our quantitative analysis, we turn our attention to analyzing the impact of two common policy interventions: subsidies to R&D and to firms' operations. We do so in our full model and an identical economy which, however, ignores customer accumulation.[4] First, we show that in our full model with customer acquisition frictions, R&D subsidies are much more successful (about twice as much) at boosting aggregate growth – confirming our analytical results.

In contrast, subsidizing firms' operations leads to a stronger decline in aggregate growth in the full model compared to an economy which ignores customer accumulation. The key reason is the different nature of business dynamism across the two economies. In a model without customer accumulation, firm growth and survival are purely productivity-driven. In the full model, however, aggregate growth is a customer-weighted average of firms' productivity growth rates. Therefore, the overall effect depends on changes to the joint distribution of firms' productivity and customer base levels.

Despite focusing only on two policy examples, our results are more general. This is because our conclusions ultimately depend on how customer accumulation shapes business dynamism and, in turn, aggregate outcomes. Therefore, using a model without customer accumulation to analyze any policies (or shocks) which affect business dynamism may provide a skewed view of their aggregate impact.

Finally, to establish the empirical relevance of the proposed mechanisms, we turn to firm-level data and document support for key model predictions. First, we show that in the cross-section of firms, marketing, R&D and markups all co-move in model-predicted

[4] We do not strive to determine optimal policies. Instead, we illustrate quantitatively that commonly used policies have a vastly different impact in environments with and without customer accumulation.

ways. This is true for both static (contemporaneous) relationships, but also for dynamic (life-cycle) patterns. Second, we propose a model-consistent proxy of the severity of customer base frictions and estimate them in the cross-section of industries. Using these estimates, we show that – as predicted by the model – stronger frictions to accumulating customers are related to weaker R&D investment.

Our paper straddles two large bodies of research. First, we build on the literature studying firm dynamics (see e.g. Jovanovic, 1982; Hopenhayn and Rogerson, 1993, for seminal contributions) and in particular on papers analyzing demand-side factors and customer acquisition (see e.g. Drozd and Nosal, 2012; Dinlersoz and Yorukoglu, 2012; Gourio, 2014; Arkolakis, 2016; Perla, 2019; Kaas and Kimasa, 2021; Rudanko, 2022; Argente et al., 2025).[5] We add to these papers the context of aggregate growth. Second, our paper connects to endogenous growth models which, however, typically ignore demand-side factors (see e.g. Klette and Kortum, 2004; Lentz and Mortensen, 2008; Akcigit and Kerr, 2018; Acemoglu et al., 2018). Exceptions include Cavenaile et al. (2021) and Einav et al. (2022). The former present a model in which advertising statically increases (perceived) product quality. The latter use Visa transaction data to document the importance of customers for sales (growth) heterogeneity in the retail sector and rationalize these findings using a model in which customers are a static function of current advertising expenditures.[6] While in both these models advertising affects sales in a static manner, our paper stresses the dynamic nature of firm-level customer accumulation and how it forms an endogenous feedback loop with firm-level productivity growth.

The paper is structured as follows. The next Section describes the benchmark framework and derives our central analytical results. Sections 3, 4 and 5, respectively, describe the full model, parameterize it and provide quantitative results. Section 6 documents support for key model predictions and the final section concludes.

2 A Benchmark Theoretical Framework

The primary goal of our theoretical analysis is to study how customer acquisition and R&D investment influence each other at the firm level and, in turn, how their interaction affects aggregate growth. We begin by considering a "benchmark" environment which allows us to derive closed-form solutions and build intuition for our proposed mechanism. As will become clear, these insights carry over to our full model which extends the benchmark along several dimensions and quantifies it using U.S. (micro-)data.

[5]Dynamic customer accumulation has also been studied in various other macroeconomic contexts not directly related to business dynamism, see e.g. Ravn et al. (2006), Nakamura and Steinsson (2011), Petrosky-Nadeau and Wasmer (2015), Shi (2016) or Fernández-Villaverde et al. (2021).

[6]A related paper is Rachel (2021), which studies leisure-enhancing technologies (and associated advertising) and how they raise firms' sales via brand equity.

2.1 The Model

In what follows, we use upper-case letters to denote aggregates and lower-case letters to denote firm-level variables. Time is discrete and we use primes to indicate next period's values.

Consumers: Preferences. We assume a representative household consisting of a continuum of members, indexed by k. Household members jointly supply a fixed amount of labor, $\overline{N}$, consume a final good, C, and invest into assets, A. They do so in a way that maximizes the life-time utility of the representative household, $\mathbb{U} = \sum_{t=0}^{\infty} \beta^t \ln C_t$, where β is the discount factor and where choices are made subject to a budget constraint described below.

Aggregate consumption – the economy's numeraire with a normalized aggregate price $P = 1$ – is a bundle of goods varieties, indexed by j and produced by individual firms. In what follows we use the subscript j interchangeably to denote firms and good varieties.

The key novelty of our framework lies in the treatment of customers for each goods variety. In particular, in order to consume good variety, j, household members first need to be aware of it. This happens by costly investment on the part of firms which we describe in detail below. Formally, aggregate consumption is a constant elasticity of substitution (CES) aggregate given by:

$$C = \left[\int_{j \in \Omega} \int_k \mathbb{1}_{j,k} c_{j,k}^{\frac{\eta-1}{\eta}} \, \mathrm{d}k \, \mathrm{d}j \right]^{\frac{\eta}{\eta-1}}, \tag{1}$$

where Ω is the set of producers and $\eta > 1$ is the elasticity of substitution between goods varieties.[7] Importantly, $\mathbb{1}_{j,k}$ is an indicator function equal to one if household member k is consuming good j and $c_{j,k}$ is the associated quantity.

For future reference, it will be useful to denote the "customer base" of firm j, i.e. all household members consuming its good, by $b_j = \int_k \mathbb{1}_{j,k} \, \mathrm{d}k$. While existing endogenous growth models typically treat the customer base of individual goods varieties as constant and common across firms, we allow them to vary endogenously. This is described in detail below.

Consumers: Optimal decisions. The representative household chooses consumption of individual goods and asset holdings to maximize life-time utility subject to the budget constraint

$$A' + \int_{j \in \Omega} \int_k p_j c_{j,k} \, \mathrm{d}k \, \mathrm{d}j = W\overline{N} + (1+R)A, \tag{2}$$

[7] We choose the CES structure for direct comparability with the vast majority of existing models of endogenous growth (see e.g. Acemoglu et al., 2018). With slight abuse of notation, Ω refers to both the set and mass of producers.

where total assets are given by $A = \int_{j\in\Omega} v_j \,\mathrm{d}j$, with v_j being the value of a firm producing variety j, p_j are variety-specific prices (relative to the aggregate price index P), W is the wage and R is the interest rate. The resulting Euler equation and demand for individual goods take on familiar forms:

$$1 = \beta\mathbb{E}\frac{C}{C'}(1+R'), \tag{3}$$

$$c_j = b_j p_j^{-\eta} C, \tag{4}$$

highlighting that the customer base, b_j, directly affects the demand for a firm's product.

Firms: Production. Each goods variety is produced by an individual firm using labor, n_j^c, and a firm-specific (endogenous) productivity level, q_j. The production technology is given by:

$$c_j = q_j n_j^c. \tag{5}$$

Firms: R&D. Firms strive to improve their production efficiency by investing n_j^x units of labor into R&D. In return, firms obtain an innovation probability x_j determining the success rate of their R&D efforts. We assume that R&D costs are convex in the innovation probability:

$$n_j^x = \overline{x} m_j^{\sigma_x} x_j^{\psi_x}, \tag{6}$$

where $\overline{x} > 0$, $\psi_x > 1$ and $\sigma_x > 0$ and where m_j is a firm-specific scaling factor which we define below. If successful, innovations lead to an increase in firm-level productivity by a factor of $(1+\lambda)$ where $\lambda > 0$. If unsuccessful, firm-level productivity remains unchanged:[8]

$$q_j' = \begin{cases} q_j(1+\lambda) & \text{with probability } x_j, \\ q_j & \text{with probability } 1 - x_j. \end{cases} \tag{7}$$

Firms: Customer acquisition. We assume that firms are allowed to invest into accumulating customers, enabling them to directly influence the demand for their products. Specifically, we consider two channels through which firms can attract new customers.

First, businesses can engage in costly "sales and marketing" activity to attract new customers. This channel builds on the tradition of models with informative advertising which have been used to study firm dynamics, business cycles or industry life-cycles (see e.g. Dinlersoz and Yorukoglu, 2012; Sedláček and Sterk, 2017; Abaluck and Adams-Prassl, 2021). Expenses on sales and marketing should be viewed as relatively broad and include not only investment into advertising, but also for instance customer service, spending

[8]Note that modeling productivity as cost-saving is purposefully conservative. Considering also product innovation – which may more directly impact customer accumulation through the creation of new product lines – is an interesting avenue for future research.

on sales force, or building brand value. In what follows, we will the terms “sales and marketing” and “marketing” interchangeably.

Second, businesses in our model can also expand their customer base through (past) sales. This channel follows existing models of “customer markets” which have been used to understand international prices, firm investment or labor market dynamics (see e.g. Ravn et al., 2006; Foster et al., 2016; Fitzgerald et al., forthcoming).

Given that each of these channels finds empirical support in the literature (see e.g. Foster et al., 2016; Piveteau, 2021; Afrouzi et al., 2021; Fitzgerald et al., forthcoming; He et al., 2025), we consider both in our framework simultaneously. Through the lens of our “full” model – presented in the next section – we then use micro-data to discipline the relative strength of these two channels for driving firms’ customer accumulation.

The flip side of customer accumulation is customer attrition, for which we also consider to distinct reasons. First, all incumbent firms face an exogenous customer separation rate, $\zeta \in (0, 1)$. Second, all customers of businesses which shut down also separate. We explain the process of firm exit and entry in detail below.

Given the above, the law of motion for the customer base of a continuing firm j is formally given by:

$$b'_j = b_j(1 - \zeta)\left(1 + \gamma a_j + (1 - \gamma)s_j\right), \tag{8}$$

where a_j is marketing-driven customer accumulation, s_j is sales-driven customer acquisition and where $\gamma \in [0, 1]$ is the relative weight on the two.

Investing into customer accumulation is costly. Similar to R&D costs, we assume that sales and marketing costs (denoted in labor units) are convex in the marketing rate:

$$n^a_j = \overline{a} m_j^{\sigma_a} a_j^{\psi_a}, \tag{9}$$

where $\overline{a} > 0$, $\sigma_a > 0$ and $\psi_a > 1$.

In contrast to marketing, the costs of attracting customers via pricing come in the form of foregone profits. Building on e.g. Foster et al. (2016), we assume that the rate at which relatively lower prices translate into new customers is given by:

$$s_j = m_j^{\sigma_s} \frac{p_j c_j}{PC}. \tag{10}$$

Firms: Entry and Exit. In the benchmark model, we assume that incumbent businesses shut down at an exogenous rate, $\delta \in [0, 1]$, and are replaced by an equal mass, $\delta\Omega$, of entrants. All startups are assumed to enter with a common initial endowment of customers, $b_e > 0$, and to obtain a random draw of initial firm-specific productivity, q_e. The latter is distributed identically and independently across firms according to the cummulative distribution function $H_q(q_e)$.[9] In our full model in Section 3, we endogenize both entry and exit decisions.

Firms: Optimal decisions. Incumbent firms choose prices, production labor, R&D and marketing in order to maximize the discounted value of all their future profits (firm values, v_j):

$$v(q_j, b_j) = \max_{p_j, n_j^c, n_j^x, n_j^a} \left\{ p_j c_j - W(n_j^c + n_j^x + n_j^a) + \frac{1-\delta}{1+R'} \begin{bmatrix} x_j v(q_j(1+\lambda), b_j') + \\ \\ (1-x_j) v(q_j, b_j') \end{bmatrix} \right\} \tag{11}$$

$$\begin{aligned} \text{s.t.} \quad c_j =& b_j p_j^{-\eta} C, \quad c_j = q_j n_j^c, \quad n_j^x = \overline{x} m_j^{\sigma_x} x_j^{\psi_x}, \quad n_j^a = \overline{a} m_j^{\sigma_a} a_j^{\psi_a}, \quad s_j = m_j^{\sigma_s} \frac{p_j c_j}{PC}, \\ b_j' =& (1-\zeta) b_j (1 + \gamma a_j + (1-\gamma) s_j), \quad x_j \in [0,1] \end{aligned}$$

where we have made explicit the dependence of firm values on the two firm-specific state variables: productivity and the customer base. For future reference, we will denote $\mathbb{E}\, v^c(q_j', b_j') = x_j v(q_j(1+\lambda), b_j') + (1-x_j) v(q_j, b_j')$ as the unconditional continuation (expected) firm value.

In this setting, firms will employ production labor to satisfy the demand from households, i.e. $n_j^c = c_j / q_j$. In addition, optimal innovation, marketing and pricing decisions are given by:

$$\psi_x W \frac{n_j^x}{x_j} = \frac{1-\delta}{1+R'} \frac{\partial \mathbb{E} v^c(q_j', b_j')}{\partial x_j}, \tag{12}$$

$$\psi_a W \frac{n_j^a}{a_j} = \frac{1-\delta}{1+R'} \gamma (1-\zeta) b_j \frac{\partial \mathbb{E} v^c(q_j', b_j')}{\partial b_j'} \tag{13}$$

$$p_j = \frac{\overline{\mu}}{1+\Xi_j} \frac{W}{q_j}, \tag{14}$$

where $\overline{\mu} = \frac{\eta}{\eta - 1}$ is the "static markup" and where $\Xi_j = \frac{1-\delta}{1+R'} b_j \frac{m_j^{\sigma_s}}{C} (1-\zeta)(1-\gamma) \frac{\partial \mathbb{E} v^c(q_j', b_j')}{\partial b_j'} > 0$.

In the above, optimal R&D investment (12) balances marginal costs – governed by the cost function (10) – with marginal benefits. The latter is given by the increase in expected firm value that follows from a successful innovation. Optimal marketing decisions (13) have an analogous structure, with marginal benefits reflecting the impact of a larger customer base on firm value. Finally, optimal prices (14) are set as markups over marginal costs. However, since sales increase a firm's customer base and the latter

[9] Note that our framework is characterized by entrants "standing on the shoulders of giants", whereby initial production efficiency, q_e, is assumed to be proportional to last period's aggregate productivity index (Q_{-1}).

is valuable, businesses have incentives to set markups *below* the static value, $\overline{\mu}$. The extent to which this happens is summarized by the term Ξ_j which – like the benefits of marketing investment – depends on the marginal increase in firm value stemming from additional customers.

Before moving on, we highlight two important features of our framework. First, all three of the above decisions are dynamic and forward-looking. Second, through their dependence on firm values, they also all depend on *both* firm-level productivity and the customer base. As we explain below, these features create an important firm-level link between R&D and customer accumulation.

Market clearing. Labor market clearing requires that all labor demanded by incumbent firms is supplied by the household:

$$\overline{N} = \int_{j\in\Omega} (n_j^c + n_j^x + n_j^a)\,\mathrm{d}j. \tag{15}$$

Finally, given that all costs are denoted in units of labor, the resource constraint amounts to output being equal to consumption, i.e. $C = Y$.

Aggregate growth. Focusing on the balanced growth path (BGP) of the economy, we note that all growing variables expand at the same rate given by $1 + G = Q'/Q$, where Q is an aggregate productivity index defined as:

$$Q = \left(\int_{j\in\Omega} \hat{b}_j q_j^{\eta-1}\,\mathrm{d}j\right)^{\frac{1}{\eta-1}}, \tag{16}$$

where $\widehat{b}_j = b_j/B$ is the customer share of firm j. Intuitively, aggregate productivity in our economy is given by the *customer-weighted* average of firms' productivity levels (adjusted for the elasticity of substitution in consumption). Finally, let us note that along the BGP we can stationarize our economy by dividing all growing variables by Q. In what follows, we denote stationarized variables with "hats", e.g. $\widehat{C} = C/Q$.

2.2 Central Theoretical Results

This section presents four sets of analytical results stemming from our benchmark model. First, we analytically characterize optimal firm decisions. Next, our propositions shed light on how customer accumulation alters firms' incentives to conduct R&D, how it impacts aggregate growth and how it changes the economy's sensitivity to R&D subsidies. We defer a formal definition of equilibrium as well as all proofs to the Appendix.

Parametric and functional form assumptions. The following assumption – which we relax in the full model described in Section 3 – allows us to derive closed form solutions.

Assumption **1.** *Assume the following parameter values and functional forms:*

(a) $\psi_x = \psi_a = 2$,

(b) $\sigma_x = \sigma_a = -\sigma_s = 1$ *and* $m_j = b_j \widehat{q}_j^{\eta-1}$ *for all* j,

(c) $(1-\delta)b'_j/b_j < 1$ *for all* j.

Part (a) of Assumption 1 is not particularly restrictive. Empirical evidence on the curvature of the R&D cost function typically points to a value of $\psi_x = 2$ (see e.g. Hall and Ziedonis, 2001; Bloom et al., 2002). In addition, recent evidence suggests similar values also for advertising costs (see e.g. Afrouzi et al., 2021). Next, as will become clear, m_j will be proportional to firms' market shares, $p_j c_j/(PC)$. Therefore, Part (b) of Assumption 1 is akin to the "perfect scaling" assumptions in existing endogenous growth models (see e.g. Akcigit and Kerr, 2018, for a discussion). Importantly for the purposes of this subsection, Assumption 1 allows for closed form solutions and leads to analytically convenient properties of the model, such as Gibrat's law. Finally, Part (c) of Assumption 1 ensures a stationary aggregate customer base (see the Appendix for further details).

Optimal choices. Before presenting our theoretical results, we first describe optimal firm-level choices under Assumption 1. In doing so, it will be convenient to define the following firm-level growth rates of productivity (g_q), customer base (g_b) and employment (g_n):

$$1 + g_q = x(1+\lambda)^{\eta-1} + 1 - x, \tag{17}$$

$$1 + g_b = (1-\zeta)(1 + \gamma a + (1-\gamma)s), \tag{18}$$

$$1 + g_n = (1+g_q)(1+g_b)(1+G)^{1-\eta}. \tag{19}$$

Proposition **1.** *Under Assumption 1, optimal firm-level innovation (x), marketing (a)*

and pricing (μ) decisions are constant, common to all firms and given by:

$$x = \widetilde{\beta}\frac{(1+\lambda)^{\eta-1}-1}{2\hat{W}\overline{x}\,(1+g_q)}\overline{v}, \tag{20}$$

$$a = \widetilde{\beta}\frac{\gamma(1-\zeta)}{2\hat{W}\overline{a}(1+g_b)}\overline{v}, \tag{21}$$

$$\frac{1}{\mu}-\frac{1}{\overline{\mu}} = \widetilde{\beta}\frac{(1-\zeta)(1-\gamma)}{\hat{C}\overline{\mu}(1+g_b)}\overline{v}, \tag{22}$$

where $\widetilde{\beta}=\beta(1-\delta)(1+g_n)$ and where $\overline{v}$ is given by

$$\overline{v} = \frac{\hat{W}^{1-\eta}\hat{C}\mu^{-\eta}(\mu-1)-\hat{W}(\overline{x}x^2+\overline{a}a^2)}{1-\widetilde{\beta}}.$$

Proposition 1 shows that our benchmark framework is characterized by Gibrat's law where firm-level employment growth, g_n, is independent of size (which in our model is driven by a combination of productivity and customer base levels). Note also that $\overline{v}$ reflects the net present value of all future profits, which are proportional to market shares, i.e. $v_j = \overline{v}m_j$.

Customer acquisition and firm-level innovation. With optimal choices at hand, we now turn to understanding how firms' R&D and customer accumulation decisions interact. We do so by analyzing comparative statics in partial equilibrium which hold all other variables (including aggregate growth) fixed. In particular, in what follows we derive firms' endogenous responses to exogenous changes in their customer base and productivity growth rates. While the former is brought about by exogenous changes to customer retention $(1-\zeta)$, the latter is due to exogenous changes in innovation step size $(1+\lambda)$.

Proposition **2.** *Under Assumption 1, and all else equal,*
(a) Firm-level innovation increases with customer base growth:

$$\epsilon_{x,1-\zeta} = \frac{\partial x}{\partial(1-\zeta)}\frac{1-\zeta}{x} > 0.$$

(b) Firm-level customer accumulation increases with innovation

$$\epsilon_{a,1+\lambda} = \frac{\partial a}{\partial(1+\lambda)}\frac{1+\lambda}{a} > 0 \quad \textit{and} \quad \epsilon_{\mu,1+\lambda} = \frac{\partial \mu}{\partial(1+\lambda)}\frac{1+\lambda}{\mu} < 0.$$

Part (a) of Proposition 2 shows that innovation depends on the speed of customer base accumulation. Specifically, an exogenous increase in a firm's customer base growth increases its future profits. Aside from this "direct" effect, firms respond optimally by increasing marketing investment (i.e. $\partial a/\partial(1-\zeta) > 0$) and lowering markups (i.e. $\partial\mu/\partial(1-\zeta) < 0$). Both of these indirect effects endogenously increase customer acquisition further. Combined, an exogenous increase in customer base growth raises firm value which provides stronger incentives to conduct R&D. We dub this effect the "firm-level market size effect." Versions of this effect have been documented in a range of empirical studies (see e.g. Acemoglu and Linn, 2004; Jaravel, 2019; Aghion et al., 2020).

Part (b) of Proposition 2 states that a similar effect also operates in the opposite direction – from innovation to customer acquisition. In particular, an exogenous increase in innovation step size leads to an endogenous rise in R&D investment as successful innovation becomes more attractive (i.e. $\partial x/\partial(1+\lambda) > 0$). Both these effects raise productivity growth and, in turn, firm value which provides stronger incentives to acquire customers. Put together, our framework features an endogenous feedback loop between customer acquisition and R&D. This feature of our model will be important for our quantitative results.[10]

Customer acquisition and aggregate growth. The previous proposition analyzed optimal firm-level R&D decisions. We now turn to understanding how these decisions influence aggregate growth. Note, however, that in doing so, we "only" aggregate firms' partial equilibrium responses and abstract away from the general equilibrium effect of aggregate growth on firms' innovation. As such, the following proposition speaks to the *impact* response of aggregate growth that follows from exogenous changes in firms' customer base levels. We leave the analysis of complete general equilibrium effects to the full model in Section 3.

