STAFF
REPORTS

NO. 1159
JULY 2025

Market Concentration and Aggregate Productivity: The Role of Demand

Jeremy Pearce | Liangjie Wu

FEDERAL RESERVE BANK of NEW YORK

Market Concentration and Aggregate Productivity: The Role of Demand

Jeremy Pearce and Liangjie Wu

Federal Reserve Bank of New York Staff Reports, no. 1159
July 2025
https://doi.org/10.59576/sr.1159

Abstract

This paper studies the relationship between market concentration and aggregate productivity when firm-level demand emerges from past marketing investments. Granular firms may invest in demand both to complement their productivity and to amplify market power—this second force can create persistent mismatch between customer capital and productivity. The importance of this mismatch depends on the relative persistence of productivity and demand. Empirically, we find that demand is more persistent than productivity, implying a sizable role for mismatch. This leads to sluggish demand-side adjustment in the face of productivity shocks in the quantified model. Policies targeting static markup distortions—such as production subsidies—can exacerbate excessive marketing and thus are subject to a tradeoff between static gains and dynamic losses.

JEL classification: O31, O32, O34, O41, D22, D43, L11, L13, L22
Key words: firm dynamics, productivity, demand, customer capital, market concentration, competition, innovation

---

Pearce: Federal Reserve Bank of New York (email: jeremy.pearce@ny.frb.org). Wu: Einaudi Institute for Economics and Finance (email: liangjie.wu@eief.it). The authors thank Tincho Almuzara, Mary Amiti, and Simone Lenzu for helpful comments.

1 Introduction

How does market concentration affect aggregate productivity? This question has received lively discussions in policy circles and recent economic literature (De Loecker et al., 2020; Olmstead-Rumsey, 2022; Akcigit and Ates, 2023; De Ridder, 2024). However, the role of demand-side characteristics of firms, which is a central driver in the firm-size distribution (Hottman et al., 2016; Foster et al., 2016; Einav et al., 2021), has been largely ignored in understanding this concentration-productivity relationship. This paper introduces endogenous demand or customer capital investment[1] (Gourio et al., 2014; Afrouzi et al., 2023) into a model of market concentration, where concentration comes from both a skewed firm size dispersion and a finite number of *granular* firms. We focus on the *dynamics* of demand to understand how it interacts with productivity at large firms. With this framework, we address three questions: Are investments in demand that drive concentration beneficial or detrimental to aggregate productivity? How does demand, as slow-moving capital, affect the response of the aggregate to changes in firm productivity? How do policies designed to undo distortions from concentration affect the dynamics of productivity and welfare?

This paper addresses these questions by making the following contributions. Theoretically, we introduce a dynamic customer capital investment decision in a model of granular firms with heterogeneous productivity and provide a computational tool to analyze equilibrium in such an environment. We match the data by modeling concentrated markets, e.g., granular firms, with dynamic demand and productivity. Granularity has two implications: (i) firms behave strategically, manifesting in variable markups (Atkeson and Burstein, 2008) and strategic interactions in marketing, leading to endogenous path dependency; (ii) firm-level productivity shocks transmit into aggregate shocks, where the transmission depends on the endogenous demand characteristics. Empirically, guided by the model, we decompose the firm size distribution into demand characteristics, cost productivity, and markups. In the data, demand is more persistent than productivity. Calibrated to the empirical regularities, we find that endogenous customer capital investment improves efficiency, while the strategic interactions on their own decrease efficiency. This endogenous investment more than doubles the time it takes for initial advantages in customer capital to vanish. Policy-wise, there is a tradeoff between

[1] In this paper, we use residual demand and customer capital interchangeably.

correcting static markup distortions and exacerbating excessive marketing and dynamic misallocation.

We start by building a theory that connects market share at the firm level to the dynamic interaction of consumer demand and productivity. The model takes an oligopolistic competition framework into a dynamic setting, where granular firms endogenously invest in customer capital to build out demand. Intuitively, the investment in customer capital can come from firms wanting to complement their productivity or their market power. The productivity of firms fluctuates exogenously, and there is exogenous decay in existing customer capital; firms must continuously invest in customer capital to maintain or expand their market share. The core differentiating feature of customer capital lies in its allocative role: customers' limited attention means that all else equal, marketing only redistributes attention, and thus marketing investment only increases aggregate productivity when it's matched to firms with higher productivity. This is in line with the *informative marketing* view. We provide a micro-foundation of such a model of marketing based on rational inattention.

This model is challenging to analyze, as firms are granular; firm-level shocks are aggregate shocks. A computational contribution is to provide a tractable algorithm to compute this model. We do a quadratic approximation of the equilibrium profits of firms around the points of equal customer capital while preserving the nonlinearity in the productivity space. This linear-quadratic approximation transforms the intractable dynamic game among firms into a linear-quadratic dynamic game. We then utilize well-established results from game theory to characterize the equilibrium dynamics.

In the equilibrium, the evolution of customer capital follows a linear system of differential equations. A novel endogenous path dependency arises from the strategic incentives of firms. Because markups are increasing in market shares, larger firms (regardless of whether they are productive or have significant customer capital) have a stronger incentive to invest in customer capital, which we refer to as the *size incentive*. Further, firms attempting to hold onto their market shares invest more in advertising due to an *escape competition* effect. The initial condition of customer capital thus has a long-lasting impact on its evolution.

Empirically, we establish the importance of demand growth and decay in driving firm market share dispersion using detailed scanner data. Our model lends itself to a decomposition of market shares into demand, productivity, and markups. We follow Hottman et al. (2016) by decomposing the observed firm size according to a nested constant elasticity of substitution demand system. Our measure of firm-level demand is the *residual* demand, which is the sales in excess of what is predicted by prices adjusted by the demand elasticity. By decomposing firm market shares into productivity, markup, and residual demand components, we find that residual demand explains approximately four times as much as productivity differences in the cross-sectional variation and growth in firm size. Our variance decomposition reveals that residual demand accounts for the majority (>90%) of market share variation, while productivity accounts for less, while remaining important ($\sim$20%).

We show that firms with higher residual demand experience a larger decay of demand from their existing products, with an AR(1) coefficient of 0.87. This speed of decay is slow when compared with the decay of measured cost productivity, with an AR(1) coefficient of 0.78. The relative persistence of these two components is the core empirical fact we bring to the model, as it governs the relative importance of firms' strategic incentives. Intuitively, when demand is more persistent than productivity, firms' current marketing investments depend more on their accumulated customer capital from the past than on the forward-looking productivity, leading to a stronger strategic incentive, and vice versa.

We quantify the welfare impact of demand on the concentration-productivity relationship. This impact is theoretically ambiguous due to two forces. First, a complementarity between productivity and customer capital encourages efficient investments in customer capital, the *productivity incentive*. Second, the *strategic incentive* from variable markups and escape competition allows less productive firms with high customer capital to persist with a large size. We quantify the relative strength of these two incentives by matching the empirical dynamics of firm-level productivity, the decay of customer capital, and the substitutability of products among firms. The endogenous decisions of firms are shaped by two core parameters: the decay rate of existing customer capital and the cost of marketing. We match the literature on marketing to revenue and use the decay of firms' existing customer capital, in proportion to their level, to calibrate these parameters. Although sparsely parametrized, our quantitative model can replicate key untargeted moments regarding concentration in the data and the persistence of the core state variables of the firm.

We highlight two quantitative findings from our model. First, granularity introduces a strong endogenous path dependency. The median half-life convergence in the calibrated model is more than 12 years, meaning it takes at least 12 years to close the gap between the initial gap of customer capital among firms to its long-run level by half, holding other conditions constant. The half-life implied by the exogenous decaying rate is 5 years, less than half of the equilibrium level and less than in models without customer capital as a state variable. Second, this endogenous path dependency leads to a mismatch between productivity and customer capital. Without size incentives, the predicted correlation between productivity and demand is 15% larger.

The path dependency of the model also affects the transmission from firm-level productivity to aggregate productivity. If a firm with high customer capital gets a positive productivity shock, its customer base is *even higher* after 10 years than an equivalent firm with low customer capital. This leads to the aggregate effect of a one standard deviation productivity shock to be 7% higher after 10 years. This dependence on initial conditions has important implications for aggregate dynamics and welfare, as there is a high dispersion in half-lives by initial states that matter for aggregate productivity.

We then combine these dynamic features of productivity with the markups to quantify the impacts on welfare. The aggregate welfare depends on both the static distortions due to markups and the dynamic matching between productivity and customer capital. We show two sets of welfare results. By comparing the calibrated equilibrium to the case where the marketing activity is eliminated, we show that endogenous marketing is welfare-enhancing; shutting down marketing leads to a welfare loss of 7%. Thus, these demand characteristics have a positive impact on welfare through complementing productivity. Although marketing exacerbates the static loss due to markups by increasing concentration, it does increase aggregate productivity. Underneath such an overall gain, there is indeed a welfare loss due to the size incentives and endogenous path dependency of 4%.

The interaction of static markups and dynamic misallocation has novel policy implications. With a social planner's solution, we show that an optimal policy that maximizes the discounted welfare should aim to correct static markups and resolve the crowding-out among firms regarding customer capital. Doing so can bring significant welfare gains from the equilibrium. A quantity subsidy, which resolves markup distortions and dynamic distortions when firms only differ in a single-dimensional characteristic (such as in Edmond et al., 2023), can exacerbate the dynamic misallocation.

The remainder of this section reviews the literature, while the rest of the paper is structured as follows. Section 2 introduces the theoretical model with oligopolistic firms with heterogeneous productivity and customer capital. Section 3 introduces the empirical framework for our study, the decomposition of demand and productivity, and the empirical dynamics of both. Section 4 estimates the model by uniting theory and empirics and studying the positive implications of our quantified model. Section 5 discusses the nature of productivity shocks and policy implications on how the quantified model changes our understanding of size-based policies, marketing, and antitrust. We start with a review of the literature.

Related Literature. There has been growing interest in the importance of demand in driving market share amongst economists (He et al., 2024). Some contributions to this topic view demand characteristics as a slow-moving object that depends on the investment of firms (Dinlersoz and Yorukoglu, 2012; Gourio and Rudanko, 2014; Ignaszak and Sedlácek, 2022; Cavenaile et al., 2023; Greenwood et al., 2025), abstracting away from strategic interactions among large firms. Other recent papers address endogenous demand in a strategic setting, but view demand as a static choice that responds to firms' productivity (Cavenaile and Roldan-Blanco, 2021; Cavenaile et al., 2025). Nord (2023) studies how the shopping efforts of heterogeneous customers drive the cyclicality of markups. Afrouzi et al. (2023) study the interaction between variable markups and customer capital, showing that markups are not necessarily increasing in firm size. Bornstein and Peter (2022) study the dynamic misallocation of customers due to nonlinear pricing decisions. The novelty of our paper is in connecting the dynamic feature of demand and its interaction with the strategic incentives of oligopolistic firms, and quantifying the novel mismatch resulting from such an environment: unproductive firms can invest more in their customer capital due to oligopolistic incentives.

There is a well-established literature on the origins of economic growth, heterogeneous firm productivity, and the transmission from firm-level changes to the aggregate economy (e.g., Jovanovic, 1982, Hopenhayn, 1992, Aghion and Howitt, 1992, Klette and Kortum, 2004, Akcigit and Kerr, 2018, Baqaee and Farhi, 2019). We build on this literature by connecting firm-level productivity to changes in aggregate productivity through heterogeneity in the firm's demand side. We focus on how the heterogeneity can drive persistence in leadership and misallocation, which builds on a rich literature on factor misallocation and reallocation (Aghion et al., 2001, Haltiwanger et al., 2014, Acemoglu et al., 2018, Peters, 2020, and Liu et al., 2022). There is a rising interest in how growth and firm heterogeneity interact with market power. Akcigit and Ates (2021, 2023) focus on the knowledge diffusion gaps between leaders and followers, driving rising concentration and falling business dynamism. Large firms may also leverage their market share to increase markups; we connect this to a new arena of dynamic decisions in the firm's demand environment, also with particular attention to market leaders and cases where firms are large. These cases relate to a literature on granularity, and the implications for the aggregate from firm-level shocks, an area of rising interest (Gabaix, 2011; Carvalho and Grassi, 2019)

Theoretically, we focus on oligopolistic competition among firms. We extend the insight from Atkeson and Burstein (2008) that variable markups may generate misallocation into a dynamic setting. This builds on a host of papers that focus on static misallocation and markups (Boar and Midrigan, 2019; Edmond et al., 2023; Mongey, 2021; Berger et al., 2019), which we also connect to the interaction between productivity and demand. The core mechanism is the role of strategic complementarities, which Amiti et al. (2019) find to be a strong force in international markets that drive variable markups. This setting connects to an extensive literature on markups and the negative and positive effects of concentration (such as De Loecker and Eeckhout, 2018; De Loecker et al., 2020; Gutiérrez and Philippon, 2017; Eggertsson et al., 2018; Hall, 2018; Autor et al., 2020; Kehrig and Vincent, 2021; Covarrubias et al., 2019; Akcigit and Ates, 2021). The role of demand-side factors in this relationship has novel implications, as most existing theories of concentration relate to productivity advantages of large firms. This insight connects to a rising interest in thinking about firm size coming from multiple dimensions (e.g., as in Salgado et al., 2024).

This paper aims to ground insights on the role of demand heterogeneity and investment in empirical frameworks in macroeconomics, trade, and industrial organization. Hottman et al. (2016) study multi-product firms and find that the "appeal" of firms, or residual demand, explains the largest share of sales variation across firms. More recently, Eslava et al. (2024) found the same result when evaluating drivers of firm plant share. This demand takes time to build, as Argente et al. (2018, 2020) and Jaravel (2018) explore how product creation and destruction are ubiquitous in product markets. Argente et al. (2021) and Einav et al. (2021) document that product sales expansion is primarily due to an expanding customer base. Bhandari and McGrattan (2020) focus on this customer base or "sweat equity" as an endogenous asset. These channels can interact and may affect the measurement of productivity (Foster et al., 2008; Syverson, 2011) and misallocation (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2017; David and Venkateswaran, 2019). We bring these studies into a dynamic macroeconomic setting to study the implications for overall productivity empirically and theoretically.

The importance of customer capital, marketing, and branding connects to a related literature on intangible assets. Firm productivity and demand are two central intangible assets to firm success and will serve a central role as the importance of intangibles continues to rise (Haskel and Westlake, 2017, Crouzet and Eberly, 2019, Olmstead-Rumsey, 2022, Syverson, 2019, De Ridder, 2024). Demand that firms hold as a state variable is an important intangible asset. Bain (1956) classically noted that "(t)he advantage to established sellers accruing from buyer preferences for their products as opposed to potential-entrant products is on the average larger and more frequent in occurrence at large values than any other barrier to entry." Theoretically, brand value can generate persistent profits in markets with imperfect information (Schmalensee, 1978, Schmalensee, 1982, and Shapiro, 1983). The power of branding has been detailed empirically as consumer brand preferences are quite persistent (e.g., in Bronnenberg et al., 2009, 2012). Firms also exchange this intangible asset, and it has persistence in value across firms (Pearce and Wu, 2022), making it an intangible asset at the center of firm value. This paper integrates it with the intangible of productivity empirically, theoretically, and quantitatively.

2 Model

We develop a dynamic model where oligopolistic firms compete with exogenous productivity and endogenous customer capital. We start by presenting the environment with one single market and later introduce other general equilibrium elements and aggregation. In this single market, a finite number of firms engage in Cournot competition and internalize their impact on the market, as in Atkeson and Burstein (2008). The main novelty of this paper is in the dynamics. Dynamically, firms invest in customer capital for both productive and strategic reasons, with the strategic element coming from variable markups due to firm size. This mechanism leads to path dependence and can amplify misallocation. The resulting disconnect between market share and productivity can have significant implications for aggregate productivity and the effectiveness of policies, which we will study quantitatively in Section 5.

Section 2.1 starts by describing the environment of households and firms. Section 2.2 characterizes the static and dynamic decisions of the firm and the approximation solution. This will connect marketing decisions to the distribution of productivity and customer capital. Section 2.3 then examines the path dependency of demand with an illustrative example. Finally, Section 2.4 discusses aggregation and extensions that enable us to ground these insights into an empirical framework that we unite in this paper.

2.1 Environment

We consider a continuous-time model. We introduce the representative household's problem and its implications for the firms' competitive environment, and then discuss the dynamics of investment in customer capital.

Households There is a representative household. The representative household maximizes utility over an infinite horizon, balancing consumption utility against disutility from endogenously supplied labor L_t. Specifically, the household solves

$$\max_{c_{it}, L_t} \int_0^\infty e^{-rt} \left(\ln \frac{C_t}{B_t} - L_t \right) dt, \tag{1}$$

subject to the flow budget constraint on total assets, $\dot{W}_t = L_t + \tilde{r}_t W_t + \Pi_t - \sum_{i=1}^{I} p_{it} c_{it}$. In the budget constraint, the household receives its labor income L_t (where we normalize the wage to be 1 throughout this paper) and financial income. The household can save in a representative portfolio of all firms, which delivers an interest rate $\tilde{r}_t$ and reimburses all profits, Π_t, to the household.[2]

In the flow utility function, B_t represents the total customer capital in the economy aggregated across firms. Consumption utility, C_t, features a constant-elasticity-substitution

[2]In the general stationary equilibrium we consider in the quantitative analysis, $\tilde{r}_t = r$. In the one-sector environment, we set the interest rate exogenously at $\tilde{r}_t = r$. As the aggregate consumption impact due to firm shocks is not the focus of our paper, this assumption is unimportant for the following discussion.

(CES) across I firms with

$$C_t = \left[\sum_{i=1}^{I} e^{b_{it}/\sigma} c_{it}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \qquad B_t = \left(\sum_{i=1}^{I} e^{b_{it}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{2}$$

$\sigma > 1$ represents the elasticity of substitution between firms. b_{it} is the customer capital of each firm that aggregates to total customer capital in B_t. The customer capital can be interpreted as the representative household's prior tendency to consume products from a firm, for example, the image of the products or the shelf availability of products. In Appendix A.4, we provide a microfoundation based on rational inattention in consumption choices, where b_{it} can be interpreted as the household's prior tendency to purchase a product before observing prices in the store.

Three results follow from our set-up on the household side. First, there is a demand curve for products of firm i: $c_{it} = e^{b_{it}} \left(\frac{p_{it}}{P_t} \right)^{-\sigma} C_t$, with the aggregate price index defined as $P_t = \left(\sum_i e^{b_{it}} p_{it}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$. Second, the equilibrium aggregate expenditure must be 1, $P_t C_t = 1$. This comes from the optimal consumption-saving choice and the optimal labor supply decision of the household. Third, due to the aggregation B_t, customer capital does not bring utility to the household on its own, but its distribution across firms matters.

Firms. The I firms differ in both their productivity and customer capital. We assume firm i produces using a production function $c_{it} = \exp\left(\frac{a_{it}}{\sigma-1} \right) l_{it}$, where l_{it} is the labor input and a_{it} measures firm i's labor productivity.[3]

We assume the productivity of firm i redraws with a Poisson arrival rate of λ. When a redraw event arrives, its productivity is drawn from a discrete distribution $\tilde{F}(a)$. We assume the distribution $\tilde{F}$ has support of a finite collection of values $supp(\tilde{F}) = \{A_1, ..., A_N\}$. Because firms are granular, the entire vector of productivity at time t becomes the relevant state variable in the dynamic problems. This productivity vector takes values in $\{A_1, ..., A_N\}^I$. We index these vectors by α, the *productivity state*. Correspondingly, we denote the conditional probability of α' becoming the productivity state after a redraw given the current state α as $F_{\alpha,\alpha'}$.

[3] For expositional simplicity, we scaled the labor productivity into the unit of market shares. This scaling will be treated explicitly in our quantitative analysis.

The customer capital of firm i evolves both endogenously and exogenously. Firm i can grow its customer capital with marketing investment η_{it}. Exogenously, the customer capital depreciates, which we model as the mean-reversion rate $\rho > 0$. We emphasize that in the primitives, our model does not impose any dependence of customer capital on productivity. Instead, we show that the dependence arises endogenously through firms' optimization. Mathematically, the law of motion for b_{jt} is

$$\dot{b}_{it} = \eta_{it} - \rho b_{it},\, b_{i0} = 0. \tag{3}$$

Marketing investment is costly. We model this as a quadratic labor cost in advertising $\kappa \frac{\eta_{it}^2}{2}$.