PROPOSITION **3.** *Under Assumption 1,*
(a) Aggregate growth is given by

$$\frac{Q'}{Q} = 1 + G = \left[(1+\lambda)^{\eta-1}x + 1 - x\right]^{\frac{1}{\eta-1}} = (1+g_q)^{\frac{1}{\eta-1}}.$$

(b) All else equal, aggregate growth increases with firm-level customer base growth

$$\frac{\partial G}{\partial(1-\zeta)} > 0.$$

First, Part (a) of Proposition 3 makes clear that aggregate economic growth is driven by firms' endogenous R&D choices. Importantly, notice that aggregate growth does not

[10]While Proposition 2 considers a change in ζ and λ which in our model are common to all firms, the same results are obtained when considering *firm-specific* increases in customer attrition (ζ_j) and innovation step size (λ_j).

directly depend on the growth of firms' customer base levels.

That said, Part (b) of Proposition 3 shows that expansions of firms' customer base levels do impact aggregate economic growth *indirectly*. This is because firm-level customer base growth increases the incentives to conduct R&D, raising innovation rates – the firm-level market size effect (see Proposition 2).

Therefore, Proposition 3 provides a new view on the driving forces of long-run economic growth. In particular, it establishes that customer base accumulation has the potential to affect aggregate growth through its impact on firm-level innovation incentives.

Foreshadowing some of our quantitative results, there is also a second potential channel through which customer base accumulation may impact aggregate growth – through the (re-)allocation of customers across firms with different productivity (growth rates). Note that in the benchmark model presented here, all such effects are absent by construction. This is because all firms grow at the same pace (see Proposition 1), leaving no room for reallocation of customers towards faster-growing firms. In fact, Proposition 3 highlights that – since all firms grow at a common and constant rate – the distribution of firms is immaterial for aggregate growth.[11] While this feature of the model is important for our analytical results, it will not hold in the full model where Assumption 1 is relaxed.

Customer acquisition and the sensitivity to R&D subsidies. Finally, we turn our attention to understanding how the presence of customer accumulation at the firm-level affects the sensitivity of the economy to R&D subsidies.

PROPOSITION **4.** *Consider R&D subsidies such that a fraction τ_x of firm-level R&D costs is paid by the government, financed by lump-sum taxes on the household. Under Assumption 1,*
(a) R&D subsidies raise innovation

$$\epsilon_{x,\tau_x} = \frac{\partial x}{\partial \tau_x}\frac{\tau_x}{x} > 0,$$

(b) The efficacy of R&D subsidies increases with customer base growth

$$\frac{\partial \epsilon_{x,\tau_x}}{\partial(1-\zeta)} > 0.$$

Part (a) of Proposition 4 states that R&D subsidies lead to an increase in innovation rates. This is intuitive as R&D investment becomes cheaper, raising incentives to innovate.

[11] Notice that despite Proposition 1 establishing that all firms make the same choices, firms are heterogeneous in their productivity and customer base *levels*. This heterogeneity, however, does not impact firms' choices.

Next, Part (b) of Proposition 4 shows that the sensitivity of innovation to R&D subsidies is higher when the growth rate of firms' customer bases is stronger. The intuition lies in firms' discounting of future profits. With higher growth in the stock of customers, firms place more weight on future periods making them more sensitive to changes in profits brought about by R&D subsidies. Overall, Proposition 4 highlights that it is important to take firm-level customer accumulation into account when gauging the efficiency of pro-growth policies.

Summary of theoretical results. Our benchmark model puts forward three key messages. First, incentives to conduct R&D depend on the speed of customer accumulation (see Proposition 2). This occurs because expansions of the customer base increase the benefits of successful innovations – the firm-level market size effect. Second, given that aggregate growth is driven by firms innovative activity, changes in customer accumulation also have the potential to affect aggregate growth (see Proposition 3). Finally, firm-level market size effects change the sensitivity with which firms (and therefore the aggregate economy) respond to the cost of R&D (see Proposition 4), e.g. due to governmental subsidies.

3 The Full Model

In this Section, we introduce our "full" model which relaxes Assumption 1 and extends the benchmark framework along several dimensions. We use the full model to quantify the role of customer acquisition for firm-level R&D and aggregate economic growth – effectively quantifying Propositions 2 to 4.

3.1 Generalizations and Their Implications

Since the full model retains the key structure of the benchmark framework – including the optimal R&D, marketing and pricing decisions (12) to (14) – we now only describe the novel features of the full model. We defer a formal definition of equilibrium to the Appendix.

Imperfect scaling. Assumption 1 in the benchmark model imposed "perfect scaling," i.e. $\sigma_x = \sigma_a = -\sigma_s = 1$. This implies that the costs of R&D and marketing, as well as the customer base benefits of past sales, all scale one-for-one with $m_j = b_j \hat{q}_j^{\eta-1}$. While analytically convenient, it is an empirical question whether these patterns hold in the data. Therefore, as we discuss in detail in Section 3.2, in our full model we build on Akcigit and Kerr (2018) and estimate the scaling parameters σ_x, σ_a and σ_s.

Endogenous firm exit. In order for our full model to be able to speak to the process of firm selection and "creative destruction", we endogenize firm entry and exit decisions. We begin by describing endogenous firm exit.

In particular, we assume that, at the beginning of each period, firms must pay a per-period cost κ_o (denoted in units of labor) in order to stay in operation. We assume that such operational costs are stochastic, distributed identically and independently across firms according to the cumulative distribution function $F(\kappa)$ with mean μ_κ and dispersion σ_κ.

If businesses choose not to pay the cost, they shut down and obtain a return of zero. In this setting, firms will choose to shut down whenever operational costs exceed a threshold level defined by

$$\kappa_j^* = \frac{v(q_j, b_j)}{W}. \tag{23}$$

Note that, unlike in existing growth models, the exit decisions depends not only on productivity, but also on firms' accumulated customers. This feature will be important for understanding firm selection in our framework and we return to this point in our quantitative analysis.

Before moving on, it will be convenient to denote the expected paid operational costs, conditional on survival, by $\widetilde{\kappa}_j = \int^{\kappa_j^*} \kappa dF(\kappa)$. In addition, we will denote firm value net of such paid operational costs and conditional on survival by

$$v_j^c = F(\kappa_j^*) \left[v_j - \widetilde{\kappa}_j \right]. \tag{24}$$

Endogenous firm entry. We assume that every period a mass of potential startups has the option to enter the economy. In order to do so, they must first pay a fixed entry cost κ_e (denoted in labor units). Upon paying the entry cost, entrants obtain initial values of customers, b_e, and productivity, q_e. In the full model, we assume that both initial productivity and customer base levels are stochastic and distributed according to a joint cumulative distribution function $H_e(q_e, b_e)$. Assuming free entry, gives rise to the following entry condition:

$$W \kappa_e = \int v^c(q_e, b_e) \, \mathrm{d}H_e(q_e, b_e). \tag{25}$$

The above entry condition implicitly defines the mass of new entrants, Ω_e. The mass of firms, Ω, is endogenous in the full model where, in equilibrium, firm entry and exit balance:

$$\Omega_e = \Omega \left(\delta + (1 - \delta) \int_j (1 - F(\kappa_j^*)) \, \mathrm{d}j \right). \tag{26}$$

Market Clearing. Given endogenous entry and exit, labor market clearing is given by

$$\overline{N} = \int_{j \in \Omega} (n_j^c + n_j^x + n_j^a + \widetilde{\kappa}_j)\, \mathrm{d}j + \Omega_e \kappa_e. \tag{27}$$

which now includes labor demand of incumbent firms (including their operational costs) and of potential entrants.

3.2 Taking the Full Model to the Data

We now describe how we map the full model to the data. The next section uses the parameterized framework for a quantitative analysis.

Measuring investment into customers. Investment into customer accumulation plays a central role in our paper. As we explained earlier, firms in our framework attract customers through two distinct channels – pricing (markups) and costly investment into sales and marketing. To measure these two activities, we lean on the existing literature.

In particular, we follow De Loecker et al. (2020) and measure markups as the inverse share of variable costs in sales, multiplied by the respective cost-output elasticity. As argued in de Ridder et al. (2024), while the level of markups requires price information, variation in markups across firms can be identified from financial data. Therefore, assuming that cost-output elasticities are constant across firms in a given year and industry, we use firms' balance sheets to estimate markups with cost-output elasticities being absorbed by a set of fixed effects in our empirical specifications.

Next, existing studies typically use information on "selling, general and administrative" (SGA) expenses to measure firms' investment into customer accumulation and intangibles more generally (see e.g. Gourio and Rudanko, 2014; Crouzet et al., 2022). However, SGA contains other variables which may not be closely related to customer investment. Therefore, some studies focus on a much narrower concept of "advertising" expenses (see e.g. Belo et al., 2014; Afrouzi et al., 2021).[12]

In this paper, we follow He et al. (2025) who explicitly measure "sales and marketing" expenses. While this investment is a sub-component of selling, general and administrative expenses, it is a broader concept than advertising expenditures alone as it includes

[12] Among others, SGA was used to measure customer acquisition investment in Drozd and Nosal (2012), Gourio and Rudanko (2014), Afrouzi et al. (2021). The latter documents that both SGA as well as its subcomponents of advertising expenses and rent expenses correlate well with the increase in the number of customers at the firm level. Argente et al. (2025) use Nielsen retail data to constructs a direct measure of customer acquisition in the food sector showing that two-thirds of sales growth after product entry is due to the extensive margin of customers. They show further that advertising rather than pricing is the main channel of customer acquisition. Alternative sources of information – such as estimates of brand value, information on workers in sales and marketing occupations or customer churn based on credit transaction data – have also been used in the literature to infer investment into customer accumulation (see e.g. Larkin, 2013; Bronnenber et al., 2022; Baker et al., 2023).

also spending on sales force, brand development, customer service or customer data acquisition and use. In what follows, we provide more details on the associated data and its treatment.

Data sources. Aside from the aggregate data released by the BLS, we make use of three primary sources of firm-level information when disciplining our model quantitatively: (i) Business Dynamic Statistics (BDS) of the Census Bureau, (ii) Compustat and (iii) Capital IQ. The key advantage of the BDS is its macroeconomic scope, covering virtually the universe of private-sector employers in the U.S. economy. For this reason, the BDS – with its underlying micro-data, the Longitudinal Business Database – is one of the main data sources for studying U.S. business dynamism (see e.g. Haltiwanger et al., 2013). However, while the BDS is broad in its coverage, its primary focus is on employment. Therefore, we complement the BDS with information from Compustat and Capital IQ which constitute other important sources of U.S. firm-level information.

While Compustat data has been used extensively in the literature to study the key concepts in our paper – firm-level markups, customer accumulation and R&D (see e.g. Gourio and Rudanko, 2014; De Loecker et al., 2020; Cavenaile et al., 2021; Afrouzi et al., 2021) – it is not without its difficulties. First, it is well known that the Compustat sample of (publicly traded) firms is not representative of the whole private sector. To address this issue, when computing model-generated moments with counterparts in Compustat, we use estimated employment-based weights which align the Compustat and BDS firm-size distributions. Second, there are certain concerns about the coverage of R&D data in Compustat. For instance, Walmart famously does not report any R&D expenditures in Compustat. Therefore, we explicitly show how parameters pinned down using Compustat data align with estimates found in other studies using alternative data sources.

Third, as we discussed in detail above, Compustat lacks precise information on firms' efforts to acquire customers. Therefore, we supplement firm level information with S&P's Capital IQ data, where direct measures of sales and marketing expenses (S&M) are available starting in 1997. We follow He et al. (2025) and link Capital IQ to Compustat. We are able to match 91% of our baseline Compustat sample to Capital IQ firms. Among these, 34% of firms report sales and marketing expenses directly. In addition, following He et al. (2025), we recover an additional 14% of firms which do not report selling and marketing expenses, but report their sub-components (either "advertising" expenses, "marketing" expenses or both). In these cases, we use either of the two or, if available, the sum of the two sub-components as a measure of sales and marketing expenses. Taken together, Capital IQ yields information on sales and marketing expenses for 48% of our sample.[13]

[13] In addition to sales and marketing expenses, Capital IQ includes the date at which firms were established and, if the former is unavailable, the incorporation date for a large subset of the firms. We

In what follows, we describe the parameterization of our model. We begin by common choices and normalizations and then describe how the remaining parameters are set using indirect inference. Note that due to the availability of Capital IQ data, targets derived from the marched Compustat-Capital IQ matched data are based on a sample between 1997 and 2019. All other targets are computed for a sample between 1979 and 2019, consistent with the BDS. All model parameters are reported in Table 2 and Figure 1 shows the model fit in terms of targeted moments.

Common choices. First, consistent with the annual frequency of our firm-level data, we assume the model period to be one year. Therefore, we set the discount factor to $\beta = 0.97$. Second, the elasticity of substitution between goods is set to $\eta = 3$. This value falls within the range of median estimates reported in Broda and Weinstein (2006). Finally, as discussed in Section 2, we set the R&D cost elasticity to $\psi_x = 2$. This value is consistent with empirical studies (see e.g. Hall and Ziedonis, 2001; Bloom et al., 2002) as well as being a common choice in a host of existing growth models (see e.g. Acemoglu et al., 2018; Akcigit and Kerr, 2018).

Normalizations. We set the fixed cost of entry, κ_e, which controls the mass of firms in the economy, such that aggregate consumption is normalized to $\widehat{C} = 1$. Next, we set aggregate labor supply, $\overline{N}$, such that the equilibrium wage is normalized to $\widehat{W} = (\eta - 1)/\eta$.

Indirect inference. The remaining parameters are estimated jointly using an indirect inference approach in the spirit of Lentz and Mortensen (2008). In particular, we compute various model-generated moments and compare them to their respective empirical counterparts to pin down the remaining parameters by minimizing

$$\min \sum_i w(i) \left(\frac{\text{model}(i) - \text{data}(i)}{\frac{1}{2}\left(\text{model}(i) + \text{data}(i)\right)} \right)^2 ,$$

where i indicates a particular moment and $w(i)$ its associated weight with $\sum_i w(i) = 1$. We choose the weights so that each group of targets (e.g. firms' life-cycle exit profiles) has a similar representation in the entire set of moments, while placing a higher importance on aggregate growth and moments pertaining to customer accumulation. We describe the estimation method and the weights in detail in Appendix C.1.

While in general all parameters affect model-implied moments, in what follows we discuss our chosen targets in conjunction with the parameters which are most closely linked to them.

use this information to construct firm age. We also use the Loughran-Ritter database (Loughran and Ritter, 2004) to supplement information of firm founding dates for some firms with missing founding date in Capital IQ. Appendices D.1 and C provide further details on the parametrization, solution method and the estimation of our size-based weights.

Firm-level R&D and aggregate growth. In our model, there are three parameters most closely related to firm-level and aggregate productivity growth: the innovation step size, λ, the level of R&D costs, $\overline{x}$, and the "scaling" of R&D costs with size, σ_x. To pin these parameters down, we require our model to match the following three targets.

First, we make our model match an annual aggregate growth rate of 1.43%, corresponding to the growth of U.S. real GDP per worker between 1979 and 2019 – consistent with the pre-pandemic BDS sample. Second, we target the aggregate ratio of R&D expenditures to sales which is 2% in our Compustat sample. This value is consistent with the national accounts ratio of real R&D expenditures to real GDP of about 2.7% in the 1979-2019 sample period.

Finally, to discipline the extent to which R&D costs scale with firm size, we follow Akcigit and Kerr (2018) and target a reduced form relationship between firm-level R&D intensity and size

$$\ln\left(1 + \text{R\&D/sales}\right)_{j,t} = \delta_{i,t} + \beta \ln \text{sales}_{j,t} + \epsilon_{j,t}, \tag{28}$$

where $\delta_{i,t}$ are industry-time fixed effects with i indicating 3-digit industries. Using Compustat data, we estimate $\beta = -0.01$, with a standard error of 0.001. Replicating regression (28) on model-simulated data and targeting the reduced form coefficient implies a value of $\sigma_x = 1.264$ which is close to that estimated in Akcigit and Kerr (2018).[14]

Customer accumulation: The depreciation rate. To discipline the customer base depreciation rate, ζ, we make use of the fact that it affects all businesses. As such, plays a key role in determining average firm level growth, see (17). Therefore, we set ζ to target average firm-level employment growth, yielding $\zeta = 0.375$.[15] This parameterization falls towards the lower end of estimates found in existing studies which range from about 0.3 (see e.g. Gourio, 2014; Afrouzi et al., 2021; Kaas and Kimasa, 2021) to 0.7 (see e.g. Foster et al., 2016; Fitzgerald et al., forthcoming).

Customer accumulation: Relative strength of the two channels. In our model, firms can attract new customers through two channels: marketing or sales (markups). The relative strength of these two channels is governed by the weight γ. To determine the value of γ, we exploit the observed relationship between firm-level growth, marketing expenditures and markups.

[14] In particular, Akcigit and Kerr (2018) estimate that costs of external innovations scale with size (the number of product lines) with $\widetilde{\sigma} = (1-\sigma)/\psi \approx 1.2$.

[15] As will become clear below, we target the entire life-cycle profile of average firm size growth in the BDS data.

In particular, up to first order, we can write firm-level sales growth as:

$$g_{j,t} \approx \Delta \ln(\text{sales}_{j,t}) = \delta_{i,t} + \beta_a \Delta \ln n^a_{j,t-1} + \beta_\mu \Delta \ln \mu_{j,t-1} + \Gamma \mathbb{X}_{j,t} + \epsilon_{j,t}, \tag{29}$$

where Δ is the first-difference operator, $\text{sales}_{j,t}$ are given by $p_{j,t}c_{j,t}$ in our model, $\delta_{i,t}$ are industry-time fixed effects accounting for sectoral and aggregate growth and $\mathbb{X}_{j,t}$ represents additional control variables.[16] The following proposition illustrates how we use regression (29) to pin down the value of γ in our model.

Proposition **5.** *Under Assumption 1, the following holds in regression (29):*

$$\beta_\mu = (1-\eta)((1-\zeta)(1-\gamma)s - 1) = \beta_0 + \beta_1 \gamma^2,$$

where $\beta_0 = (\eta - 1)(1 - g_b - \zeta)$ *and* $\beta_1 = \frac{\beta(1-\delta)(1-\zeta)^2(\eta-1)\overline{v}}{\overline{a}\hat{W}}$.

The intuition for the proposition (the proof of which is left to the Appendix) is that β_μ from (29) quantifies the contribution of changes in markups to firm-level growth. This coefficient reflects three components: (i) the elasticity with which changes in markups translate into sales, $1-\eta$, (ii) the direct effect this has on current-period sales, -1, and (iii) the indirect effect such changes have on future sales via price-driven customer accumulation, $(1-\gamma)(1-\zeta)s$.

To estimate (29), we use our Compustat sample and measure firm-level markups using the methodology in De Loecker et al. (2020). In addition, given the established importance of firm size and age for firm-level growth (see e.g. Haltiwanger et al., 2013), we also consider these variables as additional controls. Finally, we consider 3-digit industry-year fixed effects and cluster standard errors at this level (see the Appendix for further details).

The first row of Table 1 shows the results. The coefficient β_μ is estimated to be 0.04 with a standard error of 0.01. Replicating regression (29) on model-simulated data and targeting β_μ implies a value of $\gamma = 0.745$. This result suggests that while firms use markups to accumulate customers, marketing plays a stronger role. This finding is consistent with other recent evidence (see e.g. Afrouzi et al., 2021; Fitzgerald et al., forthcoming).

Customer accumulation: Marketing. Marketing-driven customer accumulation is most closely determined by the following three parameters: the level of marketing costs, $\overline{a}$, the scaling of marketing costs with size, σ_a, and the elasticity with which marketing expenditures translate into customer base growth, ψ_a.

[16] These control variables also stem from the first-order approximation of sales growth and include lagged growth rates of R&D expenditure, sales and markups. In Appendix D, we provide for further details as well as show that very similar results are obtained when using firm-level employment growth rather than sales growth as the dependent variable.

Table 1: Customer base drivers of firm-level growth

Dependent variable: $\Delta \ln \text{sales}_{j,t}$	(I)	(II)	(III)
$\Delta \ln \mu_{j,t-1}$	0.04^{***}	0.04^{***}	0.06^{***}
	(0.01)	(0.01)	(0.02)
$\Delta \ln n^a_{j,t-1}$	0.08^{***}	0.07^{***}	0.07^{***}
	(0.01)	(0.01)	(0.01)
Controls:			
Industry-year fixed effects	✓	✓	✓
$X_{j,t}$	✓	✓	✓
Firm size		✓	✓
Firm age			✓
Observations	15,502	15,502	12,452
R^2	0.24	0.24	0.27

Note: The table shows estimates of regression (29) using firms in the Compustat - Capital IQ merged sample that report R&D expenses and sales and marketing expenses between the years 1997 and 2019. Markups, μ are estimated using the methodology in De Loecker et al. (2020) and sales and marketing expenses are measured using Capital IQ. Firm size is given by log(assets) and age is measured as the log of number of years since establishing date reported in Capital IQ or the Loughran-Ritter database (Loughran and Ritter, 2004). Standard errors (in brackets) are clustered at the firm level.

First, He et al. (2025) show that firms' "selling and marketing expenses" from the Capital IQ dataset closely correlate with other measures of firms' investment in customer acquisition. Therefore, we use selling and marketing expenses in the data whenever referring to marketing costs. To determine the level of marketing costs, $\overline{a}$, we target the aggregate marketing-to-sales ratio in Capital IQ and Compustat data of about 2.6%. Next, as with R&D expenditures, we estimate the following reduced form regression to pin down σ_a:

$$\ln\left(1 + \text{S\&M/sales}\right)_{j,t} = \delta_{i,t} + \beta \ln \text{sales}_{j,t} + \epsilon_{j,t}, \tag{30}$$

where we estimate $\beta = -0.01$ with a standard error of 0.001. Targeting this reduced form regression coefficient in model simulated data yields $\sigma_a = 1.405$.