Granularity. There are two consequences of granular firms in our model. First, their shocks are aggregate shocks, with consequences on the household's consumption and a direct impact on other firms' outcomes. Second, firms are large enough to internalize their impact on the market. Recent evidence suggests this is particularly salient for large firms in product markets (Amiti et al., 2019). The natural assumption for competition is that firms engage in oligopolistic competition. We assume they compute by choosing quantity, the Cournot competition, statically.

2.2 Characterization

Firms make static pricing decisions and benefit from more productivity and customer capital due to two forces. First, more composite productivity leads to higher market share and higher profits. Second, due to variable markups, firms can extract more profit per unit when the market share is higher. In the dynamics, this informs their marketing decisions and the value of their joint customer capital and productivity. We first characterize the static pricing equilibrium and then analyze how these static outcomes affect firms' dynamic decisions. Characterizing the dynamic equilibrium involves solving a high-dimensional dynamic game. Our solution to this complexity is to approximate the profits of firms to a second order. With this approximation, the dynamic game becomes a linear-quadratic one. The tractability of the linear-quadratic games helps us to derive a linear law of motion for the evolution of customer capital.

First, we derive the outcomes from the static pricing equilibrium and write the equilibrium market shares and profits as functions of the distribution of productivity and customer capital.

Static Pricing Equilibrium. As firms engage in oligopolistic quantity competition, their equilibrium markups are given by the formula: $p_{it} = \mu_{it} \exp\left(\frac{a_{it}}{\sigma-1}\right)$, where the markup $\mu_{it} = \frac{\sigma}{\sigma-1}\frac{1}{1-s_{it}}$, where s_{it} is the equilibrium market share for the firm. Given the market shares, the firm i's profit π_{it} is given by:

$$\pi_{it} = \frac{1}{\sigma}s_{it} + \frac{\sigma-1}{\sigma}s_{it}^2.$$

Coupled with the demand curves from the CES preference, the vector of market shares $(s_{it})_{i=1}^{I_{k(i)}}$ is the solution to:

$$s_{it} = \frac{\exp(a_{it}+b_{it})(1-s_{it})^{\sigma-1}}{\sum_{j=1}^{I}\exp(a_{jt}+b_{jt})(1-s_{jt})^{\sigma-1}}. \tag{4}$$

There are a few observations here. First, we note that the system is homogeneous of degree zero in $(a_{it}, b_{it})_{i=1}^{I}$. This homogeneity also applies to profits. Economically, only the *relative* productivity and customer capital matter for profits. Further, whether market shares come from demand-side or supply-side factors of firms does not matter for profits. However, it will matter for firms' investment decisions, which we turn to next.

Dynamic Marketing Equilibrium. We focus on firms of sufficient size that significantly impact the overall product group in which they operate. In this case, the equilibrium concept we focus on is the Markov Perfect Equilibrium, where firms' strategies only depend on the payoff-relevant state. As discussed in the static pricing equilibrium, the static profits only depend on a composite of α and $\mathbf{b}$. Dynamically, due to the dynamics of productivity and customer capital, the full payoff relevant state of firms is a function of the productivity state α and the customer capital $\mathbf{b}$.

To prepare the notation for the equilibrium, we denote $\eta_i(\alpha, \mathbf{b})$ the marketing strategy of the firm i. In a Markov perfect equilibrium, firm i takes as given the strategies of other firms and maximizes:

$$V_i(\alpha, \mathbf{b}) = \max_{\eta_i \geq 0} \mathbb{E} \int_0^\infty e^{-rt} \left(\pi_i(\alpha_t, \mathbf{b}_t) - \frac{\kappa}{2} \eta_i(\alpha_t, \mathbf{b}_t)^2 \right) dt \tag{5}$$

s.t.

$$\dot{b}_{jt} = \eta(\alpha_t, \mathbf{b}_t) - \rho b_{jt}, \quad \forall j, t$$

There are three noteworthy features of the equation (5). First, we impose that all firms follow a Markov strategy: marketing investment $\eta_i(\alpha_t, \mathbf{b}_t)$ is only a function of the current payoff-relevant state $(\alpha_t, \mathbf{b}_t)$ instead of the whole history. Second, the firm forms expectations regarding both its productivity process and that of its competitors, as described above. Lastly, the firm also internalizes the law of motion of customer capital for all firms in the group. This captures the strategic interactions across firms. Using the definition of firms' problems, we now formally define the marketing equilibrium.

Definition 1 *A marketing equilibrium is marketing investment strategy $(\eta_i(\alpha_t, \mathbf{b}_t))_{i=1}^I$, where the $\eta_i(\alpha_t, \mathbf{b}_t)$ solves equation (5) for each firm i, given the other firms' strategies $(\eta_j(\alpha_t, \mathbf{b}_t))_{j \neq i}$.*

Fully analyzing a marketing equilibrium is a challenging task for two reasons. First, due to the granular features of the firms, each individual firm's productivity shock becomes an aggregate shock. The relevant state vector thus becomes the entire I-dimensional vector. In our empirical analysis, which involves a representative group of 10 firms with sizable market shares, it becomes impossible to solve this problem directly. Second, even without the dimensionality issue, characterizing the equilibrium involves finding a fixed point of firms' strategies in the functional space, which further complicates the analysis. To highlight the economics from such an environment, we start with a myopic limit $r \to \infty$, where firms only maximize the static profits while the customer capitals evolve dynamically. After understanding the economic interactions from the myopic limit, we then turn to a linear-quadratic approximation solution that is tractable for quantitative analysis.

Myopic Limit. To understand the implications of oligopolistic competition on marketing investments, we assume $r \to \infty$. The firm's problem becomes maximizing the static profits. Solving the optimal marketing investments, we find

$$\eta_{it} = \underbrace{\frac{1 + 2(\sigma - 1)s_{it}}{\sigma}}_{\text{profti-share response}} \times \underbrace{\zeta_{it}\left(1 - \frac{\zeta_i}{\sum_j \zeta_j}\right)}_{\text{share-customer response}}$$

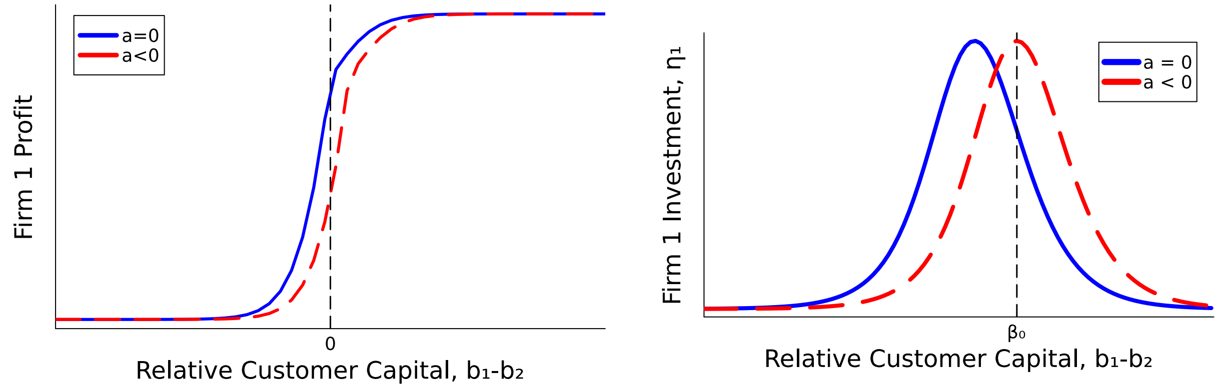
where $\zeta_j = \left(\frac{1}{\bar{s}_j} + (\sigma - 1)\frac{1}{1-\bar{s}_j}\right)^{-1}$ for all $j = 1, .., I$. From the CES demand system and the result that markups only depend on market shares, the marginal profit from an additional unit of customer capital can also be expressed as a function of the firm's current market share s_{it}. We can understand this marginal profit in two steps.

First, how does an increase in customer capital affect market share? The oligopolistic nature of our environment implies that an intermediate-sized firm has a larger increase in market share. To see this, the share-customer response term approaches 0 when the current market share of the focal firm approaches 0 or 1. When the firm is small relative to its customer, an additional unit of customer capital maps to a small increase in its market share, the *discouragement effect*. When the firm is large relative to its customer, the additional customer capital maps to mostly its own share, the *cannibalization effect*. All together, firms invest more when they are similar in size to competitors, to *escape competition* (Aghion et al., 2005). The escape competition effect means a firm's market investment can either increase or decrease when it experiences a negative productivity shock. When productivity falls for a small firm, its investment falls due to discouragement. When productivity falls for a large firm, it becomes closer to its competitors, leading to a larger marketing investment.

These two effects can be seen in Figure 1. In Figure 1a, we see the sigmoid of the profit function, which has implications for the incentives of firms. The incentives are quite different depending on the existing customer base. When a firm gets a negative productivity shock, it shifts where the profit function is steepest. This shapes the investment decisions of firms, which can be seen in Figure 1b.

There are two main messages from the Figure 1. First, the incentives to invest in customer capital vary significantly depending on where firms are in both productivity and existing customer capital. A firm could get a negative productivity shock (e.g., red dotted line) and invest *more* in customer capital. Second, granular firms internalize that an increase in their market share leads to both expansion in scale of production and an

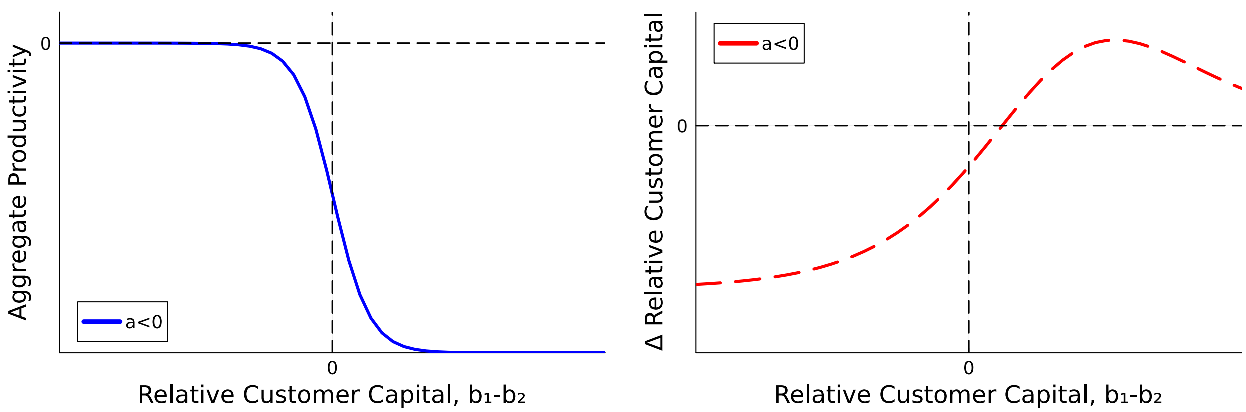
Figure 1: The Incentives for Customer Capital Investment



(a) Profit Function

(b) Firm Investment Intensity

Figure 2: Aggregate Productivity and Customer Capital



(a) Aggregate Productivity

(b) Customer Capital and Marketing

increase in markups. This means a larger firm has a stronger incentive to invest in its customer capital, regardless of its productivity.

Figure 2 presents the aggregate implications of these firm outcomes. Figure 2a plots the aggregate productivity of the economy depending on the customer capital of firm 1 against firm 2 when firm 1 is less productive than firm 2. Figure 2b plots the investment decision of firm 1 when firm 1 is less productive than firm 2. Both plots have the relative customer capital on the x-axis.

The two plots showcase the divergence in the most productive solution from the market solution depending on initial conditions. As seen in Figure 2a, the optimal solution from a productivity perspective is to allocate all customer capital to firm 2, and thus to have firm 2 only invest in marketing. Figure 2b illustrates that the firm's decisions depend crucially on initial conditions. If firm 2 has more initial customer capital than firm 1 (left quadrants), they will also invest relatively more, and the market will move to the productive outcome. If firm 1 has significantly higher initial customer capital, firm 1 invests more, and the market converges to firm 1 maintaining full market share. The intuition behind this comes from the increasing returns to size that come from firms being larger with variable markups.

The illustrations above show how suboptimal equilibria can be sustained in a world with only two myopic firms and no exogenous decay rate. Once there is an exogenous decay rate $\rho > 0$, the stability of the problem is restored. However, the intuition behind the strategic incentives in marketing that generate persistent leadership remains in the transition. We now turn to a richer framework with many firms, finite r, and decay rate $\rho > 0$.

Linear-quadratic Approximation. The oligopolistic game of many firms has naturally rich interactions that extend from the two-firm case. To gain tractability in our dynamic game, we perform a second-order Taylor expansion of firm i's profit function, $\pi_i(\alpha, \mathbf{b})$ near the point $(\alpha, \mathbf{0})$, for every i. The interpretation of such an approximation is that we capture the nonlinearity of profits in the productivity space, while keeping the nonlinearity of the profit in the customer capital to the second order. The approximated market shares can be written as,

$$\pi(s_i(\alpha, \mathbf{b})) \approx \pi(s_i(\alpha, \mathbf{0})) + \left(\frac{1}{\sigma} + 2\frac{\sigma-1}{\sigma}\bar{s}_i\right) \sum_j \frac{\partial s_i}{\partial b_j} b_j \tag{6}$$

$$+ \frac{1}{2} \sum_j \sum_\ell \left[\left(\frac{1}{\sigma} + 2\frac{\sigma-1}{\sigma}\bar{s}_i\right) \frac{\partial^2 s_i}{\partial b_j \partial b_\ell} + 2\frac{\sigma-1}{\sigma} \frac{\partial s_i}{\partial b_j} \frac{\partial s_i}{\partial b_\ell} \right] b_j b_\ell. \tag{7}$$

We understand the equation 6 from the linear to the nonlinear effects. First, focus on the linear effects. How much do changes in customer capital map into profits depends on how the market share of the focal firm i responds to customer capital and how the profit of the focal firm i responds to the change in market share. Thanks to the CES structure, the response of market share s_i to customer capital can be derived in closed form:

$$\frac{\partial s_i}{\partial b_i} = \zeta_i \left(1 - \frac{\zeta_i}{\sum_j \zeta_j}\right), \text{ and } \frac{\partial s_i}{\partial b_j} = -\frac{\zeta_i \zeta_j}{\sum_j \zeta_j},\ j \neq i,$$

where $\zeta_j = \left(\frac{1}{\bar{s}_j} + (\sigma - 1)\frac{1}{1-\bar{s}_j}\right)^{-1}$. The zero-demand market shares reflect the productivity differentials among firms. Since we approximate the profits by each $(\alpha, \mathbf{0})$, our approximation preserves the nonlinearity in the productivity dimension. Because firms are large enough to internalize impacts on the market, productivity maps differently into the marginal profits from extra customer capital, depending on the firms' size. More precisely, the marginal profits from customer capital are non-monotone in their productivity, holding competitors' productivities constant. When the firm is unproductive relative to its competitors, which maps into a small zero-demand market share $\bar{s}_i$, the marginal profit of customer capital increases in productivity.

The following effects are captured: First, there is complementarity between productivity a_i and customer capital b_i. To see this, we note the response of market share s_i to customer capital b_i can be written as

$$\frac{\partial s_i}{\partial b_i} = \zeta_i \left(1 - \frac{\zeta_i}{\sum_j \zeta_j}\right),$$

where $\zeta_j = \left(\frac{1}{s_j} + (\sigma - 1)\frac{1}{1-s_j}\right)^{-1}$.

The coefficient ζ_i captures how responsive a firm's market share is to changes in its customer capital relative to the market average. This responsiveness varies with the firm's baseline market share $\bar{s}_i(\alpha)$ and the elasticity of substitution σ. When σ approaches 1, ζ_i approaches $\bar{s}_i(\alpha)$, meaning the firm's responsiveness is proportional to its baseline market share. When σ approaches infinity, or perfect substitution, the ζ_i goes to zero.

The sigmoid function underlying the market share formulation yields an inverse U-shaped relationship between market share and responsiveness to customer capital. Firms with market shares that approximate those of the market leader ($\bar{s}_i(\alpha) \approx 0.366$ when $\sigma = 4$) exhibit the highest responsiveness to customer capital investments, while both dominant and marginal firms show diminished responsiveness. This non-monotonic relationship creates strategic investment patterns where mid-sized firms have stronger incentives to invest in customer acquisition than either market leaders or small entrants. The variable markup effect further modulates this relationship, as higher markups reduce a firm's responsiveness to customer capital changes. This again connects to the role of strategic complementarities discussed in the literature, but transports it to a dynamic setting.

Equilibrium Marketing Investment. Market shares respond to customer capital through two mechanisms. First, firms with larger $\bar{s}_i(\alpha)$, e.g., larger due to productivity, are more responsive to customer capital changes. This is the complementary force of customer capital. Second, larger firms have a strategic incentive that is summarized through ζ_i. With linearized shares and profits becoming quadratic in market share, the problem is transformed into a tractable linear-quadratic game with optimal investment as follows:

$$\eta_i(\alpha, \mathbf{b}) = \gamma_i(\alpha) + \epsilon_i(\alpha)'\mathbf{b}. \tag{8}$$

We represent the dynamics of the customer capital using the following I-dimensional linear system, which is a stacked version of equation (8). More precisely,

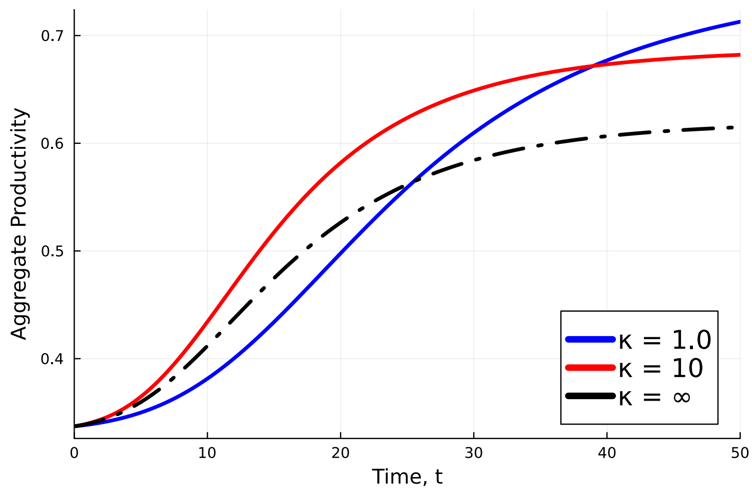
$$\dot{\mathbf{b}}(\alpha) = \Gamma(\alpha) + (\mathcal{E}(\alpha) - \rho\mathbf{I})\mathbf{b}, \tag{9}$$

where $\Gamma(\alpha) = (\gamma_1(\alpha)), ..., \gamma_I(\alpha))'$ and $\mathcal{E}(\alpha) = (\epsilon_1(\alpha), ..., \epsilon_I(\alpha))'$. Economically, $\Gamma(\alpha)$ represents the productivity incentive, and $\mathcal{E}(\alpha)$ captures the size incentive in firms' marketing decisions. The evolution of firm-level customer capital is driven by both incentives. First, there is the productivity incentive, $\Gamma(\alpha)$: more productive firms invest more in customer acquisition due to complementarity. Second, there is the size incentive, $\mathcal{E}(\alpha)$: firms with larger customer capital reinforce their advantage. This second component creates path dependency where initial conditions matter–unlike models where customer capital merely tracks productivity with a lag. These equilibrium strategies of marketing investments are solutions to a set of stacked Riccati equations, the details of which we leave to the Appendix A.2

2.3 Endogenous Path Dependency of Market Structure

The two-firm case offers analytical tractability while preserving key economic mechanisms. With parameters $I = 2$, $\lambda = 0$, $\mathbf{a} = (1, 0)$, and $\mathbf{b} = (-5, 0)$, we study how a productivity leader with initial customer disadvantage evolves over time. In this simplified environment, the discouragement and cannibalization effects exactly offset each other, yielding closed-form solutions for the Riccati equations that govern investment strategies and allowing us to isolate fundamental productivity and size incentives.

Figure 3: Aggregate Productivity w/ 2 Firms and no Shocks



The system's evolution follows $b_t = b_0 e^{-(\rho-\phi)t} + \frac{\gamma}{\rho-\phi}(1 - e^{-(\rho-\phi)t})$, where customer capital is initially distorted but converges to fundamentals in the long run. Figure 3 shows the response of aggregate productivity to a productivity increase at firm 1, which starts with lower customer capital than firm 2. These dynamics illustrate how, when a customer capital follower becomes a productivity leader, there are different transitions depending on the initial distribution and the marketing cost. With no mean reversion in customer capital, we observe how the initial misallocation resolves over time as resources gradually shift toward the more productive firm.

The illustrative two-firm case demonstrates how marketing costs (κ) critically influence the speed and pattern of aggregate productivity evolution. As shown in Figure 3, when marketing costs are low ($\kappa = 1$), aggregate productivity initially lags but eventually exceeds scenarios with higher marketing costs, reflecting more efficient long-term resource allocation. Conversely, higher marketing costs ($\kappa = 10$) produce faster initial productivity growth but ultimately reach lower steady states, while prohibitive marketing costs ($\kappa = \infty$) severely limit productivity gains. The underlying dynamics follow a formula where customer capital, though initially distorted, eventually converges toward fundamentals, with the speed and efficiency of this process directly influenced by marketing frictions.

The intuition from this two-firm case extends to the more general oligopolistic setting. The slow convergence to the world where productivity improvements transmit to the aggregate is unique to a framework where firms with significant pre-existing size have a mechanism, such as investment in customer capital, to protect their market share. This is a central component of this paper, both for positive analysis of the evolution of productivity and normative implications of policies relating to concentration and economic growth.

2.4 Extensions and Aggregation

The baseline model focuses on customer capital and productivity across firms. For aggregation and empirical analysis, we extend the baseline model to analyze the mechanics across groups. We aggregate across product groups in our data using a Cobb-Douglas specification, which yields the desirable properties of linear welfare aggregation and a well-defined stationary distribution. This approach allows us to express welfare numbers as non-discounted versions of the transitional path, providing intuitive comparisons across policy regimes. We discuss the aggregation in this environment and then turn to discuss general empirical extensions.

Aggregation and Welfare. The household makes its static consumption and labor decisions, as well as the dynamic savings decision. We assume the household has a flow utility of $\log \mathbf{C}_t - \mathbf{L}_t$. The household spends on consumption and holds assets $\mathbf{A}_t$ that evolve over time. The household can borrow and save in a representative portfolio of all firms, such that the aggregate profit $\mathbf{\Pi}_t$ is rebated to the household as a dividend. We define r_t to be the interest rate and normalize the wage to be 1. We write the household's problem as,

$$\max_{c_{ikt}, \mathbf{L}_t} \int_0^\infty e^{-\rho t} \left(\log \mathbf{C}_t - \mathbf{L}_t \right) dt,$$

s.t.

$$\log \mathbf{C_t} = \int_0^1 \phi_k \log \frac{C_{kt}}{B_{kt}},$$

$$\dot{\mathbf{W}}_t = r_t \mathbf{W}_t + \mathbf{L}_t + \mathbf{\Pi}_t - \int_0^1 \sum_{i=1}^{I} p_{ikt} c_{ikt} dk,$$

with C_{kt} as C_t given equations (1) and (2).

ϕ_k represents the share of group k in the aggregate. We now turn to the connection between firm-level choices and aggregate outcomes. In the main discussion of the model, we abstract from the group level k, and in the first step of aggregation, we start by focusing on welfare groups by group. We then aggregate over groups.

We present a heuristic discussion of overall welfare, which we expand on in Section 5. The household's utility depends on both the dispersion of markups and on whether customer capital is allocated toward the productive firm. We write out the consumption from a product group with productivity distribution α and customer capital $\mathbf{b}$.

Lemma 1 (Aggregation) *In a steady-state equilibrium, the discounted utility of the household is given by:*

$$\mathbf{W} = \int_{\alpha,\mathbf{b}} W(\alpha,\mathbf{b})dG(\alpha,\mathbf{b}), \quad W(\alpha,\mathbf{b}) = \frac{1}{\rho}\left(\log\frac{A(\alpha,\mathbf{b})}{M(\alpha,\mathbf{b})} - \frac{1}{M(\alpha,\mathbf{b})} - D(\alpha,\mathbf{b})\right)$$

where

1. $A(\alpha,\mathbf{b})$ *is the group-level labor productivity*

$$A(\alpha,\mathbf{b}) = \left(\frac{\sum_{i=1}^{I} e^{a_i+b_i}}{\sum_{i=1}^{I} e^{b_i}}\right)^{\frac{1}{\sigma-1}};$$

2. $M(\alpha,\mathbf{b})$ *is the aggregate markup*

$$M(\alpha,\mathbf{b}) = \left(\frac{\sum_{i}^{I} e^{a_i+b_i}\mu_i(\alpha,\mathbf{b})^{1-\sigma}}{\sum_{i}^{I} e^{a_i+b_i}}\right)^{\frac{1}{1-\sigma}};$$

3. $D(\alpha,\mathbf{b})$ *is the aggregate labor cost in marketing*

$$D(\alpha,\mathbf{b}) = \frac{d_0}{2}\sum_{i}^{I}\eta_i(\alpha,\mathbf{b})^2.$$

Returning to firms' reallocation decisions, there are two externalities created by firms to the representative household. First, the dispersion of markups between firms creates a misallocation of labor. Firms choose markups to maximize individual profit and not overall welfare. This misallocation reduces the productivity of labor. Second, the firms do not fully internalize the benefit of matching transferable customer capital towards more productive firms. Firms may stick with mismatch if it builds more market power. The second externality is a novel insight from our paper. This equilibrium is not necessarily efficient, which creates room for policy.

We wait to discuss optimal policies until we have set up the planner's solution in Section 5. However, the analysis so far hints at the main role of policy. A policy that aims to improve efficiency should induce firms to sort customer capital to more productive firms and reduce markups.

Extensions for Empirics. The baseline model focuses on customer capital and productivity across firms. To analyze welfare implications and conduct counterfactual analysis, we extend our baseline model in two important dimensions. First, as noted earlier, we aggregate across groups in a Cobb-Douglas fashion. Second, we explicitly model multi-product firms with constant elasticity of substitution across products and decreasing returns in production (as in Hottman et al., 2016). Formally, firm-level consumption is given by:

$$c_{it} = \left(\sum_{u=1}^{U_i} \psi_{iut} d_{iut}^{\frac{\sigma_u}{\sigma_u - 1}} \right)^{\frac{\sigma_u - 1}{\sigma_u}} , \quad \sigma_u > \sigma \tag{10}$$

where $\sigma_u > \sigma$ implies that firms set constant markups across their product lines, and production at the firm-group level exhibits decreasing returns to scale: $c_{ikt} = e^{a_{ikt}/(\sigma-1)} l_{ikt}^{\omega_k}$. This framework captures how larger firms face higher marginal costs as they expand production, affecting the efficiency of customer capital allocation. The aggregation structure delivers a tractable environment for evaluating welfare implications while maintaining the key mechanisms of our model: the mismatch between customer capital and productivity and the strategic incentives that may exacerbate this mismatch over time.

3 Data and Empirical Analysis

In this section, we connect our theory to an empirical model of firm cost productivity and customer capital, incorporating adaptations from existing literature. We use NielsenIQ Retail Measurement Services (RMS) scanner data to decompose the drivers of market share. Our empirical decomposition examines productivity and demand components, where the demand components are connected to the theoretical concept of customer capital. We identify the relative importance of demand and productivity in explaining firm size differences, with particular attention to *changes* in these objects. Finally, we construct a measure of the gross and net flows of customer capital, which we then use to calibrate the decay (ρ) and endogenous investment in customer capital (η) from the model.

3.1 Data

Our empirical analysis requires detailed data on prices, market shares, and firm characteristics over time to identify the joint evolution of customer capital and productivity. Furthermore, we leverage product-level information to decompose the gross and net flows of residual demand or customer capital. To enable the study of price and sales data, we employ detailed bar-code level data from Kilts-NielsenIQ Retail Measurement Services Data from the University of Chicago Booth School of Business. The data are large and comprehensive in the consumer product space. This dataset delivers significant coverage for products, brands, and firms. We merge GS1 company-level information to extract the firm identifier from the barcode of each product, following Argente et al. (2020).

For each product group k, we observe sales and prices at the firm-year level. Following Hottman et al. (2016), we focus on price and market share variation within product groups to control for differences across categories. Consistent with the model, we focus primarily on the largest firms in each product group, and approximately 10 firms have a market share larger than 1% of the national share, consistent with some strategic incentives. Table 1 reports the number of firms, the market shares of top firms, and the number of products in our sample.

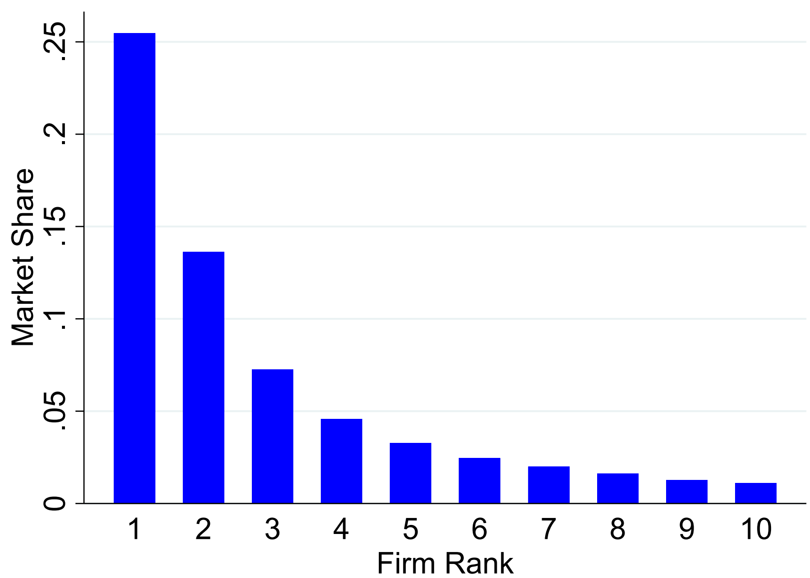
Our final sample includes 423 firms per product group on average, with substantial variation in firm size - the stark difference between mean sales (almost $6M) and median sales ($12,000) illustrates the skewness of firm size. The largest firm in a typical product group captures 27% market share, while the second largest firm captures 13%. Market share drops off quickly beyond the top firms - the 10th largest firm captures only a little over 1% share. Figure 4 shows the firm size distribution across groups weighted by group size.

Table 1: Summary Statistics

Panel A: Concentration		
	Mean	Median
Number of Firms	423	314
Average Sales (Thou)	5946	12
HHI	0.13	0.10
Top Firm Share	0.27	0.24
Top 3 Firm Share	0.50	0.46
Top 10 Firm Share	0.68	0.68
Panel B: Persistence		
	All Firms	Top 10 Firms
Sales	0.97	0.95
Price	0.76	0.43
Leadership	1.00	1.00

Notes: Both panels are weighted by product group sales.
Panel B shows AR(1) coefficients from firm-level regressions.

Figure 4: Firm Size Distribution



Notes: Average market share by firm rank, weighted by group size. Source: USPTO/RMS NielsenIQ.

One important component of our analysis is bringing a dynamic analysis to the decomposition of the firm size distribution. The panel structure of the NielsenIQ data enables us to estimate the persistence of customer capital across the firm size distribution. The centrality of leadership persistence and the drivers of this persistence will be central in this section, in line with the model. We now turn to an empirical framework and direct particular attention to large firms that hold a majority of market share and drive the bulk of market activity.

3.2 Empirical Methodology

One of the goals in this paper is to decompose the dynamics of a firm's market share driven by demand and productivity. Following Hottman et al. (2016), we first decompose firms' revenues into four static components: demand differences (customer capital), production costs (productivity), scale of production (marginal cost), and markups. This static decomposition reveals which factors matter most in explaining cross-sectional and cross-time differences in firm size. These components become crucial for understanding firms' dynamic incentives to invest in customer capital, which we study in Section 2.

Our frequency of analysis is yearly. We assume there is a representative household that spends 1 unit of expenditure on a measure 1 of product groups, indexed by $k \in [0, 1]$. Within each product group k, there are J_k firms, indexed by $j = 1, ..., J_k$.

Our empirical framework connects to Hottman et al. (2016), who provide a structural decomposition of firm size heterogeneity into various components. We focus on decomposing firm market share into three key components: productivity, residual demand (customer capital), and markups. This decomposition requires several empirical objects that we construct from NielsenIQ Retail Measurement Services scanner data. We measure prices as the geometric weighted mean across all products within each firm-group-year combination. Sales are calculated as the total revenue within the firm $\times$ group $\times$ year. Following Atkeson and Burstein (2008), we infer markups based on firms' oligopolistic pricing behavior, which allows us to back out marginal costs by combining markup and price information. Productivity is then derived by accounting for the decreasing returns to scale in production that firms face, inferred from their output and marginal cost. Finally, customer capital (or residual demand) is calculated as the residual component of sales that cannot be explained by prices, providing a measure of consumers' underlying attachment to a firm's products independent of price.

This parsimonious model allows us to write the logarithm of firm sales as an additive function of the three factors (demand, productivity, and markups) as well as group-level factors. By taking the difference of the terms with the geometric averages, we write:

$$\Delta_k \log s_{ikt} = \Delta_k \log b_{ikt} + \omega(1-\epsilon_{ik})\Delta_k \log a_{ikt} - (1-\epsilon_{ik})\mu_{ikt} \tag{11}$$

The above equation decomposes log market share differences into customer capital differences, $\Delta_k \log b_{jkt}$, productivity differences scaled by demand elasticity, $(1-\epsilon_{jk})\Delta_k \log a_{jkt}$, and markup differences, $(1-\epsilon_{jk})\mu_{jkt}$, which negatively affect market share. σ is the elasticity of substitution across firms within a group, and ω is the coefficient that governs diminishing returns at the firm level. We take these parameters from Hottman et al. (2016) in our empirical exercises.

One discrepancy in our exercises from Hottman et al. (2016) is that firms are multi-product, which requires a firm-level price aggregator. We follow Hottman et al. (2016) and assume products are aggregated through a CES. Further, firms face an upward supply curve, the estimates which we take from Hottman et al. (2016).

3.3 Demand, Productivity, and Market Share

This section relates the theoretical framework to the formation of market share at the firm level. We then turn to firm market share decompositions to ask which force is most likely to drive market share. Finally, we focus on the nature of firm growth as a function of demand and productivity. We now turn to the correlations in Table 2, where each ingredient is defined as in equation (11).

Table 2: Correlations at the Firm Level

	Group-Adj. Sales	Demand	Productivity	Marginal Cost	Markup
Group-Adj. Sales	1				
Demand	**0.669*****	1			
Productivity	0.547***	-0.255***	1		
Marginal Cost	0.664***	-0.109***	0.989***	1	
Markup	0.247***	0.144***	0.169***	0.196***	1
		Beta Variance Decomposition			
β explained (all firms)	1***	0.942***	0.191***	-0.133***	-0.001***
β explained (Top 10 firms)	1***	**0.920*****	0.273***	-0.176***	-0.020***

Notes: correlations across firms of the deviations from group-means in log terms. The β-variance decomposition takes a regression of the input variable x against log sales and coefficients sum to 1. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The correlations in Table 2 reveal several important patterns. First, demand has the strongest correlation with sales (0.669), even higher than productivity (0.547). Second, while both productivity and scale are highly correlated with sales, they are negatively correlated with customer capital, suggesting these advantages may substitute rather than complement each other. Third, markups show weaker correlations with all variables, indicating market power may play a smaller role in explaining sales variation.

While correlations show the relationships between components, they don't quantify each factor's contribution to overall market share variation. To measure these contributions precisely, we perform a β-variance decomposition where all terms sum to one, allowing us to interpret each coefficient as the share of variance of total log deviation of sales explained at the bottom of Table 2. We introduce four components into this decomposition. As earlier, we have our residual (or customer capital), which is the residual demand for the firm, conditional on price. We also have the firm-level productivity z_j and the firm-level cost scale, which comes from diminishing returns on the production function. Finally, we include the markups.

We see from the β-variance decomposition a couple of interesting results. First, residual demand is again the most significant factor driving market share, explaining more than 90%. Second, the productivity channel is also important. In the cross-section, productivity alone explains 19% of the market share. Overall, residual is about four times as strong as productivity in explaining changes in market share in the cross-section. The decreasing returns to scale from the cost curve ("Cost DRS") make it such that when demand and productivity each increase, the cost effect will diminish their overall expansion in market share due to the convexity of the cost function ($\omega < 1$).

Overall, residual demand is a core driver of firm sales. This object relates to endogenous and exogenous forces on the customer-firm relationship. We now turn to our structural decomposition of the changes in demand from these endogenous and exogenous forces.

3.4 The Dynamics of Demand

So far, we have shown that the role of demand is central to both the static firm distribution of market share and firm-level growth. The relative persistence of this demand channel suggests the centrality of customer capital. We now turn to decompose the forces driving the fluctuations in demand at the firm level. We will bring these components to the structural estimation of the law of motion of customer capital in the next section.

We start by focusing on three channels by which demand changes: entry, exit, and continuing products, and then group them into the two main drivers. First, firms introduce new products within the same group to boost attention to their basket of goods. Second, existing products exhibit churn as a result of market dynamics. Third, products exit the market. We ask how these three processes contribute to change in sales and change in overall customer capital in this section. For sales, we are interested in the evolution of firm sales from these three terms,

$$\Delta \ln \text{Sales}_{it} = \underbrace{2\frac{\text{Cont}_t - \text{Cont}_{t-1}}{\text{Sales}_t + \text{Sales}_{t-1}}}_{\text{Continuing products}} + \underbrace{2\frac{\text{Entry Sales}}{\text{Sales}_t + \text{Sales}_{t-1}}}_{\text{Entry}} \underbrace{-2\frac{\text{Exit Sales}}{\text{Sales}_t + \text{Sales}_{t-1}}}_{\text{Exit}} \tag{12}$$

This delivers the sales decomposition, which tells us how the activity of each margin affects overall sales at the firm level. We now turn to the decomposition of sales growth at the firm level in Table 3.

Table 3: Beta Decomposition of Changes in Firm-Level Sales

Component	Beta Coefficient	Std. Error	Mean Level
Continuing products	0.694	0.005	-0.064
New product entry	0.234	0.005	0.062
Product exit	0.071	0.002	-0.009

Notes: This table presents the beta decomposition of changes in firm sales into three components: continuing products, new product entry, and product exit. The sample is restricted to the top 10 firms by sales within each product group for the period 2007-2017. Beta coefficients represent each component's contribution to the total variance of changes. All estimates are weighted by product group sales.

The decomposition of sales suggests some critical margins. First, when it comes to new product entry, its role in expanding firm sales is central. New products explain approximately 23% of the variation in aggregate sales and also have a central role in firm-level residual demand, which we turn to next. Theoretically, this aligns with the endogenous force as firms introduce new products to maintain high demand. This is consistent with Argente et al. (2020), who find that constant product entry is a central driver for firms' maintaining demand.

In the model, there are two core drivers of demand changes: endogenous investments of firms and the natural decay of existing customer capital. As we can see from the sales decomposition, continuing products and exiting products both contribute negatively to firm-level sales, while entry contributes positively. For the natural creation and decay of customer capital, we group continuing and exiting products together and focus on entry as the firm's endogenous choice of investment. This can be thought of as firms refreshing shelf space or linked to their endogenous investments in customers. We present the decomposition of residual demand, which has a slightly different structure. This is presented in equation (13),

$$\Delta b_{ikt} = \underbrace{\frac{N_{entry}}{N_{ikt}} \cdot b_{entry}}_{\text{entry}} + \underbrace{\frac{N_{cont}}{N_{ikt}} \cdot (b_{cont,t} - b_{cont,t-1}) - \frac{N_{exit}}{N_{ik,t-1}} \cdot b_{exit}}_{\text{decay}} . \tag{13}$$

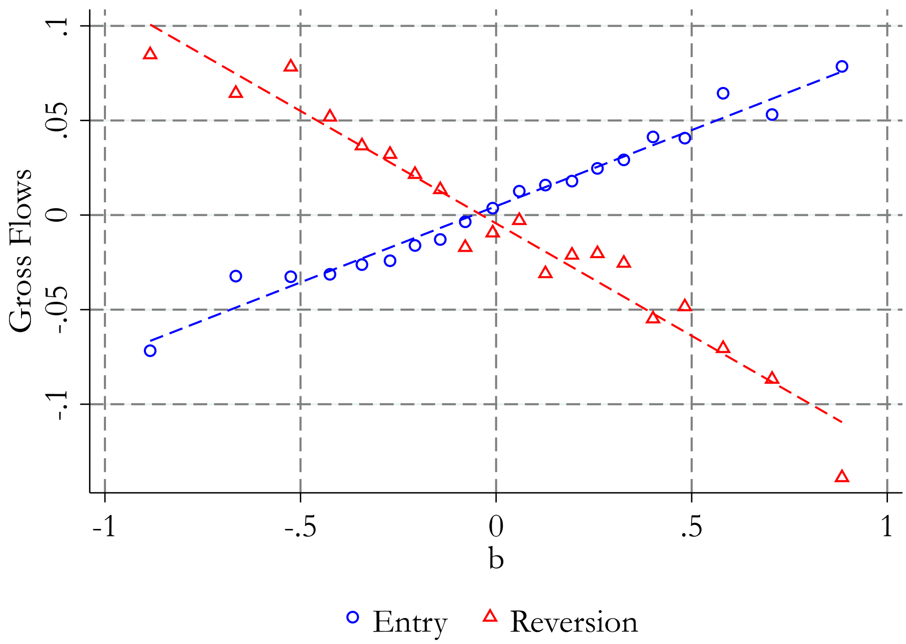
The key assumptions in our setting are that new products require an active form of endogenous investment by the firms. Given consumer limited attention, this crowds out other products, which continue or exit.