Finally, to pin down ψ_a, we once again exploit the observed relationship between firm-level growth and marketing expenditures (29). The following proposition illustrates the sources of identification of ψ_a in our model.

Proposition 6. *Under Assumption 1, the following holds in regression (29):*

$$\beta_a = \frac{1}{\psi_a}(1-\zeta)\gamma a = \frac{1}{\psi_a}\frac{\gamma^2}{\eta - 1}\beta_1,$$

where β_1 is the same as in Proposition 5.

The intuition for Proposition 6 once again lies in the fact that β_a quantifies the contribution of changes in marketing expenditure to firm-level growth. This coefficient reflects two components: (i) the elasticity with which marketing expenditures translate to customer base growth, $1/\psi_a$ and (ii) the contribution of marketing-driven customer base growth, $(1-\zeta)\gamma a$. Therefore, given all other parameters, the estimated coefficient β_a from regression (29) pins down the value of ψ_a.

The second row of Table 1 shows the results. Across our specifications, β_a is estimated at 0.07 with a standard error of 0.01. Replicating regression (29) on model-simulated data and targeting β_a implies a value of $\psi_a = 4.9$.[17]

Customer accumulation: Sales. To disciple the elasticity with which sales-driven customer accumulation scales with firms' size, σ_m, we rely on the same procedure as with similar elasticities related to marketing and R&D costs. In particular, we estimate the following reduced form regression:

$$\ln\left(1+\mu_{j,t}\right) = \delta_{i,t} + \beta \ln(\text{sales}_{j,t}) + \epsilon_{j,t}. \tag{31}$$

In our Compustat sample, and using markups computed according to De Loecker et al. (2020), we estimate $\beta = 0.02$ with a standard error of XXX. This finding is consistent with e.g. Autor et al. (2020) who also find markups to rise with firm size. In the model, we replicate this result by setting $\sigma_m = -1.150$.

Firm exit. There are three parameters which most directly control the patterns of firm exit: the exogenous exit rate (δ), the level and dispersion of stochastic fixed cost shocks (μ_κ and σ_κ). To discipline these parameters, we require our model to match the entire life-cycle profile of firm exit rates (between startup and age 20) from the BDS.

Intuitively, exogenous exit (δ) applies to all firms and, therefore, affects the overall level of exit rates. Next, as businesses grow larger over their life-cycle, fixed costs gradually become less of a threat to their operations. Therefore, the level (μ_κ) of fixed costs helps match the relatively higher exit rates of young firms. Finally, the dispersion of fixed costs (σ_κ) is closely tied to the speed with which exit drops with firm age.

Firm entry. We assume that initial values of customers and productivity are drawn independently from one another. In particular, startups obtain an initial customer endowment distributed identically and independently over time and across firms according to $\log b_e \sim N(0, \sigma_b)$. Similarly, initial firm productivity is distributed identically and

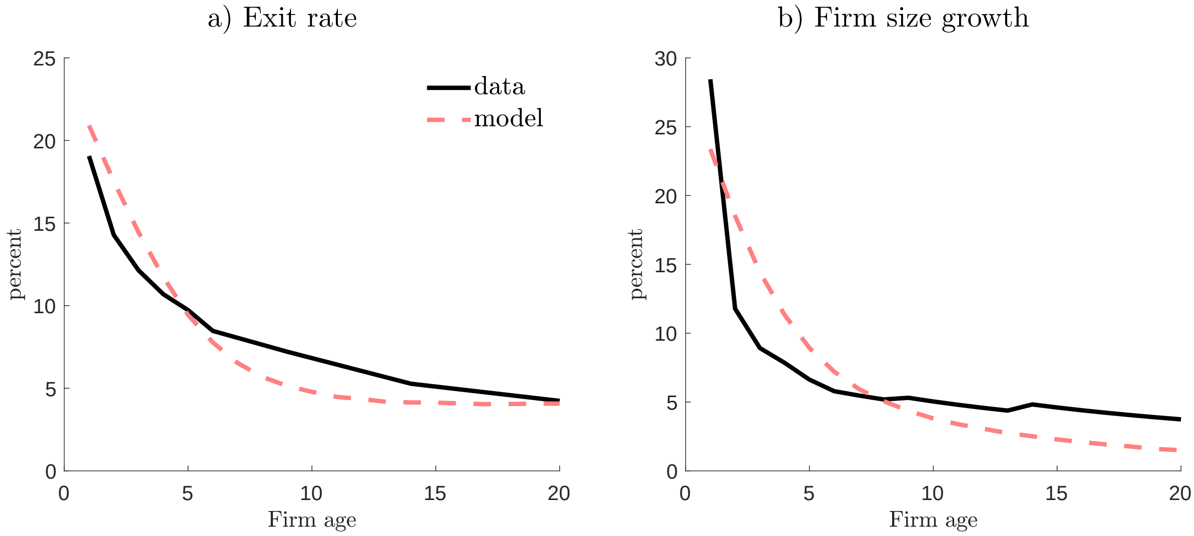
[17] For comparison, Afrouzi et al. (2021) estimate "decreasing returns to advertising" to be 0.53. Through the lens of our model, such decreasing returns amount to about $1/\psi_a \approx 0.2$. Note, however, that Afrouzi et al. (2021) use Compustat's SGA to proxy for marketing expenditures. In an earlier version of our paper, we used SGA (instead of Capital IQ's sales and marketing expenditures) to parameterize our model which resulted in $1/\psi_a \approx 0.44$.

Table 2: Parameter values and targeted moments

	Parameter	Value	Target/Source	Data	Model
β	Discount factor	0.970	Interest rate of approx. 3%		
η	Elasticity of substitution	3	Within range of Broda and Weinstein (2006)		
$\overline{N}$	Labor supply	1.392	$\widehat{W} = (\eta - 1)/\eta$, normalization		
κ_e	Entry cost	5.847	$\widehat{C} = 1$, normalization		
ψ_x	Cost elasticity of R&D	2	see e.g. Hall and Ziedonis (2001); Bloom et al. (2002)		
$\overline{x}$	Cost level of R&D	0.301	Aggregate R&D/sales	0.02	0.01
σ_x	R&D cost scaling with size	1.264	R&D-to-sales on size reg. coef. (28)	−0.02	−0.01
λ	Innovation step size	0.044	Real GDP growth rate (%)	1.43	1.43
γ	Advertising-driven customer accumulation weight	0.745	Firm growth on markup changes (29)	0.04	0.04
$\overline{a}$	Cost level of advertising	1.908	Aggregate advertising/sales	0.03	0.06
σ_a	Advertising cost scaling with size	1.405	Adv.-to-sales on size reg. coef. (30)	−0.01	−0.01
ψ_a	Cost elasticity of advertising	4.900	Firm growth on advertising changes (29)	0.08	0.11
σ_m	Markup-driven customer accumulation scaling with size	−1.150	Markups by sales reg. coef. (31)	0.02	0.03
ζ	Customer depreciation rate	0.375	Firm growth life-cycle profile, see Figure 1		
σ_b	Initial customer base, dispersion	0.860	Firm growth life-cycle profile, see Figure 1		
σ_q	Initial productivity, dispersion	0.729	Firm growth life-cycle profile, see Figure 1		
δ	Exogenous exit rate	0.040	Firm exit life-cycle profile, see Figure 1		
μ_κ	Stochastic fixed costs, mean	0.800	Firm exit life-cycle profile, see Figure 1		
$\widetilde{\sigma}_\kappa$	Stochastic fixed costs, dispersion	2.000	Firm exit life-cycle profile, see Figure 1		

Note: The table presents values for all model parameters. The first three columns describe each parameter and its parameterized value. The fourth and fifth columns indicate the most relevant target or source and the last column reports the model fit.

Figure 1: Targeted moments: Exit and size growth rates



Note: The figure shows life-cycle profiles of exit rates (panel a) and average firm size growth rates (panel b). It does so for BDS "data" (solid line) and "model" (dashed line) generated moments.

independently over time and across firms according to $\log q_e \sim N(\log Q_{-1}, \sigma_q)$, where the mean is centered around the previous period's aggregate productivity index.

Intuitively, initial conditions are informative about firms initial growth rates. The distribution of initial customer base levels among startups – which in our case depreciates at a rate of $\zeta = 0.375$ – is particularly closely related to the persistence of firm-level growth rates. Therefore, to help discipline the initial distribution of productivity and customer base levels, we require our model to replicate the entire life-cycle profile of average firm size in the BDS data between startup and age 20.[18]

3.3 Model Performance

Before moving to our quantitative results, we document how our model matches targets and a range of untargeted moments.

Targeted moments. Figure 1 and Table 2 display the model's performance in terms of targeted moments. These include the entire life-cycle profiles of firm exit and growth rates, as well as a range of moments related to R&D and marketing patterns.

Untargeted moments. In addition to targeted moments, Table 3 shows that our model is also consistent with a range of untargeted moments pertaining to the firm size distribution and job creation and destruction patterns.

Finally, we gauge the realism of our model-implied customer base dynamics. Towards

[18]Since the BDS reports values in 5 year bins for businesses older than 5 years, we use interpolated exit rate and firm size values for each firm age.

Table 3: Untargeted moments: Data and model

	Data	Model
Fraction of firms with < 20 workers	0.89	0.87
Fraction of firms with $20 - 100$ workers	0.10	0.13
Job creation rate	0.17	0.13
Job destruction rate	0.15	0.13
Firm-level demand shock, persistence	0.91	0.90
Firm-level demand shock, dispersion	1.16	0.98

Note: The table shows untargeted moments in the "data" and "model", respectively. Information on firm size and job creation and destruction rates is taken from the BDS. The last two rows show the persistence and standard deviation of "demand shocks" as estimated in Foster et al. (2008), see Table 3.

this end, we draw upon the methodology and empirical evidence presented in Foster et al. (2008). Specifically, the authors estimate firm-level "demand shocks", $\widetilde{d}_{j,t}$, as residuals from the following regression[19]

$$\ln c_{j,t} = \beta_0 + \beta_1 \ln p_{j,t} + \widetilde{d}_{j,t}. \tag{32}$$

Notice that in our framework, $\widetilde{d}_{j,t}$ exactly represent the logarithm of firms' customer base levels, $\ln b_{j,t}$, see (4). Therefore, estimating (32) on simulated data from our framework provides a direct comparison of model-implied customer base patterns with their empirical counterparts. The model estimates of the persistence of $\widetilde{d}_{j,t}$ are close to those in the data, see Table 3. The model somewhat underestimates the dispersion of of demand shocks $\widetilde{d}_{j,t}$.

4 Quantitative Analysis

In this section, we use our parameterized model for three purposes. First, we quantify how customer acquisition affects firm-level incentives to conduct R&D. Second, we investigate the impact on aggregate growth. Finally, we use our model to demonstrate that ignoring customer accumulation – as is done in existing growth models – provides significantly different conclusions about the efficacy of growth policies. These three sets of results, therefore, quantify our theoretical Propositions 2 to 4 for the U.S. economy.

4.1 Customer Accumulation and Firm-Level R&D

We begin our quantitative analysis by focusing on the impact of firm-level market size effects on R&D decisions of individual firms. The following paragraphs first explain how

[19]The authors use physical output and instrument firm-level prices with estimates of physical productivity, TFPQ. In our model, we observe prices directly. Nevertheless, following their two-step procedure delivers similar results.

we isolate the impact of customer acquisition.

Customer accumulation and firm-level R&D: Quantification. To isolate – in an accounting sense – the impact of customer base expansions on firms' R&D decisions, we compare results from our full model to a counterfactual framework. This counterfactual is identical to our full model with one exception: we let firms assume a higher customer attrition rate, $\underline{\zeta} > \zeta$, which leads to counterfactual future customer base levels, $\underline{b}'$.

This approach, therefore, mimics our theoretical results in Section 2. Unlike the marginal changes in customer attrition we consider in our theory, however, in what follows we set $\underline{\zeta}$ such that average customer base growth among continuing businesses in the full model is zero:[20]

$$\frac{1}{\Omega_C} \int_{j \in \mathbb{C}} (1 - \underline{\zeta})(1 + \gamma a_j + (1 - \gamma_j) s_j) \, \mathrm{d}j - 1 = 0,$$

where $\mathbb{C}$ (Ω_C) is the set (number) of continuing incumbent businesses.

To understand how changes in customer accumulation affect firm-level and aggregate outcomes, recall that firms' optimal choices are forward-looking (with the exception of those pertaining to production labor). Therefore, whenever $\underline{b}' \neq b'$, firms will endogenously make different decisions compared to the full model. By comparing (in partial equilibrium) results from our full model to outcomes in the counterfactual, we are able to isolate the contribution of customer accumulation ($\underline{b}' - b'$).

In doing so, we consider two different cases: (i) a "direct effect" and (ii) a "feedback loop effect." To quantify the "direct effect," we compute *only* firms' counterfactual optimal R&D decisions, holding all other choices and values unchanged at their levels in the full model. This amounts to quantifying how R&D choices in the full model depend on different levels of customer base levels (b' vs b'). Therefore, by construction, these counterfactual R&D choices – denoted by $\underline{x}_j$ – ignore the endogenous interaction between innovation and customer acquisition.

In contrast, the "feedback loop effect" aims to capture the full interaction between innovation and customer accumulation decisions. In particular, if firms choose different R&D investment in the counterfactual, they will also optimally alter their customer accumulation decisions. These, in turn, alter the incentives to conduct R&D and so on. To quantify this effect in its entirety, we resolve *all* firm-level decisions in the counterfactual. The resulting counterfactual R&D choices which account for the feedback loop effect are

[20] Note that in equilibrium, the aggregate customer base is stationary. Indeed, average customer base growth across all firms is zero due to firm exit. Note further that other counterfactual exercises are possible. For instance, if we allow counterfactual customer retention to be firm-specific, we can set $\underline{\zeta}_i$ such that customer base growth is zero for individual firms. Experimenting with such counterfactuals typically resulted in even stronger effects of customer accumulation on firm-level R&D and aggregate growth. Therefore, we view our preferred counterfactual as relatively conservative.

Table 4: Impact of customer acquisition on firm-level innovation rates (in %)

		Contrib. of customer acquisition			
		Direct effect		Feedback loop	
	Baseline	$\underline{x}$	$\frac{x-\underline{x}}{x} \times 100$	$\underline{\underline{x}}$	$\frac{x-\underline{\underline{x}}}{x} \times 100$
All firms	26.9	25.4	5.5	15.1	43.9
Young firms	33.7	31.7	6.0	13.7	59.4

Note: The first column shows "baseline" average 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 direct effect contributions of customer acquisition to innovation rates in levels ($\underline{x}$) and as contributions to the baseline ($(x-\underline{x})/x$). The fourth and fifth columns do the same, but for the feedback loop effect, $\underline{\underline{x}}$ and $(x-\underline{\underline{x}})/x$, respectively. All values are in percent.

denoted by $\underline{\underline{x}}_j$.

Customer accumulation and firm-level R&D: Results. First, let us highlight that firm-level market size effects described in theory in Proposition 2 are present also in the full model. For instance, the correlation between firms' innovation rates, x_j, and their marketing investment, a_j, is 0.73. This tight relationship is further highlighted in Table 4 which quantifies the direct and feedback loop effects of customer base accumulation on R&D investment.

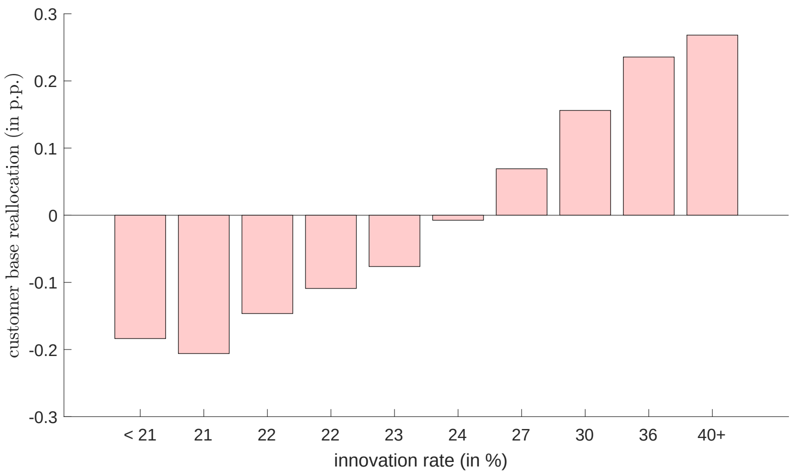
The first column of Table 4 shows average innovation rates among all and young businesses (not older than five years). The latter exhibit higher innovation rates on average, primarily owing to stronger customer base growth. The second and third columns document the direct effect of customer base accumulation – in levels (second column) and as contributions to baseline innovation rates (third column). As predicted by our theory, when businesses (counterfactually) expect customer attrition to rise, they scale back on R&D investment. Quantitatively, the direct effect accounts for "only" about 6% of average innovation rates.

The last two columns quantify the impact of customer accumulation when allowing for the feedback loop between R&D and customer base investment. As can be seen from the table, the feedback loop effect is quantitatively strong. In particular, customer acquisition accounts for over 40% of innovation rates among all firms and almost 60% when focusing on young firms only.

We further highlight this key novel channel with its visualization in Figure 2. Specifically, the bars show the difference in customer base levels in the full model relative to those implied by firms' counterfactual decisions. The figure plots this difference along the deciles of firm-level innovation rates in the full model.

As can be seen from the figure, relative to the counterfactual, customer base levels (market shares) are larger in the full model for businesses expecting fast productivity growth – often referred to as "gazelles". Our framework, therefore, suggests that the documented importance of gazelles for aggregate outcomes (see e.g. Haltiwanger et al.,

Figure 2: Feedback loop between productivity growth and customer acquisition



Note: The bars show the percentage point change between the density of firms' customer base levels, b_j, in the baseline model and in the counterfactual, $\underline{\underline{b}}_j^f$. The change is expressed as a function of firm-level innovation rate, x. Bars correspond to the deciles of the stationary distribution of firm-level innovation rates. Positive values indicate that a given decile of x accounts for a relatively higher mass of customer base in baseline relative to the counterfactual. In the baseline economy, there is more reallocation of customers towards innovative firms.

2016) may partly be the result of their customer accumulation, rather than exclusively being due to their superior productivity growth.[21]

4.2 Customer Accumulation and Aggregate Growth

Having explained the impact customer acquisition has on firms' individual R&D investment, we now turn to quantifying its impact for aggregate growth. Towards this end, the following paragraphs first introduce a decomposition of aggregate growth.

Customer accumulation and aggregate growth: A decomposition. There are two distinct channels through which changes in customer accumulation impact aggregate growth. First, because of firm-level market size effects, businesses conduct less R&D if they expect weaker customer base growth. We quantify this by computing aggregate growth implied by firms' counterfactual R&D decisions, $\underline{\underline{x}}$.

Second, in our full model customer accumulation also impacts aggregate growth through the (re-)allocation of customers across heterogeneous firms (see Figure 2). To understand this channel further, recall that the aggregate productivity index, Q, is a

[21]These results also highlight the importance of considering *dynamic* customer accumulation. It is precisely the forward-looking nature of customer base and R&D investment which is at the heart of the feedback loop with firms' innovation decisions.

"customer-weighted" average of firms' individual productivity levels:[22]

$$Q = \left(\int_{j \in \Omega} \hat{b}_j q_j^{\eta-1} \, \mathrm{d}j \right)^{\frac{1}{\eta-1}} . \tag{33}$$

The following proposition (the proof of which is left to the Appendix) describes how aggregate growth, $G = dQ/Q$, can be decomposed into contributions of productivity growth and customer (re-)allocation.

Proposition **7.** *Using a first-order Taylor expansion of Q, aggregate growth can be decomposed as*

$$G \approx \underbrace{\sum_{\mathrm{j} \in \mathbb{C}} \bar{b} \hat{q}_j^{\eta-1} \frac{\Delta q_j}{q_j}}_{\textit{productivity growth } (G_q)} + \underbrace{\sum_{\mathrm{j} \in \mathbb{C}} (\hat{b}_j - \bar{b}) \hat{q}_j^{\eta-1} \frac{\Delta q_j}{q_j}}_{\textit{customer base allocation } (G_b)} + \underbrace{\sum_{\mathrm{j} \in \mathbb{C}} \Delta \hat{b}_j \hat{q}_j^{\eta-1}}_{\textit{customer base re-allocation } (G_{\Delta b})} , \tag{34}$$

where $\bar{b} = \frac{1}{\Omega_C} \sum_{j \in \mathbb{C}} \hat{b}_j$ is the average customer base share.

Proposition 7 shows that aggregate growth can be (up to first order) decomposed into the contributions of three terms: (i) productivity growth (G_q), (ii) customer base allocation (G_b), and (iii) customer base re-allocation across firms with different productivity growth rates ($G_{\Delta b}$).[23]

Customer accumulation and aggregate growth: Firm-level R&D. First, as in standard endogenous growth models, also in our framework firm-level productivity improvements drive aggregate growth. This is captured by the "productivity growth" term, G_q. However, in our model, firms' innovation rates are partly driven by firm-level market size effects. To quantify the strength of this channel for aggregate growth, we split the overall contribution of firms' innovation on aggregate growth into two components.

The first component is based on firms' counterfactual R&D investment choices discussed above, $\underline{\underline{x}}$. We denote this contribution by $\underline{\underline{G}}_q$. The second component then captures the part of firms' innovation rates which are driven by firm-level market size effects, $x - \underline{\underline{x}}$. In our framework, this is simply given by the difference in the overall impact of firms'

[22]With variable markups, the relevant weight becomes $\hat{b}_j = b_j(\mu_j/\overline{\mu})^{1-\eta}/B$.

[23]The decomposition in Proposition 7 is approximate with a residual capturing the impact of net entry and non-linear approximation errors. Note that conventional decompositions, such as that from Foster et al. (2001) which decompose aggregate (labor productivity) growth into the "within", "between", "cross" and "net entry" contributions are not informative of the relative contributions of customer base levels and productivity. This is because observed variables such as output and employment are themselves combinations of the two firm-level state variables.