Entry and Decay of Customer Capital. Central to our study is the role of endogenous customer capital, where firms make specific investments in acquiring and maintaining customers, conditional on some decay. Here, we ask about the change in customer capital and how it varies with firms' existing level of customer capital to understand this relationship. In the next section, we use this relationship to identify the marketing cost and decay parameter.

Splitting new entry into creation and existing and exiting products into decay, Figure 5 details the changes in customer capital coming from creation and decay by the level in the previous period. We document this for the top 10 firms within a product group, where the group is weighted by its total average sales. We call "marketing intensity" the introduction of residual demand through new products. The blue dots indicate the rate of marketing intensity η relative to the group mean, whereas the red triangles indicate the relative decay intensity to the group mean. The *level* of these outcomes is not identified since it is relative to the group, but the *slope* provides intuition on the relationship between firm customer capital, creation, and decay.

We note some clear patterns in Figure 5. We find that larger firms have higher rates of marketing η relative to small firms. For a firm with one log point higher customer capital, they market 0.07 log points higher in terms of new product introduction. Conversely,

Figure 5: Gross and Net Flows of Customer Capital



Notes: This figure disaggregates the change in residual demand into continuing products (orange), exiting products (red), and entering products (blue). The decomposition is available in equation (13). Authors' calculations.

these firms also experience higher decay rates. For a firm with one log point higher customer capital, their customer capital decays 0.13 log points higher. The decay rate, ρ will be estimated from this and the product introduction will serve as an out-of-sample test for the measure of endogenous demand building.

Discussion of Empirical Results. Our empirical exercises extend standard workhorse empirical frameworks for measuring drivers of firm heterogeneity. Consistent with the literature, we find that demand differences are indeed central drivers of firm size, and they are persistent. One of the novel findings in our analysis is that firms with higher residual demand introduce more new products to increase residual demand, but also face higher decay. These demand dynamics play a central role in market share and will naturally have implications for the persistence of market share leadership. From this exercise, the nature of misallocation and the implications for aggregate productivity, welfare, and policy are less clear. This question is central to policy analysis and normative understanding of the drivers of market share and aggregate productivity. We turn to this next through the lens of the model, where we focus on the dynamic manifestation of the static distortions.

4 Estimation

In this section, we discuss the steps to estimate the parameters of our model. We proceed in two steps. First, we calibrate the model to match key empirical regularities documented in Section 3, with particular attention to the dynamics of the correlation between product creation and decay and the market share of firms. Second, we discuss the model's fit for targeted and untargeted moments.

4.1 Moments and Parameters

We interpret our empirical results as a steady-state equilibrium of the model. Some of the parameters are well estimated in the literature, and we discuss their values. Others are novel in our framework, and we discuss the moments from our empirical analysis that help identify these parameters.

Parameters from the Literature. The following parameters are directly taken from the literature. In terms of the household's preference, we calibrate the discount rate of the household to match a standard annual risk-free rate, $\rho = 0.03$. The substitution elasticity among firms is a crucial parameter and is well estimated by studies in the literature. The closest paper we build on is Hottman et al. (2016). They estimate the firm-level substitution elasticities with a similar demand system and NielsenIQ Homescan data. We take the median substitution elasticity, $\sigma = 3.9$, among product groups based on their estimates. This elasticity implies the minimum profit margin of firms is around 0.25, and the maximum is 1.0. As this is a central parameter, we also perform our quantitative analysis using alternative values of substitution elasticity to understand how this changes the counterfactuals.

Internally Calibrated Parameters. The rest of the parameters are based on our empirical analysis. Most of our parameters, except for the marketing cost, can be directly inferred from their empirical counterparts.

First, we calibrate the number of firms, $I = 10$, to be the average number of firms with market shares above 1% across product groups. Second, we estimate the parameters of the stochastic process to match the empirical persistence of estimated firm productivities. We fit a Jordà (2005) local projection regression to the firm-level productivity,

$$a_{i,t+h} = \alpha_0 + \alpha_{1,h} a_{it} + \alpha_2 a_{it-1} + \varepsilon_{it}, \tag{14}$$

where we plot out the response to a shock to a_{it} through the coefficient $\alpha_{1,h}$. We match the decay rate and find an average decay rate of 0.78. Interpreted in our model, a coefficient of 0.78 on lagged productivity implies a switching rate χ. We assume the distribution firms draw their productivity from ten possible values, and these values and their associated probability mass are chosen to mimic a normal distribution with mean 0 and the standard deviation matching the empirical volatility of the productivity process, 0.54.

Second, we calibrate the two remaining parameters for the dynamics of customer capital from Section 3.4. These are the two most important parameters for our quantitative analysis, and our empirical results provide rich content that informs these parameters. From the results in Figure 5, the correlation between the decaying flow and the current residual demand of firms is -0.13. This exactly maps into a value of $\rho = 0.13$. We then calibrate the marketing cost, κ, to match the average marketing cost as a share of revenues from the literature. Two papers from the literature provide estimates of this value. He et al. (2024) estimates that the mean of the advertising cost share in revenue is 4.1% and Cavenaile et al. (2025) estimates the share to be 2.2%. The difference in these two estimates is that He et al. (2024) also includes the personnel costs in marketing. We use the more comprehensive cost measure as we want to focus on marketing expenses in general. We report results using both estimates and will refer to the target of 4.1% as our baseline.

4.2 Goodness-of-Fit

We now discuss the data moments our model replicates. The first goodness-of-fit we perform is on the dynamics of customer capital in the data and in the model.

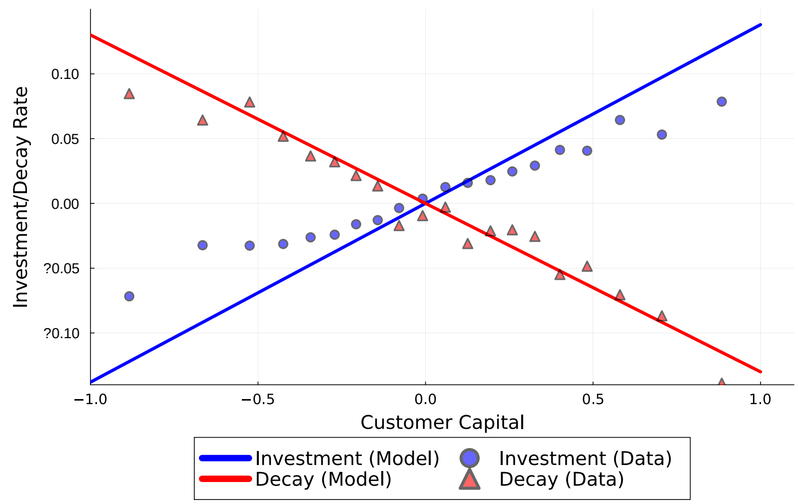
In Figure 6, we plot the average demand creation and demand decaying flow of firms against the relative customer capital, both demeaned relative to their product group av-

TABLE 4: ESTIMATION MOMENTS AND PARAMETERS

Parameter		**Value**	**Main Identification**
Interest Rate	r	0.03	Annual Risk-free Rate
Substitution Elasticity	σ	3.90	Hottman et al. (2016)
Number of Firms	I	10	Mean Number of Firms above 1% share
Productivity - Switching Rate	λ	0.78	Yearly Autocorrelation of Productivity, 0.90
Productivity - Distribution	F	Discretized $\mathcal{N}(0, 0.54)$	Yearly Volatility of Productivity, 0.54
Demand - Decay Rate	ρ	0.13	Empirical decay
Demand - Marketing Cost	κ	1.10	Average marketing cost share 4.1%

Notes: Parameters estimated separately (top panel) and jointly (bottom panel). Source: RMS NielsenIQ and author calculations.

FIGURE 6: CORRELATION OF INVESTMENT/DECAY RATE WITH CUSTOMER CAPITAL

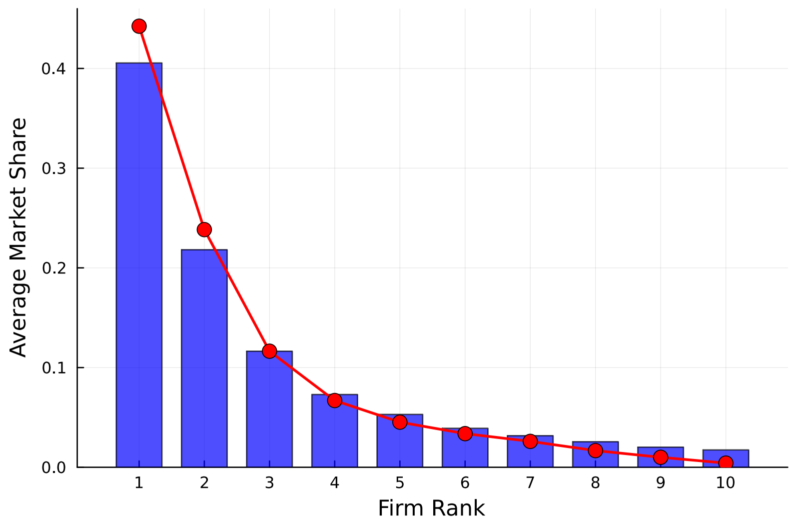


Notes: This figure plots the data (points) and model (lines). The red line is matched to the decay rate while the blue line comes from the dispersion of product residual demand creation. Authors' calculations.

erages. Our model exactly replicates the decay pattern of customer capital because it directly matches the model decay rate ρ to this slope; The correlation between customer capital and creation flows, on the other hand, is an untargetted moment. Recall that we calibrated the average costs κ to match the average marketing expenditures and the cost elasticity of marketing to be quadratic. The fact that our model replicates the empirical correlation between lagged customer capital and investment provides additional validation for our model mechanism.

Figure 6 demonstrates our model's ability to match the empirical distribution of product residual demand across firms. We calibrate two key parameters: ρ, which governs the persistence of product residual demand, and κ, which controls the strength of cannibalization effects between a firm's products. The red line represents our targeted moment coming from the decay rate of customer capital, which we use directly in our calibration procedure. The blue line shows an out-of-sample prediction: the cross-sectional distribution of the endogenous investment in customer capital. The close alignment here suggests that the targeted marketing cost can capture the size dynamics in the data.

Figure 7: Firm Size Distribution



Notes: This figure plots the model and data of the firm size distribution. The 63% of the market the top 10 firms hold is normalized to 100%. Authors' calculations.

Figure 7 demonstrates the strong alignment between our model's predicted firm size distribution (red line) and the empirical distribution observed in the NielsenIQ data (blue bars). The model successfully captures the highly skewed nature of market shares across firm ranks, with the market leader commanding approximately 40% of the market, almost double the share of the second-ranked firm (22%), conditional on restricting attention to the top 10 firms in each market. This steep decline continues as we move to lower-ranked firms, with the third firm capturing about 11%, and subsequent firms holding progressively smaller shares, approaching 2% for the tenth-ranked firm. The model's ability to match this pattern is particularly notable given that we did not directly target the full distribution in our estimation but instead focused on moments related to the dynamics of customer capital and productivity. This close fit suggests that our theoretical mechanisms—the interaction between productivity differences, strategic customer capital accumulation, and variable markups—correctly capture the forces that generate the observed concentration in consumer product markets. The granularity of competition among these top firms is essential for our analysis.

We finally turn to moments on persistence and the correlation between productivity and customer capital in Table 5.

Table 5: Untargeted Moments

Outcome of Interest	Model	Data
Market Leader Persistence	0.92	0.96
Customer Capital Leadership Persistence	0.95	0.92
Productivity Leadership Persistence	0.90	0.88
Endogenous Correlation (a, b)	0.55	0.71

Note: Author calculations. Persistence is measured as 5-year persistence annualized in the data to purge transitory components of rank.

The model successfully captures several important empirical patterns. We match the higher persistence of market leadership (0.92 in model vs. 0.96 in data) compared to productivity leadership (0.90 in model vs. 0.88 in data). This aligns with findings in Bartelsman et al. (2013) that market leaders are more persistent than would be predicted by productivity dynamics alone.

Our key contribution is connecting this pattern to customer capital dynamics. Customer capital leadership persistence (0.95 in model vs. 0.92 in data) exceeds productivity persistence, consistent with evidence from Foster et al. (2008) that demand factors exhibit greater persistence than productivity shocks. The model further generates a positive correlation between productivity and customer capital (0.55 in model vs. 0.71 in data) coming from endogenous investment, reflecting the endogenous accumulation mechanism in our framework.[4]

These moments confirm the central role of customer capital in explaining market share dynamics and persistence in environments with granular firms. The path dependency generated by strategic marketing investments allows market leaders to maintain their positions despite productivity mean reversion, creating patterns of persistence that match empirical observations.

[4]The baseline relationship between productivity and customer capital in the data is negative due to the contemporaneous shocks having a negative correlation. To isolate the endogenous component, we run local projections with controls and ask about the correlation between productivity and the fitted values of customer capital. We discuss this in detail in Appendix B.6.

5 Quantitative Analysis

This section answers the following questions. Section 5.1 asks, how do the three margins of customer capital churn (size incentive, productivity incentive, and natural decay) drive the matching between customer capital and firm productivity? Section 5.2 addresses how customer capital affects the transmission of firm-level productivity shocks to aggregate productivity and, thus, aggregate allocative efficiency. Section 5.3 asks, how do these findings change our understanding of policies that aim to alleviate the distortions due to oligopoly?

5.1 What Drives Matching between Demand and Productivity?

We have documented empirically that firms' demand does not always match productivity. Our structural model provides a decomposition of the drivers underlying this matching. To start this section, we plot the model-implied distribution of changes in group-level productivity. More precisely, since the aggregate productivity is the average of group-level productivities, we can decompose the change in aggregate productivity into four elements: (1) the change due to marketing investements increasing productivity, *productivity-enhancing investement*, $\Delta^+\mathbf{A}$; (2) the change due to marketing investments decreasing productivity, *productivity-damaging investment*, $\Delta^-\mathbf{A}$; (3) the change due to the exogenous decay of customer capital, $\Delta^\rho\mathbf{A}$; and (4) the change due to the exogenous shifts in productivities, $\Delta^a\mathbf{A}$:

$$\Delta\mathbf{A} = \Delta^+\mathbf{A} + \Delta^-\mathbf{A} + \Delta^\rho\mathbf{A} + \Delta^a\mathbf{A} \tag{15}$$

where

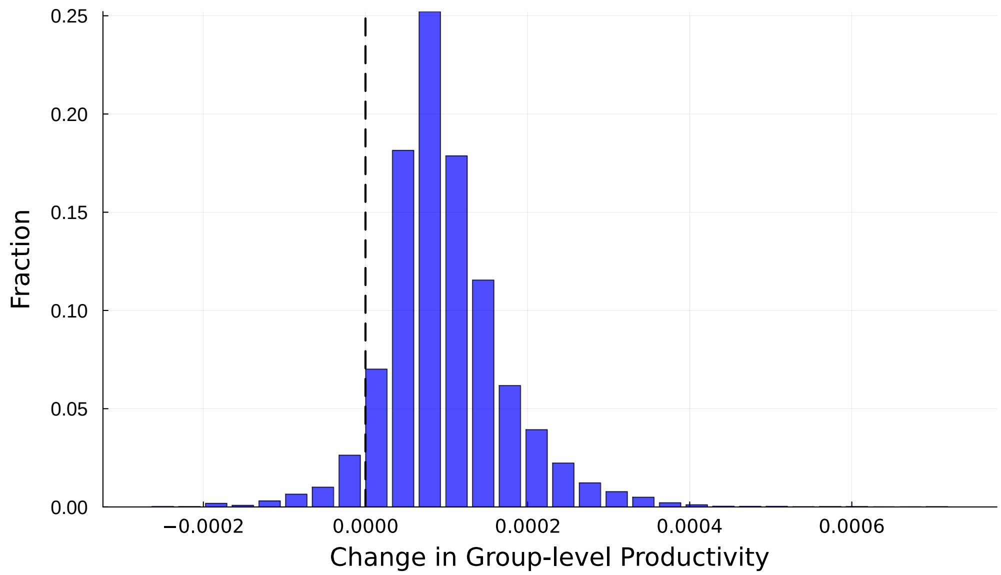
$$\Delta^+\mathbf{A} = \int_{\alpha,\mathbf{b}} \max\left\{\mathbb{COV}\left(s_i(\alpha,\mathbf{b}) - \beta_i(\mathbf{b}), \eta_i(\alpha,\mathbf{b})\right), 0\right\} dG(\alpha,\mathbf{b})$$

$$\Delta^-\mathbf{A} = \int_{\alpha,\mathbf{b}} \min\left\{\mathbb{COV}\left(s_i(\alpha,\mathbf{b}) - \beta_i(\mathbf{b}), \eta_i(\alpha,\mathbf{b})\right), 0\right\} dG(\alpha,\mathbf{b})$$

On a stationary equilibrium, the net change in the aggregate productivity is 0. This decomposition in Equation (15) offers a way to compare the magnitudes of the effects. We visualize the distribution of the values of the productivity-enhancing and the productivity-damaging investments in the following graph.

Figure 8 illustrates two central components of the drivers of the correlation between

Figure 8: Distribution of Productivity and Customer Capital Correlation

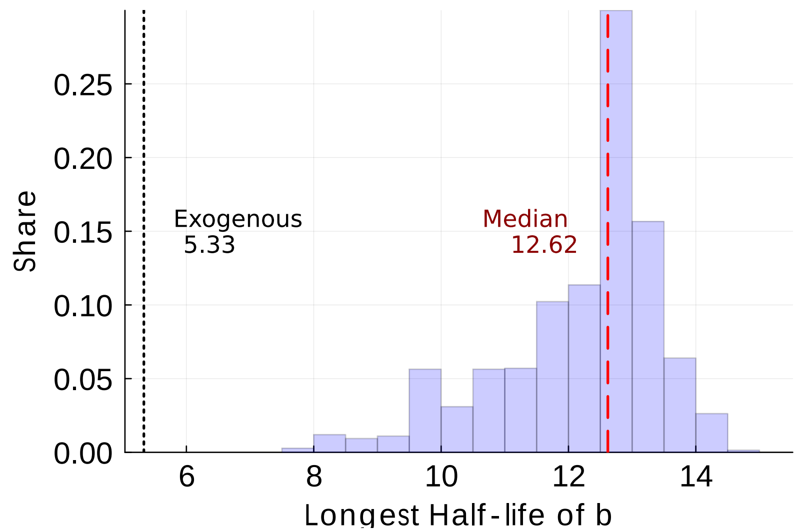


customer capital and productivity. First, the distribution reveals substantial heterogeneity in these correlations, ranging from slightly negative values to nearly perfect positive correlations, with most firms exhibiting moderate to strong positive correlations. This dispersion reflects differences in path dependency across industries, a central component of our framework. Second, the figure also highlights the impact of strategic incentives on this relationship. In our baseline model (blue), which includes both productivity and size incentives, the mean correlation is 0.546. When we remove the size incentive (red), leaving only the productivity incentive, the mean correlation increases to 0.631. This difference demonstrates how strategic considerations associated with firm size can partially decouple customer capital from productivity. When firms internalize their size incentive, they may make marketing investments that are less aligned with productivity, leading to a weaker overall correlation. This finding underscores the importance of modeling strategic interactions in markets with granular firms.

We next turn to productivity shocks and the time it takes customer capital to catch up. Figure 9 displays the distribution of the longest half-life of customer capital (b) across firms in our sample. The histogram reveals substantial heterogeneity in the persistence of demand, with values ranging from approximately 6 to 14 periods.

The median half-life in our estimated model is 12.62 periods (indicated by the ver-

Figure 9: Distribution of Half-Lives

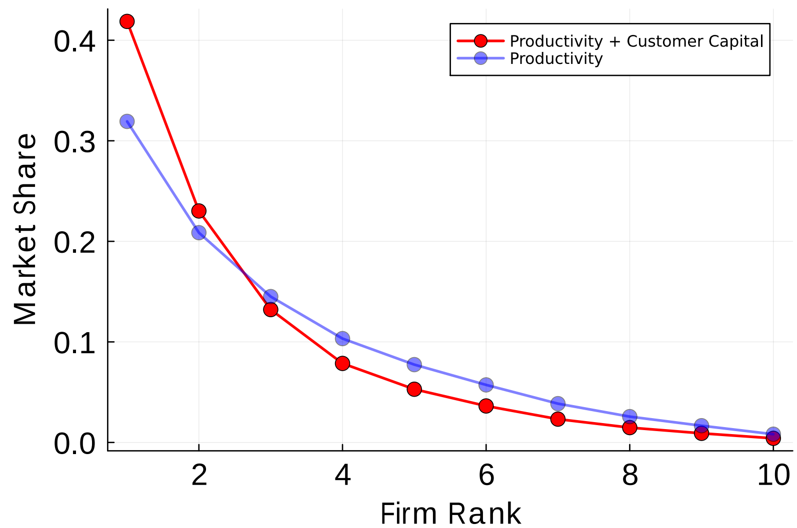


tical dashed line), which is 136% longer than the exogenous rate of 5.33 periods that would prevail if there were no size incentive in customer capital accumulation. This pronounced difference highlights the importance of endogenous marketing decisions in our framework for path dependency. Firms strategically invest in customer acquisition, extending the persistence of their demand bases well beyond what would be expected from natural decay alone. This feedback mechanism between firm size and marketing incentives generates substantial demand inertia, consistent with our theoretical prediction that the markup incentive ($\phi > 0$) amplifies path dependence and reduces the effective speed of mean reversion in the economy.

We then ask how these overall incentives shape the firm size distribution. Figure 10 decomposes market concentration by comparing market shares across firm ranks under two scenarios. The blue line represents market shares determined solely by productivity differences across firms, while the red line incorporates both productivity and accumulated customer capital. For the highest-ranked firms (ranks 1-2), the market shares, when accounting for customer capital, exceed those predicted by productivity alone, with the largest gap observed for the market leader. As we move to lower-ranked firms (ranks 3-10), this pattern reverses, and firms have less market share than their productivity would predict.

This decomposition reveals the crucial role of customer capital in granular markets. The largest firm is 31% larger due to customer capital, which comes as a mix from the

Figure 10: Decomposition of Concentration



size effect, allowing them to extract higher markups and invest more aggressively in customer acquisition, $\eta_i(\alpha, \mathbf{b})$. The fifth-ranked firm is 35% smaller. The gap between the two lines quantifies the importance of path dependency in this economy, showing how the markup incentive amplifies initial advantages for market leaders while creating additional obstacles for smaller firms, even when the latter may have comparable productivity levels. This illustrates our model's key prediction that endogenous marketing decisions can lead to persistent misallocation between demand and productivity, with important implications for market concentration and aggregate productivity.

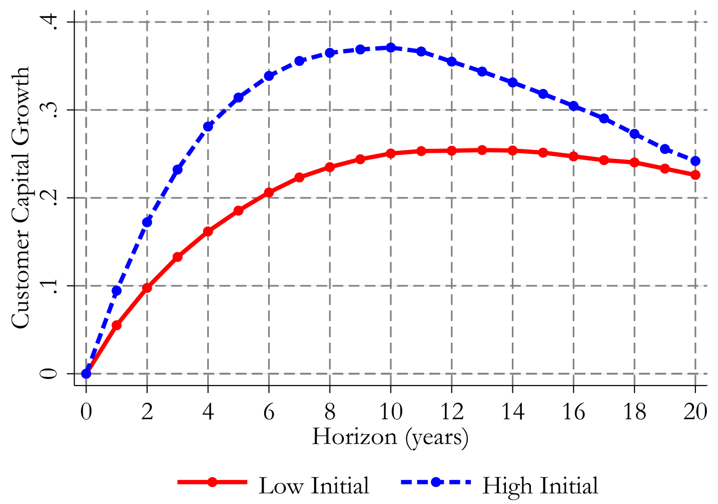
5.2 Productivity: Micro-to-Macro

In this section, we study the relationship between productivity shocks, customer capital, and aggregate dynamics. We start by discussing responses at the firm level to shocks and then discuss the transmission to the aggregate. Finally, we compare this framework to existing frameworks of firm and aggregate productivity to understand the differences.

Responses to Shocks. Firms respond differently to identical productivity shocks based on their existing customer capital. Firms with established customer bases experience amplified benefits through a multiplier effect, as their productivity improvements generate a surplus that attracts even more customers. Conversely, low-productivity firms see dampened responses because of their limited reach and the strategic incentives of larger rivals.

The amplification mechanism creates a "rich-get-richer" dynamic where initially successful firms disproportionately benefit from productivity shocks. This disproportionate reach may not be optimal for aggregate productivity. We look at this further from a local projection on the dynamical system of a shock to productivity of one standard deviation. We plot the impulse response for firms with low initial customer capital and high initial customer capital. This can be seen in Figure 11.

Figure 11: Response to Productivity Shock by Initial Customer Capital



Overall, this shows the centrality of initial conditions for firms. This heterogeneity must be accounted for when analyzing the granular origins of aggregate productivity changes, as overall productivity will move more slowly to its higher level. We can also see this explicitly when comparing the aggregate productivity to other models in the literature.

Comparison to Other Models. To understand the nature of customer capital and aggregate productivity dynamics, we next compare this model to other models standard in the literature in terms of aggregate productivity differences.

Table 6 quantifies the productivity implications of customer capital misallocation across different scenarios. The standard CES oligopoly without customer capital (no

Table 6: Aggregate Productivity Comparisons (net out markups)

	no b	most productive only	no size incentive	calibrated
Agg. Productivity	0.07	0.59	0.15	0.16

b) achieves only 0.07 in aggregate productivity, highlighting the potential importance of customer targeting. In the idealized Aghion-Howitt scenario where all customer capital flows to the productivity leader, aggregate productivity reaches 0.59—over eight times higher. Introducing customer capital with mean reversion but without strategic size incentives ("no size incentive") yields productivity of 0.15, capturing about 26% of the maximum gain. Our calibrated model with endogenous marketing decisions driven by productivity, size incentives, and mean reversion achieves a productivity level of 0.16, slightly higher than without strategic incentives. This comparison reveals that while customer capital is critical for productivity, strategic incentives have modest positive effects in our baseline calibration, primarily because the productivity-enhancing incentives slightly outweigh the distortionary size-based incentives in equilibrium. However, as our analysis of high-markup environments shows, this balance can tilt dramatically when market power is higher, suggesting that customer capital misallocation becomes more concerning in more concentrated markets.

Novel Mechanisms in Quantitative Results. The model admits rich dynamics in the manifestation of aggregate productivity. There are two key components of our model that are central to this result. First, a central mechanism in our model is the dynamic relationship between firm productivity and customer capital. Our model admits a novel framework to link these firm-level changes to aggregate productivity with slow-moving customer capital. With instantaneous adjustment, customer capital would reallocate quickly to firms with positive productivity shocks. Two key frictions impede this through the lens of our model: costly customer acquisition and the strategic incentives that distort firm marketing decisions.

Second, we study the nature of granular firms that have strategic incentives and operate with dynamic investment decisions– this is crucial for quantifying and understanding the nature of concentration to match existing firm size distributions. Monopolistic competition models, alternatively, rely on a continuum of firms. Standard aggregation in these models typically washes out firm-level shocks, which makes analysis on the link from micro dynamics to macro outcomes untenable. This section demonstrates how the strategic framework with dynamic oligopoly provides a structural foundation for understanding how firm-level shocks propagate to aggregate productivity through endogenous market share reallocation. This admits a rich decomposition of market share and more realism when it comes to the transmission of shocks at the firm level.

5.3 Welfare Implications

This paper centers on the relationship between concentration and productivity when firms endogenously invest in customer capital. We now study policies that aim to correct size distortions and ask about their effects on welfare, using the aggregate welfare metrics as in Lemma 1. More precisely, we aim to answer two questions. First, what is the welfare impact of marketing activities as a whole, and how much of the welfare incidence is due to size-based incentives? To do so, we consider the aggregate welfare under two scenarios: no marketing and no size-based incentives to our baseline calibration. Second, can the distortions in the granular economy be undone by policy interventions? We specifically compare the equilibrium outcomes to a social planner problem and a policy that aims to undo the static distortions due to granularity (Edmond et al., 2023). We show that the optimal policy faces a tradeoff between correcting the static markup distortions and correcting dynamic misallocation.

Table 7: Welfare Changes under Alternative Policies

	No Marketing	**No Size Incentive in Marketing**	**Production Subsidy**	**Efficient**
Chg.Welfare,	-7.31	3.80	10.70	41.42
due to Productivity,	-17.32	-2.32	2.01	10.21
due to Markup,	7.81	2.59	12.11	26.40
due to Marketing Cost	4.70	3.52	-3.42	-4.81

Is Marketing Welfare Enhancing? The first counterfactual welfare analysis we consider is a marketing ban. More precisely, we compare the calibrated baseline equilibrium to another equilibrium with no marketing (implemented by introducing a high tax on marketing). We then re-simulate the new equilibrium and compare the aggregate welfare under the two stationary distributions. The result is reported in the first column of Table 7.

The direct impact of a ban on marketing is the saving on marketing costs, which amounts to 4.70% of the baseline consumption equivalence. As firms reduce their marketing investments, two effects show up dynamically. Firms are no longer able to grow in size through marketing, which leads to a decline in the size of large firms and reduced markups. The overall welfare from the reduction in markup is 7.81% increase in the baseline consumption equivalence. However, there is also no longer a complementary expansion of customer capital to productivity, leading to a drop in aggregate productivity, which leads to a welfare loss of 17.3% baseline. In net, a marketing ban is welfare-reducing. In summary, marketing increases markup but also productivity in our calibrated model, with the productivity effect dominating.

The overall effect masks the distortions due to the size incentives. To isolate this welfare impact, we shut down the size incentives and recomputed the welfare incidences. Eliminating size incentives saves marketing costs more than the reduction in aggregate productivity. Thus, the marketing investment due to size incentives is welfare-reducing. Meanwhile, removing size incentives also reduces markup. In total, this leads to an improvement of 3.8% of baseline consumption.

The Planner's Solution. Before turning to the production subsidy, we discuss the nature of the planner's solution in this environment. The planner makes both static decisions and dynamic decisions to maximize the representative household's discounted utility. Statically, she chooses how much firms produce given their productivity and their customer capital. Dynamically, she chooses how much each firm should invest in customer capital. The planner aims to sort customer capital to the frontier firm and solve the static inefficiency from markups.

In the static allocation of production, the social planner chooses production (and thus consumption) given the distribution of customer capital. The optimal static labor allocation is standard. Compared to the equilibrium, there are no markup distortions in the planner's allocation. Two implications follow; the planner sets the production labor to 1, and the aggregation consumption equals the aggregate productivity $\mathbf{C} = \mathbf{A}$, where $\mathbf{A}$ is defined in Lemma 1.

We now turn to the planner's dynamic decision, which is where the role of customer capital leads to novel insights. The dynamic decision can be made group by group. This comes from (i) the real consumption from different groups being aggregated via a Cobb-Douglas aggregator and (ii) the disutility of labor from different groups being linearly additive. More precisely, we can write the optimal utility for the representative household under the planner's solution as

$$\mathbf{V}^* = \int_0^1 V_k dk, \tag{16}$$

where W_k is the expected utility from product group k:

$$V_k = \max_{\eta_{it}(\alpha_t, \mathbf{b}_t)} \mathbb{E}_0 \int_0^\infty e^{-\rho t} \left[\log A_{kt}(\alpha_t, \mathbf{b}_t) - 1 - \tau \sum_i^I \frac{\eta_i^2}{2}(\alpha_t, \mathbf{b}_t) \right] dt \tag{17}$$

Comparing the planner's problem to the firm problem in the decentralized equilibrium, we notice two types of wedges. First, the firms internalize only the profits from the static pricing equilibrium, while the planner factors in the full surplus to the representative household. Conditional on the state $(\alpha_t, \mathbf{b}_t)$, the oligopolistic equilibrium leads to distortions due to variable markups and due to the level of markup. This is the distortion well studied in the literature, such as Edmond et al. (2023). The second kind of wedge is novel in our setting when we understand firms' size through the lens of two separate dimensions. Firms become big either due to customer capital or productivity. The split of these two margins is irrelevant to the firms in equilibrium, as the profits only depend on the sum of productivity and customer capital. The social surplus, however, does differ. As an illustrative example, consider the case where all firms have the same productivity. From the planner's perspective, any configuration of demand leads to the same social surplus, as the demand heterogeneity is purely reallocative. For firms, becoming bigger in terms of demand does lead to higher profits.

The optimal policy that corrects both types of wedges thus involves correcting markups and making firms internalize the matching between productivity and demand. We now show, with alternative policy instruments, that ignoring the matching between demand and productivity can lead to further distortions. That is, a policy that ignores the split between demand and supply can actually further exacerbate the very distortions it aims to correct.

To fully characterize the planner's solution, we use a similar linearization strategy as in the equilibrium. First, we linearize the planner's period return function $\log A(\alpha, \mathbf{b})$

around the equal-demand state, $(\alpha, \mathbf{0})$.

$$v^*(\alpha, \mathbf{b}) = \frac{1}{\sigma - 1} \sum_i \left(\bar{s}_i(\alpha) - \frac{1}{I} \right) b_i + \frac{1}{\sigma - 1} \log \frac{\sum_i \exp(a_i)}{I}, \tag{18}$$

where $\bar{s}_i^*(\alpha) = \frac{\exp(a_i)}{\sum_j \exp(a_j)}$. In the planner's payoff as in Equation (18), the planner's equal-demand payoff is the simple average of productivity across all firms. The endogenous component of the planner's payoff takes a weighted sum of firm-level customer capitals, where the marginal benefit of increasing the customer capital of firm i reflects the difference between the planner's input share $\bar{s}_i^*$ and the equal input share $\frac{1}{I}$. Thus, the marginal benefit of increasing the customer capital of a productive firm (with $\bar{s}_i^* > \frac{1}{I}$) while the benefit of increasing the customer capital of an unproductive firm (with $\bar{s}_i^* < \frac{1}{I}$) is negative.

We now compare the planner's value to the equilibrium profits of firms. Since the dynamic costs are identical for the planner and for firms in the equilibrium, the difference in the dynamic decisions must be rooted in the difference in static payoffs. First, the planner does not have the markup incentives. More precisely, the linearity in the planner's static payoff means that her optimal marketing strategy only depends on the productivity of firms, not the accumulated customer capital. Second, the planner internalizes the business-stealing externality among firms. In firms' profits, an increase in their own customer capital always increases profits. Thus, the marketing investment of the firms in the equilibrium can never be zero. The planner considers the difference between the productivity gain from demand reallocation and the business-stealing externality. Since the net gain $\bar{s}_i^*(\mathbf{a}) - \frac{1}{I}$ can fall below zero, the planner will set the marketing investments for the least productive firms to be zero. This is a gap between planner and equilibrium with or without variable markups; last, there is a standard lack of appropriability problem. We compare the planner's payoff to the constant-markup firm profit to see this. As $1/\sigma < \frac{1}{\sigma - 1}$, the firms always internalize fewer gains from reallocating customer capital.

We report the welfare gains from implementing such a planner solution in Table 7. First, the planner is able to reduce all markup distortions, leading to a welfare gain of 26.4%. This number is in line with the results from Edmond et al. (2023), where our substitution elasticity corresponds to their high-markup scenario. There are additional welfare gains from the planner's solution. By factoring in the business-stealing externality, the planner is able to increase the aggregate productivity and reduce the marketing cost at the same time, leading to an additional gain of 15%. This dynamic gain is comparable to the static gain.

Production Subsidy. To highlight the tradeoff the policymakers face between resolving static markups and dynamic allocation, we consider a static subsidy in production that aims to eliminate markup distortions.

More precisely, we suppose there is a budget-balanced subsidy of $\tau > 0$ for production. With such a subsidy, the optimal production and pricing choice of the firms becomes,

$$\max_q (1+\tau) p_i q_i - q e^{a_i/(1-\sigma)},$$

s.t.

$$q_i = \frac{e^{b_i/\sigma} p_i^{-\sigma}}{\sum_j q_j p_j}.$$

Since the subsidy is uniform, it does not change the relative markups among firms but reduces the overall level of markups by a ratio $\frac{1}{1+\tau}$. This policy leaves the market shares conditional on $(\alpha, \mathbf{b})$ unchanged as in the baseline equilibrium. Dynamically, it changes firms' marketing incentives. More precisely, the profit of firms now become

$$\pi^S(\alpha, \mathbf{b}) = (1+\tau)\pi(\alpha, \mathbf{b}).$$

Thus, the production subsidy impacts welfare through two channels. Statically, it reduces welfare costs due to markups. Dynamically, it acts like a subsidy to firm size, making marketing investments more attractive. Theoretically, the dynamic effect can be both positive and negative, depending on whether the size incentive is welfare-enhancing.

Table 7 reports the welfare impact of 10% subsidy in the third column. Overall, this production subsidy brings 2% gain in consumption equivalence. Underneath this welfare gain, the static welfare gain due to markup is unambiguously positive, consistent with the literature. There is indeed a negative dynamic effect: the economy has a higher cost of marketing than the gain in productivity. On net, this induces a 1.4% loss due to the endogenous marketing response to the subsidy, indicating how these policies may backfire. These lessons are important to keep in mind for industrial policies that ineract on firm size.[5]

6 Conclusion

What is the relationship between market concentration and aggregate productivity? This paper argues that understanding the role of demand is essential for answering this question. We find that demand or customer capital is a central driver of firm market share and has a significant aggregate impact. We build a dynamic model to study the formation of aggregate productivity where granular firms invest in expanding demand and have both demand and productivity as state variables. We find that the dynamic interaction of these forces can significantly impact the transmission of productivity shocks from the firm level to the aggregate. On average, higher concentration via demand enhances productivity through positive sorting, but this is not always the case, and standard policies can backfire. For instance, standard policies for managing size-based distortions can backfire by encouraging the accumulation of customer capital at firms without a corresponding increase in aggregate productivity.

Our results highlight the importance of understanding the sources of firm performance beyond a single-dimensional productivity measure. Whether market power comes from productivity or accumulated customer capital leads to different conclusions regarding efficiency and various policies. We discuss some possible extensions of our framework here. First, we believe our two-dimensional case on customer capital and productivity can be extended to consider other forces driving firm size that interact with productivity, such as worker prestige, political connections, or location. Second, changes in customer capital may endogenously feed back to productivity through firm-level investments in technology. This could create long-run growth effects beyond the allocative efficiency dynamics discussed here. Third, the nature of firm entry and firm or brand acquisitions in markets may significantly interact with the forces discussed here. We believe these threads are fruitful extensions for further research.

[5] We discuss other potential policies of interest in Appendix A.5.

References

Acemoglu, D., U. Akcigit, H. Alp, N. Bloom, and W. Kerr: 2018, 'Innovation, Reallocation, and Growth'. *American Economic Review* **108**(11), 3450–3491.

Afrouzi, H., A. Drenik, and R. Kim: 2023, 'Concentration, Market Power, and Misallocation: The Role of Endogenous Customer Acquisition'. Technical report, National Bureau of Economic Research.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt: 2005, 'Competition and Innovation: An Inverted-U Relationship'. *Quarterly Journal of Economics* **120**(2), 701–728.

Aghion, P., C. Harris, P. Howitt, and J. Vickers: 2001, 'Competition, Imitation and Growth with Step-by-Step Innovation'. *Review of Economic Studies* **68**(3), 467–492.

Aghion, P. and P. Howitt: 1992, 'A Model of Growth through Creative Destruction'. *Econometrica* **60**(2), 323–351.

Akcigit, U. and S. T. Ates: 2021, 'Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory'. *American Economic Journal: Macroeconomics* **13**(1), 257–98.