R&D on aggregate growth and the counterfactual:

$$G_q = \underbrace{\sum_{\mathrm{J}\in\mathbb{C}} \bar{b}\hat{q}_j^{\eta-1}\frac{\Delta q_j(\underline{\underline{x}})}{q_j}}_{\substack{\text{Productivity growth contribution under} \\ \textit{0 average customer base growth } (\underline{\underline{G}}_q)}} + \underbrace{\sum_{\mathrm{J}\in\mathbb{C}} \bar{b}\hat{q}_j^{\eta-1}\left(\frac{\Delta q_j(x)}{q_j} - \frac{\Delta q_j(\underline{\underline{x}})}{q_j}\right)}_{\substack{\text{Productivity growth contribution due to} \\ \text{firm-level market size effects } (G_q - \underline{\underline{G}}_q)}},$$

where we have made explicit the dependence of firms productivity growth on their innovation rates, $\Delta q_j(x)$.

Customer accumulation and aggregate growth: (Re-)allocation of customers. Next, to the extent that firms with higher productivity growth also enjoy relatively larger customer base levels, this will increase aggregate growth. This effect is captured by the "allocation" term, G_b.

In addition, customers can also reallocate across firms. In particular, if customers flow towards firms with high productivity levels, this too will raise aggregate growth. Such an effect would be soaked up by the "re-allocation" term, $G_{\Delta b}$.

Customer accumulation and aggregate growth: Summary. Putting the above together, customer accumulation affects aggregate growth through two distinct channels: (i) firm-level market size effects and (ii) the (re-)allocation of customers across firms. Therefore, the overall contribution of customer accumulation to aggregate growth is given by the summation of the following three terms:

$$G_{\text{cust.}} = \underbrace{(G_q - \underline{\underline{G}}_q)}_{\text{firm-level market size effects}} + \underbrace{G_b + G_{\Delta b}}_{\text{customer (re-)allocation}}.$$

The first term, $G_q - \underline{\underline{G}}_q$, captures the impact of firm-level market size effects. This is precisely the mechanism identified in Proposition 3. The second and third terms are specific to our full model and capture the impact of customer (re-)allocation on aggregate growth.

Customer accumulation and aggregate growth: Results. Let us begin by discussing the effect of firms' innovation activity on aggregate growth. Overall, firms' R&D decisions (G_q) account for about 54% of aggregate growth. However, part of this is in fact driven by customer accumulation due to the presence of firm-level market size effects. In particular, the third column of Table 5 ($G_q - \underline{\underline{G}}_q$) shows that about 23% of aggregate growth can be accounted for by market size effects boosting firms' incentives to innovate.

Next, as can be seen from column four (G_b) of Table 5, the allocation of customers across firms with different productivity growth rates *dampens* aggregate growth. This is because firms displaying high productivity growth are typically younger. These firms

Table 5: Impact of customer acquisition on aggregate growth

	G	G_q	Contribution of customer acquisition ($G_{cust.}$): $G_q - \underline{\underline{G}}_q$	G_b	$G_{\Delta b}$	Total
Level	1.43	0.78	0.32	−0.31	0.61	0.63
% of baseline	100	54.4	22.7	−21.4	43.0	44.3

Note: The tables shows the decomposition of aggregate growth in (34). G is aggregate growth, and the remaining columns show the respective contributions to aggregate growth: G_q is the contribution of firms' productivity growth, $G_q - \underline{\underline{G}}_q$ is the difference between actual and counterfactual productivity growth, where $\underline{\underline{G}}_q$ is counterfactual productivity growth (based on $\underline{\underline{x}}$), G_b is the contribution of customer base allocation and $G_{\Delta b}$ is the contribution of customer base re-allocation. The first row shows contributions in levels, the second row as percent shares of baseline growth, G. All values are in percent.

had less time to accumulate customers and, therefore, have relatively less weight in the customer distribution. In contrast, larger businesses with larger customer base shares typically expand their productivity at a lower pace. Both these effects reduce aggregate growth compared to a model where firms have common customer base levels (and therefore common weights in Q).

In contrast, the re-allocation of customers towards firms with high productivity levels boosts aggregate growth. Specifically, 43% of aggregate growth can be accounted for by this channel. This effect is once again a result of the feedback loop between R&D and customer accumulation. Highly productive firms display, on average, larger firm values. This, in turn, provides greater incentives to accumulate customers and shifts the distribution of market shares in their favor.

Finally, let us summarize the overall impact of customer accumulation on aggregate growth. This is a combination of (i) firm-level market size effects and (ii) the (re-)allocation of customers across firms. Put together, these effects account for 44% of aggregate growth – see the last column of Table (5). Therefore, while aggregate growth has typically been considered a purely supply-side phenomenon, our framework puts forward a novel – and quantitatively important – notion. In particular, aggregate growth is to a large extent driven by customer acquisition of individual firms.

4.3 Implications for Growth Policies

The previous two subsections showed that changes in customer accumulation have important implications for firm-level and aggregate innovation. As such, they quantified Propositions 2 and 3 from our theoretical framework. We now turn to quantitatively investigating the extent to which ignoring customer accumulation altogether affects the policy conclusions from endogenous growth models.

In particular, we consider two popular policies – subsidies to firms' R&D investment and to their costs of operation. These should not be viewed as instruments aimed at bringing the economy closer to a first best allocation. Instead, we use them as examples to quantify the impact of customer acquisition on policy conclusions.

A "restricted" economy. In order to quantify how the presence of firm-level customer acquisition changes the responsiveness of growth to policies, we compare our full model to a "restricted" economy. The latter is identical to our full model with one crucial exception – it completely abstracts from customer accumulation. Specifically, we consider a version of our full model in which we set $b_j = b'_j = 1$ and $a_j = s_j = 0$ for all j. Importantly, however, we calibrate this restricted economy to match the same targets as described in Section 3.2 (except for moments related to customer accumulation).[24] Comparing outcomes from our full model and the restricted economy then allows us to show how the presence of customer accumulation changes quantitative predictions (in general equilibrium) regarding the macroeconomic impact of policies. We relegate a more detailed discussion of the "restricted" economy to Appendix A.4.

Policies. We introduce two policy instruments into both the full and restricted models: (i) subsidies to costs of operation, τ_κ, and (ii) R&D subsidies, τ_x. We assume that both policies are financed with lump-sum taxes on households and provided to all businesses at the same rate. Formally, firm values become:

$$v(q_j, b_j) = \max_{p_j, n_j^c, n_j^x, n_j^a, \kappa_j^*} \left\{ \begin{array}{c} p_j c_j - W(n_j^c + n_j^x(1-\tau_x) + n_j^a) + \\ \frac{1-\delta}{1+R'} \left[x_j v^c(q_j(1+\lambda), b'_j) + (1-x_j) v^c(q_j, b'_j) \right] \end{array} \right\}, \tag{35}$$

where $v_j^c = F(\kappa_j^*) \left[v_j - \widetilde{\kappa}_j (1-\tau_\kappa) \right]$. In our analysis, we consider two cases. One with positive R&D subsidies (but zero operational cost subsidies) and one with positive subsidies to costs of operation (but zero R&D subsidies). In doing so, we ensure that in both the baseline and restricted economies subsidies account for the same fraction of output.

Finally, we restrict our analysis to the long-run, abstracting from transitional dynamics. We do, however, consider general equilibrium effects. Therefore, following the introduction of the respective policies, we recompute equilibrium firm entry, aggregate spending and aggregate growth for each of the economies.

Subsidies to operational costs: Results. Let us first consider the impact of subsidizing firms' operations, $\tau_\kappa > 0$, in the restricted economy. Lower operational costs make it easier for businesses to survive. However, it is firms with relatively low productivity

[24] The restricted framework matches the same targets as the full model, with the exception of the aggregate marketing expenses to sales and the regression coefficients in marketing-to-sales ratio on size and firm revenue growth on marketing expenses growth. We provide more details on the restricted model parameterization and fit in Appendix A.5.

Table 6: Impact of policies on aggregate growth (in p.p.)

	$\Delta g \times 100$	
	Restricted model	Full model
Subsidies to operational costs, $\tau_\kappa > 0$	-0.01	-0.03
Subsidies to R&D expenditures, $\tau_x > 0$	0.03	0.06

Note: The table shows general equilibrium effects of subsidies to firms' operational costs (top row) and to R&D costs (bottom row). The table reports results for the full model and the restricted economy which abstracts from customer accumulation. The table depicts effects on aggregate growth in percentage point deviations from the respective values without policies. In all cases, the level of subsidies is chosen such that the policy cost amounts to 0.12% of GDP corresponding to the level of direct funding of R&D in the US according to the 2023 vintage of OECD BERD dataset.

which are now able to survive under the policy. Absent subsidies, these firms would be replaced by more productive entrants – a feature of growth by creative destruction. Therefore, when we help unproductive firms survive, innovation and aggregate growth decline.

The effects in the full model are qualitatively the same but quantitatively stronger. The underlying reason is the process of firm selection, which is very different in the two economies. Figure 3 shows the distribution of incumbent firms in the restricted (Panel A) and full (Panel B) models, respectively. In both models, firms decide whether or not to stay in the economy based on their firm values, see (23).

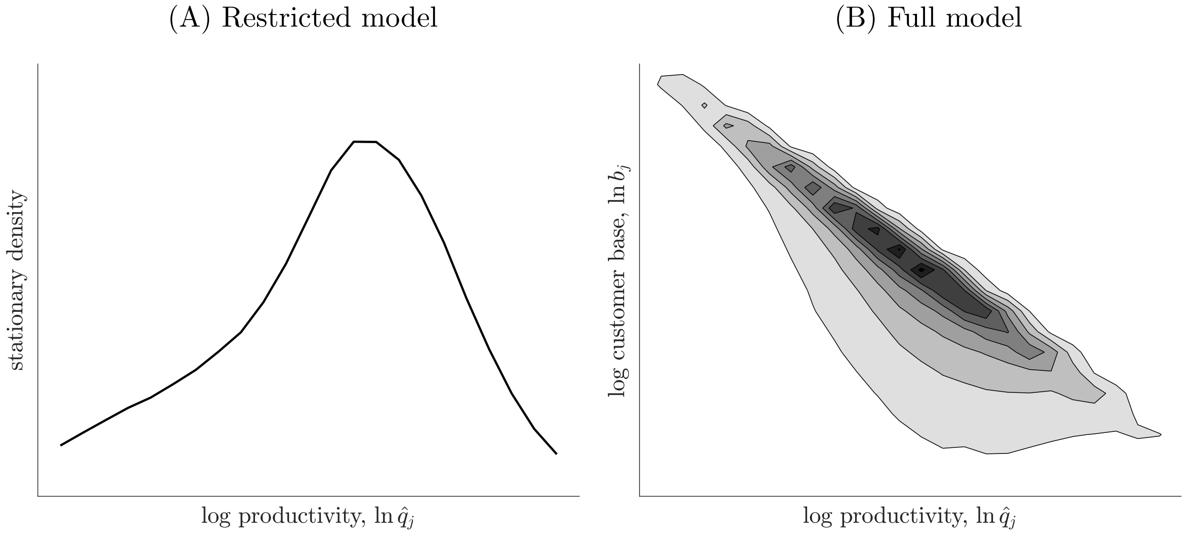
However, while in the restricted economy, firm values are determined solely by firm-specific productivity, in the full model they are driven by a combination of both productivity and customer base levels.[25] Panel B of Figure 3 shows a clear negative relationship between the two among surviving firms. Intuitively, if a firm enjoys high levels of customers, it can survive despite having relatively low productivity and vice versa. Therefore, the overall effect of subsidizing operational costs in the full model depends on how exactly it shifts the joint distribution of firm-level customer base and productivity levels among surviving businesses.

Subsidies to R&D investment: Results. As a second policy we consider subsidies to R&D costs, τ_x. Let us again begin by focusing on the effects in the restricted model. R&D subsidies make investing into productivity cheaper and, therefore, they raise firm-level innovation rates. This, in turn, directly translates into faster economic growth.

Once again, while the results are qualitatively the same in both economies, they differ quantitatively. The full model is twice as sensitive to the policy compared to the restricted economy, consistent with Proposition 4. The reason is the virtuous feedback loop between customer accumulation and innovation. Motivated by cheaper R&D costs brought about by the subsidy, firms invest more into R&D. This, in turn, raises the marginal benefit of

[25] Recall that firms face *stochastic* operational costs. Therefore, the distribution of incumbent firms does not feature a strict threshold below which all firms would shut down.

Figure 3: Distribution of incumbent firms



Note: The figure shows the distribution of incumbent firms in the restricted (Panel A) and full (Panel B) models. In the baseline model, firms select on a combination of demand stocks and productivity with darker shades indicating more densely populated areas. In the restricted model, firms select purely on productivity. There, the vertical axis marks the density in the stationary distribution. For visual clarity, in both figures we truncate the distribution while preserving 99% of total firm mass.

investing into customer accumulation. As a result, highly productive firms become even larger with market shares reallocating in their favor.

The two examples discussed in this subsection highlight that ignoring customer-side factors – as is common in existing endogenous growth models – not only skews our understanding of the drivers of growth, but also distorts our view of the effectiveness of pro-growth policies. While we have focused on two specific policies, the underlying reasons for the associated quantitative results are more general. Specifically, the presence of firm-level customer acquisition changes the nature of firm selection and the distribution of investment costs and benefits. Therefore, any policies or shocks which affect business dynamism will propagate differently through an economy which ignores customer acquisition compared to our full model which does not.

5 Empirical Support for Model Predictions

As a final step in our analysis, we provide empirical support for our model's key predictions. We do so in three steps. First, we document a connection between marketing investment and markups – the two drivers of customer accumulation in our framework. Next, we show that investment into customer acquisition is closely linked to R&D – both statically and dynamically. Finally, we propose a model-consistent way of measuring the degree of customer accumulation frictions across different industries and investigate how such frictions influence firms' innovation expenditures.

Data. As in our parameterization Section 3.2, we use firm-level data from Compustat and Capital IQ spanning 1997 to 2019. For detailed information on sample selection, estimation methodology, additional empirical findings, and robustness tests, please refer to the Appendix D.

5.1 Marketing and Markups

We begin by investigating the empirical connection between the key variables linked to customer accumulation: marketing expenditures and markups. In particular, in our framework firms accumulate customers by investing into marketing and by charging lower markups.

Indeed, both investments share common marginal benefits – an increase in firm value driven by an expansion of the customer base. In fact, combining the respective optimality conditions (13) and (14) results in an "arbitrage" condition between marketing and (inverse) markups:

$$\frac{1}{\mu} = \frac{1}{\overline{\mu}} + \frac{1-\gamma}{\gamma}\frac{\psi_a \overline{a}}{\overline{\mu}}\frac{\hat{W}}{\hat{C}} m^{\sigma_s+\sigma_a} a^{\psi_a - 1}. \tag{36}$$

The above states that, to the extent that firms use both channels of customer base accumulation (i.e. $\gamma \in (0,1)$), selling and marketing and (inverse) markups co-move positively.

To test this prediction empirically, we estimate the following reduced-form relationship:

$$\ln\left(1 + \frac{\text{S\&M}_{j,t}}{\text{sales}_{j,t}}\right) = \beta \ln\left(\frac{1}{\mu_{j,t}}\right) + \Theta \mathbb{X}_{j,t} + \epsilon_{j,t}, \tag{37}$$

where $\mathbb{X}_{j,t}$ controls for firm fixed effects, sector-year fixed effects, firm size and age. The results are in the first two columns of Table 7. They suggest that – as the model predicts – firms' marketing expenditures and the inverse of markups are positively related in the data.

5.2 Selling and marketing, and R&D

Next, we turn to the connection between selling and marketing expenses – the dominant driver of customer accumulation in the model – and R&D expenditures. First, we analyze this relationship "statically" in the cross-section of firms. Next, we focus on the "dynamic" aspect of this relationship by analyzing firms' life-cycle dynamics.

S&M and R&D: Static relationship. To formally test the static relationship between marketing and R&D investment, we estimate the following regression:

$$\ln\left(1 + \frac{\text{S\&M}_{j,t}}{\text{sales}_{j,t}}\right) = \beta \ln\left(1 + \frac{\text{R\&D}_{j,t}}{\text{sales}_{j,t}}\right) + \Theta \mathbb{X}_{j,t} + \epsilon_{j,t}, \tag{38}$$

Table 7: Marketing, markups and R&D in the data

Dep. var.: $\ln\left(1+\frac{\text{S\&M}_{j,t}}{\text{sales}_{j,t}}\right)$	(I)	(II)	(IV)	(V)
$\ln\left(\frac{1}{\mu_{j,t}}\right)$	0.03***	0.02***		
	(0.003)	(0.004)		
$\ln\left(1+\frac{\text{R\&D}_{j,t}}{\text{sales}_{j,t}}\right)$			0.15***	0.14***
			(0.01)	(0.01)
Firm fixed effects	✓	✓	✓	✓
Industry-year fixed effects	✓	✓	✓	✓
Size		✓		✓
Age		✓		✓
Observations	82,002	62,383	90,134	67,628
R^2	0.72	0.77	0.73	0.78
Within R^2	0.009	0.02	0.06	0.07

Note: The table shows estimates of regression (37) using Compustat firms between the years 1997 and 2019. S&M stands for sales and marketing expenses, markups, μ, are estimated using the methodology in De Loecker et al. (2020). Firm size is given by log(assets) and age is measured as the log of the number of years since establishing date reported in Capital IQ or the Loughran-Ritter database (Loughran and Ritter, 2004). Standard errors (in brackets) are clustered at the firm level.

where $\mathbb{X}_{j,t}$ again controls for firm fixed effects, sector-year fixed effects, firm size and age. The results are presented in the last two columns of Table 7. As the model predicts, also in the data marketing expenditures and R&D investment are positively related contemporaneously in the cross-section of firms.[26]

S&M and R&D: Dynamic relationship. A crucial aspect of our theory is the *dynamic* relationship between customer accumulation and R&D investment. This happens because expected customer base growth impacts the incentives to conduct R&D – the firm-level market size effect. In the full model, this connection gives rise to specific life-cycle patterns of marketing and R&D investment which we now analyze empirically.

Towards this end, we estimate industry-specific life-cycle profiles of marketing and R&D intensity using the following regression:

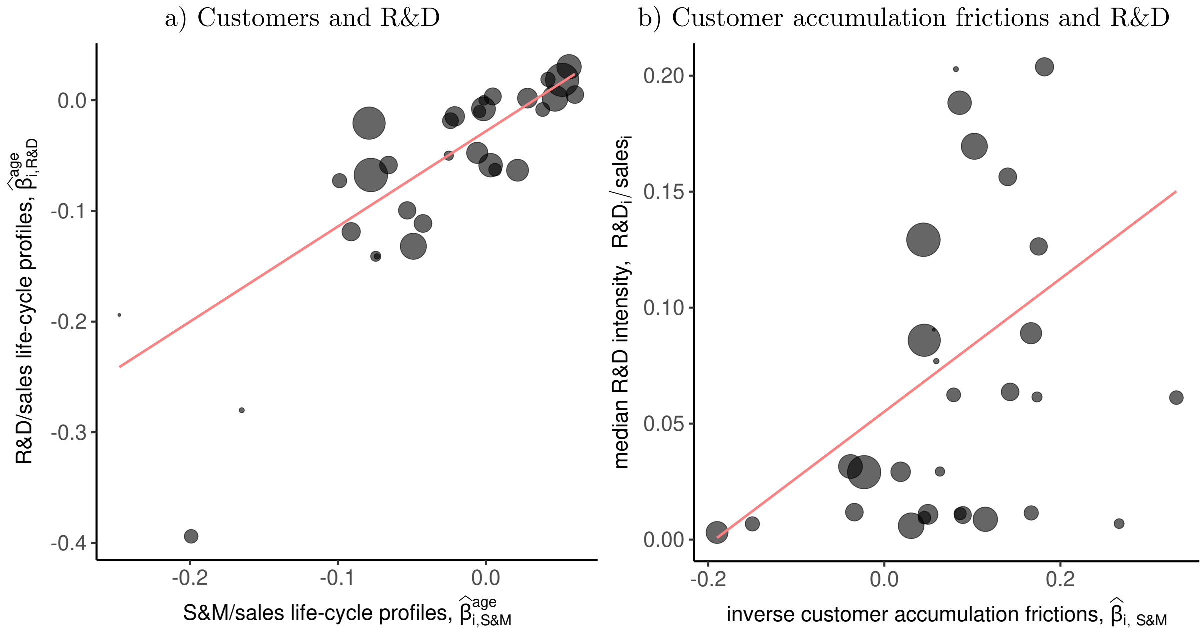
$$\ln\left(1+\frac{y_{j,t}}{\text{sales}j,t}\right)=\sum_i \mathbf{1}_{i(j)=i}\beta_{i,y}^{age}\ln\text{age}_{j,t}+\Theta_y\mathbb{X}_{j,t}+\varepsilon_{j,t}, \tag{39}$$

where y is the outcome variable of interest (either S&M or R&D), $\mathbb{X}_{j,t}$ includes time fixed effects and firm size, i indicates 3-digit NAICS industries and where $\mathbb{1}_{i(j)=i}$ is an indicator function equal to one if firm j belongs to industry i and zero otherwise.[27]

[26] In Table 7, we use assets to measure size because sales is in the denominator of the left-hand-side variable.

[27] Given our interest in a cross-sectoral comparison, we do not include industry fixed effects.

Figure 4: Marketing and R&D



Note: Panel a) presents the correlation between $\widehat{\beta}^{age}_{i,S\&M}$ and $\widehat{\beta}^{age}_{i,R\&D}$ estimated in sector-specific regressions (39). Panel b) presents the correlation between $\widehat{\beta}_{i,a}$ estimated in sector-specific regressions (29) and median R&D intensity $\frac{\text{XRD}}{\text{sales}}$ in the industry. The inverse of the estimated coefficient $\widehat{\beta}_{i,S\&M}$ is proportional to the severity of customer accumulation frictions. The size of each marker corresponds to revenue share of the particular industry. In both panels, red solid line corresponds to the fit of linear regression model. Subscript i corresponds to a 3-digit NAICS industry. We restrict attention only to industries with at least 40 firm-year observations. See Appendix D for more details on data construction and robustness of these results.

Before moving on, note that most life-cycle slopes in the cross-section of industries are negative. This is consistent with our model in which firms reduce their marketing and R&D investment (relative to sales) as they age and grow larger. Indeed, recall that our full model is parameterized to match the observed relationship between R&D (marketing) intensity and firm size.