Akcigit, U. and S. T. Ates: 2023, 'What happened to US business dynamism?'. *Journal of Political Economy* **131**(8), 2059–2124.

Akcigit, U. and W. R. Kerr: 2018, 'Growth through Heterogeneous Innovations'. *Journal of Political Economy* **126**(4), 1374–1443.

Amiti, M., O. Itskhoki, and J. Konings: 2019, 'International Shocks, Variable Markups, and Domestic Prices'. *The Review of Economic Studies* **86**(6), 2356–2402.

Arellano, M. and S. Bond: 1991, 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations'. *Review of Economic Studies* **58**(2), 277–297.

Argente, D., D. Fitzgerald, S. Moreira, and A. Priolo: 2021, 'How Do Firms Build Market Share?'. Technical report. SSRN Working Paper.

Argente, D., M. Lee, and S. Moreira: 2018, 'Innovation and Product Reallocation in the Great Recession'. *Journal of Monetary Economics* **93**, 1–20. Carnegie-Rochester-NYU Conference on Public Policy held at the Stern School of Business at New York University.

Argente, D., M. Lee, and S. Moreira: 2020, 'The Life Cycle of Products: Evidence and Implications'. Technical report.

Atkeson, A. and A. Burstein: 2008, 'Pricing-to-Market, Trade Costs, and International Relative Prices'. *American Economic Review* **98**(5), 1998–2031.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen: 2020, 'The Fall of the Labor Share and the Rise of Superstar Firms*'. *The Quarterly Journal of Economics* **135**(2), 645–709.

Bain, J. S.: 1956, *Barriers to New Competition: Their Character and Consequences in Manufacturing Industries*. Harvard University Press.

Baqaee, D. R. and E. Farhi: 2019, 'Productivity and Misallocation in General Equilibrium*'. *The Quarterly Journal of Economics* **135**(1), 105–163.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta: 2013, 'Cross-Country Differences in Productivity: The Role of Allocation and Selection'. *American Economic Review* **103**(1), 305–34.

Berger, D. W., K. F. Herkenhoff, and S. Mongey: 2019, 'Labor Market Power'. Working Paper 25719, National Bureau of Economic Research.

Bhandari, A. and E. R. McGrattan: 2020, 'Sweat Equity in U.S. Private Business*'. *Quarterly Journal of Economics* **136**(2), 727–781.

Boar, C. and V. Midrigan: 2019, 'Markups and inequality'. Technical report, NBER Working Paper.

Bornstein, G. and A. Peter: 2022, 'Nonlinear Pricing and Misallocation'. Technical report, Working Paper.

Bronnenberg, B. J., S. K. Dhar, and J.-P. Dubé: 2009, 'Brand History, Geography, and the Persistence of Brand Shares'. *Journal of Political Economy* **117**(1), 87–115.

Bronnenberg, B. J., J.-P. Dubé, and M. Gentzkow: 2012, 'The Evolution of Brand Preferences: Evidence from Consumer Migration'. *American Economic Review* **102**(6), 2472–2508.

Carvalho, V. M. and B. Grassi: 2019, 'Large Firm Dynamics and the Business Cycle'. *American Economic Review* **109**(4), 1375–1425.

Cavenaile, L., M. A. Celik, J. Perla, and P. Roldan-Blanco: 2023, 'A Theory of Dynamic Product Awareness and Targeted Advertising'. *Working Paper*.

Cavenaile, L., M. A. Celik, P. Roldan-Blanco, and X. Tian: 2025, 'Style over substance? advertising, innovation, and endogenous market structure'. *Journal of Monetary Economics* **149**, 103683.

Cavenaile, L. and P. Roldan-Blanco: 2021, 'Advertising, Innovation, and Economic Growth'. *American Economic Journal: Macroeconomics* **13**(3), 251–303.

Covarrubias, M., G. Gutiérrez, and T. Philippon: 2019, 'From Good to Bad Concentration? U.S. Industries over the past 30 years'. Working Paper 25983, National Bureau of Economic Research.

Crouzet, N. and J. C. Eberly: 2019, 'Understanding weak capital investment: The role of market concentration and intangibles'. Technical report, National Bureau of Economic Research.

David, J. M. and V. Venkateswaran: 2019, 'The Sources of Capital Misallocation'. *American Economic Review* **109**(7), 2531–67.

De Loecker, J. and J. Eeckhout: 2018, 'Global Market Power'. National Bureau of Economic Research Working Paper 24768.

De Loecker, J., J. Eeckhout, and G. Unger: 2020, 'The Rise of Market Power and the Macroeconomic Implications*'. *Quarterly Journal of Economics* **135**(2), 561–644.

De Ridder, M.: 2024, 'Market power and innovation in the intangible economy'. *American Economic Review* **114**(1), 199–251.

Dinlersoz, E. M. and M. Yorukoglu: 2012, 'Information and Industry Dynamics'. *American Economic Review* **102**(2), 884–913.

Edmond, C., V. Midrigan, and D. Y. Xu: 2023, 'How Costly Are Markups?'. *Journal of Political Economy* **131**(7), 1619–1675.

Eggertsson, G. B., J. A. Robbins, and E. G. Wold: 2018, 'Kaldor and Piketty's Facts: The Rise of Monopoly Power in the United States'. National Bureau of Economic Research Working Paper 24287.

Einav, L., P. Klenow, J. D. Levin, and M.-G. Ravid: 2021, 'Customers and Retail Growth'. Working Paper.

Eslava, M., J. Haltiwanger, and N. Urdaneta: 2024, 'The size and life-cycle growth of plants: The role of productivity, demand, and wedges'. *Review of Economic Studies* **91**(1), 259–300.

Foster, L., J. Haltiwanger, and C. Syverson: 2008, 'Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?'. *American Economic Review* **98**(1), 394–425.

Foster, L., J. Haltiwanger, and C. Syverson: 2016, 'The Slow Growth of New Plants: Learning about Demand?'. *Economica* **83**(329), 91–129.

Gabaix, X.: 2011, 'The granular origins of aggregate fluctuations'. *Econometrica* **79**(3), 733–772.

Gourio, F., T. Messer, and M. Siemer: 2014, 'What Is the Economic Impact of the Slowdown in New Business Formation?'. *Chicago Fed Letter* (Sep).

Gourio, F. and L. Rudanko: 2014, 'Customer Capital'. *The Review of Economic Studies* **81**(3), 1102–1136.

Greenwood, J., Y. Ma, and M. Yorukoglu: 2025, '"You Will:" A Macroeconomic Analysis of Digital Advertising'. *Review of Economic Studies* **92**(3), 1837–1881.

Gutiérrez, G. and T. Philippon: 2017, 'Declining Competition and Investment in the U.S.'. National Bureau of Economic Research Working Paper 23583.

Hall, R. E.: 2018, 'New Evidence on Market Power, Profit, Concentration, and the Role of Mega-Firms in the U.S. Economy'. Standofrd University, mimeo.

Haltiwanger, J., S. Scarpetta, and H. Schweiger: 2014, 'Cross Country Differences in Job Reallocation: The Role of Industry, Firm Size and Regulations'. *Labour Economics* **26**, 11–25.

Haskel, J. and S. Westlake: 2017, *Capitalism without capital: the rise of the intangible economy*. Princeton University Press.

He, B., L. I. Mostrom, and A. Sufi: 2024, 'Investing in Customer Capital'. Technical report, National Bureau of Economic Research.

Hopenhayn, H. A.: 1992, 'Entry, Exit, and Firm Dynamics in Long Run Equilibrium'. *Econometrica* **60**(5), 1127–1150.

Hottman, C. J., S. J. Redding, and D. E. Weinstein: 2016, ' Quantifying the Sources of Firm Heterogeneity *'. *The Quarterly Journal of Economics* **131**(3), 1291–1364.

Hsieh, C.-T. and P. J. Klenow: 2009, 'Misallocation and Manufacturing TFP in China and India'. *Quarterly Journal of Economics* **124**(4), 1403–1448.

Ignaszak, M. and P. Sedlácek: 2022, 'Customer acquisition, business dynamism and aggregate growth'. Technical report, Working Paper.

Jaravel, X.: 2018, 'The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector*'. *The Quarterly Journal of Economics* **134**(2), 715–783.

Jordà, Ò.: 2005, 'Estimation and inference of impulse responses by local projections'. *American economic review* **95**(1), 161–182.

Jovanovic, B.: 1982, 'Selection and the Evolution of Industry'. *Econometrica* **50**(3), 649–70.

Kehrig, M. and N. Vincent: 2021, 'The micro-level anatomy of the labor share decline'. *The Quarterly Journal of Economics* **136**(2), 1031–1087.

Klette, T. J. and S. Kortum: 2004, 'Innovating Firms and Aggregate Innovation'. *Journal of Political Economy* **112**(5), 986–1018.

Liu, E., A. Mian, and A. Sufi: 2022, 'Low Interest Rates, Market Power, and Productivity Growth'. *Econometrica* **90**(1), 193–221.

Mongey, S.: 2021, 'Market structure and monetary non-neutrality'. Technical report, National Bureau of Economic Research.

Nord, L.: 2023, 'Shopping, Demand Composition, and Equilibrium Prices'. Available at SSRN: https://ssrn.com/abstract=4178271 or http://dx.doi.org/10.2139/ssrn.4178271.

Olmstead-Rumsey, J.: 2022, 'Market concentration and the productivity slowdown'. *Working Paper*.

Pearce, J. and L. Wu: 2022, 'Brand Reallocation, Concentration, and Growth'. *Available at SSRN 4279983*.

Peters, M.: 2020, 'Heterogeneous Markups, Growth, and Endogenous Misallocation'. *Econometrica* **88**(5), 2037–2073.

Restuccia, D. and R. Rogerson: 2017, 'The Causes and Costs of Misallocation'. *Journal of Economic Perspectives* **31**(3), 151–74.

Salgado, S., S. Ozkan, J. Hubmer, G. Hong, and M. Chan: 2024, 'Scalable vs. Productive Technologies'. *Productive Technologies*(July 11, 2024)*.

Schmalensee, R.: 1978, 'A Model of Advertising and Product Quality'. *Journal of Political Economy* **86**(3), 485–503.

Schmalensee, R.: 1982, 'Product Differentiation Advantages of Pioneering Brands'. *The American Economic Review* **72**(3), 349–365.

Shapiro, C.: 1983, 'Premiums for High Quality Products as Returns to Reputations*'. *The Quarterly Journal of Economics* **98**(4), 659–679.

Syverson, C.: 2011, 'What Determines Productivity?'. *Journal of Economic Literature* **49**(2), 326–65.

Syverson, C.: 2019, 'Macroeconomics and Market Power: Facts, Potential Explanations, and Open Questions'. *Brookings Institution Report*. Accessed January 21, 2019. https://www.brookings.edu/research/macroeconomics-and-market-power-facts-potential-explanations-and-open-questions/.

Wu, L.: 2024, 'Partially directed search in the labor market'. *Journal of Economic Theory* **220**, 105858.

Appendix

The Appendix contains three sections. Appendix A discusses the theoretical proofs and expands on the firm's dynamic problem. Appendix B discusses the estimation and robustness.

A Theoretical Appendix

A.1 Approximation of the profits

We start by approximating the profit of firms. More specifically, the market shares are solutions to the following system of equations, given $(\alpha, \mathbf{b})$:

$$s_i = \frac{\exp(a_i + b_i)(1-s_i)^{\sigma-1}}{\sum_j \exp(a_j + b_j)(1-s_j)^{\sigma-1}}.$$

Denoting this system of equations by $H(\mathbf{s}; \alpha, \mathbf{b})$, the shares are the solution to $H(\mathbf{s}; \alpha, \mathbf{b}) = 0$. The profit for firm i is

$$\pi(s_i(\alpha, \mathbf{b})) = \frac{1}{\sigma} s_i(\alpha, \mathbf{b}) + \frac{\sigma - 1}{\sigma} s_i(\alpha, \mathbf{b})^2.$$

We approximate this function around point $(\alpha, \mathbf{0})$ to a second order:

$$\begin{aligned}\pi(s_i(\alpha, \mathbf{b})) \approx & \pi(s_i(\alpha, \mathbf{0})) + \left(\frac{1}{\sigma} + 2\frac{\sigma-1}{\sigma}\bar{s}_i\right) \nabla_{\mathbf{b}} s_i(\alpha, \mathbf{0})^T \mathbf{b} \\ & + \frac{1}{2}\mathbf{b}^T \left[\left(\frac{1}{\sigma} + 2\frac{\sigma-1}{\sigma}\bar{s}_i\right) \nabla^2_{\mathbf{b}} s_i(\alpha, \mathbf{0}) + 2\frac{\sigma-1}{\sigma} \nabla_{\mathbf{b}} s_i(\alpha, \mathbf{0})^T \nabla_{\mathbf{b}} s_i(\alpha, \mathbf{0})^T \right] \mathbf{b}\end{aligned}$$

where we use $\nabla_{\mathbf{b}} s_i(\alpha, \mathbf{0})$ to denote the jacobian at point $(\alpha, \mathbf{0})$ of s_i with respect to $\mathbf{b}$, and $\nabla^2_{\mathbf{b}} s_i(\alpha, \mathbf{0})$ for the heissian.

Starting from the Jacobian, we obtain the result by totally differentiating the system H with respect to a perturbation db_i, holding everything else constant. This purterbation leads to a shift in market shares $\{s_i\}_{i=1,\dots,I}$ according to
(firm i)

$$\frac{ds_i}{\bar{s}_i} = db_i - (\sigma - 1)\frac{ds_i}{1-\bar{s}_i} - \sum_{\ell=1}^{I} \bar{s}_\ell \left[db_\ell - (\sigma - 1)\frac{ds_\ell}{1-\bar{s}_\ell} \right]$$

(firm $j \neq i$)

$$\frac{ds_j}{\bar{s}_j} = -(\sigma-1)\frac{ds_j}{1-\bar{s}_j} - \sum_{\ell=1}^{I} \bar{s}_\ell \left[db_\ell - (\sigma-1)\frac{ds_\ell}{1-\bar{s}_\ell} \right]$$

We can guess and verify that there is a closed form for the change in shares,
(firm i)

$$\frac{ds_i}{db_i} = \zeta_i \left(1 - \frac{\zeta_i}{\sum_\ell \zeta_\ell}\right)$$

(firm $j \neq i$)

$$\frac{ds_j}{db_i} = \zeta_j \left(-\frac{\zeta_i}{\sum_\ell \zeta_\ell}\right)$$

where we define $\zeta_i = \left(\frac{1}{\bar{s}_i} + (\sigma-1)\frac{1}{1-\bar{s}_i}\right)^{-1}$. With this closed form, we verify that

$$\sum_j \frac{ds_j}{db_i} = \zeta_i - \frac{\zeta_i \sum_j \zeta_j}{\sum_j \zeta_j} = 0.$$

Now we derive the heissian of the system by differentiating the first-order terms. There are three scenarios. First, we consider the second-order effect of change in customer capital on the focal firm

$$\frac{\partial^2 s_i}{(\partial b_i)^2} = \frac{\partial \zeta_i}{\partial b_i} - \frac{2\zeta_i}{\sum_\ell \zeta_\ell}\frac{\partial \zeta_i}{\partial b_i} + \frac{\zeta_i^2}{(\sum_\ell \zeta_\ell)^2}\sum_\ell \frac{\partial \zeta_\ell}{\partial b_i}$$

Second, we consider the cross effects from the other firm ($j \neq i$) to the focal firm (i)

$$\frac{\partial^2 s_i}{\partial b_i \partial b_j} = \frac{\partial \zeta_i}{\partial b_j} - \frac{2\zeta_i}{\sum_\ell \zeta_\ell}\frac{\partial \zeta_i}{\partial b_j} + \frac{\zeta_i^2}{(\sum_\ell \zeta_\ell)^2}\sum_\ell \frac{\partial \zeta_\ell}{\partial b_j}$$

We wan to verify that $\sum_j \frac{\partial^2 s_i}{\partial b_i \partial b_j} = 0$. To do so, we note from the definition of ζ_i:

$$-\frac{\partial \zeta_i}{\zeta_i^2 \partial b_i} = \left(-\frac{1}{s_i^2} + (\sigma-1)\frac{1}{(1-s_i)^2}\right)\frac{\partial s_i}{\partial b_i} = \left(-\frac{1}{s_i^2} + (\sigma-1)\frac{1}{(1-s_i)^2}\right)\zeta_i\left(1 - \frac{\zeta_i}{\sum_j \zeta_j}\right)$$

Thus:

$$\frac{\partial \zeta_i}{\partial b_i} = -v_i \zeta_i^3 \left(1 - \frac{\zeta_i}{\sum_j \zeta_j}\right),$$

where we define $v_i = \left(-\frac{1}{s_i^2} + (\sigma-1)\frac{1}{(1-s_i)^2}\right)$. Similarly,

$$-\frac{\partial \zeta_i}{\zeta_i^2 \partial b_j} = -\left(-\frac{1}{s_i^2} + (\sigma-1)\frac{1}{(1-s_i)^2}\right)\zeta_i\left(\frac{\zeta_j}{\sum_j \zeta_j}\right),$$

this implies

$$\frac{\partial \zeta_i}{\partial b_j} = \zeta_i^3 v_i \left(\frac{\zeta_j}{\sum_j \zeta_j}\right)$$

Now we verify the second-order effects add up to zero, when we focus on $\frac{\partial s_i}{\partial b_i}$.

$$\sum_j \frac{\partial^2 s_i}{\partial b_i \partial b_j} = \left(1 - \frac{2\zeta_i}{\sum_\ell \zeta_\ell}\right)\sum_j \frac{\partial \zeta_i}{\partial b_j} + \frac{\zeta_i^2}{(\sum_\ell \zeta_\ell)^2}\sum_j \sum_\ell \frac{\partial \zeta_\ell}{\partial b_j}$$

We first unpack the term $\sum_j \frac{\partial \zeta_i}{\partial b_j}$. Using the earlier results

$$\sum_j \frac{\partial \zeta_i}{\partial b_j} = \zeta_i^2 \nu_i \left(1 - \frac{\sum_j \zeta_j}{\sum_j \zeta_j}\right) = 0.$$

This statement holds for $\frac{\partial \zeta_\ell}{\partial b_j}$ as well. Thus $\sum_j \frac{\partial^2 s_i}{\partial b_i \partial b_j} = 0$.