Panel a) of Figure 4 plots the estimated life-cycle profile slopes ($\widehat{\beta}^{age}_{i,S\&M}$ and $\widehat{\beta}^{age}_{i,R\&D}$) against each other in the cross-section of industries. The figure shows a tight positive (and statistically significant) relationship between the two, confirming the model-predicted dynamic connection. In other words, markets in which incentives to invest into marketing drop quickly with firm age are also markets where businesses spend relatively less on R&D as they mature.

5.3 Customer Accumulation Frictions and Innovation

Customer acquisition frictions are likely to be more severe in some industries than in others. Through the lens of our model, sectors with particularly severe customer accumulation frictions should also be sectors where firms have fewer incentives to invest into

R&D. We now test this model prediction.

A model-consistent measure of customer accumulation frictions. First, we propose a model-consistent measure of customer accumulation frictions which we estimate at the industry-level. Recall that in our framework, the ease with which firms can accumulate new customers depends on three key parameters: (i) the elasticity of marketing costs, ψ_a, (ii) the customer depreciation rate, ζ and (iii) the level of marketing costs, $\overline{a}$. In particular, higher values of any of these parameters make customer base expansions more expensive, decreasing the speed of customer accumulation.

To measure customer accumulation frictions in the cross-section of industries, we lean on the approach used in our parameterization Section 3.2. In particular, we estimate industry-specific coefficients $\beta_{i,a}$ from regression (29) at the 3-digit industry level. These elasticities represent model-consistent measures of customer acquisition frictions, because they are proportional to $\frac{(1-\zeta)^2}{\overline{a}\psi_a}$ – see Proposition 6. Specifically, higher values of $\beta_{i,a}$ correspond to weaker customer acquisition frictions.

Customer accumulation frictions and R&D investment. Panel b) of Figure 4 shows that our model predictions hold up in the data. Specifically, in the cross-section of industries, there is a clear positive (and statistically significant) relationship between the ease with which firms accumulate customers, $\widehat{\beta}_{i,a}$, and median R&D intensity. Through the lens of our model, this is because lower frictions associated with customer accumulation (higher $\widehat{\beta}_{i,a}$) provide more incentives to invest into R&D through stronger firm-level market size effects.

6 Concluding Remarks

Firm life-cycle dynamics are a central force in endogenous growth models. In existing growth models, such dynamics are typically determined purely by productivity differences between firms. However, recent and expanding evidence suggests that demand-side factors, in particular active customer accumulation, are crucial for driving firm-level outcomes. In this paper, we build on this evidence and propose a novel endogenous growth model featuring customer acquisition. Changes in customer acquisition increase firms' incentives to conduct R&D and endogenously reallocates market shares and resources towards businesses experiencing fast productivity growth. Combined, these effects serve as a quantitatively important driver of aggregate growth. Moreover, ignoring customer acquisition changes the quantitative response of the macroeconomy to policy interventions.

We believe that our framework opens the door to several intriguing questions which we have left for future research. For instance, what new tools could policy markers use to spur aggregate growth? While demand-side growth policies have been debated in policy circles (see e.g. European Commission, 2003), they have been largely absent from systematic academic analysis within state-of-the-art models of endogenous growth.

Our results also suggest that economies with different firm life-cycle growth profiles and consumption spending patterns may have very different aggregate growth dynamics. What does this imply for the efficacy of pro-growth policies in developing economies, in which markets are potentially more segmented and customer accumulation more frictional compared to developed countries? Or can changes in income inequality affect long-run growth because of associated movements in consumption expenditure allocations?

Finally, our framework studies cost-saving (process) innovations. In the presence of customer accumulation, it would be particularly interesting to also study "product" innovation and understand how these two types of technological progress interact with each other and with firms' life-cycle incentives to attract customers.

References

Abaluck, Jason and Abi Adams-Prassl (2021) "What do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses," *Quarterly Journal of Economics*, 136 (3), 1611–1663.

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William Kerr (2018) "Innovation, Reallocation and Growth," *American Economic Review*, 108 (11), 3450–3491.

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.

Akcigit, Ufuk and William Kerr (2018) "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 126 (4), 1374–1443.

Argente, David, Doireann Fitzgerald, Sara Moreira, and Anthony Priolo (2025) "How Do Entrants Build Market Share? The Role of Demand Frictions," *American Economic Review: Insights*, 7 (2), 160–76, 10.1257/aeri.20230289.

Arkolakis, Costas (2016) "A Unified Theory of Firm Selection and Growth," *Quarterly Journal of Economics*, 131 (1), 89–155.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen (2020) "The Fall of the Labor Share and the Rise of Superstar Firms*," *The Quarterly Journal of Economics*, 135 (2), 645–709, 10.1093/qje/qjaa004.

Baker, Scott, Brian Baugh, and Marco Sammon (2023) "Customer Churn and Intangible Capital," *Journal of Policital Economy: Macroeconomics*, 1 (3), 447–505.

Belo, Frederico, Xiaoji Lin, and Maria Ana Vitorino (2014) "Brand Capital and Firm Value," *Review of Economic Dynamics*, 17 (1), 150–169.

Bernard, Andrew, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes (2022) "The Origins of Firm Heterogeneity: A Production Network Approach," *Journal of Political Economy*, 130, 1765–1804.

Bloom, Nicholas, Rachel Griffith, and John Van Reenen (2002) "Do R&D Tax Credits Work?" *Journal of Public Economics*, 85 (1), 1–31.

Broda, Christian and David Weinstein (2006) "Globalization and the Gains form Variety," *Quarterly Journal of Economics*, 121 (2), 541–585.

Bronnenber, Bart, Jean-Pierre Dube, and Chad Syverson (2022) "Marketing Investment and Intangible Brand Capital," *Journal of Economic Perspectives*, 36 (3), 53–74.

Campbell, John Y. and Robert J. Shiller (1988) "Stock Prices, Earnings, and Expected Dividends," *The Journal of Finance*, 43 (3), 661–676.

Cavenaile, Laurent, Murat Alp Celik, Pau Roldan-Blanco, and Xu Tian (2021) "Style over Substance? Advertising, Innovation, and Endogenous Market Structure," mimeo.

Crouzet, Nicolas, Janice Eberly, Andrea Eisfeldt, and Dimitris Papanikolaou (2022) "The Economics of Intangible Capital," *Journal of Economic Perspectives*, 36 (3), 29–52.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger (2020) "The Rise of Market Power and the Macroeconomic Implications*," *The Quarterly Journal of Economics*, 135 (2), 561–644, 10.1093/qje/qjz041.

Dinlersoz, Emin and Mehmet Yorukoglu (2012) "Inofrmation and Industry Dynamics," *American Economic Review*, 102 (2), 884–913.

Drozd, Lukasz A and Jaromir B Nosal (2012) "Understanding international prices: Customers as capital," *American Economic Review*, 102 (1), 364–95.

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

Eslava, Marcela and John Haltiwanger (2021) "The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand and Wedges," NBER Working Paper, 27184.

Fernández-Villaverde, Jesús, Federico Mandelman, Yang Yu, and Francesco Zanetti (2021) "The "Matthew effect" and market concentration: Search complementarities and monopsony power," *Journal of Monetary Economics*, 121, 1177–1216.

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

Foster, Lucia, John C Haltiwanger, and Cornell John Krizan (2001) "Aggregate productivity growth: lessons from microeconomic evidence," in *New developments in productivity analysis*, 303–372: University of Chicago Press.

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, Francois (2014) "Financial Distress and Endogenous Uncertainty," mimeo, Federal Reserve Bank of Chicago.

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

Hall, Bronwyn and Rosemarie Ziedonis (2001) "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-1995," *RAND Journal of Economics*, 101–128.

Haltiwanger, John (2012) "Job Creation and Firm Dynamics in the U.S.," *Innovation Policy and the Economy*, 17–38.

Haltiwanger, John, Ron Jarmin, Robert Kulick, and Javier Miranda (2016) "High Growth Young Firms: Contribution to Job, Output and Productivity Growth," mimeo.

Haltiwanger, John, Ron Jarmin, and Javier Miranda (2013) "Who Creates Jobs? Small Versus Large Versus Young," *Review of Economics and Statistics*, 95 (2), 347–361.

He, Bianca, Lauren Mostrom, and Amir Sufi (2025) "Investing in Customer Capital," NBER Working Paper 33171.

Hopenhayn, Hugo A. and Richard Rogerson (1993) "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101 (5), 915–938, 10.1086/261909.

Hottman, Collin, Stephen Redding, and David Weinstein (2016) "Quantifying the Sources of Establishment Heterogeneity," *Quarterly Journal of Economics*, 131 (3), 1291–1364.

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

Jones, Chad and Jihee Kim (2018) "A Schumpeterian Model of Top Income Inequality," *Journal of Political Economy*, 126 (5), 1785–1826.

Jovanovic, Boyan (1982) "Selection and the Evolution of Industry," *Econometrica*, 50 (3), 649–670.

Kaas, Leo and Bihemo Kimasa (2021) "Firm Dynamics with Frictional Product and Labor Markets," *International Economic Review*, 62 (3), 1281–1317.

Kaas, Leo and Phillip Kircher (2015) "Efficient Firm Dynamics in a Frictional Labor Market," *American Economic Review*, 105 (10), 3030–3060.

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

Larkin, Yelena (2013) "Brand Perception, Cash Flow Stability, and Financial Policy," *Journal of Financial Economics*, 110 (1), 232–253.

Lentz, Rasmus and Dale T Mortensen (2008) "An empirical model of growth through product innovation," *Econometrica*, 76 (6), 1317–1373, Publisher: Wiley Online Library.

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`.

Melitz, Marc (2003) "The Impact of International Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 6 (71), 1695–1725.

Nakamura, Emi and Jón Steinsson (2011) "Price setting in forward-looking customer markets," *Journal of Monetary Economics*, 58 (3), 220–233.

Ottonello, Pablo and Thomas Winberry (2020) "Financial Heterogeneity and the Investment Channel of Monetary Policy," *Econometrica*, 88 (6), 2473–2502.

Perla, Jesse (2019) "A Model of Product Awareness and Industry Life Cycles," mimeo.

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

Petrosky-Nadeau, Nicolas and Etienne Wasmer (2015) "Macroeconomic dynamics in a model of goods, labor, and credit market frictions," *Journal of Monetary Economics*, 97–113.

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

Rachel, Lukasz (2021) "Leisure-Enhancing Technological Change," mimeo.

Ravn, Morten, Stephanie Schmidt-Grohe, and Martin Uribe (2006) "Deep Habits," *Review of Economic Studies*, 73, 195–218.

de Ridder, Maarten, Basile Grassi, and Giovanni Morzenti (2024) "The Hitchhiker's Guide to Markup Estimation: Assessing Estimates from Financial Data," mimeo.

Rudanko, Leena (2022) "Price Setting with Customer Capital: Sales, Teasers and Rigidity."

Sedláček, Petr and Vincent Sterk (2017) "The Growth Potential of Startups over the Business Cycle," *American Economic Review*, 107 (10), 3182–3210.

Shi, Shouyong (2016) "Customer relationships and sales," *Journal of Economic Theory*, 166 (C), 483–516.

Spencer, Adam (2022) "Policyt Effects of International Taxation on Firm Dynamics and Capital Structure," *Review of Economic Studies*, 89 (4), 2149–2200.

Online Appendix

Customer Acquisition, Business Dynamism and Aggregate Growth

Marek Ignaszak and Petr Sedláček

Contents

A Model Details

This appendix provides additional details regarding the three model versions in the main text: (i) benchmark model of Section 2, (ii) full model of Section 4 and (iii) restricted model of Section 4.

A.1 Benchmark model: Equilibrium

The benchmark model is described in its entirety in the main text. Here we provide a formal definition of the associated equilibrium.

Benchmark model: Equilibrium definition. *A balanced growth path equilibrium consists of the following tuple in every period with $j \in \Omega$: c_j, n_j^c, q_j, b_j, v_j, p_j, x_j, a_j, C, R, W, Q, G such that (i) firm-level output, c_j, and production employment, n_j^c, are given by (4) and (5), (ii) firms' productivity, q_j, and customer base levels, b_j, evolve according to (7) and (8), (iii) firm values are given by (11), (iv) optimal innovation rates, x_j, marketing rates, a_j, and prices, p_j, satisfy (12) to (14), (v) aggregate consumption, C, is given by (1), (vi) the interest rate, R, and wage rate, W satisfy (3) and (15), and (vii) the aggregate productivity index, Q, and it's growth rate, G, are given by (16) and $G = Q'/Q - 1$.*

A.2 Full model: Details

In this appendix, we provide details pertaining to the full model used in quantitative analysis in Section 4.

Households. The representative household remains the same as in the benchmark model. Preferences are given $U = \sum_t \beta^t \log C_t$ where C_t is defined in (1). The household is subject to the budget constraint (2). The resulting household's Euler equation and demand for goods varieties j also remain unchanged and are, therefore, given by (3) and (4).

Firms. Firms choose prices, R&D expenditures, marketing investment, and decide whether to continue operating or exit. They do so by maximizing firm value, formally given by

$$v(q,b) = \max_{p,n^x,n^a} \left\{ bp^{1-\eta}C - W(bp^{-\eta}C/q + n^x + n^a) + \frac{1-\delta}{1+R'} \begin{bmatrix} xv^c(q(1+\lambda), b')+ \\ \\ (1-x)v^c(q,b') \end{bmatrix} \right\} \tag{A1}$$

$$
\begin{aligned}
\text{s.t.}\quad c =&bp^{-\eta}C, \quad c = qn^c, \quad n^x = \overline{x}m^{\sigma_x}x^{\psi_x}, \quad n^a = \overline{a}m^{\sigma_a}a^{\psi_a}, \quad s = m^{\sigma_s}\frac{pc}{PC}, \\
b' =&(1-\zeta)b(1+\gamma a+(1-\gamma)s), \quad x \in [0,1] \\
v^c =&F(\kappa^*)\left[v - W\widetilde{\kappa}\right].
\end{aligned}
$$

where $v^c(q,b)$ is the continuation value which depends on the survival probability $F(\kappa^*) = P(\kappa \leq \kappa^*)$ and firm value net of expected operational costs conditional on survival ($v - \hat{W}\widetilde{\kappa}$). In the former, F is a logistic distribution with mean μ_κ and scale σ_κ and where firms' optimal exit threshold is given by

$$
\kappa^* = \frac{v(q_j, b_j)}{W} \tag{A2}
$$

In the latter, expected operational costs conditional on survival are given by $\widetilde{\kappa} \equiv \mathbb{E}\left[\kappa \mid \kappa \leq \kappa^*\right]$.

The mass of surviving entrants Ω_e is determined by the free entry condition:

$$
W\kappa_e = \int v^c(q_e, b_e)\,\mathrm{d}H_e(q_e, b_e), \tag{A3}
$$

where H_e is the initial distribution of entrants over productivity and customer base levels, q_e and b_e. Note that firms do not produce upon entry. Rather, they start producing next period in which they become part of the set of active varieties Ω.

With a slight abuse of notation let $\Omega \equiv \int\int \mathrm{d}\Omega(q,b)$ be the mass of firms. Along the balanced growth the total mass of firms is constant, satisfying:

$$
\Omega = (1-\delta)\int\int F(\kappa^*)\,\mathrm{d}\Omega + \Omega_e \tag{A4}
$$

Labor market Clearing. Given endogenous entry and exit, labor market clearing is given by

$$
\overline{N} = \int\int (n^c(q,b) + n^x(q,b) + n^a(q,b) + \widetilde{\kappa}(q,b))\Omega(\,\mathrm{d}q,\,\mathrm{d}b) + \Omega_e\kappa_e. \tag{A5}
$$

which now includes labor demand of incumbent firms (including their operational costs) and of potential entrants.

A.3 Full model: Equilibrium definition

Given the description of the full model above, the associated equilibrium can be defined as follows:

Full model: Equilibrium definition. *A balanced growth path equilibrium consists of the following tuple in every period with* $j \in \Omega$*:* c_j, n_j^c, q_j, b_j, v_j, p_j, x_j, a_j, κ_j^*, C, R, W, Q, G, Ω, Ω_e *such that (i) firm-level output,* c_j*, and production employment,* n_j^c*, are given by (4) and (5), (ii) firms' productivity,* q_j*, and customer base levels,* b_j*, evolve according to (7) and (8), (iii) firm values are given by (A1), (iv) optimal innovation rates,* x_j*, marketing rates,* a_j*, and prices,* p_j*, satisfy (12) to (14), (v) firms' exit choices,* κ_j^**, are given by (23), (vi) aggregate consumption,* C*, is given by (1), (vii) the interest rate,* R*, and wage rate,* W *satisfy (3) and (27), (viii) the aggregate productivity index,* Q*, and it's growth rate,* G*, are given by (16) and* $G = Q'/Q - 1$*, and (ix) the mass of firms,* Ω*, and entrants,* Ω_e*, are given by (26) and (25).*

A.4 Restricted Model: Equilibrium definition

As mentioned in the main text, the restricted "productivity-only" version of full model is obtained by assuming $b' = b = 1$, i.e. $a_j = s_j = 0$ for all j in every period.[1] The associated equilibrium is then given by the following:

Restricted model: Equilibrium definition. *A balanced growth path equilibrium consists of the following tuple in every period with* $j \in \Omega$*:* c_j, n_j^c, q_j, b_j, v_j, p_j, x_j, κ_j^*, C, R, W, Q, G, Ω, Ω_e *such that (i) firm-level output,* c_j*, and production employment,* n_j^c*, are given by (4) and (5), (ii) firms' productivity,* q_j*, and customer base levels,* b_j*, evolve according to (7) and* $b_j = 1$ *for all* j*, (iii) firm values are given by (A1), (iv) optimal innovation rates,* x_j *and prices,* p_j*, satisfy (12) and (14), (v) firms' exit choices,* κ_j^**, are given by (23), (vi) aggregate consumption,* C*, is given by (1), (vii) the interest rate,* R*, and wage rate,* W *satisfy (3) and (27) where* $n_j^a = 0$ *for all* j*, (viii) the aggregate productivity index,* Q*, and it's growth rate,* G*, are given by (16) and* $G = Q'/Q - 1$*, and (ix) the mass of firms,* Ω*, and entrants,* Ω_e*, are given by (26) and (25).*

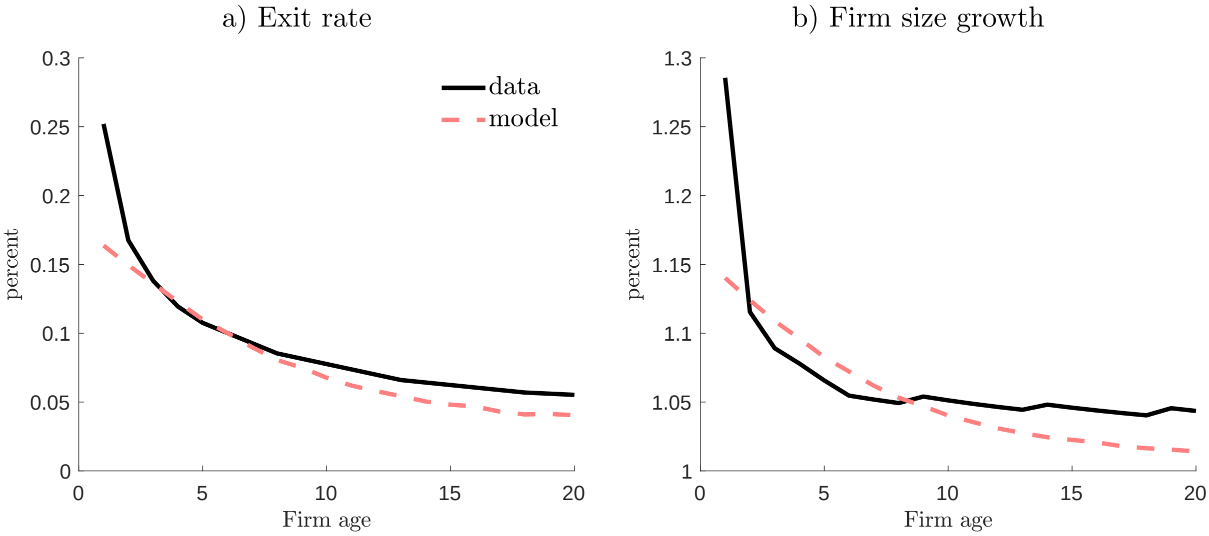
A.5 Restricted model: Parameterization

We parametrize the restricted model in exactly the same way as the our full model, with the exception that we do not target moments related to marketing or markups.

Towards this end, we externally set/normalize β, η, ψ_x, κ_e, and $\overline{N}$ in exactly the same way as described in Section 3.2. Next, parameters pertaining to innovation, firm exit and entry (λ, $\overline{x}$, σ_x, δ, μ_κ, σ_κ, σ_b, σ_q) are set according to the same indirect inference procedure as described in Section 3.2. Finally, all remaining which pertain to customer accumulation (γ, $\overline{a}$, σ_a, ψ_a, σ_m, ζ) are irrelevant in the restricted model. All model

[1] In the absence of customer accumulation motives, the optimal pricing decision (14) boils down to prices being set with a static markup over marginal costs, i.e. $p = \eta/(\eta - 1)W/q$.

Figure A1: Targeted moments in the restricted model: Exit and size growth rates



Note: The figure shows life-cycle profiles of exit rates (panel a) and average firm size growth rates (panel b). It does so for BDS "data" (solid line) and "model" (dashed line) generated moments.

parameters, their targets, and the respective fit of the restricted model are presented in Table A1 and Figure A1.

Table A1: Parameter values and targeted moments, restricted model

	Parameter	Value	Target/Source	Data	Model
β	Discount factor	0.97	Interest rate of approx. 3%		
η	Elasticity of substitution	3	Within range of Broda and Weinstein (2006)		
$\overline{N}$	Labor supply	22.5	$W = 1/\mu$, normalization		
ψ_x	Cost elasticity of R&D	2	see e.g. Hall and Ziedonis (2001); Bloom et al. (2002)		
$\overline{x}$	Cost level of R&D	4.53	Aggregate R&D/sales	0.02	0.03
σ_x	R&D cost scaling with size	1.4	R&D-to-sales on size reg. coef. (28)	-0.01	-0.01
λ	Innovation step size	0.117	Real GDP growth rate (%)	1.43	1.43
μ_q	Log initial productivity, mean	-1.26	Firm growth life-cycle profile, see Figure A1		
σ_q	Log initial productivity, dispersion	0.906	Firm growth life-cycle profile, see Figure A1		
δ	Exogenous exit rate	0.03	Firm exit life-cycle profile, see Figure A1		
μ_κ	Stochastic fixed costs, mean	0.175	Firm exit life-cycle profile, see Figure A1		
$\widetilde{\sigma}_\kappa$	Stochastic fixed costs, dispersion	0.475	Firm exit life-cycle profile, see Figure A1		

Note: The table presents values for all model parameters. The first three columns describe each parameter and its parameterized value. The fourth and fifth columns indicate the most relevant target or source and the last column reports the model fit.

B Proofs

In this appendix, we provide proofs of all the propositions in the main text. Towards this end, it will be useful to define the stationarized firm problem, which we do next. To ease the exposition, we drop the firm-level subscript j whenever it does not create confusion and we use primes ($'$) to denote next period values.

Stationarizing the firm problem. Using household's demand for individual goods (4) and the firm-level production function (5), firm value evaluated at optimal choices can be written as

$$\begin{aligned} v =& pc - W(n^c + n^x + n^a) + \frac{1-\delta}{1+R'}\left(xv(q(1+\lambda), b') + (1-x)v(q, b')\right) \\ =& c\frac{W}{q}(\mu - 1) - W(n^x + n^a) + \frac{1-\delta}{1+R'}\left(xv(q(1+\lambda), b') + (1-x)v(q, b')\right) \\ =& bq^{\eta-1}W^{1-\eta}C\mu^{-\eta}(\mu-1) - W(n^x + n^a) + \frac{1-\delta}{1+R'}\mathbb{E}v'. \end{aligned}$$

Dividing both sides with the aggregate productivity index, Q, and using the household Euler equation (3), we can write the stationarized firm value as

$$\begin{aligned} \frac{v}{Q} &= b\hat{q}^{\eta-1}\hat{W}^{1-\eta}\frac{C}{Q}\mu^{-\eta}(\mu-1) - \frac{W}{Q}(n^x + n^a) + \frac{1-\delta}{1+R'}\frac{Q'}{Q}\frac{\mathbb{E}v'}{Q'} \\ \widehat{v} &= b\hat{q}^{\eta-1}\frac{\widehat{C}}{\widehat{W}^{\eta-1}}\frac{\mu-1}{\mu^{\eta}} - \widehat{W}(n^x + n^a) + \beta(1-\delta)\mathbb{E}\hat{v}'. \end{aligned} \tag{A6}$$

A stationary aggregate customer base. Under Parts (a) and (b) of Assumption 1, the aggregate customer base, B, in the benchmark model can be written as

$$B = \int_j b_j dj = \sum_a b_a \mu_a = \sum_a [(1-\zeta)(1+g_b)]^a b_e (1-\delta)^a \mu_0 = \frac{b_e \delta \Omega}{1-(1-\delta)(1-\zeta)(1+g_b)},$$

where μ_a is the mass of firms of age a, b_e is the customer base entry endowment and Ω is the mass of firms. In the above, we have used the fact that under Assumption 1 all firms choose the same customer base growth rates. In this setting, B is well defined if $(1-\delta)(1-\zeta)(1+g_b) < 1$.

B.1 Proof of Proposition 1

Firm value. Using Assumption 1, and guessing that $\hat{v} = \overline{v}m$ (and, therefore, $\mathbb{E}\hat{v}' = \mathbb{E}\overline{v}m'$), we can write stationarized firm value (A6) as:

$$\widehat{v} = b\hat{q}^{\eta-1}\frac{\widehat{C}}{\widehat{W}^{\eta-1}}\frac{\mu-1}{\mu^{\eta}} - \widehat{W}(n^x + n^a) + \beta(1-\delta)\mathbb{E}\hat{v}'$$

$$\widehat{v} = b\hat{q}^{\eta-1}\left(\frac{\widehat{C}}{\widehat{W}^{\eta-1}}\frac{\mu-1}{\mu^{\eta}} - \widehat{W}(\overline{x}x^{\psi_x} + \overline{a}a^{\psi_a})\right) + \beta(1-\delta)\overline{v}b'\overbrace{\frac{q^{\eta-1}(x(1+\lambda)^{\eta-1}+1-x)}{(Q')^{\eta-1}}}^{\mathbb{E}(q')^{\eta-1}}$$

$$\widehat{v} = b\hat{q}^{\eta-1}\left(\frac{\widehat{C}}{\widehat{W}^{\eta-1}}\frac{\mu-1}{\mu^{\eta}} - \widehat{W}(\overline{x}x^{\psi_x} + \overline{a}a^{\psi_a}) + \beta(1-\delta)\overline{v}\frac{b'}{b}(x(1+\lambda)^{\eta-1}+1-x)\left(\frac{q/Q'}{q/Q}\right)^{\eta-1}\right)$$

$$\widehat{v} = b\hat{q}^{\eta-1}\left(\frac{\widehat{C}}{\widehat{W}^{\eta-1}}\frac{\mu-1}{\mu^{\eta}} - \widehat{W}(\overline{x}x^{\psi_x} + \overline{a}a^{\psi_a}) + \beta(1-\delta)\overline{v}\underbrace{(1+g_b)(1+g_q)(1+G)^{1-\eta}}_{=1+g_n}\right)$$

Combining with the results above, we obtain the following expression for stationarized firm value which confirms our initial guess:

$$\widehat{v} = b\widehat{q}^{\eta-1}\underbrace{\left[\frac{\widehat{C}}{\widehat{W}^{\eta-1}}\frac{\mu-1}{\mu^{\eta}} - \widehat{W}\left(\overline{x}x^2 + \overline{a}a^2\right) + \beta(1-\delta)(1+g_n)\overline{v}\right]}_{=\overline{v}}. \tag{A7}$$

Therefore, the firm value constant is implicitly given by:

$$\overline{v} = \frac{\widehat{W}^{1-\eta}\widehat{C}\mu^{-\eta}(\mu-1) - \widehat{W}\left(\overline{x}x^2 + \overline{a}a^2\right)}{1-\beta(1-\delta)(1+g_n)}. \tag{A8}$$

Optimal R&D, marketing and pricing decisions. The above result allows us to express the marginal benefits of optimal R&D, marketing and pricing choices. In particular, using the fact that $\widehat{v}(q(1+\lambda), b_j') - \widehat{v}(q_j, b_j') = \overline{v}b(1+g_b)\widehat{q}^{\eta-1}((1+\lambda)^{\eta-1}-1)$ and $\frac{\partial \mathbb{E}v(q',b')}{\partial b'} = \frac{b\widehat{q}^{\eta-1}(1+g_n)\overline{v}}{1+g_b}$, we can write firms' optimality conditions (12) to (14) as

$$x = \widetilde{\beta}\frac{(1+\lambda)^{\eta-1}-1}{2\widehat{W}\overline{x}(1+g_q)}\overline{v} \tag{A9}$$

$$a = \widetilde{\beta}\frac{(1-\zeta)\gamma}{2\widehat{W}\overline{a}(1+g_b)}\overline{v} \tag{A10}$$

$$\frac{1}{\mu} - \frac{1}{\overline{\mu}} = \widetilde{\beta}\frac{(1-\zeta)(1-\gamma)}{1+g_b}\frac{\overline{v}}{\hat{C}\overline{\mu}}, \tag{A11}$$

where $\widetilde{\beta} = \beta(1-\delta)(1+(1+g_q)(1+g_b)(1+G)^{1-\eta})$.

Note that with $\overline{v}$ constant and common across firms, all firm-level choices are also constant and common across firms which verifies our initial guess that $\widehat{v} = b\widehat{q}^{\eta-1}\overline{v} = m\overline{v}$.

For future reference, also notice that combining (A10) and (A11) we obtain an "arbitrage" condition between marketing and markups:

$$\frac{1}{\mu} = \frac{1}{\overline{\mu}} + \frac{1-\gamma}{\gamma}\frac{2\hat{W}\overline{a}}{\hat{C}\overline{\mu}}a \tag{A12}$$

The above states that, to the extent that firms use both channels of customer base accumulation (i.e. $\gamma < 1$), marketing and (inverse) markups co-move.

B.2 Proof of Proposition 2

Part a) of Proposition 2: Firm-level innovation increases with customer base growth. In what follows, we consider the impact of an exogenous change in the customer retention rate, $1-\zeta$, on optimal innovation, marketing and markups. First, we focus on optimal innovation using (A9):

$$\begin{aligned} dx = & \overbrace{\frac{x}{1-\zeta}d(1-\zeta) + \frac{x}{\overline{v}}\frac{\partial \overline{v}}{\partial(1-\zeta)}}^{\text{direct effect (on impact and in future periods)}} \\ & + \underbrace{\frac{x}{1+g_b}\gamma(1-\zeta)\frac{\partial a}{\partial(1-\zeta)}d(1-\zeta)}_{\text{indirect effect via marketing}} + \underbrace{\frac{x}{1+g_b}(1-\gamma)(1-\zeta)\frac{\partial s}{\partial \mu}\frac{\partial \mu}{\partial(1-\zeta)}d(1-\zeta)}_{\text{indirect effect via markups}} \\ \epsilon_{x,1-\zeta} = & \frac{dx}{x}\frac{1-\zeta}{d(1-\zeta)} = \frac{1}{1-\widetilde{\beta}} + \frac{(1-\zeta)\gamma a}{1+g_b}\epsilon_{a,1-\zeta} + (1-\eta)\frac{(1-\zeta)(1-\gamma)s}{1+g_b}\epsilon_{\mu,1-\zeta}, \end{aligned} \tag{A13}$$

where we have used the fact that $\frac{\partial \overline{v}}{\partial x} = 0$ (Envelope theorem) and $\frac{\partial \overline{v}}{\partial(1-\zeta)} = \frac{\overline{v}}{1-\widetilde{\beta}}\frac{\widetilde{\beta}}{1-\zeta}$. Similarly, we derive the impact elasticities for marketing and markups using (A10) and (A11):

$$\begin{aligned} da = & \overbrace{\frac{a}{1-\zeta}d(1-\zeta) + \frac{a}{\overline{v}}\frac{\partial \overline{v}}{\partial(1-\zeta)}d(1-\zeta)}^{\text{direct effect (on impact and in future periods)}} + \underbrace{\frac{a}{1+g_q}\left((1+\lambda)^{\eta-1}-1\right)\frac{\partial x}{\partial(1-\zeta)}d(1-\zeta)}_{\text{indirect effect via innovation}} \\ \epsilon_{a,1-\zeta} = & \frac{da}{a}\frac{1-\zeta}{d(1-\zeta)} = \frac{1}{1-\widetilde{\beta}} + \frac{g_q}{1+g_q}\epsilon_{x,1-\zeta}. \end{aligned} \tag{A14}$$

$$
\begin{aligned}
-\frac{1}{\mu}\frac{d\mu}{\mu} =& \overbrace{\left(\frac{1}{\mu}-\frac{1}{\overline{\mu}}\right)\frac{d(1-\zeta)}{1-\zeta}+\left(\frac{1}{\mu}-\frac{1}{\overline{\mu}}\right)\frac{\partial\overline{v}}{\partial(1-\zeta)}d(1-\zeta)}^{\text{direct effect (on impact and in future periods)}} \\
&+\underbrace{\left(\frac{1}{\mu}-\frac{1}{\overline{\mu}}\right)\frac{(1+\lambda)^{\eta-1}-1}{1+g_q}\frac{\partial x}{\partial(1-\zeta)}d(1-\zeta)}_{\text{indirect effect via innovation}} \\
\epsilon_{\mu,1-\zeta} =& \frac{d\mu}{\mu}\frac{1-\zeta}{d(1-\zeta)} = -\left(1-\frac{\mu}{\overline{\mu}}\right)\left[\frac{1}{1-\widetilde{\beta}}+\frac{g_q}{1+g_q}\epsilon_{x,1-\zeta}\right]. \qquad \text{(A15)}
\end{aligned}
$$

Combining (A13) to (A15), we can write

$$
\epsilon_{x,1-\zeta} = \frac{1}{1-\widetilde{\beta}}\frac{\overbrace{1+\frac{(1-\zeta)\gamma a}{1+g_b}+(\eta-1)\left(1-\frac{\mu}{\overline{\mu}}\right)\frac{(1-\zeta)(1-\gamma)s}{1+g_b}}^{>0}}{\underbrace{1-\frac{g_q}{1+g_q}\left(\frac{(1-\zeta)\gamma a}{1+g_b}+(\eta-1)\left(1-\frac{\mu}{\overline{\mu}}\right)\frac{(1-\zeta)(1-\gamma)s}{1+g_b}\right)}_{>0}} > 0, \qquad \text{(A16)}
$$

where we have used the fact that the denominator of the right-most fraction can be written as[2]

$$
\begin{aligned}
&\frac{1}{1+g_b}\left(1+g_b-\frac{g_q}{1+g_q}\left((1-\zeta)\gamma a+(\eta-1)\left(1-\frac{\mu}{\overline{\mu}}\right)(1-\zeta)(1-\gamma)s\right)\right) \\
=&\frac{1-\zeta}{1+g_b}\left(1+\gamma a\underbrace{\left(1-\frac{g_q}{1+g_q}\right)}_{>0}+(1-\gamma)s\underbrace{\left(1-(\eta-1)\left(1-\frac{\mu}{\overline{\mu}}\right)\right)}_{>0}\right) > 0.
\end{aligned}
$$

Therefore, R&D increases in response to an exogenous rise in customer retention rates. This happens both because of a direct effect and because of an indirect contribution through stronger endogenous customer accumulation. The latter is a weighted average of both marketing- and markup-driven customer accumulation, where the weights are the relative contributions of the two forces to customer base growth.

[2]Note that since $\mu > 1$, we can write

$$
\begin{aligned}
1-(\eta-1)\left(1-\frac{\mu}{\overline{\mu}}\right) =& 1-(\eta-1)\left(\frac{\frac{\eta}{\eta-1}-\mu}{\frac{\eta}{\eta-1}}\right) \\
=&1-(\eta-1)\frac{\eta-\mu(\eta-1)}{\eta} = 1-\eta\underbrace{\frac{1-\mu}{\overline{\mu}}}_{<0}-\underbrace{\frac{\mu}{\overline{\mu}}}_{<1} > 0.
\end{aligned}
$$

Part b) of Proposition 2: Firm-level customer accumulation increases with innovation. Using the same approach as above, we can write the following responses of R&D, marketing and markups to an exogenous increase in innovation step size, $1+\lambda$:

$$\begin{aligned}\epsilon_{x,1+\lambda} =&(\eta-1)\frac{(1+\lambda)^{\eta-1}}{(1+\lambda)^{\eta-1}-1}\left[1+\frac{\widetilde{\beta}}{1-\widetilde{\beta}}\frac{x}{1+g_q}\right]\\ &+\frac{(1-\zeta)\gamma a}{1+g_b}\epsilon_{a,1+\lambda}+(1-\eta)\frac{(1-\zeta)(1-\gamma)s}{1+g_b}\epsilon_{\mu,1+\lambda}>0,\end{aligned}$$

$$\epsilon_{a,1+\lambda} =(\eta-1)\frac{x(1+\lambda)^{\eta-1}}{1+g_q}\frac{1}{1-\widetilde{\beta}}+\frac{g_q}{1+g_q}\epsilon_{x,1+\lambda}>0,$$

$$\epsilon_{\mu,1+\lambda} =-\left(1-\frac{\mu}{\overline{\mu}}\right)\left[\frac{(\eta-1)x(1+\lambda)^{\eta-1}}{1+g_q}\frac{1}{1-\widetilde{\beta}}+\frac{g_q}{1+g_q}\epsilon_{x,1+\lambda}\right]<0.$$

B.3 Proof of Proposition 3

Part a) of Proposition 3: Aggregate growth. Let us 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 common customer base growth 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

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

where in the above we guess that $Q' = (x(1+\lambda)^{\eta-1} + 1 - x)^{\frac{1}{\eta-1}} Q$, which we subsequently confirm. Therefore, $1 + G = Q'/Q = (x(1+\lambda)^{\eta-1} + 1 - x)^{\frac{1}{\eta-1}}$.

Part b) of Proposition 3: All else equal, aggregate growth increases with firm-level customer base growth. The proof of this part of the proposition directly follows from Part a) of Proposition 3 and from Proposition 1, where we have established that $\epsilon_{x,1-\zeta} > 0$. Combining the two, we can write

$$\frac{dG}{d(1-\zeta)} = \frac{(1+G)^{2-\eta}}{\eta - 1}((1+\lambda)^{\eta-1} - 1)\frac{\partial x}{\partial(1-\zeta)} > 0.$$

B.4 Proof of Proposition 4

Part a) of Proposition 4: R&D subsidies raise innovation. First, observe that the impact of R&D subsidies τ_x is isomorphic to the effect of changes in R&D cost scaling $\overline{x}$. 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. $\partial x/\partial \tau_x = -\partial x/\partial \overline{x}$. Taking G as given and differentiating (A9) with respect to $\overline{x}$ gives:

$$\epsilon_{x,\overline{x}} = \frac{dx}{x}\frac{\overline{x}}{d\overline{x}} = -\left(1 + \frac{\widehat{W}\overline{x}x^2}{\widehat{W}^{1-\eta}\widehat{C}\mu^{-\eta}(\mu - 1) - \widehat{W}(\overline{x}x^2 + \overline{a}a^2)}\right) < 0.$$

Therefore, intuitively, cheaper R&D leads to firms optimally choosing more of it, i.e. $\epsilon_{x,\tau_x} > 0$.

Part b) of Proposition 4: The efficacy of R&D subsidies increases with customer base growth. We now compute how the magnitude of the elasticity of R&D with respect to subsidies, $|\epsilon_{x,\tau_x}|$, changes with $1-\zeta$. To ease the notation, in what follows, we denote $\widetilde{\pi} = \widehat{W}^{1-\eta}\widehat{C}\mu^{-\eta}(\mu - 1) - \widehat{W}(\overline{x}x^2 + \overline{a}a^2)$. Next, we show that $\widetilde{\pi}$ decrease with $1 - \zeta$:

$$\begin{aligned}\frac{\partial\widetilde{\pi}}{\partial(1-\zeta)} =& \widehat{W}^{1-\eta}\widehat{C}\mu^{-\eta}\frac{\partial\mu}{\partial(1-\zeta)}(1 - \eta + \eta\mu^{-1}) - 2\widehat{W}\left(\overline{x}x\frac{\partial x}{\partial(1-\zeta)} + \overline{a}a\frac{\partial a}{\partial(1-\zeta)}\right)\\ =& \frac{1}{1-\zeta}\left[\widehat{W}^{1-\eta}\widehat{C}\mu^{-\eta}\underbrace{(\mu(1-\eta)+\eta)}_{>0}\underbrace{\epsilon_{\mu,1-\zeta}}_{<0} - 2\widehat{W}\underbrace{(\overline{x}x^2\epsilon_{x,1-\zeta} + \overline{a}a^2\epsilon_{a,1-\zeta})}_{>0}\right] < 0,\end{aligned}$$

where we have used the fact that $\mu(\eta - 1) < \eta$. This can be seen from the fact that $\mu < \overline{\mu} = \eta/(\eta - 1)$. Therefore, $\mu(1 - \eta) + \eta > 0$. Intuitively, as customer retention rates increase, this incentivizes firms to invest more into customer base growth and R&D which, in turn, reduces (scaled) profits, $\widetilde{\pi}$.