A.2 Derivation of Equilibrium Dynamics

In this section, we derive the equilibrium conditions by solving the dynamic problem for firms. From the approximation step, we have already obtained the following quadratic formula for the profit of firm i

$$\pi(\alpha, \mathbf{b}) = \pi_0(\alpha) + P_i(\alpha)^T \mathbf{b} + \frac{1}{2}\mathbf{b}^T \tilde{Q}_i(\alpha)\mathbf{b}$$

We start by re-writing firm i's dynamic problem, given the competitors follow the strategy $\eta_j(\alpha, \mathbf{b}) = \gamma_j(\alpha) + \epsilon_j(\alpha)^T \mathbf{b}$, for $j \neq i$.

$$\max_{\eta_i(\alpha,\mathbf{b})} \mathbb{E}_0 \int_0^\infty e^{-rt}\left(\pi_0(\alpha) + P_i(\alpha)^T \mathbf{b} + \frac{1}{2}\mathbf{b}^T \tilde{Q}_i(\alpha)\mathbf{b} - \frac{\kappa}{2}\eta_i(\alpha, \mathbf{b})^2\right) dt, \tag{A1}$$

s.t.
(for i)

$$\dot{b}_i(\alpha) = \eta_i(x) - \rho b_i$$

(for $j \neq i$)

$$\dot{b}_j(\alpha) = \gamma_j(\alpha) + \sum_{\ell} \epsilon_{j\ell} b_\ell$$

We write out the HJB equation from this problem:

$$rV_i(\alpha, \mathbf{b}) = \max_{\eta_i(\alpha,\mathbf{b})} \hat{\pi}(\alpha, \mathbf{b}) - \kappa \frac{\eta_i(\alpha, \mathbf{b})^2}{2} + \sum_j \frac{\partial V_i}{\partial b_j}(\alpha, \mathbf{b}) \dot{b}_j(\alpha) + (\lambda + \delta) I \sum_{\alpha'} (V_i(\alpha', \mathbf{b}) - V_i(\alpha, \mathbf{b})) F_{\alpha,\alpha'}.$$

We now guess and verify that the value function follows the quadratic form. More precisely, we guess that,

$$V_i(\alpha, \mathbf{b}) = v_i(\alpha) + \kappa \left(\sum_j g_{ij}(\alpha) b_j + \frac{1}{2} \sum_{j,\ell} e_{ij\ell}(\alpha) b_j b_\ell \right),$$

$$\gamma_i = g_{ii}, \ \ \epsilon_{ij} = e_{iij}.$$

where g_{ij} and $e_{ij\ell}$ are the unknowns we need to pin down. Based on this guess, the first-order condition with respect to $\eta_i(\alpha, \mathbf{b})$ requires that

$$\eta_i(\alpha, \mathbf{b}) = \frac{\partial V_i}{\partial b_i}(\alpha, \mathbf{b}) = g_{ii}(\alpha) + \sum_j e_{iij} b_j.$$

With this optimal solution, we verify that the value function is indeed quadratic:

$$\begin{aligned}
&(r + (\lambda + \delta) I) \left[v_i(\alpha) + \kappa \left(\sum_j g_{ij}(\alpha) b_j + \frac{1}{2} \sum_{j,\ell} e_{ij\ell}(\alpha) b_j b_\ell \right) \right] \\
=& \pi_0(\alpha) + \sum_j p_{ij} b_j + \frac{1}{2} \sum_{j,\ell} q_{ij\ell} b_j b_\ell - \frac{\kappa}{2} \left(g_{ii}(\alpha) + \sum_j e_{iij} b_j \right)^2 \\
&+ \sum_j \left(g_{ij}(\alpha) + \sum_\ell e_{ij\ell} b_\ell \right) \left(\gamma_j(\alpha) + \sum_\ell e_{jj\ell} b_\ell - \rho b_j \right) \\
&+ (\lambda + \delta) I \sum_{\alpha'} \left[v_i(\alpha) + \kappa \left(\sum_j g_{ij}(\alpha) b_j + \frac{1}{2} \sum_{j,\ell} e_{ij\ell}(\alpha) b_j b_\ell \right) \right] F_{\alpha,\alpha'}.
\end{aligned}$$

This verifies the quadratic guess, and we end up with three equations for the matrices. We focus on the unknowns that matter for allocation. This means $v_i(\alpha)$ is not of interest since it does not affect allocation. For fixed i, we match the coefficients for term e_j:

$$(r+\rho+(\lambda+\delta)I)g_{ij} = \frac{p_{ij}}{\kappa} - g_{ii}e_{iij} + \sum_{\ell} g_{i\ell}e_{\ell\ell i} + \sum_{\ell} \gamma_\ell e_{i\ell i} + (\lambda+\delta)I\sum_{\alpha'} g_{ii}(\alpha')F_{\alpha,\alpha'}$$

Starting with $g_{iij}(\alpha)$, we match the terms involving b_ib_j on the two sides of the equation:

$$(r+\rho+(\lambda+\delta)I)e_{iij} = \frac{q_{iij}}{\kappa} - e_{iii}e_{iij} + \sum_m (e_{imi}e_{mmj} + e_{imj}e_{jji}) + (\lambda+\delta)I\sum_{\alpha'} e_{iij}(\alpha')F_{\alpha,\alpha'}$$

We only care about the element g_{ii}. By definition, it is equivalent to say $\gamma_i = g_{ii}$. Similarly, $\epsilon_{ij} = e_{iij}$.

A.3 Details of Two-firm Case

We discuss the details in the derivation of the results in the two-firm case. More specifically, we assume the productivities take value in $\{0, \lambda\}$. In this case, we can write the approximated market share as

$$\hat{s}_j(\mathbf{a},\mathbf{b}) = \bar{s}_j(\mathbf{a}) + \zeta_j(\mathbf{a})\bar{s}_{-j}(\mathbf{a})(b_j - b_{-j}),$$

where we used the fact $\bar{s}_j(\mathbf{a}) + \bar{s}_{-j}(\mathbf{a}) = 1$. Using this result, we can write the approximated profit as

$$\hat{\pi}_j(\mathbf{a},\mathbf{b}) = \bar{\pi}_j(\mathbf{a}) + \left(\frac{1}{\sigma} + \frac{\sigma-1}{\sigma}2\bar{s}_j(\mathbf{a})\right)\zeta_j(\mathbf{a})\bar{s}_{-j}(\mathbf{a})(b_j - b_{-j}) + \frac{\sigma-1}{\sigma}\zeta_j(\mathbf{a})^2\bar{s}_{-j}(\mathbf{a})^2(b_j - b_{-j})^2.$$

We guess that

$$V_i(\alpha,\mathbf{b}) = \nu_i(\alpha) + \kappa\gamma_i(\alpha)(b_i - b_{-i}) + \frac{\kappa}{2}\epsilon_i(\alpha)(b_i - b_{-i})^2.$$

With this guess, we take the partial derivatives,

$$\frac{\partial V_i}{\partial b_i}(\alpha,\mathbf{b}) = \kappa(\gamma_i(\alpha) + \epsilon_i(\alpha)(b_i - b_{-i})),$$

and

$$\frac{\partial V_i}{\partial b_{-i}}(\alpha, \mathbf{b}) = -\kappa(\gamma_i(\alpha) + \epsilon_i(\alpha)(b_i - b_{-i})).$$

Taking the first-order condition with respect to the marketing investment:

$$\eta_i(\alpha, \mathbf{b}) = \kappa(\gamma_i(\alpha) + \epsilon_i(\alpha)(b_i - b_{-i}).$$

$$\eta_{-i}(\alpha, \mathbf{b}) = \kappa(\gamma_{-i}(\alpha) + \epsilon_{-i}(\alpha)(b_{-i} - b_i).$$

The net value from marketing is

$$\eta_j(\mathbf{a}, \mathbf{b})\frac{\partial V_j}{\partial b_j}(\mathbf{a}, \mathbf{b}) - \frac{\gamma}{2}\eta_j(\mathbf{a}, \mathbf{b})^2 = \frac{1}{2\gamma}\left(v_j(\mathbf{a}) + \omega_j(\mathbf{a})(b_j - b_{-j})\right)^2$$

We plug these decisions to verify the guess into the HJB equation:

$$\begin{aligned}
& r\left(\nu_i(\alpha) + \kappa\gamma_i(\alpha)(b_i - b_{-i}) + \frac{\kappa}{2}\epsilon_i(\alpha)(b_i - b_{-i})^2\right) \\
=& \bar{\pi}_i(\alpha) + \left(\frac{1}{\sigma} + \frac{\sigma-1}{\sigma}2\bar{s}_i(\alpha)\right)\zeta_i(\alpha)\bar{s}_{-i}(\alpha)(b_i - b_{-i}) + \frac{\sigma-1}{\sigma}\zeta_i(\alpha)^2\bar{s}_{-i}(\alpha)^2(b_i - b_{-i})^2 \\
& + \frac{\kappa}{2}\left(\gamma_i(\alpha) + \epsilon_i(\alpha)(b_i - b_{-i})\right)^2 - \kappa\rho\left(\gamma_i(\alpha) + \epsilon_i(\alpha)(b_i - b_{-i})\right)(b_i - b_{-i}) \\
& - \kappa\left(\gamma_i(\alpha) + \epsilon_i(\alpha)(b_i - b_{-i})\right)\left(\gamma_{-i}(\alpha) + \epsilon_{-i}(\alpha)(b_{-i} - b_i)\right) + \mathcal{A}(V_j)
\end{aligned}$$

Matching terms, we find the guessed form solves the HJB equation, and we now have the following separate HJB equations for the coefficients:

$$\rho v_j(\mathbf{a}) = \bar{\pi}_j(\mathbf{a}) + \frac{1}{2\gamma}v_j(\mathbf{a})^2 - \frac{1}{\gamma}v_j(\mathbf{a})v_{-j}(\mathbf{a}) + \mathcal{A}v_j(\mathbf{a})$$

$$\rho v_j(\mathbf{a}) = \left(\frac{1}{\sigma} + \frac{\sigma-1}{\sigma}2\bar{s}_j(\mathbf{a})\right)\zeta_j(\mathbf{a})\bar{s}_{-j}(\mathbf{a}) + \frac{1}{\gamma}v_j(\mathbf{a})\omega_j(\mathbf{a}) - \frac{1}{\gamma}\left(\omega_j(\mathbf{a})v_{-j}(\mathbf{a}) - v_j(\mathbf{a})\omega_{-j}(\mathbf{a})\right) + \mathcal{A}v_j(\mathbf{a})$$

$$(r + 2\rho + \lambda)\epsilon_i(\alpha) = \frac{2}{\kappa}\frac{\sigma-1}{\sigma}\zeta_i(\alpha)^2\bar{s}_{-i}(\alpha)^2 + \epsilon_i(\alpha)^2 + 2\epsilon_i(\alpha)\epsilon_{-i}(\alpha) + \lambda\sum_{\alpha'}F_{\alpha,\alpha'}\epsilon_i(\alpha')$$

Identical Productivity. For this case, we set $\lambda = 0$. Starting with the higher-order term. The common response to the customer capital gap is the solution to:

$$\frac{2}{\kappa}\frac{\sigma-1}{\sigma}\zeta^2\bar{s}^2 + 3\epsilon^2 - (r + 2\rho)\epsilon = 0$$

$$\dot{b} = (\epsilon - \rho)b$$

In this case, knowing $\bar{s} = \frac{1}{2}$, we can further simplify this equation to

$$3\epsilon^2 - (r + 2\rho)\epsilon + \frac{\sigma - 1}{8\kappa\sigma^3} = 0$$

This equation has real solutions when

$$(r + 2\rho)^2 > \frac{3(\sigma - 1)}{2\kappa\sigma^3}$$

When this condition is satisfied, we have the solutions in the form of

$$\frac{(r + 2\rho) + \sqrt{(r + 2\rho)^2 - \frac{3(\sigma-1)}{2\kappa\sigma^3}}}{6}$$

and

$$\frac{(r + 2\rho) - \sqrt{(r + 2\rho)^2 - \frac{3(\sigma-1)}{2\kappa\sigma^3}}}{6}$$

Only the small one satisfies the transversality condition.

Heterogeneous Productivity with $\lambda = 0$. Denoting $q_i = \frac{2}{\kappa}\frac{\sigma-1}{\sigma}\zeta_i(\alpha)^2\bar{s}_{-i}(\alpha)^2$, we have a two-equation system:

$$(r + 2\rho)\epsilon_1 = q_1 + \epsilon_1^2 + 2\epsilon_1\epsilon_2$$

$$(r + 2\rho)\epsilon_2 = q_2 + \epsilon_2^2 + 2\epsilon_1\epsilon_2$$

Defining $\epsilon^s = \epsilon_1 + \epsilon_2$,

A.4 Rational Inattention and Customer Base

With multiple firms, we use a rational inattention framework to map the consumer's choice across firms' product bundles to the concept of customer capital. In this framework, adapted from Wu (2024), households face cognitive limitations and must optimally allocate their limited attention across competing options. Formally, each household h chooses a probability distribution q_{hi} of purchasing from each firm i by maximizing the following:

$$\max_{q_{hi}, c_{hi}} \sum_i q_{hi} \log c_{hi} - \psi \sum_i q_{hi} \log \frac{q_{hi}}{\bar{q}_i},$$

where c_{hi} represents consumption and

$$\bar{q}_i = \frac{\exp(b_i)}{\sum_j \exp(b_j)}$$

is the default attention allocation based on customer capital b_i. This optimization is subject to two key constraints: the limited attention constraint $\sum_i q_{hi} = 1$, reflecting that probabilities must sum to one, and the limited budget constraint $c_{hi} = \frac{1}{p_i}$, indicating that consumption is inversely related to price.

Solving the household's optimization problem yields an expression for the expenditure share

$$q_i = \frac{\exp(b_i) c_i^{1/\psi}}{\sum_k \exp(b_j) c_j^{1/\psi}} = \frac{\exp(b_i) p_i^{-1/\psi}}{\sum_k \exp(b_j) p_j^{-1/\psi}},$$

where ψ represents the cost of households directing their choice towards the more productive firms. This formulation generates expected utility $\log\left(\frac{\sum_i e^{b_i} c_i^{1/\psi}}{\sum_i e^{b_i}}\right)^{\psi}$, which is equivalent to the baseline CES utility when $\psi = \frac{\sigma}{\sigma-1}$, yielding

$$\log\left(\frac{\sum_i e^{b_i} c_i^{\frac{\sigma-1}{\sigma}}}{\sum_i e^{b_i}}\right)^{\frac{\sigma}{\sigma-1}}.$$

The model further relates consumption to firm productivity through $c_{it} = \exp\left(\frac{a_{it}}{\sigma-1}\right) l_{it}$, where a_{it} represents firm-specific productivity and l_{it} is labor input.

The firm's customer capital b_i can be interpreted as the result of marketing and brand-building investments that influence the default attention allocation $\bar{q}_i$. By investing in customer capital, firms can shift the default probability distribution in their favor, effectively reducing consumers' cognitive costs of choosing their products. This mechanism explains why firms engage in marketing even when it does not directly affect product quality or characteristics—it strategically influences the attention allocation process of boundedly rational consumers. Consequently, firms with larger customer capital can maintain higher market shares and potentially charge premium prices, highlighting the economic value of marketing and brand recognition beyond productivity and in line with a growing literature on the centrality of customers in market share.

A.5 Policies: Extension

We start by considering a size-based subsidy when the policymakers have access to both demand and productivity differences of firms, separately. The optimal policy that corrects the distortions due to markups removes the revenue incentives of firms and reimburses the firms with the full social surplus they created. In practice, this involves setting a gross transfer:

$$T(q) = \frac{\sigma}{\sigma-1}\frac{e^{b_i/\sigma} q^{\frac{\sigma-1}{\sigma}}}{\sum_j \exp(a_j+b_j)} - \frac{e^{b_i/\sigma} q_j^{\frac{\sigma-1}{\sigma}}}{\sum_{j\neq i}\exp(b_j+a_j) + e^{b_i/\sigma} q^{\frac{\sigma-1}{\sigma}}}.$$

Given that the other firms all price at the marginal cost, the post-transfer optimization of the focal firm j becomes:

$$\max_q \frac{\sigma}{\sigma-1}\frac{e^{b_i/\sigma} q^{\frac{\sigma-1}{\sigma}}}{\sum_j \exp(b_j+a_j)} - q e^{a_i/(1-\sigma)}.$$

The optimal choice of quantity is $q_i^* = \frac{e^{a_i+b_i\frac{\sigma}{\sigma-1}}}{\sum_j \exp(b_j+a_j)}$ and the price is the marginal cost $p_i^* = e^{a_i/(1-\sigma)}$. This implements the static optimal allocation. In the equilibrium with such a subsidy, every firm receives the share of surplus it created:

$$T_j^* = \frac{1}{\sigma-1}\frac{e^{a_i+b_i}}{\sum_k \exp(b_i+a_i)}.$$

We again consider a linearization of such a post-transfer payoff, around the equal-demand state, $(\alpha, \mathbf{0})$:

$$T_j^* \approx \frac{1}{\sigma-1}\bar{s}_i^*(\alpha)\left(b_i - \sum_j \bar{s}_j^*(\alpha) b_j\right) + \frac{1}{\sigma-1}\frac{\exp(a_i)}{\sum_j \exp(a_j)}.$$

B Estimation Appendix

This appendix details our two-step estimation procedure for recovering the structural parameters of our model. We first estimate the autoregressive parameters using Arellano-Bond dynamic panel methods, then recover variance components through GMM moment matching.

B.1 Model Structure

Our empirical model decomposes both productivity and customer capital into persistent and transitory components. For firm j at time t:

$$a_{jt} = a_{jt}^P + \zeta_{jt}^a + \alpha_j$$
$$b_{jt} = b_{jt}^P + \zeta_{jt}^b + \beta_j$$

where ζ_{jt}^a and ζ_{jt}^b are the transitory shocks, they are independent of the current and past persistent shocks; α_j and β_j are the fixed effects; ν_{jt}^a and ν_{jt}^b are the persistent shocks. In this note, we assume that we have already recovered ρ_{aa}, ρ_{ba}, ρ_{bb} from Arellano-Bond and we want to recover the variances $\left(\sigma_{aT}^2, \sigma_{bT}^2, \sigma_{aP}^2, \sigma_{aP}^2\right)$. We will recover these parameters from the auto-covariance function of the first difference. More precisely, we denote

$$\Sigma_k^x = Cov\left(x_{jt} - x_{jt-1}, x_{jt-k} - x_{jt-k-1}\right), \quad x = a, b$$

It is useful to write out the values in terms of past realizations. More precisely, for any $k \geq 1$

$$a_{jt}^P = \rho_{aa}^k a_{jt-k}^P + error$$

where the error is independent to a_{jt-k}^P. We can do the same with b:

$$b_{jt}^P = \rho_{bb}^k b_{jt-k}^P + \left(\sum_{l=0}^{k-1} \rho_{bb}^l \rho_{ba} \rho_{aa}^{k-l-1}\right) a_{jt-k}^P + error$$

For notation simplicity, let us define

$$R_k = \sum_{l=0}^{k-1} \rho_{bb}^l \rho_{ba} \rho_{aa}^{k-l-1}$$

Thus

$$b_{jt}^P = \rho_{bb}^k b_{jt-k}^P + R_k a_{jt-k}^P + error$$

Productivity To start, we impose stationarity and denote the stationary variance of a^P as $Var(a^P)$:

$$Var\left(a^P\right) = \rho_{aa}^2 Var\left(a^P\right) + \sigma_{aP}^2 \implies Var\left(a^P\right) = \frac{\sigma_{aP}^2}{1-\rho_{aa}^2}$$

We start with $k=0$:

$$\begin{aligned}
\Sigma_0^a = Var\left(a_{jt} - a_{jt-1}\right) &= Var\left(a_{jt}^P - a_{jt-1}^P + \zeta_{jt}^a - \zeta_{jt-1}^a\right) \\
&= Var\left(a_{jt}^P - a_{jt-1}^P\right) + Var\left(\zeta_{jt}^a - \zeta_{jt-1}^a\right) \\
&= Var\left(\rho_{aa} a_{jt-1}^P + \nu_{jt}^a - a_{jt-1}^P\right) + 2\sigma_{aT}^2 \\
&= (1-\rho_{aa})^2 Var\left(a^P\right) + \sigma_P^2 + 2\sigma_T^2
\end{aligned}$$

where the first equality uses the definition, the second equality uses the fact ζ^a are independent of the permanent components, the third equality uses the definition of the variance of the transitory variance, and the last equation uses fact ν_{jt}^a is independent of past values. We denote the stationary value of the persistent variance as $Var(a^P)$. Plugging the stationary variance, we have

$$\Sigma_0^a = \frac{2}{1+\rho_a}\sigma_{aP}^2 + 2\sigma_{aT}^2$$

Now we move on to $k=1$:

$$\begin{aligned}
\Sigma_0^a =& Cov\left(a_{jt}^P - a_{jt-1}^P + \zeta_{jt}^a - \zeta_{jt-1}^a, a_{jt-1}^P - a_{jt-2}^P + \zeta_{jt-1}^a - \zeta_{jt-2}^a\right) \\
=& Cov\left(a_{jt}^P - a_{jt-1}^P, a_{jt-1}^P - a_{jt-2}^P\right) + Cov\left(\zeta_{jt}^a - \zeta_{jt-1}^a, \zeta_{jt-1}^a - \zeta_{jt-2}^a\right) \\
=& Cov\left(a_{jt}^P - a_{jt-1}^P, a_{jt-1}^P - a_{jt-2}^P\right) - \sigma_{aT}^2 \\
=& Cov\left(a_{jt}^P, a_{jt-1}^P\right) - Var\left(a^P\right) - Cov\left(a_{jt}^P, a_{jt-2}^P\right) + Cov\left(a_{jt-1}^P, a_{jt-2}^P\right) - \sigma_{aT}^2
\end{aligned}$$

where the first equality uses the fact ζ^a is independent of the persistent values, the second equality writes out the variance, the third equality expands the terms. Now we inspect the terms one by one. Using the sequential form we derived:

We now unpack the terms

$$Cov\left(a_{jt}^P, a_{jt-1}^P\right) = Cov\left(a_{jt-1}^P, a_{jt-2}^P\right) = \rho_{aa} Var(a^P) = \rho_{aa} Var(a)$$

$$Cov\left(a_{jt}^P, a_{jt-2}^P\right) = \rho_{aa}^2 Var(a_{jt-2}^P) = \rho_{aa}^2 Var(a)$$

Plugging these values we have

$$\Sigma_1^a = -\frac{1-\rho_a}{1+\rho_a}\sigma_{aP}^2 - \sigma_{aT}^2$$

Now we move on the $k \geq 2$

$$\begin{aligned}
\Sigma_k^a =& Cov\left(a_{jt}^P - a_{jt-1}^P + \zeta_{jt}^a - \zeta_{jt-1}^a, a_{jt-k}^P - a_{jt-k-1}^P + \zeta_{jt-k}^a - \zeta_{jt-k-1}^a\right) \\
=& Cov\left(a_{jt}^P - a_{jt-1}^P, a_{jt-k}^P - a_{jt-k-1}^P\right) \\
=& \rho_{aa}^k Var\left(a^P\right) - \rho_{aa}^{k-1} Var\left(a^P\right) - \rho_{aa}^{k+1} Var\left(a^P\right) + \rho_{aa}^k Var\left(a^P\right) \\
=& -\rho_{aa}^{k-1}\frac{1-\rho_{aa}}{1+\rho_{aa}}\sigma_{aP}^2
\end{aligned}$$

This gives us the general formula

$$\Sigma_k^a = \begin{cases} \frac{2}{1+\rho_a}\sigma_{aP}^2 + 2\sigma_{aT}^2 & k=0 \\ -\frac{1-\rho_a}{1+\rho_a}\sigma_{aP}^2 - \sigma_{aT}^2 & k=1 \\ -\rho_{aa}^{k-1}\frac{1-\rho_{aa}}{1+\rho_{aa}}\sigma_{aP}^2 & k>1 \end{cases}$$

Demand Similarly, we start by imposing stationarity:

$$\left(1-\rho_{bb}^2\right) Var\left(b^P\right) = \rho_{ba}^2 Var\left(a^P\right) + \rho_{ab}\rho_{ab} Cov\left(a^P, b^P\right) + \sigma_{aP}^2$$

$$Cov(a, b^P) = \rho_{aa}\rho_{ba} Var(a^P) + \rho_{aa}\rho_{bb} Cov(a^P, b^P)$$

With $Var(a)$ known from earlier steps. These two equations pin down $Var(b)$ and $Cov(a,b)$. From here on, we treat the two values as known.

$k = 0$

$$
\begin{aligned}
\Sigma_0^b &= Var\left(b_{jt}^P - b_{jt-1}^P + \zeta_{jt}^b - \zeta_{jt-1}^b\right) \\
\text{(independence of } \zeta^b) &= Var\left(b_{jt}^P - b_{jt-1}^P\right) + Var\left(\zeta_{jt}^b - \zeta_{jt-1}^b\right) \\
\text{(expand terms)} &= Var\left(\rho_{ba} a_{jt-1}^P + (\rho_{bb} - 1) b_{jt-1}^P + \nu_{jt}^b\right) + 2\sigma_{bT}^2 \\
\text{(expand terms)} &= \rho_{ba}^2 Var(a^P) + (1-\rho_{bb})^2 Var\left(b^P\right) + \sigma_{bP}^2 - 2\rho_{ba}(1-\rho_{bb}) Cov(a^P, b^P) + 2\sigma_{aT}^2
\end{aligned}
$$

For higher order terms, it is convenient to find a sequential form for b_{jt}^P. Using the definition, for any $k \geq 1$, we can write:

where the error terms are independent of the values from $t-k$.

$k = 1$

$$
\begin{aligned}
\Sigma_1^b =& Cov\left(b_{jt}^P - b_{jt-1}^P + \zeta_{jt}^b - \zeta_{jt-1}^b, b_{jt-1}^P - b_{jt-2}^P + \zeta_{jt-1}^b - \zeta_{jt-2}^b\right) \\
\text{(independence of } \zeta^b) =& Cov\left(b_{jt}^P - b_{jt-1}^P, b_{jt-1}^P - b_{jt-2}^P\right) - \sigma_T^2 \\
\text{(expand terms)} =& Cov\left(b_{jt}^P, b_{jt-1}^P\right) - Var\left(b_{jt-1}^P\right) - Cov\left(b_{jt}^P, b_{jt-2}^P\right) + Cov\left(b_{jt-1}^P, b_{jt-2}^P\right) - \sigma_T^2
\end{aligned}
$$

Now we can use the sequential form.

$$
\begin{aligned}
Cov\left(b_{jt}^P, b_{jt-1}^P\right) = Cov\left(b_{jt-1}^P, b_{jt-2}^P\right) &= Cov\left(\rho_{bb} b_{jt-1}^P + R_1 a_{jt-1}^P + error, b_{jt-1}^P\right) \\
&= R_1 Cov\left(a^P, b^P\right) + \rho_{bb} Var\left(b^P\right)
\end{aligned}
$$

$$
\begin{aligned}
Cov\left(b_{jt}^P, b_{jt-2}^P\right) &= Cov\left(\rho_{bb}^2 b_{jt-2}^P + R_2 a_{jt-2}^P + error, b_{jt-2}^P\right) \\
&= R_2 Cov\left(a^P, b^P\right) + \rho_{bb}^2 Var\left(b^P\right)
\end{aligned}
$$

Plugging in

$$
\Sigma_1^b = -\left(1-\rho_{bb}\right)^2 Var\left(b^P\right) + \left(2R_1 - R_2\right) Cov\left(a^P, b^P\right) - \sigma_T^2
$$

$k > 1$

$$
\begin{aligned}
\Sigma_k^b =& Cov\left(b_{jt}^P - b_{jt-1}^P + \zeta_{jt}^b - \zeta_{jt-1}^b, b_{jt-k}^P - b_{jt-k-1}^P + \zeta_{jt-k}^b - \zeta_{jt-k-1}^b\right) \\
\text{(independence of } \zeta^b\text{)} =& Cov\left(b_{jt}^P - b_{jt-1}^P, b_{jt-k}^P - b_{jt-k-1}^P\right) \\
=& Cov\left(b_{jt}^P, b_{jt-k}^P\right) - Cov\left(b_{jt-1}^P, b_{jt-k}^P\right) - Cov\left(b_{jt}^P, b_{jt-k-1}^P\right) + Cov\left(b_{jt-1}^P, b_{jt-k-1}^P\right) \\
=& \left(2\rho_{bb}^k - \rho_{bb}^{k-1} - \rho_{bb}^{k+1}\right) Var\left(b^P\right) + \left(2R_k - R_{k-1} - R_{k+1}\right) Cov(a,b) \\
=& -\rho_{bb}^{k-1}\left(1-\rho_{bb}\right)^2 Var(b^P) + \left(2R_k - R_{k-1} - R_{k+1}\right) Cov(a,b)
\end{aligned}
$$

This gives us the general formula

$$
\Sigma_k^b = \begin{cases}
\rho_{ba}^2 Var(a) + (1-\rho_{bb})^2 Var(b) + \sigma_{bP}^2 - 2\rho_{ba}(1-\rho_{bb})Cov(a,b) + 2\sigma_{aT}^2 & k=0 \\
-\left(1-\rho_{bb}\right)^2 Var\left(b^P\right) + \left(2R_1 - R_2\right) Cov\left(a^P, b^P\right) - \sigma_T^2 & k=1 \\
-\rho_{bb}^{k-1}\left(1-\rho_{bb}\right)^2 Var(b^P) + \left(2R_k - R_{k-1} - R_{k+1}\right) Cov(a,b) & k>1
\end{cases}
$$

B.2 First Stage: Arellano-Bond Estimation

We first estimate the autoregressive parameters $(\rho_{aa}, \rho_{ba}, \rho_{bb})$ using the Arellano-Bond GMM estimator. The key insight of this approach is using lagged levels as instruments for first differences to address the correlation between fixed effects and regressors.

B.2.1 Productivity Persistence

For productivity, we estimate:

$$\Delta a_{jt} = \rho_{aa}\Delta a_{jt-1} + \Delta\epsilon_{jt}$$

where $\Delta\epsilon_{jt} = \Delta\nu_{jt}^a + \Delta\zeta_{jt}^a$. The moment conditions are:

$$E[a_{jt-s}\Delta\epsilon_{jt}] = 0 \quad \text{for } s \geq 2$$

B.2.2 Brand Capital Dynamics

For customer capital, we estimate:

$$\Delta b_{jt} = \rho_{ba}\Delta a_{jt-1} + \rho_{bb}\Delta b_{jt-1} + \Delta\eta_{jt}$$

where $\Delta\eta_{jt} = \Delta\nu_{jt}^b + \Delta\zeta_{jt}^b$. The moment conditions are:

$$E[x_{jt-s}\Delta\eta_{jt}] = 0 \quad \text{for } x \in \{a, b\}, s \geq 2$$

B.2.3 Implementation and Results

We implement the estimator using firms with at least 4 consecutive years of data. The baseline specification uses up to 4 lags as instruments and a two-step efficient GMM estimator with Windmeijer-corrected standard errors.

Our estimates are:

$$\hat{\rho}_{aa} = 0.73 \quad (0.04)$$
$$\hat{\rho}_{ba} = 0.24 \quad (0.05)$$
$$\hat{\rho}_{bb} = 1.02 \quad (0.01)$$

Standard errors in parentheses. Specification tests support the validity of our instruments:

- Hansen J-test fails to reject overidentifying restrictions (p = 0.42)
- AR(2) test finds no evidence of second-order serial correlation in residuals (p = 0.38)

The high estimate of ρ_{bb} suggests customer capital is highly persistent, while the positive ρ_{ba} indicates productivity improvements help build customer capital. These first-stage estimates are treated as known in the second stage of estimation. Before turning to the second stage, we discuss identification concerns.

Identification Concerns. Our estimation strategy faces several potential threats to identification that we consider here. We organize these challenges into four main categories and detail our approaches to addressing each concern.

A. Reverse Causality. The primary identification concern is potential feedback from customer capital to productivity. While our baseline specification assumes productivity evolves independently, firms might adjust their production processes in response to brand-related shocks. We address this by estimating an expanded system that allows for bidirectional effects:

$$\text{Baseline:} \quad a_{jt} = \rho_{aa} a_{jt-1} + \zeta_j + \nu_{ajt} \tag{A2}$$

$$\text{Alternative:} \quad a_{jt} = \rho_{aa} a_{jt-1} + \rho_{ab} b_{jt-1} + \zeta_j + \nu_{ajt} \tag{A3}$$

Through vector autoregression estimation and systematic Granger causality testing, we examine the timing and direction of relationships between productivity and customer capital innovations. The results, detailed in Section 4.2, suggest limited evidence of reverse causality affecting our main estimates.

*B. Unobserved Heterogeneity.*A second concern is that time-varying shocks might simultaneously affect both productivity and customer capital, violating our moment conditions. For example, quality improvements could drive both measures. We implement an instrumental variables approach that leverages industry-level variation:

$$\text{First Stage:} \quad \Delta a_{jt} = \pi_1 Z_{jt} + \eta_{jt} \tag{A4}$$

$$\text{Second Stage:} \quad \Delta b_{jt} = \rho_{ba} \Delta \hat{a} jt - 1 + \rho bb \Delta b_{jt-1} + \epsilon_{jt} \tag{A5}$$

where Z_{jt} includes industry-level shifters of productivity that are plausibly exogenous to firm-specific customer capital. We further examine heterogeneity across market structures to validate our identification strategy.

C. Validity of Lag Instruments. The Arellano-Bond approach relies crucially on the validity of lagged levels as instruments. We conduct extensive specification testing through Hansen tests of overidentifying restrictions and serial correlation tests in first-differenced errors. Additionally, we examine sensitivity to instrument set construction by varying lag depth and implementing collapsed instrument matrices. These robustness checks support the validity of our identification strategy.

D. Non-linear Dynamics. Our linear AR(1) specification may miss important non-linearities in the evolution of productivity and customer capital. We estimate an expanded specification allowing for state-dependent parameters:

$$\Delta b_{jt} = \rho_{ba}(\phi(s_{jt-1}))\Delta a_{jt-1} + \rho_{bb}(\phi(s_{jt-1}))\Delta b_{jt-1} + \Delta \nu_{bjt} \tag{A6}$$

where $\phi(s_{jt-1})$ is a flexible function of firm market share. We complement this analysis with threshold regression models and non-parametric tests for non-linear dependence. The results suggest our baseline linear specification captures the first-order dynamics while missing limited higher-order effects.

B.3 Second Stage: GMM Moment Matching

Given these first-stage estimates, we recover the variance parameters $(\sigma^2_{aT}, \sigma^2_{bT}, \sigma^2_{aP}, \sigma^2_{bP})$ by matching theoretical and empirical autocovariance functions.

B.3.1 Autocovariance Functions

We denote the autocovariance of first differences as:

$$\Sigma^x_k = Cov(x_{jt} - x_{jt-1}, x_{jt-k} - x_{jt-k-1}), \quad x = a, b$$

The theoretical autocovariance functions are:
For productivity $(k \geq 0)$:

$$\Sigma^a_k = \begin{cases} \frac{2}{1+\rho_a}\sigma^2_{aP} + 2\sigma^2_{aT} & k = 0 \\ -\frac{1-\rho_a}{1+\rho_a}\sigma^2_{aP} - \sigma^2_{aT} & k = 1 \\ -\rho^{k-1}_{aa}\frac{1-\rho_{aa}}{1+\rho_{aa}}\sigma^2_{aP} & k > 1 \end{cases}$$

For customer capital $(k \geq 0)$:

$$\Sigma^b_k = \begin{cases} \rho^2_{ba} Var(a) + (1-\rho_{bb})^2 Var(b) + \sigma^2_{bP} - 2\rho_{ba}(1-\rho_{bb})Cov(a,b) + 2\sigma^2_{bT} & k = 0 \\ -(1-\rho_{bb})^2 Var(b^P) + (2R_1 - R_2)Cov(a^P, b^P) - \sigma^2_{bT} & k = 1 \\ -\rho^{k-1}_{bb}(1-\rho_{bb})^2 Var(b^P) + (2R_k - R_{k-1} - R_{k+1})Cov(a,b) & k > 1 \end{cases}$$

where R_k captures cross-persistence effects:

$$R_k = \sum_{l=0}^{k-1} \rho^l_{bb}\rho_{ba}\rho^{k-l-1}_{aa}$$

B.3.2 Implementation Details

Key numerical considerations:

- Near unit root in customer capital ($\rho_{bb} \approx 1$) requires careful handling of variance terms
- Regularization in matrix inversions:

$$A = \begin{bmatrix} 1 - \rho^2_{bb} + 1e-6 & -\rho_{ba} \\ -\rho_{aa}\rho_{ba} & 1 - \rho_{aa}\rho_{bb} \end{bmatrix}$$

- Treatment of initial conditions

B.3.3 Results

Our second-stage estimates are:

$$\sigma_{aP} = 0.5534 \text{ (Productivity innovation std)}$$
$$\sigma_{aT} = 0.3348 \text{ (Transitory productivity shock std)}$$
$$\sigma_{bP} = 2.1619 \text{ (Customer capital innovation std)}$$
$$\sigma_{bT} = 0.0034 \text{ (Transitory brand shock std)}$$

The estimates reveal three main messages. Customer capital innovations have larger variance than productivity innovations. Productivity has significant transitory components while customer capital is primarily persistent. The model fits productivity autocovariance well but shows some deviation in customer capital autocovariance at longer lags.

B.4 Identification Discussion

The separate identification of variance components comes from a few forces. First, transitory shocks $(\sigma^2_{aT}, \sigma^2_{bT})$ primarily identified by $k = 1$ autocovariances. Second, persistent shock variances $(\sigma^2_{aP}, \sigma^2_{bP})$ identified by decay rates at higher lags. Third, cross-persistence ρ_{ba} helps identify relative contribution of productivity versus customer capital shocks.

B.5 Different Selection Criteria

For our main results, we use all firms and do not condition on a balanced panel. Qualitatively, we find very similar results when we

MORE ON AB:

To address the incidental parameters problem arising from firm fixed effects and the dynamic structure, we employ first differences:

$$\Delta a_{jt} = \rho_{aa} \Delta a_{jt-1} + \Delta \nu_{ajt} \quad \text{(A7)}$$
$$\Delta b_{jt} = \rho_{ba} \Delta a_{jt-1} + \rho_{bb} \Delta b_{jt-1} + \Delta \nu_{bjt} \quad \text{(A8)}$$

Table B1: Persistence Parameters by Firm Size and Time Period

Size	Year	ρ_{aa}	ρ_{ba}	ρ_{bb}
$1/10^2$	0	0.641	0.186	0.841
$1/10^2$	7	0.671	0.164	0.853
$1/10^2$	13	0.569	0.194	0.853
$1/10^3$	0	0.577	0.397	1.043
$1/10^3$	7	0.616	0.384	1.025
$1/10^3$	13	0.776	0.244	0.955
$1/10^4$	0	0.735	0.249	1.056
$1/10^4$	7	0.749	0.285	1.037
$1/10^4$	13	0.813	0.161	0.966
$1/10^5$	0	0.733	0.248	1.039
$1/10^5$	7	0.755	0.292	1.034
$1/10^5$	13	0.819	0.143	0.945
$1/10^6$	0	0.732	0.241	1.021
$1/10^6$	7	0.745	0.292	1.022
$1/10^6$	13	0.803	0.157	0.932

Notes: ρ_{aa} represents productivity persistence, ρ_{ba} represents the effect of lagged productivity on demand, and ρ_{bb} represents demand persistence. Size indicates minimum average firm size, and Year indicates minimum number of years a firm must be present in the sample (13 years is a balanced panel).

The first-differenced specification eliminates the firm fixed effects but introduces correlation between Δa_{jt-1} and $\Delta\nu_{ajt}$ through the shared ν_{ajt-1} term. Following Arellano and Bond (1991), we construct moment conditions using lagged levels as instruments:

$$E[a_{jt-s}\Delta\nu_{ajt}] = 0 \quad \text{for } s \geq 2 \tag{A9}$$

$$E[b_{jt-s}\Delta\nu_{bjt}] = 0 \quad \text{for } s \geq 2 \tag{A10}$$

These moment conditions exploit the assumption that productivity and customer capital levels from $t-2$ and earlier are uncorrelated with the differenced errors. We estimate the system using two-step GMM with optimal weighting matrix and windmeijer-corrected standard errors to account for potential finite-sample bias. The validity of our estimation approach relies on two key assumptions. First, sequential exogeneity, as follows,

$$E[\nu_{ajt}|a_{j1}, ..., a_{jt-1}, \zeta_j] = 0 \tag{A11}$$

$$E[\nu_{bjt}|b_{j1}, ..., b_{jt-1}, a_{j1}, ..., a_{jt-1}, \xi_j] = 0, \tag{A12}$$

and no serial correlation in the error terms:

$$E[\nu_{ajt}\nu_{ajt-s}] = 0 \quad \text{for } s \geq 1 \tag{A13}$$

$$E[\nu_{bjt}\nu_{bjt-s}] = 0 \quad \text{for } s \geq 1. \tag{A14}$$

We test these assumptions using the Arellano-Bond test for serial correlation in the first-differenced errors and the Hansen test of overidentifying restrictions.

B.6 Local Projections for Correlations

In this analysis, we examine the relationship between customer capital (b) and cost productivity (a) using local projections. Local projections provide a flexible approach to estimating impulse response functions without imposing the dynamic restrictions inherent in vector autoregressions.

We estimate the following regression for each horizon h:

$$b_{i,t+h} = \beta_0^h + \beta_1^h a_{i,t} + \beta_2^h b_{i,t-1} + \beta_3^h a_{i,t-1} + \beta_4^h b_{i,t-2} + \beta_5^h a_{i,t-2} + \varepsilon_{i,t}^h$$

where $b_{i,t}$ represents customer capital and $a_{i,t}$ represents cost productivity for firm i at time t. The regression is weighted by the firm's share of the product group (*pg_share*).

After estimating these local projections, we construct predicted values of customer capital based on the estimated coefficients:

$$\hat{b}_{i,t+h} = \begin{cases} \hat{\beta}_1^h a_{i,t} & \text{if } t = 0 \\ \hat{\beta}_1^h a_{i,t} + \hat{\beta}_2^h \hat{b}_{i,t+h-1} & \text{if } t > 0 \end{cases}$$

The correlation between the predicted customer capital ($\hat{b}_{i,t+1}$) and the realized future cost productivity ($a_{i,t+1}$) is 0.71, indicating a strong positive relationship between the predicted values of customer capital and future cost productivity. This framework allows us to purge the contemporaneous negatively correlated shocks between the two to focus on the dynamics of their relationship consistent with the model.