Returning to the elasticity of ϵ_{x,τ_x} w.r.t. $(1-\zeta)$, we can write:

$$\begin{aligned}\frac{\partial \epsilon_{x,\tau_x}}{\partial(1-\zeta)} =& \frac{\partial\left(1+\frac{\widehat{W}\overline{x}x^2}{\widetilde{\pi}}\right)}{\partial(1-\zeta)} = 2\frac{\widehat{W}\overline{x}x^2}{\widetilde{\pi}}\frac{\epsilon_{x,1-\zeta}}{1-\zeta} - \frac{\epsilon_{x,\tau_x}-1}{\widetilde{\pi}}\frac{\partial\widetilde{\pi}}{\partial(1-\zeta)} \\ =& \frac{\epsilon_{x,\tau_x}-1}{1-\zeta}\left(2\epsilon_{x,1-\zeta} - \epsilon_{\widetilde{\pi},1-\zeta}\right) > 0.\end{aligned}$$

B.5 Proof of Proposition 5 and 6

First, recall that sales in our model are given by $pc = bq^{\eta-1}(\mu W)^{1-\eta}C$. Therefore, we can approximate firm-level sales growth as

$$\begin{aligned}\Delta\log(pc) =& \Delta\log(b) + (\eta-1)(\Delta\log(q) - \Delta\log(\mu)) + \Delta(1-\eta)\log(W) + \Delta\log(C) \\ g \approx& g_b + (\eta-1)(g_q - g_\mu) + (2-\eta)G,\end{aligned}$$

Recall further that we have the following expressions for customer base and productivity growth:

$$\begin{aligned}g_b =& (1-\zeta)(1+\gamma a + (1-\gamma)s) - 1 \\ g_q =& x(1+\lambda)^{\eta-1} - x\end{aligned}$$

In addition, using the definitions of s, n_a and n_x, we can write:

$$a = \left(\frac{n_a}{\overline{a}m^{\sigma_a}}\right)^{\frac{1}{\psi_a}}, \quad x = \left(\frac{n_x}{\overline{x}m^{\sigma_x}}\right)^{\frac{1}{\psi_x}}, \quad s = \left(\mu\widehat{W}\right)^{1-\eta}.$$

Therefore, changes in firm-level sales growth can be written as

$$\begin{aligned}dg \approx& (1-\zeta)\gamma a\frac{1}{\psi_a}\left[\frac{dn_a}{n_a} - \sigma_a\frac{dm}{m}\right] + (1-\zeta)(1-\gamma)s(1-\eta)\frac{d\mu}{\mu} \\ &+ (\eta-1)((1+\lambda)^{\eta-1}-1)x\frac{1}{\psi_x}\left[\frac{dn_x}{n_x} - \sigma_x\frac{dm}{m}\right] + (1-\eta)dg_\mu + \text{time f.e.}\end{aligned}$$

Before moving on, note that it will be more convenient to rewrite $dg = g_t - g_{t-1} \approx RHS_{t-1}$ as $g_t \approx RHS_{t-1} + g_{-1}$. In addition, we note that $dm/m = d(pc/(PC))/(pc/(PC)) = g + \text{time fixed effects}$. Finally, note the timing on markup growth is $dg_\mu = \Delta\log(\mu_t) - \Delta\log(\mu_{t-1})$, since *current* period markups directly influence sales. Therefore, mapping the above to our observables (i.e. $g = \Delta\log(sales)$, $dn_x/n_x = \Delta log(XRD)$ etc) and explicitly stating timing, we get

$$\begin{aligned}\Delta\log(sales_t) \approx & \overbrace{(1-\zeta)\gamma a\frac{1}{\psi_a}}^{\beta_a}\Delta\log(S\&M_{t-1}) + \overbrace{((1-\zeta)(1-\gamma)s-1)(1-\eta)}^{\beta_\mu}\Delta\log(\mu_{t-1}) \\ & +(\eta-1)g_q\frac{1}{\psi_x}\Delta\log(XRD_{t-1}) \\ & -\left[(1-\zeta)\gamma a\frac{\sigma_a}{\psi_a}+(\eta-1)g_q\frac{\sigma_x}{\psi_x}-1\right]\Delta\log(sales_{t-1}) \\ & +(1-\eta)\Delta\log(\mu_t) + \text{time f.e.}\end{aligned}$$

Therefore, we have

$$\begin{aligned}\beta_a =& (1-\zeta)\gamma a\frac{1}{\psi_a} \\ \beta_\mu =& ((1-\zeta)(1-\gamma)s-1)(1-\eta).\end{aligned}$$

Finally, Assumption 1 and using (13) and (17), we can write:

$$\begin{aligned}\beta_a =& \frac{1}{\psi_a}\frac{\gamma^2}{\eta-1}\beta_1, \\ \beta_\mu =& \beta_0+\gamma^2\beta_1,\end{aligned}$$

where $\beta_0 = (\eta-1)(1-g_b-\zeta)$ and $\beta_1 = \frac{(1-\zeta)^2(\eta-1)}{\bar{a}}\frac{\beta(1-\delta)\overline{v}}{2\hat{W}}(\eta-1)$. For future reference, note that the term $\frac{(1-\zeta)^2}{\bar{a}\psi_a}$ is an overall measure of customer accumulation frictions.

B.6 Proof of Proposition 7

Aggregate productivity index is defined as

$$Q = \left(\int\int bq^{\eta-1}\Omega(\,\mathrm{d}q,\,\mathrm{d}b)\right)^{\frac{1}{\eta-1}}$$

In what follows, we use short-hand notation with firm index j for the sake of clarity. For example we write, $\int_\Omega x_j\,\mathrm{d}j$ instead of $\int\int x(q,b)\,\Omega(\,\mathrm{d}q,\,\mathrm{d}b)$. This makes the following derivations easier to follow.

Special case: Exact decomposition. We follow the decomposition used in Foster et al. (2001) and consider the change in Q, where we take last period's values as the baseline. First, we consider the special case of $\eta = 2$. Let $\Delta x_t \equiv x_t - x_{t-1}$. Let Ω_s denote the set of surviving incumbent firms, Ω_{ex} the set of firms that exit between $t-1$ and t, and $\Omega_{en} \equiv \Omega_e H$ the set of entrants.

$$
\begin{aligned}
\Delta Q_t =& \int_{\Omega_s} b_{j,t} q_{j,t} \,\mathrm{d}j + \int_{\Omega_{en}} b_{j,t} q_{j,t} \,\mathrm{d}j - \int_{\Omega_s} b_{j,t-1} q_{j,t-1} - \int_{\Omega_{ex}} b_{j,t-1} q_{j,t-1} \,\mathrm{d}j \\
=& \int_{\Omega_s} (b_{j,t} - b_{j,t-1}) q_{j,t-1} \,\mathrm{d}j + \int_{\Omega_s} b_{j,t-1} (q_{j,t} - q_{j,t-1}) \,\mathrm{d}j + \int_{\Omega_s} (b_{j,t} - b_{j,t-1})(q_{j,t} - q_{j,t-1}) \,\mathrm{d}j \\
&+ \int_{\Omega_{en}} b_{j,t} q_{j,t} \,\mathrm{d}j - \int_{\Omega_{ex}} b_{j,t-1} q_{j,t-1} \,\mathrm{d}j \\
=& \underbrace{\int_{\Omega_s} \Delta b_j q_j \,\mathrm{d}j + \int_{\Omega_s} b_j \Delta q_j \,\mathrm{d}j + \int_{\Omega_s} \Delta b_j \Delta q_j \,\mathrm{d}j}_{\text{surviving incumbents}} + \underbrace{\int_{\Omega_{en}} b_j q_j \,\mathrm{d}j - \int_{\Omega_{ex}} b_j q_j \,\mathrm{d}j}_{\text{net entry}}
\end{aligned}
$$

The first equality follows from accounting identity our definition of our sets Ω_x. Notice that the above is an *exact* decomposition. Also, since last period's entrants become this period's incumbents (either continuing to the next period or exiting), then entry contributes positively with the (customer-base-)weighted average of firm-level productivity. With the similar logic, exiting firms contribute negatively with their respective (customer-base-)weighted average productivity.

Generalization: Approximate decomposition. In the full model, $\eta > 2$ and, therefore, we use a first-order approximation to the above. However, a 1^{st} order approximation cannot account for discontinuous changes in firm distribution, such as entry and exit. Therefore, we will consider the impact of net entry as a residual, Γ:

$$
\Delta Q = \frac{1}{\eta - 1} Q^{2-\eta} \left(\int_{\Omega_s} \Delta b_j q_j^{\eta-1} + (\eta - 1) \int_{\Omega_s} b_j \frac{\Delta q_j}{q_j} q_j^{\eta-1} \right) + \Gamma
$$

Rewriting the above, gives the following expression for aggregate growth:

$$
G = \frac{\Delta Q}{Q} = \frac{1}{\eta - 1} Q^{1-\eta} \left(\int_{\Omega_s} \Delta b_j q_j^{\eta-1} + (\eta - 1) \int_{\Omega_s} b_j \frac{\Delta q_j}{q_j} q_j^{\eta-1} \right) + \Gamma
$$

Therefore, we can compute the value of Γ as follows:

$$
\Gamma = G - \frac{1}{\eta - 1} Q^{1-\eta} \left(\int_{\Omega_s} \Delta b_j q_j^{\eta-1} + (\eta - 1) \int_{\Omega_s} b_j \frac{\Delta q_j}{q_j} q_j^{\eta-1} \right)
$$

Finally, we highlight the reallocation channels within incumbent survivors' growth:

$$
\begin{aligned}
G =& \frac{\Delta Q}{Q} = \frac{1}{\eta - 1} Q^{1-\eta} \left(\int_{\Omega_s} \Delta b_j q_j^{\eta-1} + (\eta - 1) \int_{\Omega_s} b_j \frac{\Delta q_j}{q_j} q_j^{\eta-1} \right) + \Gamma \\
=& \frac{1}{\eta - 1} Q^{1-\eta} \left(\int_{\Omega_s} \Delta b_j q_j^{\eta-1} + (\eta - 1) \left(\int_{\Omega_s} \bar{b} \frac{\Delta q_j}{q_j} q_j^{\eta-1} + \int_{\Omega_s} (b_j - \bar{b}) \frac{\Delta q_j}{q_j} q_j^{\eta-1} \right) \right) + \Gamma
\end{aligned}
$$

C Model Estimation, Solution and Simulation

In this appendix we describe how we estimate the general model, our solution method, and how we simulate our model and how we construct counterfactuals in Section 4.

C.1 Estimation

We estimate the model by the Method of Simulated Moments (MSM). Let $d(i)$ for $i = 1, \ldots 49$ denote i-th data moment. Using the method described in Section C.3 below, we simulate the corresponding model moment $m(i)$. We minimize the weighted percentage deviation between the model and data moments as

$$\min \sum_i \tilde{w}(i) \left(\frac{m(i) - d(i)}{\frac{1}{2}\left(d(i) + m(i)\right)} \right)^2 . \tag{A17}$$

We set the weights $\tilde{w}(i)$ to (i) account for the fact that firms' life-cycle targets (on growth and exit) have 20 moments each and (ii) put more emphasis on aggregate growth and moments pertaining to customer accumulation.[3] Specifically, we set the weights on firms' growth and exit rates to $w(i) = 1/20$, while setting $w(i) = 2$ for i corresponding to aggregate growth, S&M/sales and coefficients in (29). All the remaining weights are set to $w(i) = 1$. Then, we normalize the weights $\tilde{w}(i) = w(i)/\sum_i w(i)$.

C.2 Solution method

In the full model, the firm problem is characterized by two state variables: productivity q and customer base b. We discretize the state space using 41 grid points for productivity 61 points for customer base. Both grids are equidistant in the log of variables and span -1 to 1 for log productivity and -3 to 6 for log customers base. These relatively large grids allow us to capture the full extent of the firm life-cycle dynamics induced by the evolution of productivity and demand. Experimenting with a wider grid and a larger number of grid points did not significantly affect the results.

The quantitative model is challenging to solve. A key challenge is that firm's pricing, marketing, and R&D investment problems are highly non-linear and interdependent. Indeed, a key feature of our model is the endogenous feedback loop between customer accumulation and R&D. This non-linearity calls for global solution method that can capture accurately the lifecycle dynamics of businesses. The dependence of all endogenous choices then requires a gradual approach to parameterizing the model.

[3] Note that we impose a penalty value of 10 for qualitative differences between model and data moments, i.e. when $d(i)m(i) < 0$.

Solution algorithm. To solve the full model numerically, we use value function iteration. At each step of the iteration, we employ a nonlinear solver to determine the optimal marketing, markup, and R&D investment. We discretize value functions using grid-based interpolation. If the value functions fall off the grid points described above, but within the grid range, we use linear interpolation. To evaluate off-grid points outside of the grid range - below the lowest or above the highest grid point - we use linear interpolation.

For policy functions we use linear interpolation for off-grid values within the grid range, but nearest on-grid value for values outside the grid range. The reason is that, unlike firm values, policy functions have known upper and lower bounds. These are 0 and 1 for innovation probability x and 0 and z^{max} for marketing intensity z.[4] Finally, optimal prices cannot be below zero. By solving the first-order conditions on-grid, we can guarantee that these bounds hold at each grid point and - by construction - for each linear interpolant between them.

The solution algorithm is as follows:

1. Given a guess for value function $V^{(i)}\colon \mathbb{R}^2 \to \mathbb{R}$, determine firms' exit thresholds $\kappa^*\colon \mathbb{R}^2 \to \mathbb{R}$ defined in (23). Compute $v^c = F(\kappa^*)\left[v - \hat{W}\widetilde{\kappa}\right]$. This is available in closed form, because under the logistic distribution, the survival probability is given by
$$F(\kappa^*) = \left[1 + \exp\left(\frac{-(\kappa^* - \mu_\kappa)}{\sigma_\kappa}\right)\right]^{-1}.$$
and conditional expectation is given by
$$\widetilde{\kappa} = \mathbb{E}\left[\kappa \mid \kappa \leq \kappa^*\right] = \mu_\kappa + \sigma_\kappa\left(\frac{1 - F(\kappa^*)}{F(\kappa^*)}\ln(1 - F(\kappa^*)) + \ln(F(\kappa^*))\right)$$

2. Given v^c, solve the first-order conditions (12), (13), and (14) to obtain policy functions $x\colon R^2 \to [0,1]$, $z\colon R^2 \to [0, z^{max}]$, and $p\colon R^2 \to [0,\infty)$.

3. Given all optimal choices, use equation (A1) to determine an updated guess for the value function $V^{(i+1)}$.

4. Go back to 1 and iterate until $\sup_{q,b} \frac{|V^{(i)} - V^{(i+1)}|}{1+V^{(i)}} < 10^{-4}$.

Solving for the balanced growth path equilibrium. In the initial equilibrium, we can normalize the total spending in the economy to unity, $\hat{C} = 1$, using the mass of entrants as a normalizing constant. Given our definition of Q, we can normalize $\hat{W} = \frac{1}{\mu}$ which implicitly determines the labor supply.[5] Finally, we can normalize the mean of the

[4] The value z^{max} is such that for all $z \geq z^{max}$ the firm value falls below outside option and the firm would rather exit.

[5] This is a particularly convenient normalization of the wage level since it corresponds to the equilibrium wage in a model with constant markup, $\mu(q,b) \equiv \overline{\mu}$.

initial customer base of firms to unity and solve for the value of entry costs that make the first-order condition hold. The remaining equilibrium variable - the aggregate growth rate G - is one of the parameterization targets described in Section 3.2.

In the policy experiments in Section 4, we introduce R&D subsidies or operational cost subsidies and solve for the new BGP equilibrium. Towards this end, we solve for the new equilibrium values of $\hat{C}$, $\hat{W}$, B, and G as follows

1. guess the new values G^{guess} and C^{guess}
2. using the above guesses, solve the firm problem and simulate the economy
3. use free entry condition (25) to back out implied aggregate growth (G^{impl})
4. use labor market clearing (27) to back out the mass of firms
5. given equilibrium firm mass, compute implied C^{impl}
6. stop if $\max\left\{|G^{guess} - G^{impl}|, |(C^{guess} - C^{impl})/C^{guess}|\right\} < 0.001$, otherwise go back to 1.

The values of the entry cost and the labor supply are implied by the normalizations employed in the initial equilibrium.[6]

C.3 Simulation

To estimate moments pertaining to firm dynamics and aggregate variables, we need to approximate the stationary firm distribution. Towards this end, we use off-grid stochastic simulation.

Simulating the stationary distribution. We draw 10^5 entrants from the initial distribution $H_e(q_e, b_e)$ over the idiosyncratic states. From then on, in each period, firms face transitory exit and innovation shocks, and make decisions regarding pricing, employment, investment into R&D, and exit. Based on these choices, firms accumulate customers and improve productivity. The implied values may lie off-grid, since we use grid-based interpolation as the non-parametric functional for of the policy functions.

In every period, we replace exiters with entrants drawn from the initial distribution H_e. We repeat this period over 250 periods. We disregard first 150 periods and use remaining 100 to compute moments from the stationary distribution such as the aggregate growth rate or auxiliary regressions corresponding to moments estimated in Compustat.

More specifically, while the firm is alive, in each period we simulate the following steps:

[6]Note that in the iteration described above, it is computationally more convenient to use the free entry condition to back out initial customer base levels, μ_{bini}, and then iterate until $\mu_{bini}^{impl} = 1$ (the normalization in the full model).

- before any shock is realized, firms decide whether to pay the fixed cost and continue operating or avoid paying the cost and exit
- exiting firms are replaced by new entrants
- all firms make their endogenous choices, given their relative productivity and customer base levels
- given optimal decisions, firms' customer base levels evolve and all firms observe the realization of the innovation shock which determines next period's productivity
- firms transition into the next period and the steps repeat.

Simulating life-cycle profiles. When generating moments pertaining to firm life-cycle dynamics, we mimic the design of the BDS. Starting from an initial distribution, we simulate the life-cycle of a single "representative" cohort of 10^5 firms. The procedure is identical to the one described above pertaining to the stationary distribution with the exception that firms which shut down are *not* replaced by new entrants. We follow the cohort for 20 years and use the resulting simulation to generate moments in Figure 1 in the main text.

Isolating the influence of customer acquisition. Finally, let us describe the counterfactual R&D decisions which allow us to isolate the contribution of customer acquisition on aggregate growth. In particular, we consider a "counterfactual" environment in which firms expect an increase in customer attrition, $\underline{\zeta} = \zeta + \epsilon$, where $\epsilon \geq 0$. Specifically, we consider a value for ϵ such that *average* customer base growth among continuing businesses is zero in the baseline economy. When computing the counterfactual decisions, all equilibrium variables and structural parameters other than ζ remain at the values of the full model:

$$v(q,b) = \max_{p,n^x,n^a} \left\{ bp^{1-\eta}C - W(bp^{-\eta}C/q + n^x + n^a) + \frac{1-\delta}{1+R'} \left[\begin{array}{c} xv^c(q(1+\lambda), b') + \\ \\ (1-x)v^c(q,b') \end{array} \right] \right\} \tag{A18}$$

$$\begin{aligned} \text{s.t.} \quad c =& bp^{-\eta}C, \quad c = qn^c, \quad n^x = \overline{x}m^{\sigma_x}x^{\psi_x}, \quad n^a = \overline{a}m^{\sigma_a}a^{\psi_a}, \quad s = m^{\sigma_s}\frac{pc}{PC}, \\ b' =& (1-\underline{\zeta})b(1+\gamma a + (1-\gamma)s), \quad x \in [0,1] \\ v^c =& F(\kappa^*)\left[v - \hat{W}\widetilde{\kappa}\right]. \end{aligned}$$

Note that the only change with respect to (A1) is the use of $\underline{\zeta}$. The "direct effect" is calculated by solving (A18) for new value of n^x while keeping all other parameters, policy functions, and value functions unchanged at their levels in the full model. In contrast, the "feedback loop effect" is calculated by solving (A18) for all policy functions (including exit) and value functions.

When simulating the counterfactual economy, we follow the same procedure as in the baseline model. Importantly, we use the same series of stochastic shocks to innovation and operating costs. This allows us to compare – for the same values of the state-variables – optimal and counterfactual decisions. Note that when computing "direct effect" counterparts, we *only* use firms' counterfactual R&D decisions. In particular, we consider the same set of firms as in the full model (i.e. we do not use the exit rule implied by firms counterfactual firm values).

Finally, using the counterfactual R&D decisions (and thus productivity values) we can compute the counterfactual evolution of firm-level customer base using (8). 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 full and the counterfactual model. See Figure 2 in the main text for the illustration of the reallocation channel.

D Empirical results: Data and robustness

In this appendix, we provide details of our empirical results. This includes a description of our data sources and sample selection, but also several robustness exercises pertaining to our parameterization and validation in Sections 3.2 and 5.

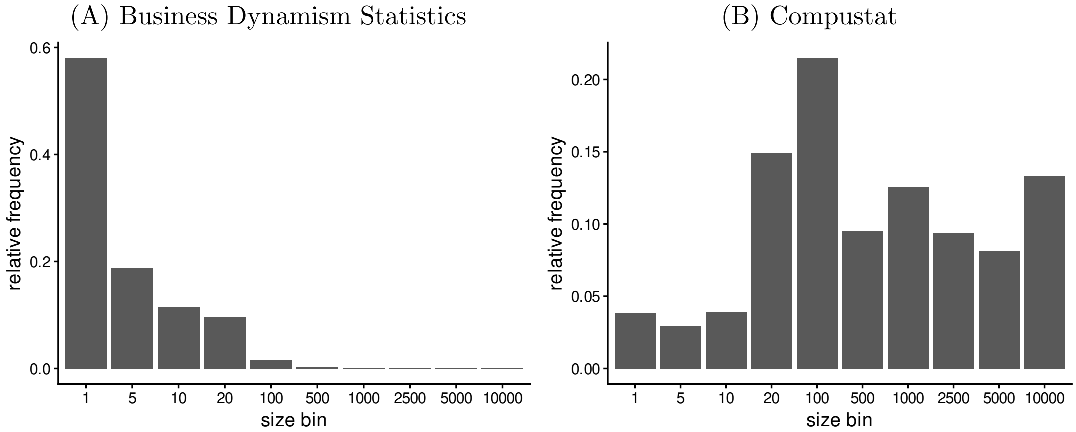
D.1 Data and Sample Selection

Throughout the paper, we use three main data sources for information on firms and business dynamims: (i) Business Dynamics Statistics (BDS), (ii) Compustat and (iii) Capital IQ. In what follows, we describe each of these datasets and how we use them in our analysis.

Business Dynamic Statistics. The BDS database is an aggregated version of administrative data on U.S. firms covering about 98 percent of private employment. As such, it is our primary source of information on aggregate business dynamism patterns. We use information on the number of existing firms, number of exiting firms, and their employment from 1979 to 2019, avoiding the COVID-19 pandemic.

Importantly, the BDS offers also information on firm age, which allows us to construct firms' average life-cycle patterns of exit and employment growth. Since the BDS bins together firms aged 6-10, 11-15, 16-20 years, we linearly interpolate values within these bins.

Figure A2: 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.

Compustat and Capital IQ. We use the Compustat fundamental annual data from 1997 to 2019. The starting point is chosen to align the sample with the availability of sales and marketing expenses in Capital IQ. The Compustat sample consists only of publicly listed companies. This sample differs significantly from the universe of firms in the U.S., i.e. from the BDS data and our model, which matches the latter well. Figure A2 illustrates the differences in the size distributions between BDS and Compustat. In the latter, the mass of firms is shifted towards the largest firms.

Therefore, 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, we weigh our regressions in the model simulated data with employment based weights from Compustat.

Specifically, we pool all observation across all years 1997 - 2019 and calculate the empirical distribution of firm employment in Compustat in each BDS size bin. Since the BDS corresponds to the universe of firms in the US economy, we can calculate "selection probability" into Compustat in each size bin. This selection probability is defined as the ratio of the number of firms in Compustat in the given size bin to the total number of firms in this size bin in the US. For all auxiliary calibration regressions that are based on Compustat sample, we weigh the model simulated data with the selection probabilities to mimic the selection into the Compustat sample.

Sample selection. In our analysis, we exclude observations of financial firms (SIC classification between 6000 and 6999), utilities (SIC classification 4900-4999), mining and agriculture (SIC $\leq$ 1400).[7] In addition, we consider only companies incorporated in the

[7] As is standard in the literature, we remove financial firms because the balance reporting conventions are very different for those companies. We drop utilities because it is a non-representative and regulated sector. We removed mining and agriculture because of very few firms undertaking R&D leading to

U.S. and traded in the U.S. stock exchange.

To deal with measurement error and outliers, we first remove observations with clear errors, such as negative sales (variable `sale` in Compustat) or cost of goods sold (`cogs`). Next, we remove the observations in which the ratio of R&D to sales or S&M to sales exceeds 1000%.

Annual growth rates tend to be noisy measures of the underlying fundamentals of public firms (Campbell and Shiller, 1988). To deal with this issue, we winsorize at the top and bottom 1% of all growth rates in regression (29). Our goal is to account for extreme growth rates while preserving as much information as possible. he percentiles are calculated individually in each 3-digit-NAICS $\times$ year cell to avoid over-weighting certain sectors or years.[8] All nominal variables are deflated with the GDP deflator.

Finally, we drop top and bottom 1% observations in terms of COGS-to-sales ratio as in De Loecker et al. (2020); Autor et al. (2020). The reason is to remove excessive weight of firms with very low (or very high) COGS and consequently very larger (low) measured markups. TAgain, the percentiles are calculated individually in each 3-digit-NAICS $\times$ year cell.

Definitions. As for the Compustat sample, R&D intensity is defined as the ratio of the R&D expenses, `xrd` in Compustat, relative to `sale`. We follow De Loecker et al. (2020) and define the markup as the ratio of sales to the cost of goods sold (COGS) multiplied by the sector specific output elasticity estimated in De Loecker et al. (2020). We refer the reader to De Loecker et al. (2020) for the details of the markup estimation.

After the cleaning procedure, the Compustat data sample consists of 13,710 unique firms and 120,130 firm-year observations.[9]

Next, we merge the Compustat sample with Capital IQ. We use Capital IQ data from 1997-2019 which allows us to match 12,535 firms (91%) to our baseline Compustat sample. Among these firms, in 34% of firm-year observation we measure selling and marketing expenses (S&M) directly. In addition, following He et. al. (2025), we recover an additional 14% of firm-year observations which do not report "selling and marketing expenses," but report their sub-components (either "advertising" expenses, "marketing" expenses or both). In these cases, we use the sum of the two sub-components as a measure of sales and marketing. Overall, this yields information on sales and marketing expenses for 48% of the sample.

outlying observations in our cross-industry regressions.

[8]The results continue to hold in the non-winsorized sample. However, we strive to make sure that our results are not driven by a small number of extreme observations.

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

D.2 Further Details: Empirical Support for Model Predictions

In this section, we provide more details pertaining to the results pertaining to Section 5 and, especially, those underlying Figure 4.

Data and sample selection. The empirical analysis is based on the merged Compustat-Capital IQ data described in the previous Section D.1. To estimate industry fixed effects, we use the 3-digit NAICS classification. The data is pooled across all years 1997-2019. Firm age is defined as the number of years since founding determined by merging Capital IQ and the Loughran-Ritter database (Loughran and Ritter, 2004).[10]

The results are based on firm-level regressions in which we allow the coefficients of interest to vary at the level of 3-digit NAICS industries. Finally, to estimate the industry-level coefficients reliably, we restrict attention to industries with at least 40 firm-year observations with positive R&D. We choose a conservative threshold in the baseline to avoid the risk that imprecisely estimated coefficients are driving our result. Below, we show robustness to the choice of minimum observations threshold (see Figure A3).

Regression specifications. In Panel a) of Figure 4, we estimate two regressions indicated in (39) by letting $y = R\&D$ and $y = S\&M$, and report industry-specific regression coefficients $\widehat{\beta}^{age}_{i,R\&D}$ on the horizontal axis and $\widehat{\beta}^{age}_{i,S\&M}$ on the vertical axis.

In Panel b) of Figure 4, we report coefficients $\widehat{\beta}_{i,a}$ estimated in regression

$$\Delta \ln(\text{sales}_{j,t}) = \delta_i + \delta_t + \sum_i \mathbf{1}_{i(j)=i}\beta_{i,a}\Delta \ln n^a_{j,t-1} + \beta_\mu \Delta \ln \mu_{j,t-1} + \Gamma \mathbb{X}_{j,t} + \epsilon_{j,t},$$

where $\mathbf{1}_{i(j)=i}$ is the indicator variable equal to one if firm j belongs to industry i and zero otherwise. δ_i and δ_t mark industry and time fixed effects, respectively. Controls X includes the lagged dependent variable, lagged growth rate of R&D expenses, current and lagged growth of markups, and current size measured in terms of sales.[11] On the vertical axis in Panel b) in Figure 4, we show industry i's median R&D-to-sales ratio.

The red lines in both panels in Figure 4 correspond to the revenue-weighted OLS regression in which the variable presented on the vertical axis is projected on a constant and the variable on the horizontal axis. The regression weights (and the size of the markers in the Figure) correspond to industry revenue shares. In Panel a) of Figure 4, the regression coefficient between our estimates of $\widehat{\beta}^{\text{age}}_{i,a}$ and $\widehat{\beta}^{\text{age}}_{i,R\&D}$ is 0.66 with standard error of 0.11 (i.e., p-value of 0.1%) and R^2 of 0.56. In Panel b) of Figure 4, the relevant

[10] Capital IQ provides the year the firm was established and also the year it was incorporated. If the former is missing, we use the latter if available. The Loughran-Ritter database provides founding date we use if the Capital IQ information is missing for the particular firm.

[11] Note that the cross-industry regression (29) estimates 26 industry-specific $\widehat{\beta}_{i,a}$ coefficients in addition to 5 coefficients on controls and 23 year fixed effects.

Table A2: Customer base drivers of firm-level growth

Dependent variable: $\Delta \ln \text{employment}_{j,t}$	(I)	(II)	(III)
$\Delta \ln \mu_{j,t-1}$	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
$\Delta \ln n^a_{j,t-1}$	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)
Controls:			
Industry-year fixed effects	✓	✓	✓
$X_{j,t}$	✓	✓	✓
Firm size (employment)		✓	✓
Firm age			✓
R^2	0.14	0.14	0.14
Observations	13,875	13,875	9,914

Note: The table shows estimates of regression (29) using Compustat firms reporting R&D expenditures between the years 1979 and 2019. As the outcome variable, we use employment growth rather than revenue growth. Markups, μ are estimated using the methodology in De Loecker et al. (2020) and selling and marketing expenditures are measured with the eponymous variable in Capital IQ. Firm size is given by log(employment) and age is measured as years since IPO. Standard errors (in brackets) are clustered at the 3-digit industry-year level.

regression coefficient between the median R&D to sales ratio in the industry and our estimates $\widehat{\beta}_{i,a}$ is 0.29 with standard errorr of 0.12 (i.e., p-value of 5%) and R^2 of 0.18.

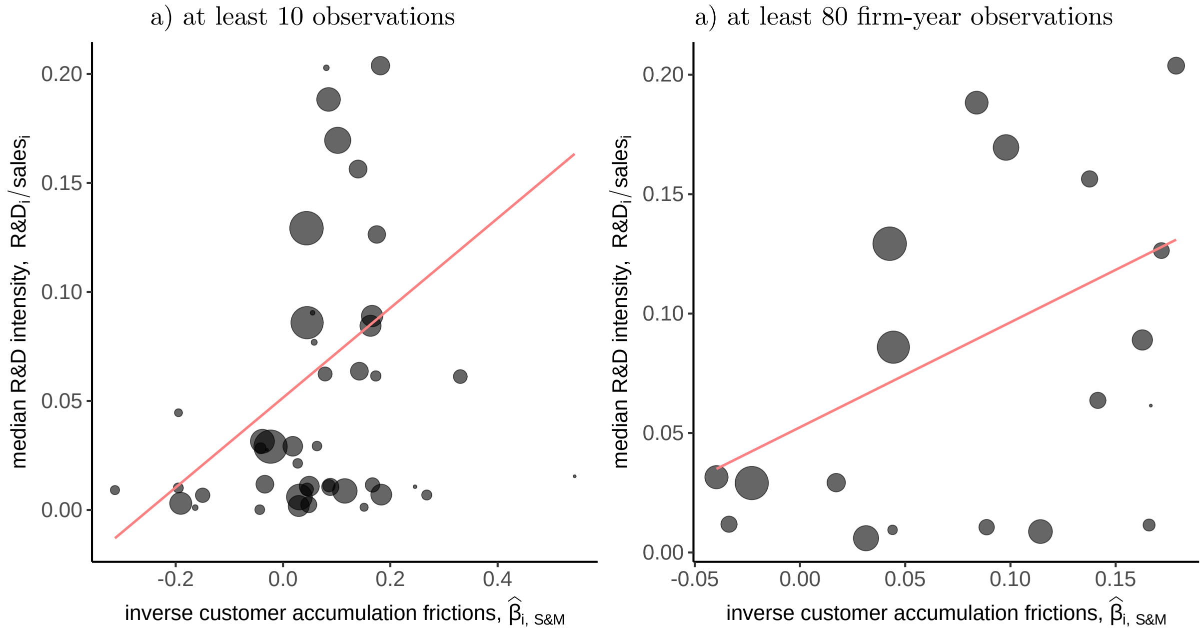
D.3 Robustness: Parameterization

A key component of our quantification is disciplining the relative importance of advertising- and sales-based customer accumulation, γ, and the cost elasticity of advertising expenditures, ψ_a. As explained in Propositions 5 and 6, we discipline these parameters by targeting the impact of markup and advertising growth have on firm-level growth (29).

Our preferred specification in the main text uses firm-level revenue growth. In this appendix, we present results based on firm-level employment growth as the dependent variable. According to our theory, Propositions 5 and 6 continue to apply because the customer base enters both employment ($n_c = bp^{-\eta}C/q$) and sales ($pc = bp^{1-\eta}C$) in the same log-linear way.

Table A2 shows that using employment growth instead of revenue growth as we do in the main text delivers qualitatively similar results. The sales and marketing expenses remain the dominant force driving firm growth. Quantitatively, the effects are lower to the extent that markups loose statistical significance. The results suggests that sales-based calibration provides an upper bound on the importance of markups for customer acquistion.

Figure A3: Customer accumulation frictions: sensitivity to sample selection



Note: Figure presents the correlation between $\widehat{\beta}_{i,a}$ estimated in sector-specific regressions (29) and industry-level aggregate R&D intensity $\frac{\text{XRD}}{\text{sales}}$. Each Panel corresponds to a different restriction on the minimum number of firm-year observation with a positive R&D in an industry to use this industry in the regression. In the baseline sample underlying Panel b) in Figure 4, the cutoff is 40 firm-year observations.

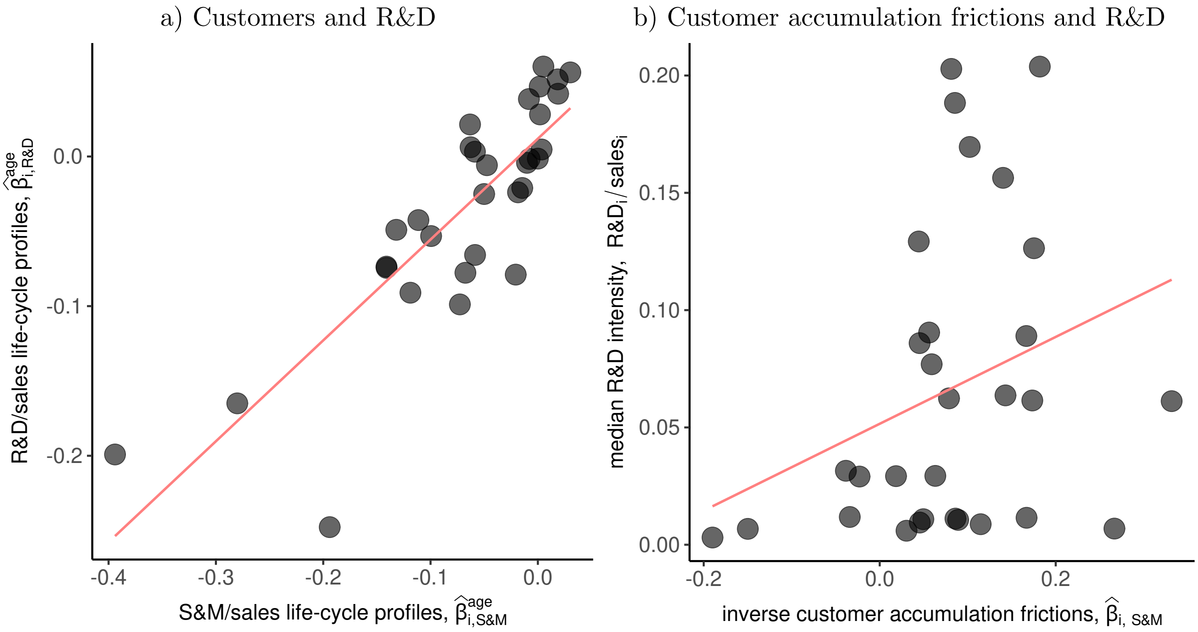
D.4 Robustness: Empirical Support for Model Predictions

In this appendix, we show robustness of our empirical results pertaining to the model mechanism. In particular, we consider three different alternatives: (i) different sample selection, (ii) using unweighted regressions and (iii) using employment growth instead of sales growth.

Robustness: Sample selection. In our baseline specification of cross-industry regression (29), we restrict attention to industries with at least 40 firm-year observations in regression (29). In this Appendix, we report results for two additional cutoffs: 10 observation and 80 observations. Figure A3 shows that the results are robust to these choices.

Robustness: unweighted regressions. In this appendix, we show that our validation results in Section 5 are robust to reporting unweighted regression results. Figure A4 shows the results. In Panel a), the regression coefficient rises to 0.67 with standard error of 0.08 and $R^2 = 0.70$. In panel b), the coefficient remains the same, 0.19 with std. error of 0.11 and $R^2 = 0.06$. This suggests that the relationship between customer accumulation and R&D predicted by the model is a robust feature of the Compustat data.

Figure A4: S&M and R&D, Unweighted Regressions



Note: Panel a) presents the correlation between $\widehat{\beta}^{age}_{i,a}$ and $\widehat{\beta}^{age}_{i,R\&D}$ estimated in sector-specific regressions (39). Panel b) presents the correlation between $\widehat{\beta}_{i,a}$ estimated in sector-specific regressions (29) and industry-level aggregate R&D intensity $\frac{\text{XRD}}{\text{sales}}$. The inverse of the estimated coefficient $\widehat{\beta}_{i,a}$ is proportional to the severity of customer accumulation frictions. In both figures, the red solid lines correspond to a fit of a sales-weighted linear regression model. Subscript i corresponds to a 3-digit NAICS industry. We restrict attention only to industries with at least 40 firm-year observations.

References

Abaluck, Jason and Abi Adams-Prassl (2021) "What do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses," *Quarterly Journal of Economics*, 136 (3), 1611–1663.

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William Kerr (2018) "Innovation, Reallocation and Growth," *American Economic Review*, 108 (11), 3450–3491.

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.

Akcigit, Ufuk and William Kerr (2018) "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 126 (4), 1374–1443.

Argente, David, Doireann Fitzgerald, Sara Moreira, and Anthony Priolo (2025) "How Do Entrants Build Market Share? The Role of Demand Frictions," *American Economic Review: Insights*, 7 (2), 160–76, 10.1257/aeri.20230289.

Arkolakis, Costas (2016) "A Unified Theory of Firm Selection and Growth," *Quarterly Journal of Economics*, 131 (1), 89–155.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen (2020) "The Fall of the Labor Share and the Rise of Superstar Firms*," *The Quarterly Journal of Economics*, 135 (2), 645–709, 10.1093/qje/qjaa004.

Baker, Scott, Brian Baugh, and Marco Sammon (2023) "Customer Churn and Intangible Capital," *Journal of Policital Economy: Macroeconomics*, 1 (3), 447–505.

Belo, Frederico, Xiaoji Lin, and Maria Ana Vitorino (2014) "Brand Capital and Firm Value," *Review of Economic Dynamics*, 17 (1), 150–169.

Bernard, Andrew, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes (2022) "The Origins of Firm Heterogeneity: A Production Network Approach," *Journal of Political Economy*, 130, 1765–1804.

Bloom, Nicholas, Rachel Griffith, and John Van Reenen (2002) "Do R&D Tax Credits Work?" *Journal of Public Economics*, 85 (1), 1–31.

Broda, Christian and David Weinstein (2006) "Globalization and the Gains form Variety," *Quarterly Journal of Economics*, 121 (2), 541–585.

Bronnenber, Bart, Jean-Pierre Dube, and Chad Syverson (2022) "Marketing Investment and Intangible Brand Capital," *Journal of Economic Perspectives*, 36 (3), 53–74.

Campbell, John Y. and Robert J. Shiller (1988) "Stock Prices, Earnings, and Expected Dividends," *The Journal of Finance*, 43 (3), 661–676.

Cavenaile, Laurent, Murat Alp Celik, Pau Roldan-Blanco, and Xu Tian (2021) "Style over Substance? Advertising, Innovation, and Endogenous Market Structure," mimeo.

Crouzet, Nicolas, Janice Eberly, Andrea Eisfeldt, and Dimitris Papanikolaou (2022) "The Economics of Intangible Capital," *Journal of Economic Perspectives*, 36 (3), 29–52.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger (2020) "The Rise of Market Power and the Macroeconomic Implications*," *The Quarterly Journal of Economics*, 135 (2), 561–644, 10.1093/qje/qjz041.

Dinlersoz, Emin and Mehmet Yorukoglu (2012) "Inofrmation and Industry Dynamics," *American Economic Review*, 102 (2), 884–913.

Drozd, Lukasz A and Jaromir B Nosal (2012) "Understanding international prices: Customers as capital," *American Economic Review*, 102 (1), 364–95.

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

Eslava, Marcela and John Haltiwanger (2021) "The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand and Wedges," NBER Working Paper, 27184.

Fernández-Villaverde, Jesús, Federico Mandelman, Yang Yu, and Francesco Zanetti (2021) "The "Matthew effect" and market concentration: Search complementarities and monopsony power," *Journal of Monetary Economics*, 121, 1177–1216.

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

Foster, Lucia, John C Haltiwanger, and Cornell John Krizan (2001) "Aggregate productivity growth: lessons from microeconomic evidence," in *New developments in productivity analysis*, 303–372: University of Chicago Press.

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, Francois (2014) "Financial Distress and Endogenous Uncertainty," mimeo, Federal Reserve Bank of Chicago.

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

Hall, Bronwyn and Rosemarie Ziedonis (2001) "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-1995," *RAND Journal of Economics*, 101–128.

Haltiwanger, John (2012) “Job Creation and Firm Dynamics in the U.S.,” *Innovation Policy and the Economy*, 17–38.

Haltiwanger, John, Ron Jarmin, Robert Kulick, and Javier Miranda (2016) “High Growth Young Firms: Contribution to Job, Output and Productivity Growth,” mimeo.

Haltiwanger, John, Ron Jarmin, and Javier Miranda (2013) “Who Creates Jobs? Small Versus Large Versus Young,” *Review of Economics and Statistics*, 95 (2), 347–361.

He, Bianca, Lauren Mostrom, and Amir Sufi (2025) “Investing in Customer Capital,” NBER Working Paper 33171.

Hopenhayn, Hugo A. and Richard Rogerson (1993) “Job Turnover and Policy Evaluation: A General Equilibrium Analysis,” *Journal of Political Economy*, 101 (5), 915–938, 10.1086/261909.

Hottman, Collin, Stephen Redding, and David Weinstein (2016) “Quantifying the Sources of Establishment Heterogeneity,” *Quarterly Journal of Economics*, 131 (3), 1291–1364.

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

Jones, Chad and Jihee Kim (2018) “A Schumpeterian Model of Top Income Inequality,” *Journal of Political Economy*, 126 (5), 1785–1826.

Jovanovic, Boyan (1982) “Selection and the Evolution of Industry,” *Econometrica*, 50 (3), 649–670.

Kaas, Leo and Bihemo Kimasa (2021) “Firm Dynamics with Frictional Product and Labor Markets,” *International Economic Review*, 62 (3), 1281–1317.

Kaas, Leo and Phillip Kircher (2015) “Efficient Firm Dynamics in a Frictional Labor Market,” *American Economic Review*, 105 (10), 3030–3060.

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

Larkin, Yelena (2013) “Brand Perception, Cash Flow Stability, and Financial Policy,” *Journal of Financial Economics*, 110 (1), 232–253.

Lentz, Rasmus and Dale T Mortensen (2008) “An empirical model of growth through product innovation,” *Econometrica*, 76 (6), 1317–1373, Publisher: Wiley Online Library.

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.

Melitz, Marc (2003) “The Impact of International Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 6 (71), 1695–1725.

Nakamura, Emi and Jón Steinsson (2011) “Price setting in forward-looking customer markets,” *Journal of Monetary Economics*, 58 (3), 220–233.

Ottonello, Pablo and Thomas Winberry (2020) "Financial Heterogeneity and the Investment Channel of Monetary Policy," *Econometrica*, 88 (6), 2473–2502.

Perla, Jesse (2019) "A Model of Product Awareness and Industry Life Cycles," mimeo.

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

Petrosky-Nadeau, Nicolas and Etienne Wasmer (2015) "Macroeconomic dynamics in a model of goods, labor, and credit market frictions," *Journal of Monetary Economics*, 97–113.

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

Rachel, Lukasz (2021) "Leisure-Enhancing Technological Change," mimeo.

Ravn, Morten, Stephanie Schmidt-Grohe, and Martin Uribe (2006) "Deep Habits," *Review of Economic Studies*, 73, 195–218.

de Ridder, Maarten, Basile Grassi, and Giovanni Morzenti (2024) "The Hitchhiker's Guide to Markup Estimation: Assessing Estimates from Financial Data," mimeo.

Rudanko, Leena (2022) "Price Setting with Customer Capital: Sales, Teasers and Rigidity."

Sedláček, Petr and Vincent Sterk (2017) "The Growth Potential of Startups over the Business Cycle," *American Economic Review*, 107 (10), 3182–3210.

Shi, Shouyong (2016) "Customer relationships and sales," *Journal of Economic Theory*, 166 (C), 483–516.

Spencer, Adam (2022) "Policyt Effects of International Taxation on Firm Dynamics and Capital Structure," *Review of Economic Studies*, 89 (4), 2149–2200.